RNA (1999), *5*:318–329. Cambridge University Press. Printed in the USA.

Photocrosslinking of 4-thio uracil-containing RNAs supports a side-by-side arrangement of domains 5 and 6 of a group II intron

MIRCEA PODAR[1] and PHILIP S. PERLMAN

Department of Molecular Biology and Oncology, University of Texas Southwestern Medical Center, Dallas, Texas 75235-9148, USA

ABSTRACT

Previous studies suggested that domains 5 and 6 (D5 and D6) of group II introns act together in splicing and that the two helical structures probably do not interact by helix stacking. Here, we characterized the major Mg^{2+} ion- and salt-dependent, long-wave UV light-induced, intramolecular crosslinks formed in 4-thiouridine-containing D56 RNA from intron 5γ (al5γ) of the COXI gene of yeast mtDNA. Four major crosslinks were mapped and found to result from covalent bonds between nucleotides separating D5 from D6 [called J(56)] and residues of D6 near and including the branch nucleotide. These findings are extended by results of similar experiments using 4-thioU containing D56 RNAs from a mutant allele of al5γ and from the group IIA intron, al1. *Trans*-splicing experiments show that the crosslinked wild-type al5γ D56 RNAs are active for both splicing reactions, including some first-step branching. An RNA containing the 3-nt J(56) sequence and D6 of al5γ yields one main crosslink that is identical to the most minor of the crosslinks obtained with D56 RNA, but in this case in a cation-independent fashion. We conclude that the interaction between J(56) and D6 is influenced by charge repulsion between the D5 and D6 helix backbones and that high concentrations of cations allow the helices to approach closely under self-splicing conditions. The interaction between J(56) and D6 appears to be a significant factor establishing a side-by-side (i.e., not stacked) orientation of the helices of the two domains.

Keywords: 4-thio uracil; crosslinking; group II introns; RNA structure; self-splicing

INTRODUCTION

A great deal of research has been directed toward identifying and understanding the functional importance of secondary and tertiary structure elements responsible for the formation of the three dimensional structure of self-splicing group II intron RNAs (Michel & Ferat, 1995; Jacquier, 1996; Pyle, 1996). The secondary structures of group II introns consist of six highly conserved substructures called domains (D1 through D6). Most of the known tertiary interactions were proposed based on sequence covariation analysis. Subsequent studies have confirmed or assessed their importance, mostly using the self-splicing intron al5γ from yeast mtDNA. The secondary structure domains are separated by short joining sequences, some of which are involved in tertiary interactions (Michel & Ferat, 1995; Jacquier, 1996; Pyle, 1996). Experiments using direct structure-probing approaches have begun to reveal new elements of tertiary structure of group II intron RNA (e.g., Chanfreau & Jacquier, 1994, 1996; Podar et al., 1995, 1998; Jestin et al., 1997; Costa et al., 1998).

Domain 5 (D5) is an essential constituent of the catalytic core of group II introns that has been the subject of intense research over the last several years. It contains several nucleotides whose functional groups appear to influence catalysis directly (Chanfreau & Jacquier, 1994; Boulanger et al., 1995; Peebles et al., 1995; Abramovitz et al., 1996; Schmidt et al., 1996; Konforti et al., 1998). It has been known for some time that D5 interacts with other regions of the intron, and that those tertiary interactions are strong enough to allow D5 to function *in trans* (Jarrell et al., 1988). *Cis* deletions and *trans*-splicing experiments showed that D1 contains one or more components of the binding site for D5, and that binding to D1 is sufficient for D5 to function in the first step of splicing (Koch et al., 1992; Michels & Pyle, 1995). The ζ-ζ' tertiary interaction was shown to occur

Reprint requests to: Philip S. Perlman, Department of Molecular Biology and Oncology, University of Texas Southwestern Medical Center, Dallas, Texas 75235-9148, USA; e-mail: perlman@utsw.swmed.edu.

[1]Present address: Biology Department, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA.

between the GAAA loop of D5 and an internal loop in D1 (Costa & Michel, 1995). Additionally, nucleotides of the lower helix and the 2-nt bulge of D5 also appear involved in as yet undefined tertiary contacts in the ground-state active-site structure, possibly involving part of D3 (Jestin et al., 1997). Recent studies of long-wave UV light-induced crosslinks between 4-thioU containing D5 RNA and several unmodified substrate molecules showed that the lower helix of D5 is positioned near the J2,3 region of the intron (Podar et al., 1998).

D6 contains the attacking nucleophile for wild-type first-step reactions (van der Veen et al., 1986). Intron RNA deleted for D6 retains substantial activity for the first splicing reaction, but uses an alternative hydrolysis pathway (Koch et al., 1992). That deletion strongly inhibits both the extent of second-step reactions and the accuracy of 3′-splice-site selection. Introns with mutations of the branch nucleotide of D6 are blocked for first-step branching (Gaur et al., 1997) but remain highly reactive for first-step hydrolysis and for second-step splicing (Liu et al., 1997; Chu et al., 1998). In vivo, D6 point mutations permit enough of the first-step hydrolysis reaction to support growth (Podar et al., 1998). A tertiary interaction between the tetraloop capping D6 and a substructure in D2 has been proposed (Costa et al., 1997). However, that interaction appears to serve more of a modulatory role in the transition between the two splicing reactions than to provide energy for the ground-state interaction preceding the first step of splicing (Chanfreau & Jacquier, 1996).

D6 is not self-sufficient for its own positioning in the active site of the intron, and it functions poorly in *trans* splicing when separated from D5 (Dib-Hajj et al., 1993; Jarrell et al., 1988). Our finding that D6 functions in *trans*-splicing experiments most efficiently when provided in *cis* with D5 suggested that D5 and D6 interact somehow or that their orientation relative to one another is important for first-step splicing with transesterification. The very low level of transesterification obtained in *trans*-splicing experiments where the reaction is initiated by addition of D5 RNA to a substrate RNA containing both exons and all intron domains except for D5 indicates that any D5–D6 interaction must be very weak (Dib-Hajj et al., 1993).

D5 and D6 are separated by a short spacer sequence, termed J(56), which is 2 nt long in subgroup IIA introns and 3–4 nt long in most subgroup IIB introns (Michel et al., 1989). Mutational analysis of J(56) of aI5γ showed that the length of J(56) is critical both in vitro and in vivo (Boulanger et al., 1996). Its complete deletion (which may force the two domains to stack coaxially) completely blocks branching in vitro and relegates self-splicing to occur only by first-step hydrolysis followed by inaccurate second-step reactions. A 5-nt-long J(56) was compatible with some self-splicing with branching, but splicing in vivo was blocked by that mutation. Spontaneous revertants of the J(56)5 mutant had a shorter joiner, of 4 nt, that permitted some accurate splicing, but with a pronounced defect in the second splicing reaction. That study demonstrated the importance of the J(56) region for splicing of a group II intron, and suggested that D5 and D6 probably do not stack coaxially in the active intron structure. Despite those findings, the nature of the interaction between D5 and D6 and the structural role of J(56) remained uncharacterized.

Here we have used photocrosslinking of RNAs containing 4-thio uracil (s^4U) to probe the structure of the D56 region of aI5γ. Crosslinks were detected between J(56) and the bottom helix of D6 and the crosslinked RNAs were found to be functional, suggesting that they have trapped active configurations of the RNA. It appears that the J(56) sequence interacts directly with the D6 helix, perhaps forming a triple helix structure. This interaction may be a determining factor in the positioning of D5 and D6 with respect to each other. Crosslinking of an RNA lacking D5 showed that D5 contributes to this arrangement of the two domains. These findings argue strongly against there being coaxial stacking of D5 and D6, and favor a side-by-side orientation of the helices of the two domains. Such a mode of interaction is further supported by the finding of direct crosslinks between nucleotides of D5 and D6 using a D56 RNA from a group IIA intron.

RESULTS

Intramolecular crosslinks within D56 RNA from aI5γ

End-labeled D56 RNAs from aI5γ that contain approximately one s^4U residue per RNA were irradiated with a long-wave UV light source, under several buffer conditions. After irradiation, denaturing polyacrylamide gel electrophoresis, which resolves various crosslinked species based on the locations of the joined nucleotides, was used to separate the RNAs. A heterogeneous array of crosslinked RNAs with a few discrete species is formed in the absence of salt or Mg^{2+} ion. Several of the discrete bands were analyzed (Fig. 1, lane 2) and found to result from short-range crosslinks within D6 (data not shown, but see Fig. 7 for a specific example of such intra-D6 crosslinks in another RNA). Because those crosslinked species are detected under conditions that do not support self-splicing (and are only minor products under self-splicing conditions), they were not analyzed further.

