RNA (1998), *4*:151–166. Cambridge University Press. Printed in the USA.

Domain 5 binds near a highly conserved dinucleotide in the joiner linking domains 2 and 3 of a group II intron

MIRCEA PODAR,[1,3] JIN ZHUO,[2] MINCHENG ZHANG,[2,4] JAMES S. FRANZEN,[2] PHILIP S. PERLMAN,[1] and CRAIG L. PEEBLES[2]

[1]Department of Molecular Biology & Oncology, University of Texas, Southwestern Medical Center, Dallas, Texas 75235-9148, USA
[2]Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania 15260-7700, USA

ABSTRACT

Photocrosslinking has identified the joiner between domains 2 and 3 [J(23)] as folding near domain 5 (D5), a highly conserved helical substructure of group II introns required for both splicing reactions. D5 RNAs labeled with the photocrosslinker 4-thiouridine (4sU) reacted with highly conserved nucleotides G_{588} and A_{589} in J(23) of various intron acceptor transcripts. These conjugates retained some ribozyme function with the lower helix of D5 crosslinked to J(23), so they represent active complexes. One partner of the $\gamma\cdot\gamma'$ tertiary interaction ($A_{587}\cdot U_{887}$) is also in J(23); even though $\gamma\cdot\gamma'$ is involved in step 2 of the splicing reaction, D5 has not previously been found to approach $\gamma\cdot\gamma'$. Similar crosslinking patterns between D5 and J(23) were detected both before and after step 1 of the reaction, indicating that the lower helix of D5 is positioned similarly in both conformations of the active center. Our results suggest that the purine-rich J(23) strand is antiparallel to the D5 strand containing U_{32} and U_{33}. Possibly, the interaction with J(23) helps position D5 correctly in the ribozyme active site; alternatively, J(23) itself might participate in the catalytic center.

Keywords: active site; 4-thiouridine; group II intron; photocrosslink; ribozyme; RNA tertiary structure

INTRODUCTION

Domain 5 (D5) is the most highly conserved primary sequence and secondary structure within group II introns (Michel et al., 1989; Michel & Ferat, 1995; see Fig. 1). These introns are organized into six secondary structure domains (D1–D6), and some members of this group are capable of self splicing (Peebles et al., 1986; van der Veen et al., 1986; Michel et al., 1989; Michel & Ferat, 1995). D5 contributes functionally critical residues to the catalytic center of group II introns both in vivo (Boulanger et al., 1995) and in vitro (Chanfreau & Jacquier, 1994; Peebles et al., 1995; Abramovitz et al., 1996). An active D5 is also required for intron mobility (Zimmerly et al., 1995). D5 functions in various *trans*-reaction systems as a separate, folded RNA (Jarrell et al., 1988a) with a K_d of approximately 0.5 μM for intron RNAs (Franzen et al., 1993; Pyle & Green, 1994; Peebles et al., 1995). D5 has a stable secondary structure, so that its binding as a separate RNA to the complementing intron transcript in various bimolecular *trans*-splicing and *trans*-hydrolysis systems (Jarrell et al., 1988a; Koch et al., 1992; Dib-Hajj et al., 1993) depends on multiple tertiary interactions rather than a stretch of Watson–Crick base pairing (Franzen et al., 1993; Peebles et al., 1995). Experiments with intron RNAs deleted for various domains showed that a minimal ribozyme, which included D1 plus D5, was capable of 5′ splice junction hydrolysis (5′SJ hydrolysis), a model step 1 reaction (Koch et al., 1992). D5 also plays a role in positioning D6 for branching and the 3′ splice junction (3′SJ) for step 2 of splicing (Dib-Hajj et al., 1993).

On D5 itself, potential partners for binding interactions have been identified by chemical modification-interference (Chanfreau & Jacquier, 1994) and by kinetic and physical analyses of D5 variants with nucleotide replacements (Peebles et al., 1995; Abramovitz et al.,

Reprint requests to: Craig L. Peebles, Department of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania 15260-7700, USA; e-mail: cpeebles@vms.cis.pitt.edu.
[3]Present address: Biology Department, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA.
[4]Present address: Department of Structural Biology, Fairchild Center, Stanford University School of Medicine, Stanford, California 94305-5400, USA.

1996). These binding-interaction sites on D5 presumably make specific contacts elsewhere on the intron. For example, the D5 tetraloop (5′-$G_{15}A_{16}A_{17}A_{18}$) participates in a tertiary interaction, designated $\zeta \cdot \zeta'$, with a tetraloop-binding motif that consists of an internal loop and its adjacent helix in D1, subdomain D (Costa & Michel, 1995). Another D5 interaction that was identified by modification-interference involves D3 and depends on the two-nucleotide bulge region of D5 (Jestin et al., 1997). Probably, D5 contacts additional intron elements that remain to be identified and characterized. Those D5 contacts are most likely nonadjacent intron residues that fold close together within the secondary and tertiary structure of the RNA to create a suitable receptor surface. In any case, participation in the active site for *trans*-reactions requires that D5 achieve specific binding to the rest of the intron RNA.

Several highly conserved D5 residues have been implicated in ribozyme catalytic function (Chanfreau & Jacquier, 1994; Peebles et al., 1995; Abramovitz et al., 1996). The candidates for active-site residues include those nucleotides in D5 where chemical modifications or sequence substitutions yield catalytically inactive D5 variants that still retain near-normal binding ability for