Additional crosslinks are formed in the presence of $MgCl_2$, with or without additional salts of monovalent cations (Fig. 1, lanes 3–8), with at least four distinct bands being readily identifiable (termed X1 through X4). The set of the crosslinked species was evident after as little as 30 s of irradiation, accumulated with approximately linear kinetics, and reached an apparent maxi-

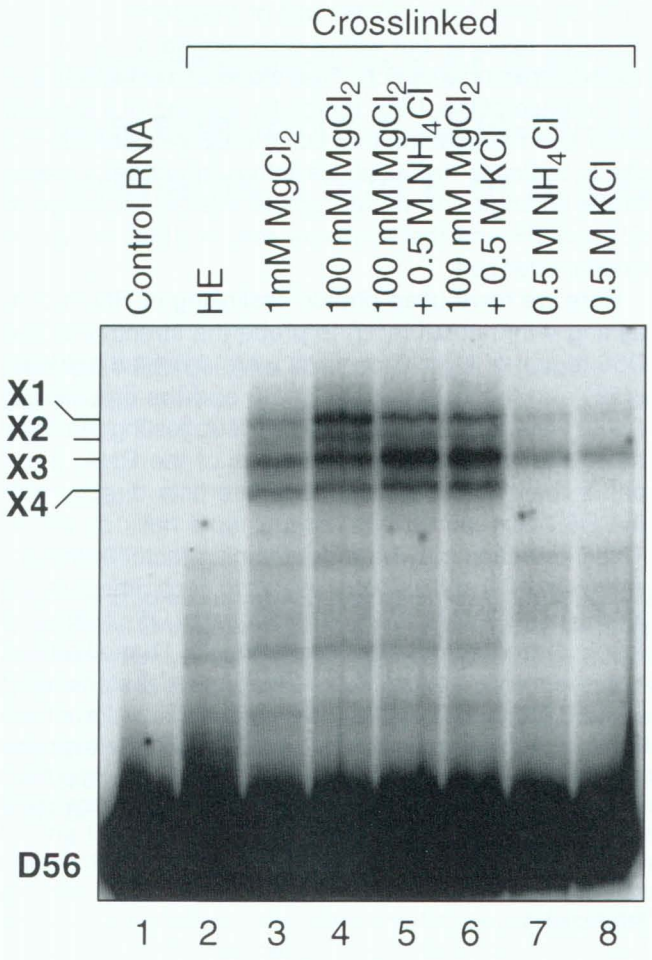

FIGURE 1. UV crosslinking of al5γ D56 RNA randomly substituted with 4-thio uracil. The 5′ end-labeled D56 RNA was irradiated with 360 nm UV light for 20 min at room temperature in 10 mM HEPES, 1 mM EDTA (HE) or in HE buffer containing Mg^{2+} plus or minus another salt, as indicated above the lanes. Control RNA is nonirradiated. The main crosslinked species formed have a reduced mobility on this 10% polyacrylamide gel and are identified as X1–X4.

mum (~10% of RNA input, combined) at about 20 min of irradiation. Similarly high yields of these species were obtained in the presence of 100 mM $MgCl_2$ with or without 0.5 M NH_4Cl or KCl (Fig. 1, lanes 4–6, respectively). Figure 1 shows that the most abundant crosslinked product was X3 and the least abundant product was X2. The samples containing salts of monovalent cations but with no $MgCl_2$ added (Fig. 1, lanes 7 and 8) supported the formation of lower levels of X1 and X3, but formation of X2 and X4 was $MgCl_2$-dependent.

The crosslinked products X1–X4 (from both 5′ and 3′ end-labeled preparations) were purified and subjected to limited alkaline hydrolysis followed by gel electrophoresis alongside of sequence markers obtained by enzymatic sequencing of control D56 RNA (Fig. 2A,B). The nucleotides connected by these crosslinks were identified based on the position of the gap in the alkaline hydrolysis ladder. All of the crosslinks present in these RNAs were between nucleotides of the joiner sequence, J(56), and the region of D6 including the branchpoint A (Fig. 2C). Nucleotide U850 of J(56) forms three different crosslinks with nt 879–881 (5′-UAU) of D6 (in X1, X3, and X4, respectively). Because X3 is the most abundant crosslink under splicing conditions, U850 appears to be situated near the branch nucleotide (A880) much of the time. Here the D56 RNA appears to have a somewhat dynamic structure so that nt U850 of J(56) crosslinks detectably to three adjacent nucleotides of D6. The least frequent crosslink, X2, between nt A851 and U879 of D6, appears to be mutually exclusive with the other three, again consistent with there being some fluidity to the structure of this RNA under these conditions. Both X1 and X4 contain a crosslink between two U residues and we do not know whether s^4U derivatives of both can initiate formation of these crosslinks.

The crosslinked D56 RNAs are functional in a *trans*-splicing assay

Next we determined whether the crosslinked species are functional in self-splicing assays that depend on the specific participation of both D5 and D6. Low concentrations of purified crosslinked D56 RNAs (3′ end labeled with 5′-^{32}P-pCp [*pCp]) were incubated with an excess (1 μM) of unlabeled E1D1-3 RNA for 20 min in self-splicing buffer containing 1 M NH_4Cl, and the products were analyzed on a polyacrylamide gel (Fig. 3). The control reaction (C) employed end-labeled D56 RNA that had not been irradiated and it yielded two abundant products. The slowly migrating product is the Y-shaped branched reaction intermediate from which E1 has been released (D1-3xD56*pCp); this product was first characterized in group II intron reactions like these by Jarrell et al. (1988). The much faster migrating product is spliced exon RNA (E1*pCp) resulting from second-step reactions in which the *pCp served as the second exon. Previous research has shown that a second exon as short as one phosphate functions in assays like this (Mueller et al., 1991). The input-labeled D56 RNA is much smaller than E1*pCp and has run off the gel. These control reactions are diagramed in Figure 3 below the gel.

Reactions using crosslinked X1, X3, and X4 RNAs also yielded two products: some of the spliced exon product is evident in those reactions, as is a doublet of material migrating a bit slower than the branched intermediate of the control reaction. We conclude that the latter product is branched and crosslinked reaction intermediate and refer to that material as D1-3xD56(X)*pCp. It probably migrates somewhat more slowly than the product obtained in the control reaction because it contains an internal loop due to the crosslink. In our experience with gel-purified branched or crosslinked RNAs, some of the "circles" break during handling. On these gels, the broken material migrates

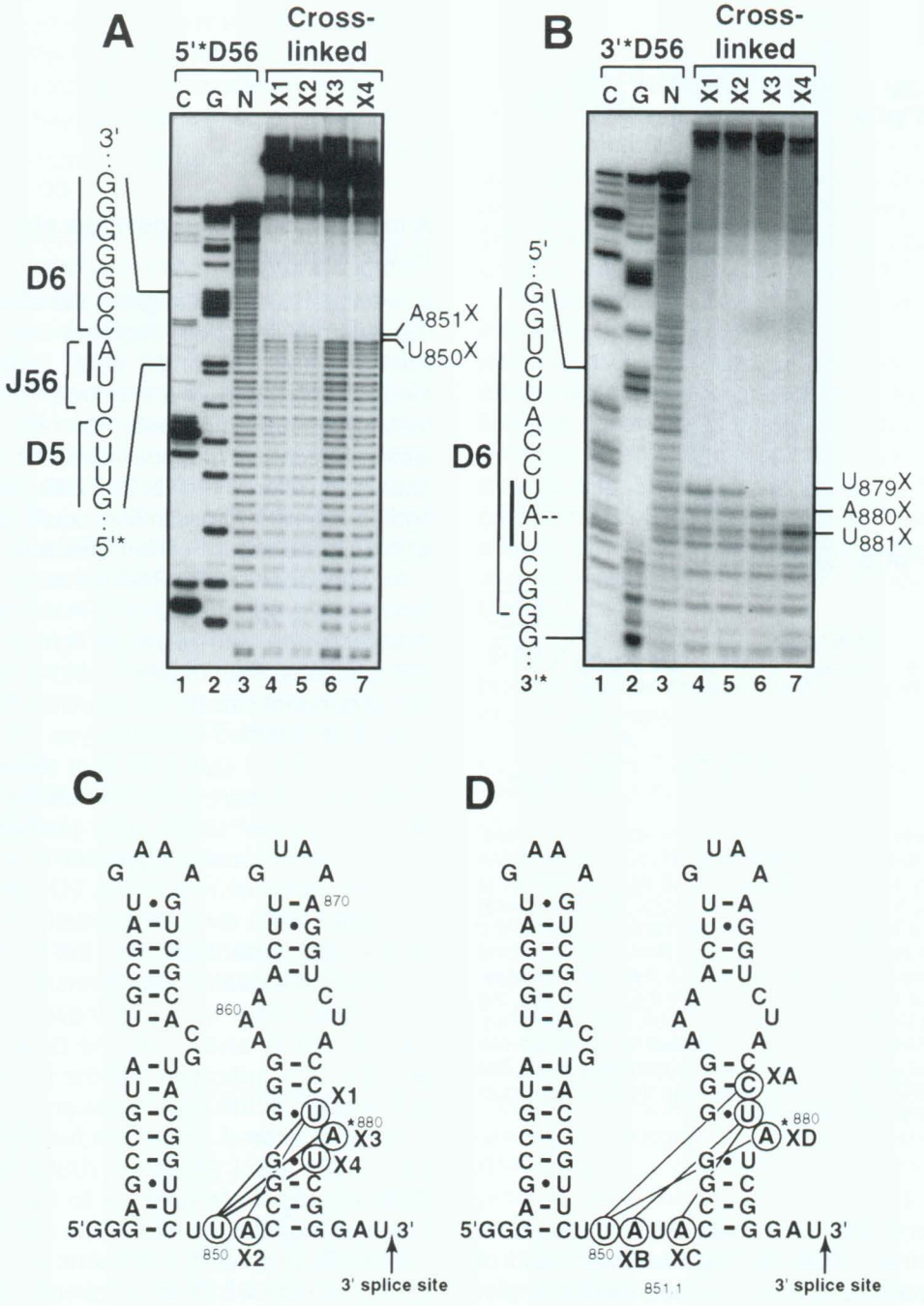

FIGURE 2. Mapping of the crosslinks in D56 Xn RNAs. **A**: Digestion patterns of 5′ end-labeled D56 and D56 Xn RNAs. Purified end-labeled D56 RNAs were partially digested with RNase CL3 (C-specific, lane 1), RNase T1 (G-specific, lane 2), or subjected to partial alkaline hydrolysis (lane 3) and fractionated on a 20% polyacrylamide sequencing gel. End-labeled crosslinked RNAs (D56 X1, X2, X3, and X4) from a 30-min crosslinking reaction containing 100 mM $MgCl_2$ (on ice) (resembling Fig. 1, lane 4) were subjected to partial alkaline hydrolysis (lanes 4–7) and the resulting fragments were separated on the gel. A partial sequence of D56 RNA is shown on the left side of the gel to help map the positions of the crosslinks. The gap in the nucleotide ladders of X1 through X4 identifies the nucleotide involved in the crosslink. **B**: Digestion patterns of 3′ end-labeled D56 and D56 Xn RNAs. 3′ end-labeled RNAs were prepared and aliquots were crosslinked for 5 min on ice in HE buffer containing 100 mM $MgCl_2$ and analyzed as in **A**. The gap in the nucleotide ladders of X1 through X4 identifies the nucleotide involved in the crosslink. **C**: Diagram of the secondary structure of D56 RNA of aI5γ. The characterized crosslinks (X1–X4) are indicated as lines connecting the circled nucleotides involved in each crosslink. Each crosslink involves at least one uridine, which, upon substitution with s^4U, provides the photoreactive group. The crosslinks in products X1 and X4 are between two uracils and the crosslinked RNAs were not analyzed to learn whether they contain a mixture of the two possible donor/acceptor pairs (or whether only one of the two s^4U residues initiated most crosslinks). The asterisk indicates the branch site nucleotide, A880. **D**: Diagram of the secondary structure of D56 J(56)5 mutant RNA. End-labeled samples of D56 J(56)5 RNA were prepared, crosslinked for 20 min on ice, and analyzed as in **A** and **B** (not shown). The nucleotides involved in the characterized crosslinks (A through D) are circled and lines connect the crosslinked pairs.

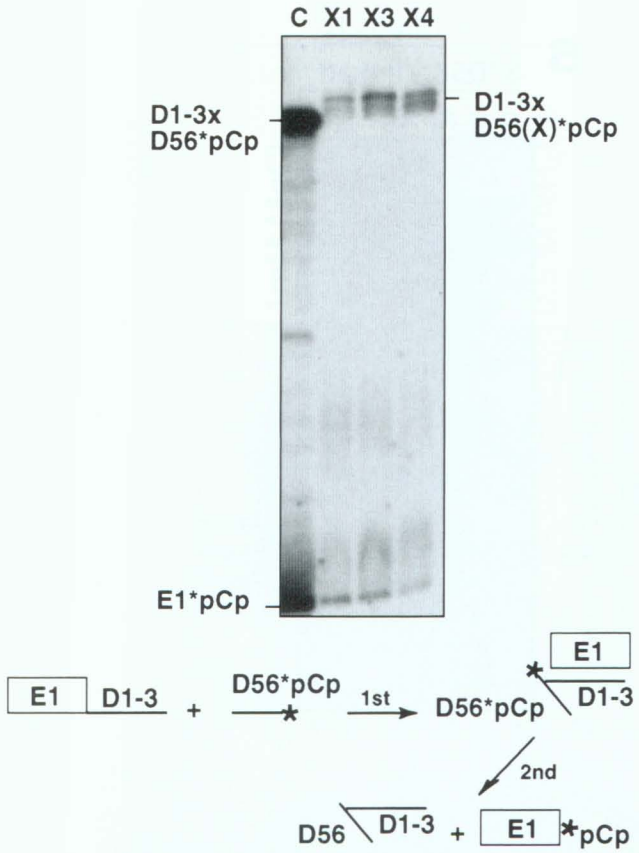

FIGURE 3. Reactivity of crosslinked D56 Xn RNAs in a *trans*-splicing assay. Purified, 3′ end-labeled D56 X1, X3, and X4 RNAs were incubated with excess unlabeled E1D1-3 RNA for 20 min at 45 °C in self-splicing buffer containing 1 M NH_4Cl. The RNAs were then separated on a 6% polyacrylamide gel. Lane C represents a control reaction with noncrosslinked D56 RNA. Both steps of splicing can occur in this reaction; in the first step, a Y-shaped molecule results from the *trans*-branching reaction, and in the second step, the labeled *pCp group functions as the second exon. The scheme under the gel shows the reaction products of the two steps, the asterisk being the radioactive label that makes some of them detectable. The D56 and D56 Xn RNAs are much smaller than the E1pCp product and so were run off the end of this gel.

somewhat more rapidly (like linear RNA of the same length; Podar et al., 1998) so that the lower band of each doublet probably is branched intermediates with nicks in the crosslinked D56 RNA. Clearly, all three crosslinked RNAs are reactive in this assay, showing that crosslinking has trapped functional forms of D56 RNA; thus, these crosslinks are compatible with correct positioning of both D5 and D6 in the active site. The low amounts of these crosslinked RNAs limit these experiments to qualitative estimates of activity.

A single time-course experiment was conducted using 5′ end-labeled samples of these crosslinked D56 RNAs, reacted with a saturating amount of unlabeled EID1-3, as above. The apparent first-order reaction rates for the first-step transesterification reaction were approximately tenfold lower than the rate obtained in a parallel experiment using native D56 RNA. This lower reactivity could be explained by a change in the flexibility of each D56 RNA, caused in each case by the crosslink trapping a single structure, each of which lowers its reactivity in this *trans* assay. Those data are consistent with the single time point shown in Figure 3 and, so, are not shown.

A mutant D56 RNA crosslinks similarly to the wild type

A previous study investigated the effects of increasing the length of the J(56) segment connecting D5 to D6 from 3 to 5 nt (Boulanger et al., 1996). That mutation blocked splicing in vivo and lowered the in vitro reactivity by about tenfold with aberrant 3′ splice-site selection. Because the positioning of D5 and D6 was apparently altered by that 2-nt insertion, we next investigated whether crosslinking could provide some insight into the nature of that structural change.