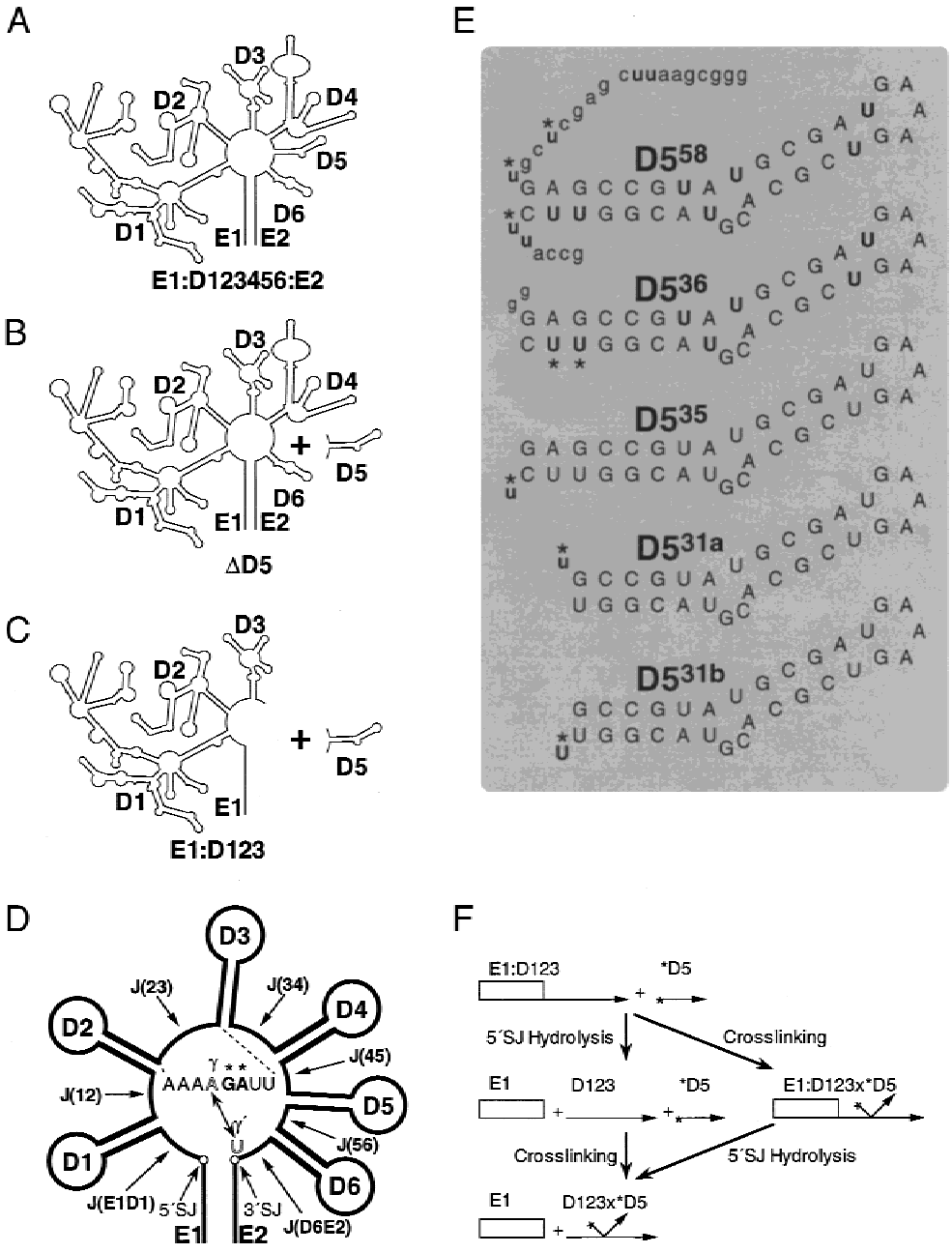

FIGURE 1. (*Legend on facing page.*)

the D5 site on the rest of the intron transcript. Binding-proficient but catalytically inactive D5 variants, such as base substitutions at G_3 or C_4, exhibit competitive inhibition against the wild-type D5 (Peebles et al., 1995). Some of the catalytically crucial D5 residues occur in the highly conserved D5 trinucleotide (5′-$A_2G_3C_4$; see Fig. 1; Chanfreau & Jacquier, 1994; Peebles et al., 1995; Abramovitz et al., 1996). The rest of the critical residues in D5 are near the two-nucleotide bulge (Chanfreau & Jacquier, 1994; Abramovitz et al., 1996).

The two steps of the group II intron splicing pathway have long been recognized as being chemically distinct (Peebles et al., 1986; van der Veen et al., 1986; Padgett et al., 1994). The step 1 reaction uses the internal 2′-OH at the branch site A_{880} in D6 to attack the 5′ splice junction (5′SJ) phosphodiester bond, to yield the 2′-5′ phosphodiester bond of the branch. An alternative pathway initiates splicing by hydrolysis at the 5′SJ in vitro (Jarrell et al., 1988b) and in vivo (Podar et al., 1997). The terminal 3′-OH of exon 1 (E1), after release by branching or hydrolysis, is required by the step 2 reaction to attack the phosphodiester bond at the 3′SJ, to yield the 3′-5′ phosphodiester bond connecting E1 to exon 2 (E2) (Peebles et al., 1986; van der Veen et al., 1986; Jarrell et al., 1988b). One currently favored model for the group II intron ribozyme invokes a single active center that sequentially attains the two distinct conformational states needed to perform these two different reactions (Chanfreau & Jacquier, 1994, 1996). D5 evidently takes part in both states of the ribozyme active site, because D5 is required for both reactions of splicing (Jarrell et al., 1988a, Chanfreau & Jacquier, 1994).

Group II intron active site elements might be embedded in secondary structure domains other than D5 or displayed by the joining segments between the folded domains of the intron. Candidates for active site residues outside of D5 might be recognized as conserved elements of primary sequence where base substitution and chemical modification severely impair splicing. Furthermore, it seems likely that genuine active site components of the intron, including the critical bases in D5, will be near each other and also near the sites where the chemical reactions of splicing occur in the three-dimensional structure of the active form of the intron. [These splicing reaction sites include the 5′SJ and 3′SJ plus A_{880}, the branchpoint in domain 6 (D6).] Therefore, we sought to find elements that are close to D5 within the folded group II intron as potential ribozyme active site residues or as likely binding partners for unidentified D5 tertiary interactions.

To pursue this goal, photocrosslinking was performed on acceptor RNA transcripts containing exon 1 and the first three intron domains (E1:D123) from donor D5 RNAs labeled with the nucleoside analogue 4-thio-uridine (4sU). UV light near 330 nm is absorbed by donor RNAs containing 4sU, and excitation of 4sU leads to covalent crosslinking to suitable nearby acceptors, including all four of the natural bases of RNA, by a free-radical mechanism (reviewed in Favre, 1990; Favre & Fourrey, 1995).