An end-labeled D56 RNA containing the J(56)5 mutation was prepared with s^4U residues, as above, and irradiated with long-wave UV light in several different salt solutions. That material yielded a pattern of Mg^{2+} ion-dependent crosslinked species of similar complexity to that obtained with wild-type RNA with similar effects of different salts (data not shown; Podar, 1997). The most abundant bands formed by irradiation of this RNA in 100 mM $MgCl_2$ were purified from gel slices and the main crosslink present in each sample was mapped (summarized in Fig. 2D). Interestingly, all of these crosslinks involved the nucleotides of the J(56)5 spacer and nucleotides near the branch nucleotide of D6. Thus, the overall arrangement of J(56) and D6 in this mutant RNA is not changed, but the crosslinks extend further away from the base of D6 (compare with Fig. 2C), indicating that the fine positioning of D5 with respect to the branch site and the 3′ splice site is somewhat altered. There are two UA dinucleotides in the mutant J(56) sequence (U850 A851 and U851.1 A851.2), and there appears to be some competition between them for the interaction with D6. For example, both A851 and A851.2 crosslink with U879; however, U850 but not U851.1 crosslinked to D6. The crosslink involving U850 (to C878) (XA of Fig. 2D) is the most abundant crosslink in this set and is the only crosslink detected in this study that involves C878.

We infer that there is some flexibility in the interaction between J(56) and D6 in this (and the wild-type) intron fragment under self-splicing conditions. Some (or all) of the detected conformations of the wild-type RNA are presumably compatible with self-splicing; some of the conformations detected with the J(56)5 RNA are probably inactive and, so, may account for the reduced activity of the mutant allele (Boulanger et al., 1996). If the D56 regions of both mutant and wild-type RNAs are less dynamic in vivo then it may follow that different conformations predominate in vivo where the wild-type

intron is fully active whereas the mutant intron is fully inactive. It is likely that the structure of D56 in the intact intron reflects the interactions between it and other regions of the intron.

Crosslinking between J(56) and D6 in the absence of D5

The formation of crosslinks between the J(56) region and D6 does not explain the role (if any) of D5 in generating the conformations trapped by this crosslinking, nor does it help decide whether J(56) itself plays an active role. To study those points, we synthesized a s^4U-containing RNA consisting of only J(56) and D6 (called J(56)-D6 RNA). End-labeled preparations of that material were UV-irradiated under various salt conditions and analyzed on polyacrylamide gels as above. As shown in Figure 4, a single, efficient crosslinked species (called J(56)-D6 X) was obtained under every condition analyzed, including buffer lacking Mg^{2+} ion, salt of a monovalent cation, or both. The formation of this crosslink was not influenced by the presence of a tenfold molar excess of D5 RNA (Podar, 1997), indicating that the influence of D5 on the interaction between J(56) and D6 depends on D5 being in *cis* with the other sequences.

The crosslink in J(56)-D6 X RNA was mapped using 5′ and 3′ end-labeled crosslinked RNAs (Fig. 5A,B), and was found to be between the sole A of J(56) (A851, see Fig. 1C) and the U nucleotide 5′ to the branch site in D6, U879 (summarized in Fig. 5C). This crosslink is between the same 2 nt that were joined in the X2 species, the most minor of the four crosslinked species

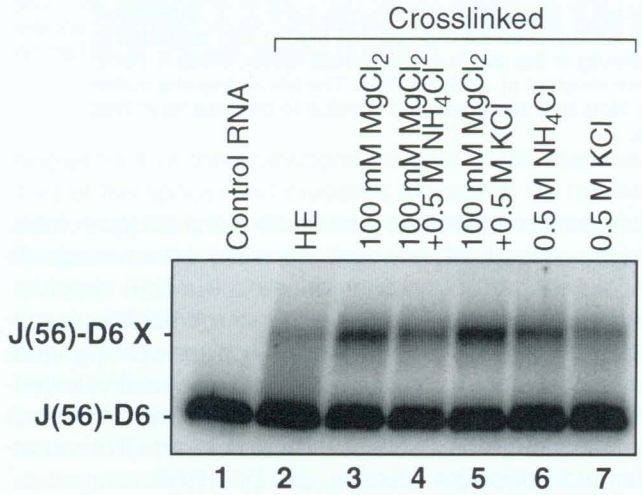

FIGURE 4. UV crosslinking of J(56)-D6 RNA randomly substituted with 4-thio uracil. The 5′ end-labeled RNA was irradiated with 360 nm UV light at 45 °C for 5 min in 10 mM HEPES, 1 mM EDTA (HE) or in HE buffer containing different concentrations of $MgCl_2$ and/or salts, as indicated. Control RNA was not irradiated. The crosslinked species, J(56)-D6 X, has a reduced mobility on this 10% polyacrylamide gel, as compared to J(56)-D6.

analyzed in irradiated wild-type D56 RNA. These results suggest that the 3-nt-long J(56) plays an active role in the tertiary structure of this part of the intron, possibly interacting in a triple helix-like fashion with the D6 helix. This interaction could be a major determinant in positioning D5 and D6 relative to each other. However, it should be stressed that D5 and salt appear to be necessary to form the structures in the wild-type D56 RNA that were captured by crosslinks between U850 and U879, A880 and U881 of D6.

The unique and salt-independent crosslink that forms efficiently in the absence of D5 indicates that the interaction of J(56) with D6 can be quite stable. We attempted to detect this interaction *in trans*, but no crosslink was observed under any conditions of salt or temperature when a high concentration (200 μM) of the trinucleotide 5′-UUA, was added *in trans* to a s^4U-containing D6 RNA lacking an attached J(56) (data not shown). This suggests that the interaction between J(56) and D6 is weak and depends on their covalent association.

Intramolecular crosslinks formed in D56 RNA of another intron

Domains 5 and 6 are highly conserved secondary structures of group II introns. They are important parts of the catalytic core, the overall architecture of which is likely to be substantially conserved in all group II introns. It is known that both transesterification reactions catalyzed by group II introns, first-step branching and second-step exon ligation, depend on a specific, yet little understood, tertiary structure involving D5 and D6. Although there may be features specific to each individual intron and subclass, the positioning of D5 and D6 relative to each other is probably conserved among all group II introns. As an initial test of this inference, we carried out analogous crosslinking experiments with D56 RNA from al1, another self-splicing group II intron from yeast mtDNA. This intron belongs to the subclass IIA, all members of which have a 2-nt-long J(56) segment, usually with the sequence GG (Michel et al., 1989).

An al1 D56 RNA was prepared containing randomly incorporated s^4U, end labeled, and irradiated under various salt conditions (Fig. 6). As was the case with D56 RNA from al5γ, a number of fast-migrating crosslinked species formed in the absence of salt or Mg^{2+} ion (Fig. 6, lane 2, Xi). When analyzed, they did not generate a readable gap in the hydrolysis ladder, suggesting that they contain very short-range crosslinks. Numerous more slowly migrating crosslinked products were obtained in the presence of salts and at least two products (X1 and X3) strictly required Mg^{2+} ion for their formation. The major crosslinked products were purified from a sample like that shown in Figure 6, lane 5 and mapped (Fig. 7A,B), and the findings are summarized in Figure 7C. The bold lines denote the three