FIGURE 1. Intron-related and domain 5 transcripts for photocrosslinking. **A:** Diagram of the secondary structure of the group II intron aI5γ of yeast mtDNA. The major substructures (D1, D2, D3, D4, D5, and D6) and the two exons (E1 and E2) are labeled on this diagram of a full-length group II intron self-splicing transcript (E1:D123456:E2). Additional details are illustrated in panel D. **B:** Diagram of RNA molecules used in *trans*-reactions and photocrosslinking experiments. This pair of transcripts, one RNA with D5 deleted (ΔD5; Dib-Hajj et al., 1993) and a complementing D5 RNA ($D5^{58}$) (Jarrell et al., 1988b), illustrates one system for *trans*-reactions and photocrosslinking. **C:** Diagram of RNA molecules used in *trans*-reactions and photocrosslinking experiments. This pair of transcripts, an RNA with E1 and part of the intron (E1:D123) and a complementing $D5^{58}$ RNA, represents another system used for *trans*-reactions and photocrosslinking (Jarrell et al., 1988b). **D:** Diagram of joiner segments between domains, the sequence of J(23), and the $\gamma \cdot \gamma'$ interaction. This drawing emphasizes the joining segments [J(E1D1, J(12), J(23), J(34), J(45), J(56), and J(D6E2)] (reviewed in Michel et al., 1989) between the domains. The 5′ and 3′ splice junctions ("5′SJ" and "3′SJ") are marked with small open circles to delimit the E1 and E2 ("**E1**" and "**E2**") segments of the transcript. The primary sequence for the J(23) region of aI5γ (5′-AAAAGAUU) is centered in the diagram. J(23) contains A_{587} ("γ," shown as "A" in outline font), the residue that forms a Watson–Crick base pair with U_{887} ("γ'," shown as "U" in outline font), the last base of this intron (Jacquier & Michel, 1990). Highly conserved nucleotides G_{588} and A_{589} (shown in bold font, "**G**" and "**A**") in J(23) are marked with asterisks to indicate that they were defined in this study as major targets of photocrosslinking from various D5 donors. Numbering of intron positions is relative to G_1 at the 5′ end of wild-type aI5γ. **E:** Primary sequences and secondary structures of the principal D5 photocrosslinking donors. Sequence locations that were substituted as 4sU in D5 donor transcripts are marked in bold ("**u**" or "**U**"); uppercase lettering in the donors ($D5^{58}$, $D5^{36}$, $D5^{35}$, $D5^{31a}$, and $D5^{31b}$) emphasizes the residues of the helical D5 substructure itself and lowercase lettering identifies leader and trailer segments. Those sites shown in this report to serve as major donors of photocrosslinking are marked with asterisks. The superscript designations ("58," "36," "35," "31a," and "31b") signify the length and specific type of D5 derivative. The numbering of D5 transcripts is relative to the G at the 5′ end of wild-type D5 (for example, G_{-2}, U_{-1}, G_1, A_2, G_3, …, U_{32}, U_{33}, C_{34}, U_{+1}, U_{+2}, …). G_1 of D5 is G_{815} of aI5γ. Both $*D5^{31a}$ and $*D5^{31b}$ were made from $D5^{30}$, a variant lacking base pairs $G_1 \cdot C_{34}$ and $A_2 \cdot U_{33}$ from the minimum-size active D5 (Peebles et al., 1995). **F:** Reaction pathway leading to crosslinked products. The diagram shows the reaction pathway for D5-catalyzed 5′ splice junction hydrolysis (5′SJ hydrolysis) of E1:D123 along with the photocrosslinking process induced by UV-light irradiation of *D5 variants containing 4sU. The pathway emphasizes that E1:D123 RNA photocrosslinks to radioactive 4sU-containing D5 (*D5) to yield a radioactive adduct (E1:D123x*D5). The intron product of 5′SJ hydrolysis (D123) photocrosslinks to yield another radioactive adduct (D123x*D5). The alternative pathway shows that photocrosslinking before 5′SJ hydrolysis generates the same products. A second cleavage reaction occurs after position 567 in D2 of D123 (Koch et al., 1992), following the sequence (5′-AAUUUUC) that resembles IBS1 of aI5γ (5′-CAUUUUC). The asterisk marks the radioactivity in 5′-^{32}P-labeled 4sU-D5. The diagram is not drawn to scale.

Numerous intramolecular RNA crosslinks and intermolecular adducts have been characterized for RNA–RNA and RNA–protein interactions (reviewed by Favre, 1990; Favre & Fourrey, 1995; for example, Tanner et al., 1988; Wyatt et al., 1992; Sontheimer & Steitz, 1993). Crosslinked products derived from 4sU-containing RNAs are often sufficiently stable to recover after gel electrophoresis, permitting their further analysis. The donor and acceptor sites are mapped by deoxyoligonucleotide-directed RNase-H digestion, primer extension by reverse transcriptase (RTase), or direct RNA sequencing (for example, Lemaigre-Dubreuil et al., 1991; Woissard et al., 1994; Podar et al., 1995). The 4sU-derived photocrosslinks are interpreted as close interactions, because the size of 4sU limits its range to a few angstroms. Thus, we set out to catalogue the set of nucleotides near D5 donors of photocrosslinking, with the expectation that these acceptors would include splicing reaction sites and possibly other active-site elements. The splicing reaction sites are readily recognizable, of course, whereas other tests would be needed to determine whether any additional residues identified by photocrosslinking to D5 might actually serve specific functions that are important for splicing.

We report here the results of photocrosslinking between the partners of a *trans*-reaction system consisting of intron-containing acceptor transcripts and D5 donor RNAs labeled with 4sU and radioactivity (Fig. 1). The prototype *trans*-reaction system used an intron transcript with E1 and intron domains 1–3 (E1:D123) plus a separate D5 RNA; E1:D123 undergoes 5′SJ hydrolysis promoted by a functional D5 during incubation under suitable reaction conditions (Jarrell et al., 1988a). Here, D5 was labeled with an average of less than one 4sU per molecule by incorporating a moderate fraction of 4sUTP during transcription (Lemaigre-Dubreil et al., 1991). This allows each D5 position that is normally a U to be surveyed as a potential 4sU photocrosslinking donor to suitable nearby acceptors, whereas the limited 4sU content of D5 restricts photocrosslinking to single events attaching one D5 to an individual E1:D123 molecule. We also constructed and analyzed D5 donors with single 4sU residues incorporated at unique positions. In this case, every photocrosslink represents this unique 4sU label on D5 reacting with one acceptor residue. We find that the principal acceptor sites of photocrosslinking are the highly conserved intron nucleotides G_{588} and A_{589}, located in the joining segment between domains 2 and 3 [J(23)], and the main D5 donor sites are the 4sU residues in or near the lower helix of D5. Because J(23) participates in the $\gamma\cdot\gamma'$ tertiary interaction that helps to define the 3′SJ, these results indicate that D5 binds quite near the 3′SJ. The view that D5 serves directly as part of the active center within group II intron ribozymes is buttressed by the proximity of D5 to this reaction substrate.

RESULTS

Photolabeled D5 RNAs crosslink to E1:D123 and form adducts that retain ribozyme activity

The nucleotide analogue 4sUTP was randomly incorporated into 58-mer transcripts of D5 ($D5^{58}$) to make a mixed photocrosslinking donor (4sU-$D5^{58}$) with less than one 4sUMP residue per $D5^{58}$ molecule. This 4sU-$D5^{58}$ was radioactively end-labeled prior to crosslinking (*$D5^{58}$; the asterisk denotes a radioactively labeled crosslinking donor). Thirteen different isomers with just one 4sU are in this *$D5^{58}$ mixture (Fig. 1). This *$D5^{58}$ also contains small amounts of many multiply substituted 4sU-$D5^{58}$ isomers and some unsubstituted $D5^{58}$. The *$D5^{58}$ was incubated under splicing reaction conditions with unlabeled E1:D123 and irradiated with long-wavelength UV. After separation, several crosslinked adducts were detected by the radioactivity from *$D5^{58}$ (Fig. 2A). These crosslinked RNAs migrated much slower than *$D5^{58}$ and some were slower than E1:D123. Crosslinked adducts to E1:D123 (E1:D123x*$D5^{58}$; branched, 1,066 nt) should migrate slower than E1:D123 (linear, 1,008 nt), and adducts to splicing product D123 (D123x*$D5^{58}$; branched, 770 nt) should migrate slower than D123 (linear, 712 nt).

The appearance of these low-mobility radioactive RNAs depended absolutely on three conditions: (1) incorporating 4sU into the D5 donor, (2) adding intron-related acceptor transcripts, and (3) exposing the RNA to UV light for at least a few minutes (data not shown). The final yield and pattern of the photocrosslinked adducts was affected by the time and intensity of UV-light exposure, the percentage of 4sU incorporated into *$D5^{58}$, the incubation temperature, and the splicing reaction conditions provided before, during, and after UV-light irradiation. Furthermore, photocrosslinking requires sufficient magnesium salt to support the *trans*-reactions (data not shown). This is not surprising, because photocrosslinking must occur in the noncovalent E1:D123·D5 complex, and D5 binding by E1:D123 transcripts requires magnesium cations (Pyle & Green, 1994).

Several adduct fractions ($X1^{58}$–$X5^{58}$) were recovered from a preparative reaction that was done in splicing buffer with NH_4Cl (Fig. 2A). The more slowly migrating adducts ($X1^{58}$ and $X2^{58}$) probably consist primarily of isomers of E1:D123x*$D5^{58}$, because $X1^{58}$ and $X2^{58}$ were the lowest-mobility adducts, and they were relatively abundant after crosslinking in a splicing buffer with $(NH_4)_2SO_4$ that supports little 5′SJ hydrolysis in this *trans*-reaction system. Also, the $X1^{58}$ and $X2^{58}$ adducts were rare after crosslinking in a splicing buffer with KCl that supports an efficient 5′SJ hydrolysis reaction (Fig. 2A; Jarrell et al., 1988b). Based on the product formation profiles, the more-rapidly migrating adducts ($X3^{58}$, $X4^{58}$, and $X5^{58}$) probably contain isomers of D123x*$D5^{58}$ (Fig. 2A), because they were abundant after crosslinking in the splicing buffer with

KCl and rare after crosslinking in the splicing buffer with $(NH_4)_2SO_4$.

The prominent crosslinked species ($X1^{58}$–$X5^{58}$) were recovered and tested for self-cleaving activity to see whether they could achieve the active conformation (Fig. 2B). The $X1^{58}$ and $X2^{58}$ adduct fractions probably contain E1:D123x*$D5^{58}$, where crosslinking has occurred to the intron part of the transcript, because these adducts mostly release large radioactive products after incubation under splicing conditions (Fig. 2B). Moreover, the splicing reaction products of crosslinked species $X1^{58}$ and $X2^{58}$ migrate with other crosslinked species, including $X3^{58}$–$X5^{58}$. The $X3^{58}$–$X5^{58}$ species thus appear to contain forms of D123x*$D5^{58}$. The reactive adducts in fractions $X1^{58}$–$X5^{58}$ also release radioactive products that migrate faster than E1 (Fig. 2B). Apparently, the intron-derived D123x*$D5^{58}$ products undergo a second cleavage in D2 at C_{567} after a sequence that resembles IBS1 (5′SJ* cleavage) (Koch et al., 1992). Because 5′SJ* cleavage in D2 produces a short ′D23 fragment (linear, 145 nt), D123x*$D5^{58}$ species could release a labeled ′D23x*$D5^{58}$ fragment (branched, 203 nt). Because prominent reaction products are evident at this mobility, particularly for the D123x*$D5^{58}$ species, most of the *$D5^{58}$ crosslinking has evidently targeted the ′D23 region (the 3′-terminal 145 nt) of the E1:D123 transcript.

Mapping the sites of photocrosslinking between D5 and E1:D123

The photocrosslinking donor sites in the *$D5^{58}$ segment of the E1:D123x*$D5^{58}$ species were mapped in parallel with the input *$D5^{58}$ by partial hydrolysis with alkali (Fig. 2C). The five different fractions of E1:D123x*$D5^{58}$ reveal at least three different sites of crosslinking relative to *$D5^{58}$, all in the single-stranded extensions close to the base of *$D5^{58}$ (see Fig. 1 or Fig. 5A). The two most active photocrosslinking donor sites were those immediately adjacent to the first and last nucleotides of D5; the other detected site was three residues away in the 5′ leader. This result shows that the 4sU residues in the single-stranded extensions of *$D5^{58}$ are positioned near the crosslinking acceptor intron transcript. By implication, the attached lower helix of D5 is near the intron acceptor transcript as well. In contrast, the 4sU residues throughout the rest of *$D5^{58}$ are much less active in *trans* as crosslinking donors, perhaps because they are positioned outside the folded acceptor transcript or are constrained to photocrosslink within *$D5^{58}$ itself.

Next, RTase primer extension on the intron segment of the E1:D123x*$D5^{58}$ adducts was used to identify crosslinking acceptor sites on the E1:D123 transcript. The most prominent sites were in the 3′ half of J(23), the unstructured joiner between D2 and D3 (Fig. 2D; see also Fig. 1). The strongest stop for RTase aligns with the DNA sequencing product for A_{589} of aI5γ; thus, the main crosslinking acceptor site was assigned as G_{588}. Although G_{588} was the most prominent site, some crosslinking also targeted the adjacent residues A_{587} and A_{589} in J(23). The simplest interpretation of these results is that *$D5^{58}$ binds with its single-stranded extensions near J(23) to allow crosslinking to G_{588} and its adjacent residues. (These results are summarized schematically in Fig. 5A.) It appears that the most effective photocrosslinking donors are the 4sU residues in the single-stranded regions of *$D5^{58}$, as predicted from studies of the crosslinking performance of several other structured RNAs (Favre & Fourrey, 1995).

A smaller D5 transcript crosslinks from its lower helix to J(23) of E1:D123

To get crosslinks from donor sites in the helical parts of D5, 4sUTP was incorporated into a 36-mer D5 RNA ($D5^{36}$) by transcribing an oligonucleotide template (Franzen et al., 1993) to make a mixture of 4sU-$D5^{36}$ crosslinking donors. This 4sU-$D5^{36}$ mixture contains seven distinct singly substituted isomers plus small amounts of several multiply substituted forms and some unsubstituted $D5^{36}$ (see Fig. 1E). This type of D5 has no unpaired 4sU residues, so the extent of photocrosslinking was expected to be lower than for $D5^{58}$. The 5′-^{32}P-end-labeled 4sU-$D5^{36}$ (*$D5^{36}$) was photocrosslinked to E1:D123 RNA in splicing reaction buffer, and two main crosslinked adducts appeared (Fig. 3A). Both *$D5^{36}$ adducts migrated much slower than *$D5^{36}$ itself, with crosslinked species $X1^{36}$ migrating somewhat slower than E1:D123 and $X2^{36}$ migrating slower than D123. The requirements for crosslinking were essentially the same as for *$D5^{58}$ (data not shown). As expected, the extent of photocrosslinking from *$D5^{36}$ was far less than from *$D5^{58}$. (Compare Fig. 3A to Fig. 2A; these images are from the same photographic exposure of the same gel with *$D5^{36}$ and *$D5^{58}$ donors labeled to similar specific radioactivities and for trials done in parallel with the same E1:D123.)

The two main *$D5^{36}$ adducts, $X1^{36}$ and $X2^{36}$, were isolated from a preparative reaction (Fig. 3A) and tested for self-cleaving activity. The slower-mobility crosslinked *$D5^{36}$ adduct $X1^{36}$ reacted to give a product that migrated with the faster-migrating *$D5^{36}$ crosslinked adduct $X2^{36}$, whereas $X2^{36}$ itself was largely unreactive (Fig. 3B). These results suggest that $X1^{36}$ is E1:D123x*$D5^{36}$, whereas $X2^{36}$ is D123x*$D5^{36}$. To confirm this assignment, photocrosslinking of *$D5^{36}$ to the acceptor transcript D123 was done; this directly generates a photocrosslinked D123x*$D5^{36}$. Because directly made D123x*$D5^{36}$ migrates with $X2^{36}$ and is unreactive, like $X2^{36}$ (data not shown), our provisional assignment for $X2^{36}$ is supported (see below for a further discussion of these species). Significantly, the successful reaction of E1:D123x*$D5^{36}$ shows that *$D5^{36}$

remains in the normal D5 binding site. Because no labeled reaction products migrated close behind E1, and because the reaction product derived from X1[36] comigrates with the directly made D123x*D5[36], the *D5[36] is apparently crosslinked to the intron part of the E1:D123 acceptor transcript (Fig. 3B).