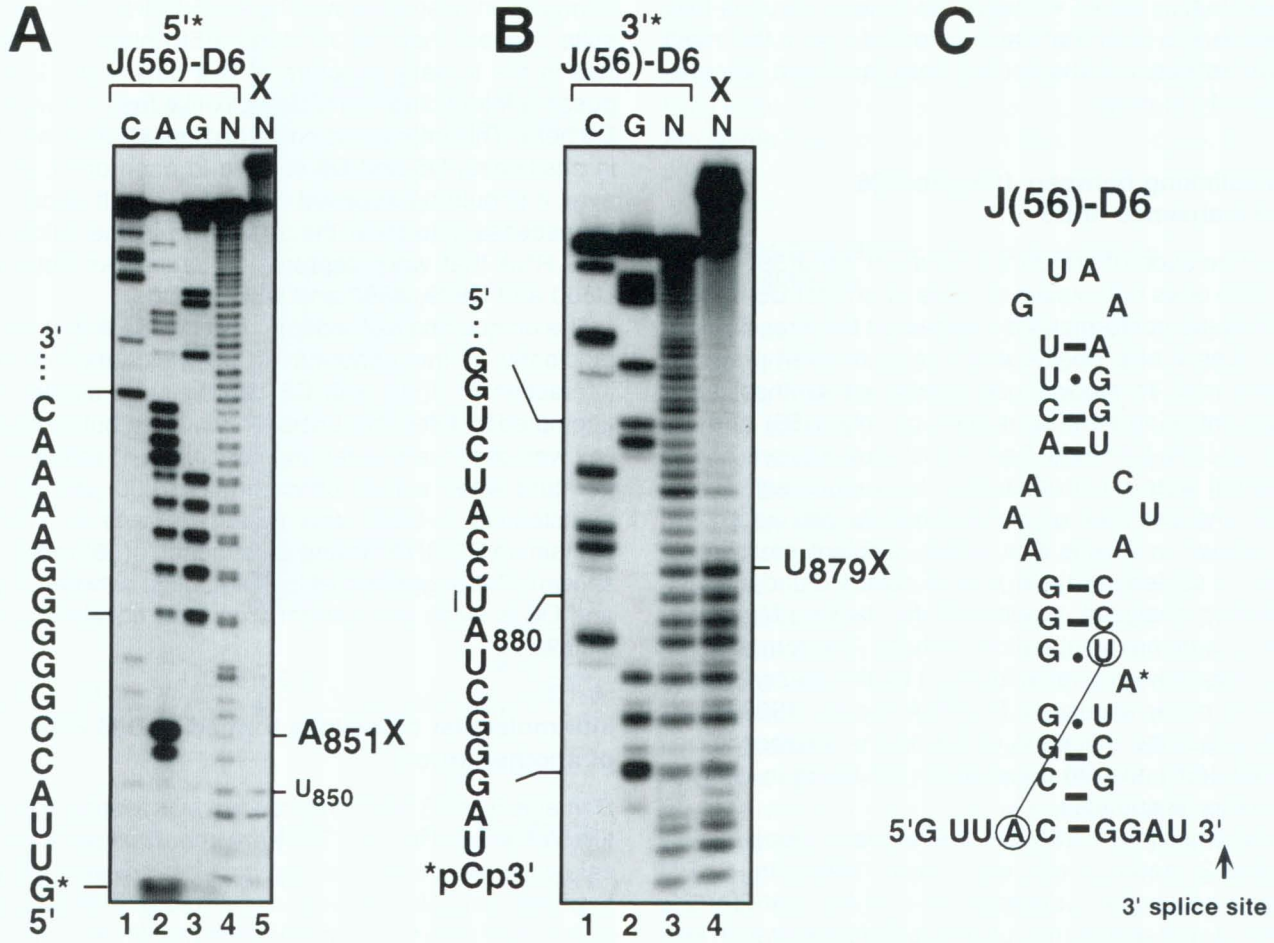

FIGURE 5. Mapping of the crosslink in J(56)-D6 X RNA. **A**: Mapping of the 5′ nucleotide involved in the crosslink. 5′ end-labeled J(56)-D6 RNA was partially digested with RNase CL3 (C-specific, lane 1), RNase PhyM (A-specific, lane 2), RNase T1 (G-specific, lane 3), and subjected to partial alkaline hydrolysis (lane 4). 5′ end-labeled J(56)-D6 X RNA (crosslinked for 10 min at room temperature in HE buffer containing 100 mM $MgCl_2$) was subjected to partial alkaline hydrolysis (lane 5). The RNA fragments were separated on a 20% polyacrylamide sequencing gel. The gap in the nucleotide ladder of J(56)-D6 X RNA shows that the 5′ nucleotide of the crosslink is A851, the 3′ nucleotide of the J(56) sequence. **B**: Mapping of the 3′ nucleotide involved in the crosslink. 3′ end-labeled J(56)-D6 RNA was partially digested with RNase CL3 (C-specific, lane 1), RNase T1 (G-specific, lane 2), or subjected to partial alkaline hydrolysis (lane 3). 3′ end-labeled J(56)-D6 X RNA (crosslinked for 10 min at room temperature in HE buffer containing 100 mM $MgCl_2$) was subjected to partial alkaline hydrolysis (lane 4). The band that marks the beginning of the gap in the nucleotide ladder, which is U879, indicates the crosslinked nucleotide. **C**: Diagram of the secondary structure of J(56)-D6 RNA. The line connecting nucleotides A851 and U879 indicates the crosslink characterized here. Note that this crosslink is identical to crosslink X2 in D56 RNA (Fig. 2C). The asterisk indicates the branch site nucleotide.

crosslinks that are most prominent in samples containing Mg^{2+} ion (X1, X2, and X3); X4 is present under all of these conditions, and X5 and X6 are the main crosslinks of this mobility class only in the sample lacking Mg^{2+} ion and salt (Fig. 6, lane 2).

Crosslinks X1–X4 are between residues of D5 and D6, and, so, detect a direct side-by-side arrangement of the two domains. Although in the case of al5γ no direct crosslinks were identified between the D5 and D6 domains, the contact between J(56) and D6 is compatible with various sorts of side-by-side orientations of the two helices in which the lower part of the D5 helix is the branch region of D6. In both al5γ and al1, the crosslinks found suggest an axial displacement of the two helices, such that the first several base pairs of D5 are positioned around the middle of the D6 lower helix. In the case of al1, however, it is still unknown what role (if any) the J(56) segment plays in the D56 structure. Attempts to investigate whether the crosslinked species of al1 D56 are functional in *trans*-splicing were unsuccessful. *Trans* assays for al1 are relatively inefficient for branching, and it was not possible to obtain large enough amounts of the crosslinked RNAs to attempt to drive the reaction with D56 RNA.

DISCUSSION

Our experiments show that in D56 RNA of al5γ incubated under self-splicing conditions, several nucleotides of the J(56) spacer crosslink efficiently to

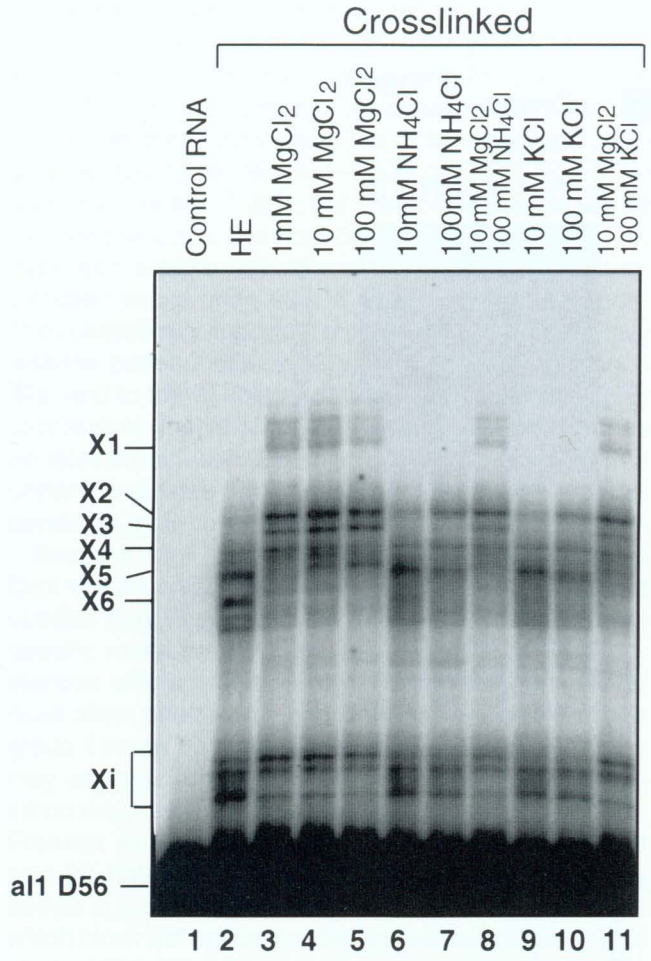

FIGURE 6. UV crosslinking of aI1 D56 RNA randomly substituted with 4-thio uracil. The 5′ end-labeled D56 RNA was irradiated with 360 nm UV light for 10 min on ice in 10 mM HEPES, 1 mM EDTA (HE, lane 2) or in HE buffer containing Mg^{2+} ion or salts of monovalent cations as indicated (lanes 3–11). Lane 1 was control, nonirradiated RNA. The crosslinked species were identified after electrophoresis on a 10% polyacrylamide gel, based on their reduced mobility. Numerous crosslinked products are formed, but only the ones indicated on the left (X1–X6) were analyzed further.

nucleotides of the branchpoint region of D6. Because 1 nt of the spacer can crosslink to several D6 nucleotides, it appears that the tertiary structure in that region is somewhat flexible under these conditions. The nature of these crosslinks and the evidence that the crosslinked D56 RNAs retain branching activity in vitro effectively rule out coaxial stacking of the D5 and D6 helices. Our whole data set suggests a side-by-side orientation in which the end of the D5 helix is displaced upward from the base of the D6 helix. Our data cannot distinguish among various specific side-by-side orientations that can be envisioned (e.g., parallel, crossed, or partially parallel helices). The mapped crosslinks in D56 RNA depended on the presence of salts of a monovalent cation or Mg^{2+} ion, a finding that may indicate that positive charges are required to neutralize charge repulsion between the phosphodiester backbones of the D5 and D6 helices. The crosslinks formed in the J(56)5 D56 RNA also support these views; it appears that the longer spacer reaches higher on the D6 helix, but the overall arrangement of the two domains relative to one another appears to be the same.

The crosslinked D56 RNAs of aI5γ analyzed here were found to be capable of participating in *trans* self-splicing reactions. Even D56 X3, which is the most abundant crosslinked species and contains a covalent bond to the branch nucleotide, participates in first-step transesterification reactions. However, the first-step branching reaction observed with these RNAs was relatively inefficient, probably because of several factors. First, each crosslink captures a particular structure that may be only a minor configuration in the native molecule. Second, the crosslinks should constrain the possible structures of the RNA, and it is possible that a somewhat flexible structure of the D56 unit may be needed to achieve its proper orientation in the active site. Nevertheless, it is clear that these crosslinks permit D5 and D6 to be positioned adequately relative to each other and to other components of the active site.

It has been reported that mutations of the branch nucleotide of D6 strongly inhibit, and some abolish, first-step branching, and that little or no cryptic branching reactions occur (Gaur et al., 1997; Liu et al., 1997; Chu et al., 1998; Podar et al., 1998). It is somewhat surprising that the crosslinked RNAs retained any branching activity at all, especially the relatively abundant one involving the branch nucleotide. Direct mapping of the branch sites formed in these reactions using crosslinked D56 RNAs was not possible using standard methods, because nucleotides at or near the natural branch nucleotide are crosslinked. While we think it is unlikely, it remains possible that some branching by these crosslinked D56 RNAs may involve a nucleophile other than A880.

The finding that J(56) crosslinks efficiently to D6 in the absence of D5 was a surprise. Unlike the situation using D56 RNA, J(56)-D6 RNA forms only a single product in which A851 is crosslinked to U879. The A851–U879 crosslink is also formed when D5 is present, but in that case it is the most minor species (X2) analyzed. Also, using D56 RNA, formation of the X2 crosslink requires Mg^{2+} ions, while the same crosslink forms more efficiently with J(56)-D6 RNA in several buffers lacking Mg^{2+} ions (see Fig. 4). As suggested above, it appears that charge repulsion between the backbones of the D5 and D6 helices interferes with the folding of D56 RNA, and that the Mg^{2+} ions and monovalent cations neutralize the charges, allowing the helices to approach closely. It is possible that charge neutralization may also be needed for the folding of D56 RNA in vivo; alternatively or in addition, one or more nuclear-encoded proteins may also contribute to the folding of D5 relative to D6 (Lambowitz et al., 1999; Schmidt et al., 1998).

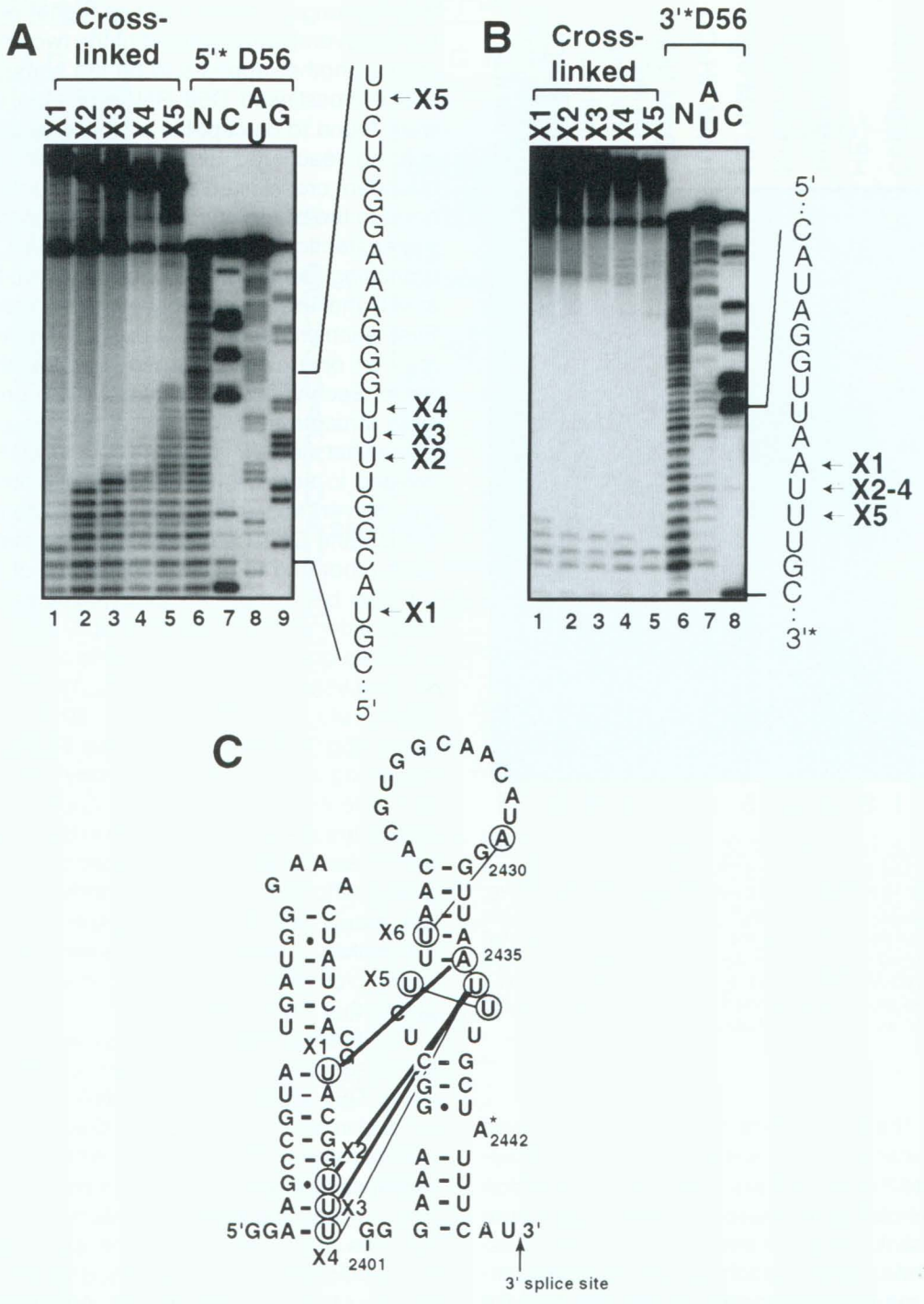

FIGURE 7. Mapping of the major crosslinks in al1 D56 Xn RNAs. Purified end-labeled al1 D56 and D56 Xn RNAs were sequenced as described in Figures 2 and 5 and in Materials and Methods. The RNA fragments were separated on 20% sequencing gels. A partial sequence of al1 D56 RNA is shown on the right side of each gel to define the positions of the crosslinks. **A**: Analysis of 5′ end-labeled al1 D56 and al1 D56 Xn RNAs. These RNAs were obtained by crosslinking on ice for 10 min in HE buffer containing 100 mM $MgCl_2$. The gap in the nucleotide ladders of X1–X5 identifies the nucleotide involved in the crosslinks. The crosslinked nucleotides are noted in the sequence shown on the right side. **B**: Analysis of 3′ end-labeled al1 D56 and al1 D56 Xn RNAs. These RNAs were obtained by crosslinking on ice for 10 min in HE buffer containing 100 mM $MgCl_2$. The crosslinked nucleotides are indicated in the sequence shown on the right side. **C**: Diagram of the secondary structure of D56 RNA of al1. Lines connect the circled nucleotides joined by each characterized crosslink (X1–X6). (Primary data for X6 are not shown in **A** and **B**.) As before, some of these crosslinks (X2–X5) connect two U residues and it is not known which one was the s^4U that initiated the event.