RNA sequencing was done to map the donor sites in *D5[36] (Fig. 1C). E1:D123x*D5[36] consists mostly of photocrosslinks originating from the $U_{33} \rightarrow$ 4sU isomer in the *D5[36] population (Fig. 3C). The uniformity of this alkaline hydrolysis ladder suggests that few if any other sites within *D5[36] contribute significantly to the major crosslinking adduct X1[36]. The same analysis was done for X2[36] and for directly made D123x*D5[36] with the same result (data not shown). These results establish that virtually all of the crosslinks were donated by $U_{33} \rightarrow$ 4sU of *D5[36] to the various acceptor RNAs.

Primer extension with RTase identified the crosslinking acceptor sites within the intron regions of the crosslinked adduct E1:D123x*D5[36]. Several oligonucleotide primers were used in a complete survey (data not shown), but only one region reproducibly gave prominent stops for RTase on the crosslinked product (Fig. 3D). The strongest RTase stop aligned with the DNA sequencing product for A_{589} of aI5γ, so the crosslink acceptor residue was assigned to position G_{588}. Less prominent RTase stops occurred at DNA sequence positions G_{588}, T_{590}, and T_{591}; those secondary primer extension stops probably identify minor forms of the crosslinked adduct with *D5[36] crosslinked to A_{587}, A_{589}, or U_{590}, three residues near the major site at G_{588} (Fig. 3D; diagrammed in Fig. 5B). These results show that $U_{33} \rightarrow$ 4sU in the lower helix of D5 is close to the 3′ half of J(23) in the correctly folded ribozyme.

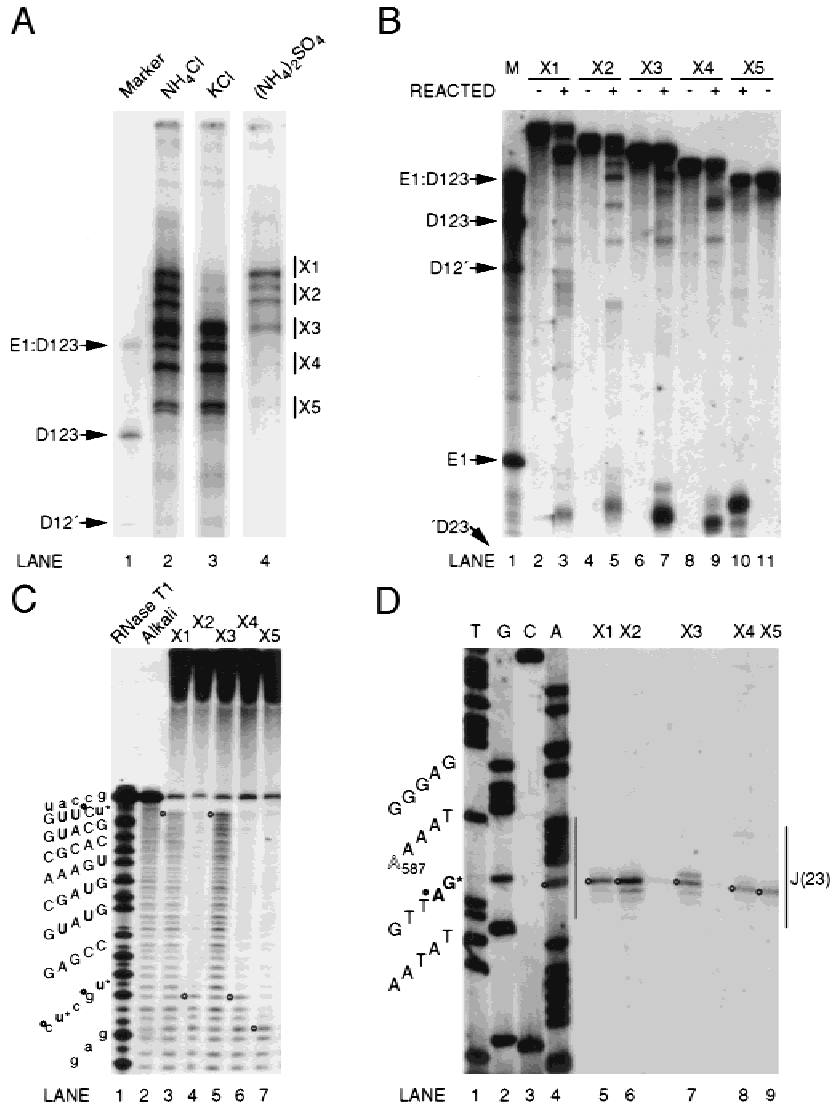

FIGURE 2. (*Legend on facing page.*)

Crosslinking from sequence variant D5 crosslinkers orients J(23) of E1:D123 relative to D5

More crosslinking experiments were done with point-mutant sequence variants of *D5^{36} (Peebles et al., 1995). This approach was used to establish unequivocally that the *D5^{36} crosslinking originated from $U_{33} \rightarrow$ 4sU, instead of the adjacent residue $U_{32} \rightarrow$ 4sU. Active variants $U_{32} \rightarrow$ A-*D5^{36} and $U_{33} \rightarrow$ A-*D5^{36} were made as photocrosslinking donors (Fig. 4). For the $U_{32} \rightarrow$ A-*D5^{36} variant, 4sU-dependent photocrosslinking cannot arise from the A at position 32, whereas crosslinking from the 4sU at position 33 should be unaffected. Note that the $U_{32} \rightarrow$ A-substitution has only a small effect on D5 function (Peebles et al., 1995). The $U_{32} \rightarrow$ A-*D5^{36} was photocrosslinked and analyzed as described for wild-type *D5^{36} (Fig. 4). The $U_{32} \rightarrow$ A substitution did not reduce the extent of photocrosslinking; instead, the crosslinking efficiency seemed to increase somewhat. The adduct E1:D123x$U_{32} \rightarrow$ A-*D5^{36} retained some hydrolysis activity to yield a product that comigrates with X2^{36} that may be D123x$U_{32} \rightarrow$ A-*D5^{36} (Fig. 4A). Furthermore, the crosslinking donor site remained at $U_{33} \rightarrow$ 4sU in $U_{32} \rightarrow$ A-*D5^{36} (Fig. 4B), whereas the primary crosslinking acceptor site stayed at G_{588} (Fig. 4C; see also the schematic diagrams of Fig. 5B,C).