The inference that charge neutralization is important for arranging D5 relative to D6 strengthens the view that D5 and D6 are positioned in a side-by-side fashion, although their precise orientation cannot be deduced from these data. Weak interactions between J(56) and D6 leading to the A851–U879 crosslink do not require salts or Mg^{2+} ions, but are somewhat enhanced by their presence. It is possible that the 3-nt-long J(56) may form a triple helix-like structure with D6; that interaction would bring the D5 bottom helix, which contains catalytically important residues, into close proximity with the bottom helix of D6, which contains the branch site, and to the 3′ splice junction. The inability of J(56) to crosslink when it is presented *in trans* to D6 either as an isolated trinucleotide or attached to a D5 RNA (not shown), indicates that this interaction is weak and depends on their covalent association *in cis*.

Because Mg^{2+} ions are required for D56 RNA to form the X4 crosslink (between U850 and U881), that species may depend on the presence of one or more specific metal ion bridges. This is one of the few instances where salt and Mg^{2+} ions are observed to have direct effects on a specific structural element of a group II intron RNA (see Qin & Pyle, 1997). This finding may explain, in part, why efficient self-splicing of this intron requires high ionic strength (Jarrell et al., 1988; Franzen et al., 1994; Podar et al., 1995). It should also be noted that increased salt concentrations are known to suppress inhibitory effects of some mutations which block self-splicing under low-salt conditions (Koch et al., 1992; Chanfreau & Jacquier, 1994; Boulanger et al., 1995; Schmidt et al., 1996).

A side-by-side orientation of D5 and D6 is supported by intramolecular crosslinking of D56 RNA from another group II intron, aI1. That intron belongs to the subclass IIA, in which J(56) is the dinucleotide, GG. Crosslinking of aI1 D56 RNA produced numerous species, which, as was the case with aI5γ, could be divided into salt-independent, salt or Mg^{2+} ion-dependent, and absolutely Mg^{2+} ion-dependent types. Several crosslinks that can form in the absence of salt were mapped and were found to be intradomain 6 (summarized in Fig. 7C). The other crosslinks however are between nucleotides of the bottom helix of D5 and the internal loop of D6. Their existence suggests that, as in aI5γ, the bottom helix of D5 is displaced (vertically as drawn in Fig. 7C) with respect to D6. Also, the structure appears fluid, as two nucleotides from D5 crosslink to the same nucleotide in D6 and some mutually exclusive crosslinks were present. Unlike aI5γ, none of the crosslinks in aI1 D56 RNA involved the J(56) sequence, so we cannot infer a specific role for the J(56) RNA of that intron in organizing the D56 region. Nevertheless, the relative orientation of the two domains appears similar in both aI5γ and aI1.

Our data suggest that J(56) provides some of the specificity and energy responsible for the basic arrangement of D5 relative to D6 in the RNAs analyzed here. D5 also contributes to the organization of these D56 RNAs. It should be noted that some of the detected fluidity of these D56 RNAs could result from our having separated the two domains from the rest of the intron. In the context of the whole intron, where both D5 and D6 also make tertiary contacts with other substructures (Costa & Michel, 1995; Costa et al., 1997; Jestin et al., 1997), the structure of the D56 region could be restricted from adopting some of the configurations that were sampled here by crosslinking. Also, a flexible D56 region may be an important factor during the initial folding of the intron RNA as well as during the conformational change that occurs between the two steps of splicing. Further structural studies are needed to refine this initial picture of the interaction between these two important domains of group II introns.

MATERIALS AND METHODS

Transcription templates

The DNA templates used in transcription were synthesized by PCR and contained an introduced T7 RNA polymerase promoter (Milligan & Uhlenbeck, 1989; Perlman & Podar, 1996). For the wild-type aI5γ D56 or the J(56)5 mutant, plasmids pJD20 or pJD20-J(56)5 (Boulanger et al., 1996) were used as templates in the PCR reactions, using the oligonucleotides T7GGD5 (5′TAATACGACTCACTATAGGGAGCCG UAUGCGAUGAAAG3′) and anti-3′ss (5′ATCCCGATAGGT AGACCTTTACAAG3′). pJD20 was also used to synthesize the template for a J(56)D6 RNA, in a PCR reaction with oligonucleotides T7GD6 (5′TAATACGACTCACTATAGUUAC CGGGGGAAAACUU3′) and anti-3′ss. Plasmid pSH2 (Hebbar et al., 1992) was used to synthesize the template for aI1 D56 RNA, using oligonucleotides aI1-T7GGD5 (5′TAATACG ACTCACTATAGGGAAGCCGTATGATGGG3′) and aI1-anti-3′ss (5′ATGAAATAGCAAATTAACCTATGTTGC3′). Plasmid pJDI3′-673 (Jarrell et al., 1988), linearized with *Hin*dIII, was used to transcribe an aI5γ E1:D123 RNA.

In vitro transcription and RNA labeling

4-thio UTP (s^4UTP) was synthesized by phosphorylation of 4-thio UDP (Sigma), according to a published method (Bartholomew et al., 1986). D56 RNAs were transcribed using T7 RNA polymerase, according to standard procedures (e.g., Perlman & Podar, 1996). To incorporate approximately one s^4U per RNA transcript, a low amount of s^4UTP (5 or 10% of the total UTP content) was included in the transcription reaction in a ratio of 10:1 to 20:1, depending on the transcript to be synthesized. The transcription reaction also contained a small amount (2–5 μCi) of [α-^{32}P]-UTP, which served for detection and quantitation of the RNA. After transcription, the RNAs were purified by gel electrophoresis, eluted, and precipitated. The purified RNAs were 5′ or 3′ end labeled with T4 polynucleotide kinase and [γ-^{32}P]-ATP or T4 RNA ligase and ^{32}P-pCp, purified again, and dissolved in 10 mM Tricine-KOH, 1 mM EDTA, pH 7. Unlabeled E1:D123 RNA, tran-

scribed using T7 RNA polymerase and the pJDI3′-673-*Hin*dIII template (Jarrell et al., 1988), was gel purified and quantitated by spectrophotometry.

Crosslinking of RNAs containing s^4U

For crosslinking experiments, several picomoles of RNAs containing s^4U were heated at 85 °C for 2 min, cooled to room temperature, and added to HE buffer solutions containing different concentrations of salts or/and $MgCl_2$, depending on the experiment. The RNAs were incubated at 42 °C for 5 min, and then were placed as 10–20 μL drops on Parafilm. Crosslinking was done using a 330–360 nm UV lamp (Black Ray, UVP), placed at ~5 cm from the samples (Podar et al., 1998). To eliminate residual short wavelengths, a plastic Petri dish was placed between the samples and the light source. Irradiation times of 10–30 min at room temperature provided a good yield of crosslinked products; irradiations at 0 °C or 42 °C yielded similar levels and arrays of products and were used in some experiments. Varying the concentration of the RNAs from 10–50 nM had no detectable effect on the pattern of crosslinked products. The crosslinked products were separated by electrophoresis on 40-cm-long, 0.5-mm thick, 20% polyacrylamide gels containing TBE and 8 M urea for 24 h at 1,000 V and detected by autoradiography. For crosslink mapping and in vitro reactions using purified samples of specific crosslinked RNAs, larger samples of end-labeled RNAs were crosslinked in a solution containing 100 mM Mg Cl_2 and HEPES buffer (see Fig. 1, lane 4; Fig. 4, lane 3; and Fig. 6, lane 5). Individual crosslinked species were eluted from preparative gel slices and used in further analysis.

Enzymatic RNA sequencing

End-labeled (5′ or 3′) RNAs were sequenced by partial enzymatic digestions and limited alkaline hydrolysis using an enzymatic RNA sequencing kit (Amersham-USB), according to manufacturer's instructions. The nucleases that were used were RNase T1 (G-specific), RNase CL3 (C-specific), and PhyM (A- and U-specific).

In vitro self-splicing reactions

Trace amounts of end-labeled aI5γ D56 and purified crosslinked product RNAs were reacted with excess (2 μM) unlabeled E1:D123 RNA, in buffer containing 40 mM HEPES, pH 7.5, 100 mM $MgCl_2$, and 1 M NH_4Cl at 42 °C. The products were separated by electrophoresis on denaturing 6% polyacrylamide gel and visualized by autoradiography or exposure to a PhosphorImager screen. The overall strategy for the self-splicing assays was recently reviewed (Perlman & Podar, 1996).