Because $U_{33} \rightarrow$ 4sU is the major crosslinking donor in wild-type *D5^{36}, introducing the $U_{33} \rightarrow$ A substitution will abolish 4sU-dependent crosslinking from D5 position 33. Then, detectable photocrosslinking could originate from $U_{32} \rightarrow$ 4sU or from some other 4sU position in D5. Note that the $U_{33} \rightarrow$ A replacement reduces D5 catalytic function significantly, despite somewhat improved binding to the transcript (Peebles et al., 1995). The $U_{33} \rightarrow$ A-*D5^{36} variant was crosslinked and analyzed as before (Fig. 4). Surprisingly, the photocrosslinking efficiency was not reduced by the $U_{33} \rightarrow$ A change. However, the adduct E1:D123x$U_{33} \rightarrow$ A-*D5^{36} showed no self-cleavage activity (Fig. 4A); this result was not surprising, because the function of this weakly active mutant form of D5 might be further impaired as a consequence of crosslinking from a residue so close to the crucial G_3. As expected, the crosslinking donor site shifted away from the wild-type D5 position at $U_{33} \rightarrow$ 4sU; the new donor site was $U_{32} \rightarrow$ 4sU in $U_{33} \rightarrow$ A-*D5^{36} (Fig. 4B). The major crosslinking acceptor site simultaneously shifted from G_{588} for wild-type *D5^{36} to A_{589} for $U_{33} \rightarrow$ A-*D5^{36} (Fig. 4C; summarized in Fig. 5B,D).

These results establish convincingly that the major photocrosslinking donor site for wild-type *D5^{36} is $U_{33} \rightarrow$ 4sU. Furthermore, these findings verify that it is $U_{33} \rightarrow$ 4sU of wild-type *D5^{36} that crosslinks to G_{588} of E1:D123, because the crosslinking acceptor site on the intron shifts by one base and that change is accompa-

---

FIGURE 2. Analysis of photocrosslinking between E1:D123 and D5^{58}. **A:** Photocrosslinking reactions. Radioactive, 5′-end-labeled 4sUTP-containing D5^{58} donor transcripts (*D5^{58}) were incubated with unlabeled E1:D123 at 45 °C in various splicing reaction buffers for 20 min before UV-light irradiation and crosslinking for 20 min (lanes 2, 3, and 4). A *trans*-reaction with randomly labeled radioactive E1:123 plus D5^{36} is shown as a marker (lane 1); this sample was incubated at 45 °C for 30 min, but it was not exposed to UV light. The reactions were done in splicing buffers with 1 M NH_4Cl (lanes 1 and 2), 1 M KCl (lane 3), or 1 M $(NH_4)_2SO_4$ (lane 4). Five adduct fractions ("X1" to "X5") were resolved after a preparative crosslinking reaction and recovered as shown in lanes 2, 4, 6, 8, and 11 of panel B. E1:D123 is the target transcript, D123 is the intron product of 5′SJ hydrolysis, E1 is the exon product of 5′SJ hydrolysis, D12′ and ′D23 are the 5′ and 3′ intron products of site-specific 5′SJ* cleavage in D2 (Koch et al., 1992). X1^{58} and X2^{58} are mostly E1:D123x*D5^{58}, and X3^{58}–X5^{58} are mostly D123x*D5^{58} (see Fig. 1F). **B:** Splicing-related reactions by crosslinked RNAs. The five crosslinked fractions ("X1" to "X5") from a large-scale reaction were incubated at 45 °C in splicing reaction buffer with 1 M NH_4Cl for either 0 min (lanes 2, 4, 6, 8, and 11) or 45 min (lanes 3, 5, 7, 9, and 10). The radioactivity in lanes 2–11 originated from the *D5^{58} donor of crosslinking. The marker (M, lane 1) is the same as panel A, lane 1, and the labels are defined there. **C:** Mapping the photocrosslinking donor sites of the *D5^{58} adducts by alkaline hydrolysis. Fractions X1^{58}–X5^{58} (lanes 3–7) were digested in parallel with *D5^{58} (lane 2) with alkali (90 °C for 5 min in 0.1 M Na_2CO_3 pH 9.5 ("Alkali"). The RNase T1 products were made from *D5^{58} (lane 1) with 10 units of RNase T1 (10 min at 50 °C in 4 M urea, 50 mM sodium acetate, pH 5.0; "RNase T1"). The smallest missing bands (marked by asterisks in the sequence at the left) in the alkaline hydrolysis ladders at the discontinuities found for adducts X1^{58}–X5^{58} (lanes 3–7) mark the major D5 donor sites present in these fractions, specifically U_{+1} of X1^{58}, U_{-1} of X2^{58}, U_{+1} of X3^{58}, U_{-1} of X4^{58}, and U_{-4} of X5^{58}. The largest retained bands are marked on the photograph and on the sequence with small circles. The equivalent of about 10 nt near the 5′ end of *D5^{58} was cropped from the bottom of the photograph, because this region contained no features important to this analysis. Note that the 5′-proximal *D5^{58} fragments bear the radioactive label and migrate normally through the gel; the distal *D5^{58} fragments remain attached to the much-larger heterogeneous fragments derived from E1:D123 that migrate slowly through the gel. Fractions in lanes 3, 4, 5, 6, and 7 contain mixtures of crosslinked species (see panel A), but the most abundant components are E1:D123x*D5^{58}, E1:D123x*D5^{58}, D123x*D5^{58}, D123x*D5^{58}, and D123x*D5^{58}, respectively. **D:** Mapping the *D5^{58} photocrosslinking acceptor sites by primer extension. The adduct fractions X1^{58}–X5^{58} (lanes 5–9) were tested by RTase primer extension with various 5′-^{32}P-oligonucleotides complementary to E1:D123. The only prominent specific differences from the control reaction with E1:D123 transcript were found in the J(23) region (vertical bar) shown here. Calibration markers were DNA sequencing reactions of plasmid pJD-I3′-673 with the same 5′-^{32}P-primer and various termination mixes (lane 1, "T," ddATP; lane 2, "G," ddCTP; lane 3, "C," ddGTP; and lane 4, "A," ddTTP). The DNA sequence near J(23) is listed on the left. The γ nucleotide at position A_{587} is marked ("A" in outline font); position G_{588} is highlighted ("**G**" in bold font plus an asterisk) as the main crosslinking acceptor; position A_{589} is indicated ("**A**" in bold font plus a small circle) as the most prominent stop for RTase.