ACKNOWLEDGMENTS

This research was supported by NIH research grant GM31480 to P.S.P. M.P. was a fellow of the Robert A. Welch Foundation (Grant I-1211) during part of this research. We thank Anna Pyle (Columbia University College of Physicians and Surgeons) and Craig Peebles (University of Pittsburgh) for thoughtful comments on a draft of this paper.

Received September 18, 1998; returned for revision October 14, 1998; revised manuscript received November 16, 1998

REFERENCES

Abramovitz DL, Friedman RA, Pyle AM. 1996. Catalytic role of 2′-hydroxyl groups within a group II intron active site. *Science 271*:1413–1416.

Bartholomew B, Dahmus ME, Meares CF. 1986. RNA contacts subunits IIo and IIc in HeLa RNA polymerase II transcription complexes. *J Biol Chem 261*:14226–14231.

Boulanger SC, Belcher SM, Schmidt U, Dib-Hajj SD, Schmidt T, Perlman PS. 1995. Studies of point mutants define essential nucleotides in the domain 5 substructure of a group II intron. *Mol Cell Biol 15*:4479–4488.

Boulanger SC, Faix PH, Yang H, Zhuo J, Franzen JS, Peebles CL, Perlman PS. 1996. Length changes in the joining segment between domains 5 and 6 of a group II intron inhibit self-splicing and alter 3′ splice site selection. *Mol Cell Biol 16*:5896–5904.

Chanfreau G, Jacquier A. 1994. Catalytic site components common to both splicing steps of a group II intron. *Science 266*:1383–1387.

Chanfreau G, Jacquier A. 1996. An RNA conformational change between the two chemical steps of group II self-splicing. *EMBO J 15*:3466–3476.

Chu VT, Liu Q, Podar M, Perlman PS, Pyle AM. 1998. More than one way to splice an RNA: Branching without a bulge and splicing without branching in group II introns. *RNA 4*:1186–1202.

Costa M, Christian EL, Michel F. 1998. Differential chemical probing of a group II self-splicing intron identifies bases involved in tertiary interactions and supports an alternative secondary structure model of domain V. *RNA 4*:1055–1068.

Costa M, Deme E, Jacquier A, Michel F. 1997. Multiple tertiary interactions involving domain II of group II self-splicing introns. *J Mol Biol 267*:520–536.

Costa M, Michel F. 1995. Frequent use of the same tertiary motif by self-folding RNAs. *EMBO J 14*:1276–1285.

Dib-Hajj SD, Boulanger SC, Hebbar SK, Peebles CL, Franzen JF, Perlman PS. 1993. Domain 5 interacts with domain 6 and influences the second transesterification reaction of group II intron self-splicing. *Nucleic Acids Res 21*:1797–1804.

Franzen JS, Zhang M, Chay TR, Peebles CL. 1994. Thermal activation of a group II intron ribozyme reveals multiple conformational states. *Biochemistry 33*:11315–11326.

Gaur RK, McLaughlin LW, Green MR. 1997. Functional group substitutions of the branchpoint adenosine in a nuclear pre-mRNA and a group II intron. *RNA 3*:861–869.

Hebbar SK, Belcher SM, Perlman PS. 1992. A maturase-encoding group IIA intron of yeast mitochondria self-splices in vitro. *Nucleic Acids Res 20*:1747–1754.

Jacquier A. 1996. Group II introns: Elaborate ribozymes. *Biochimie 78*:474–487.

Jarrell KA, Dietrich RC, Perlman PS. 1988. Group II intron domain 5 facilitates a *trans*-splicing reaction. *Mol Cell Biol 8*:2361–2366.

Jarrell KA, Peebles CL, Dietrich RC, Romiti SL, Perlman PS. 1988. Group II intron self-splicing: Alternative reaction conditions yield novel products. *J Biol Chem 263*:3432–3439.

Jestin JL, Deme E, Jacquier A. 1997. Identification of structural elements critical for interdomain interactions in a group II self-splicing intron. *EMBO J 16*:2945–2954.

Koch JL, Boulanger SC, Dib-Hajj SD, Hebbar SK, Perlman PS. 1992. Group II introns deleted for multiple substructures retain self-splicing activity. *Mol Cell Biol 12*:1950–1958.

Konforti BB, Abramovitz DL, Duarte CM, Karpeisky A, Beigelman L, Pyle AM. 1998. Ribozyme catalysis from the major groove of group II intron domain 5. *Mol Cell 1*:433–441.

Lambowitz AM, Caprara MG, Zimmerly S, Perlman PS. 1999. Group I and group II ribozymes as RNPs: Clues to the past and guides to the future. In: Gesteland R, Cech TR, Atkins J, eds. *The RNA world*, 2nd ed. Cold Spring Harbor New York: Cold Spring Harbor Laboratory Press. In press.

Liu Q, Green J, Khodadadi A, Haeberli P, Beigelman L, Pyle A. 1997. Branch-site selection in a group II intron mediated by active rec-

ognition of the adenine amino group and steric exclusion of non-adenine functionalities. *J Mol Biol 267*:163–171.
Michel F, Ferat JL. 1995. Structure and activities of group II introns. *Ann Rev Biochem 64*:435–461.
Michel F, Umesono K, Ozeki H. 1989. Comparative and functional anatomy of group II catalytic introns—a review. *Gene 82*:5–30.
Michels WJ Jr, Pyle AM. 1995. Conversion of a group II intron into a new multiple-turnover ribozyme that selectively cleaves oligonucleotides: Elucidation of reaction mechanism and structure/function relationships. *Biochemistry 34*:2965–2977.
Milligan JF, Uhlenbeck OC. 1989. Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol 180*:51–62.
Mueller MW, Stocker P, Hetzer M, Schweyen RJ. 1991. Fate of the junction phosphate in alternating forward and reverse self-splicing reactions of group II intron RNA. *J Mol Biol 222*:145–154.
Peebles CL, Zhang M, Perlman PS, Franzen JF. 1995. Catalytically critical nucleotides in domain 5 of a group II intron. *Proc Natl Acad Sci USA 92*:4422–4426.
Perlman P, Podar M. 1996. Reactions catalyzed by group II introns in vitro. *Methods Enzymol 264*:66–86.
Podar M. 1997. Biochemical, molecular, and genetic investigations of the structure and mechanism of a group II intron catalytic RNA. Ph.D. Dissertation, University of Texas Southwestern Medical Center, Dallas, Texas.
Podar M, Chu VT, Pyle AM, Perlman PS. 1998. Group II intron splicing in vivo by first-step hydrolysis. *Nature 391*:915–918.
Podar M, Dib-Hajj S, Perlman PS. 1995. A UV-induced, Mg^{2+}-dependent crosslink traps an active form of domain 3 of a self-splicing group II intron. *RNA 1*:828–840.
Podar M, Perlman PS, Padgett RA. 1995. Stereochemical selectivity of group II intron splicing, reverse splicing, and hydrolysis reactions. *Mol Cell Biol 15*:4466–4478.
Podar M, Zhuo J, Zhang M, Franzen JS, Perlman PS, Peebles CL. 1998. Domain 5 binds near a highly conserved dinucleotide in the joiner linking domains 2 and 3 of a group II intron. *RNA 4*:151–166.
Pyle A. 1996. Catalytic reaction mechanisms and structural features of group II intron ribozymes. In: Eckstein F, Lilley DMJ, eds. *Nucleic Acids and Molecular Biology*. Vol. 10. New York: Springer Verlag. pp 75–107.
Qin PZ, Pyle AM. 1997. Stopped-flow fluorescence spectroscopy of a group II intron ribozyme reveals that domain 1 is an independent folding unit with a requirement for specific Mg^{2+} ions in the tertiary structure. *Biochemistry 36*:4718–4730.
Schmidt U, Maue I, Lehmann K, Belcher SM, Stahl U, Perlman PS. 1998. Mutant alleles of the *MRS2* gene of yeast nuclear DNA suppress mutations in the catalytic core of a mitochondrial group II intron. *J Mol Biol 282*:525–541.
Schmidt U, Podar M, Stahl U, Perlman PS. 1996. Mutations of the two-nucleotide bulge of D5 of a group II intron block splicing in vitro and in vivo: Phenotypes and suppressor mutations. *RNA 2*:1161–1172.
van der Veen R, Arnberg AC, van der Horst G, Bonen L, Tabak HF, Grivell LA. 1986. Excised group II introns in yeast mitochondria are lariats and can be formed by self-splicing in vitro. *Cell 44*:225–234.