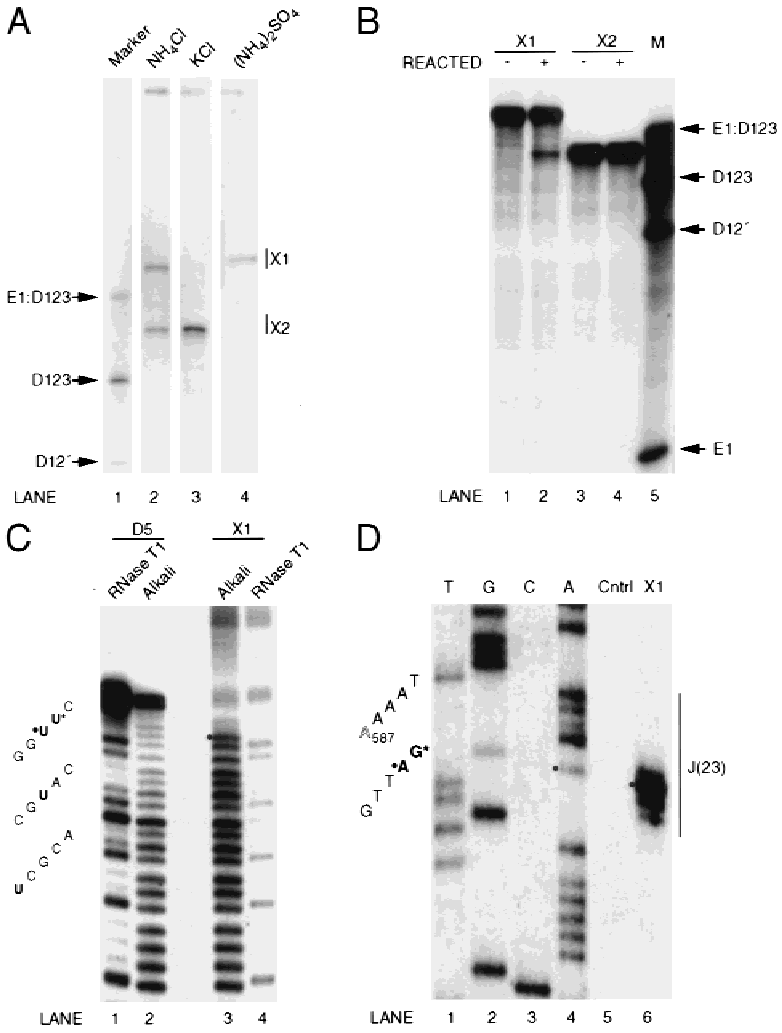

FIGURE 3. Analysis of photocrosslinking between E1:D123 and *D5[36]. **A:** Photocrosslinking reactions. Radioactive, 5′-end-labeled 4sUTP-containing D5[36] donor transcripts (*D5[36]) were incubated with unlabeled E1:D123 at 45 °C in various splicing reaction buffers for 20 min before UV-light irradiation and crosslinking for 20 min (lanes 2, 3, and 4). The marker sample (lane 1) is the same as lane 1 of Figure 2A, and the labels are defined there. The reactions were done in splicing reaction buffers with 1 M NH_4Cl (lanes 1 and 2), 1 M KCl (lane 3), or 1 M $(NH_4)_2SO_4$ (lane 4). Two main adduct fractions ("X1" and "X2") were recovered as shown after preparative crosslinking. X1[36] is E1:D123x*D5[36], and X2[36] is D123x*D5[36]. This panel was assembled from the same photograph and exposure as panel A of Figure 2. **B:** Splicing-related reactions by crosslinked RNAs. Two crosslinked fractions ("X1" and "X2") from a large-scale crosslinking reaction were incubated at 45 °C in splicing reaction buffer with 1 M NH_4Cl for either 0 min (lanes 1 and 3) or 45 min (lanes 2 and 4). In lanes 1–4, the radioactivity originated in the *D5[36] crosslinking donor. The marker sample (lane 1) is the same as lane 1 of Figure 2A, and the labels are defined there. **C:** Mapping photocrosslinking donor sites in the *D5[36] adduct X1[36] by alkaline hydrolysis. Fraction X1[36] (lanes 3 and 4) was digested in parallel to *D5[36] (lanes 1 and 2) with alkali (90 °C for 5 min in 0.1 M Na_2CO_3, pH 9.5 ("Alkali," lanes 2 and 3) or with 10 units of RNase T1 (10 min at 50 °C in 4 M urea, 50 mM sodium acetate, pH 5.0; "RNase T1," lanes 1 and 4). The smallest missing band (marked by an asterisk in the sequence at the left) in the alkaline hydrolysis ladder at the discontinuity found for adduct X1[36] (lane 3) marks the major D5 donor site present in this fraction, specifically U_{33} of D5. The largest retained band is marked on the photograph and on the sequence with a small circle. The equivalent of about 15 nt near the 5′ end of *D5[36] was cropped from the bottom of the photograph, because this region contained no features important for this analysis. The 5′-proximal *D5[36] fragments carry the radioactive label and migrate normally through the gel; the distal *D5[36] fragments remain attached to the much-larger heterogeneous fragments derived from E1:D123 that migrate slowly through the gel. There is a larger-than-usual step between U_{33} and C_{34} of *D5[36] (full-length *D5[36] is marked as C_{34}), because alkaline hydrolysis products have 5′-phosphoryl and 3′-phosphoryl (either 2′-P or 3′-P) ends, whereas the remaining *D5[36] has 5′-phosphoryl and 3′-hydroxyl ends. This reduces the relative charge on *D5[36] and allows it to interact with the borate ions in the TBE gel buffer; both effects reduce the relative mobility of *D5[36]. **D:** Mapping the *D5[36] photocrosslinking acceptor sites by primer extension. Fraction X1[36] (lane 6) was tested by RTase primer extension with several 5′-^{32}P-oligonucleotides complementary to E1:D123. The only prominent specific differences from the control reaction with E1:D123 transcript (lane 5) were found in the J(23) region shown here (vertical bar). Calibration markers were DNA sequencing reactions of plasmid pJD-l3′-673 with the same 5′-^{32}P-oligonucleotide and various termination mixes (lane 1, "T," ddATP; lane 2, "G," ddCTP; lane 3, "C," ddGTP; and lane 4, "A," ddTTP). The DNA sequence near J(23) is listed on the left. The γ nucleotide at position A_{587} is marked ("A" in outline font); position G_{588} is highlighted ("**G**" in bold font plus an asterisk) as the main crosslinking site; position A_{589} is indicated ("**A**" in bold font plus a small circle) as the most prominent stop for RTase.

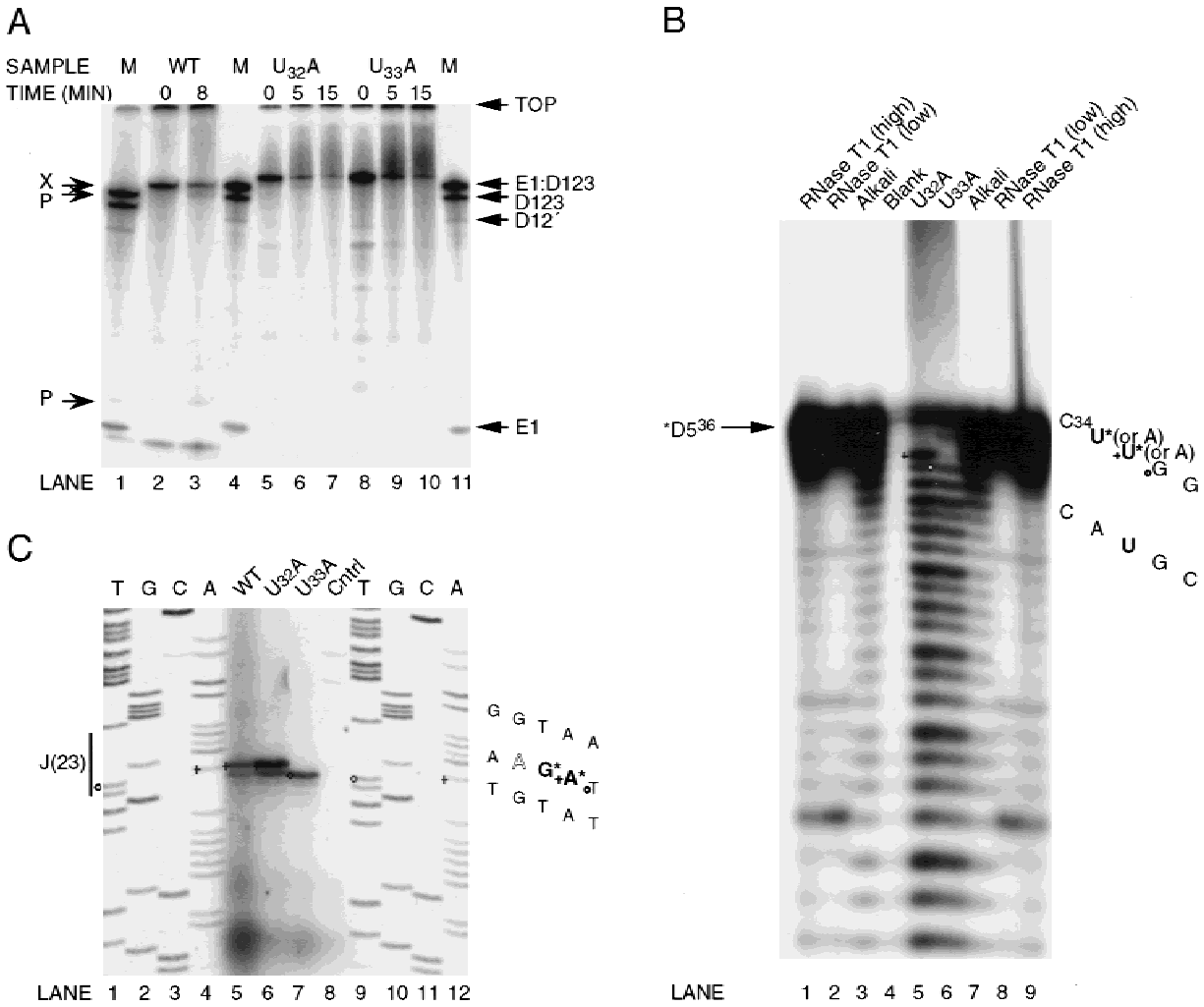

FIGURE 4. Analysis of photocrosslinking to E1:D123 by sequence variants of *D5^{36}. Two active variants of D5^{36} ($U_{32} \rightarrow A$ and $U_{33} \rightarrow A$) were prepared as radioactive photocrosslinking-donor 4sUTP-labeled transcripts ($U_{32} \rightarrow$ A-*D5^{36} and $U_{33} \rightarrow$ A-*D5^{36}). These two variant D5s and wild-type *D5^{36} were incubated with unlabeled E1:D123 at 45 °C in splicing reaction buffer with 1 M NH_4Cl for 5 min before UV-light irradiation and crosslinking for 10 min. **A:** Splicing-related reactions by crosslinked RNAs. The crosslinked adducts corresponding to X1^{36} of D5 variants ("WT" is derived from wild-type D5, "U_{32}A" is from $U_{32} \rightarrow$ A-*D5^{36}, and "U_{33}A" is from $U_{33} \rightarrow$ A-*D5^{36}) were prepared and incubated at 45 °C in splicing reaction buffer with 1 M KCl for 0 min (lanes 2, 5, and 8), 5 min (lanes 6 and 9), 8 min (lane 3), or 15 min (lanes 7 and 10). In lanes 2–3, 5–7, and 8–10, the radioactivity originated from the radioactive, 4sUTP-labeled *D5^{36} transcripts used for crosslinking. The marker sample is a *trans*-reaction of the shortened E1:D123 used for crosslinking, but labeled radioactively and incubated with D5^{36} in splicing reaction buffer with 1 M KCl for 10 min at 45 °C (M, lanes 1, 4, and 11). This E1 is 70 nt, and the intron segment is 688 nt. The positions of the crosslinked adduct (X) and its two main reaction products (P) are marked at the left. The labels (right) refer to the marker and are defined in panel A of Figure 2. **B:** Mapping the donor sites of photocrosslinking in the *D5^{36} adducts by alkaline hydrolysis. The variant *D5^{36} forms of fraction X1^{36} ("U32A" and "U33A," lanes 5 and 6) were digested in parallel to *D5^{36} (lanes 3 and 7) with alkali [90 °C for 2 min in 0.1 M Na_2CO_3, pH 9.5 ("Alkali")] or with RNase T1 [either 1 unit ("RNase T1 (high)," lanes 1 and 9) or 0.1 unit ("RNase T1 (low)," lanes 2 and 8) for 10 min at 50 °C in 4 M urea, 50 mM sodium acetate, pH 5.0]. The smallest missing band (marked by an asterisk in the D5 sequence listed at the left) in the alkaline hydrolysis ladder (lanes 5 and 6) identifies the major D5 donor site present in each fraction, specifically U_{33} of $U_{32} \rightarrow$ A-*D5^{36} and U_{32} of $U_{33} \rightarrow$ A-*D5^{36}. The largest retained band is marked on the picture and sequence with a cross (lane 5) or with a small circle (lane 6). **C:** Mapping the *D5^{36} photocrosslinking acceptor sites by primer extension. The X1^{36} adducts of wild-type and two variants of *D5^{36} were tested by RTase primer extension (lanes 5, 6, and 7) with several 5′-^{32}P-oligonucleotides complementary to E1:D123. The only prominent specific differences from the control reaction with E1:D123 transcript (lane 8) were found in the J(23) region shown here (vertical bar on the left). DNA sequencing reactions of pJD-I3′-673 plasmid were prepared as calibration markers with the same 5′-^{32}P-primer and various termination mixes (lanes 1 and 9, "T," ddATP; lanes 2 and 10, "G," ddCTP; lanes 3 and 11, "C," ddGTP; and lanes 4 and 12, "A," ddTTP). The DNA sequence in the vicinity of J(23) is listed on the right. The γ nucleotide at position A_{587} is marked ("A" in outline font); position G_{588} is highlighted ("**G**" in bold font plus an asterisk) as the main crosslinking acceptor from U_{33} of wild-type (lane 5) and $U_{32} \rightarrow$ A-*D5^{36} (lane 6); and position A_{589} is indicated ("**A**" in bold font plus an asterisk) as the main crosslinking acceptor from U_{32} of $U_{33} \rightarrow$ A-*D5^{36} (lane 7). The strong stops at position A_{589} are marked with a cross (lanes 4, 5, 6, and 12, wild-type *D5^{36} and $U_{32} \rightarrow$ A-*D5^{36}) on the sequence and picture. The strong stops at position U_{590} are marked with a small circle (lanes 1, 7, and 9; $U_{33} \rightarrow$ A-*D5^{36}) on the sequence and picture.

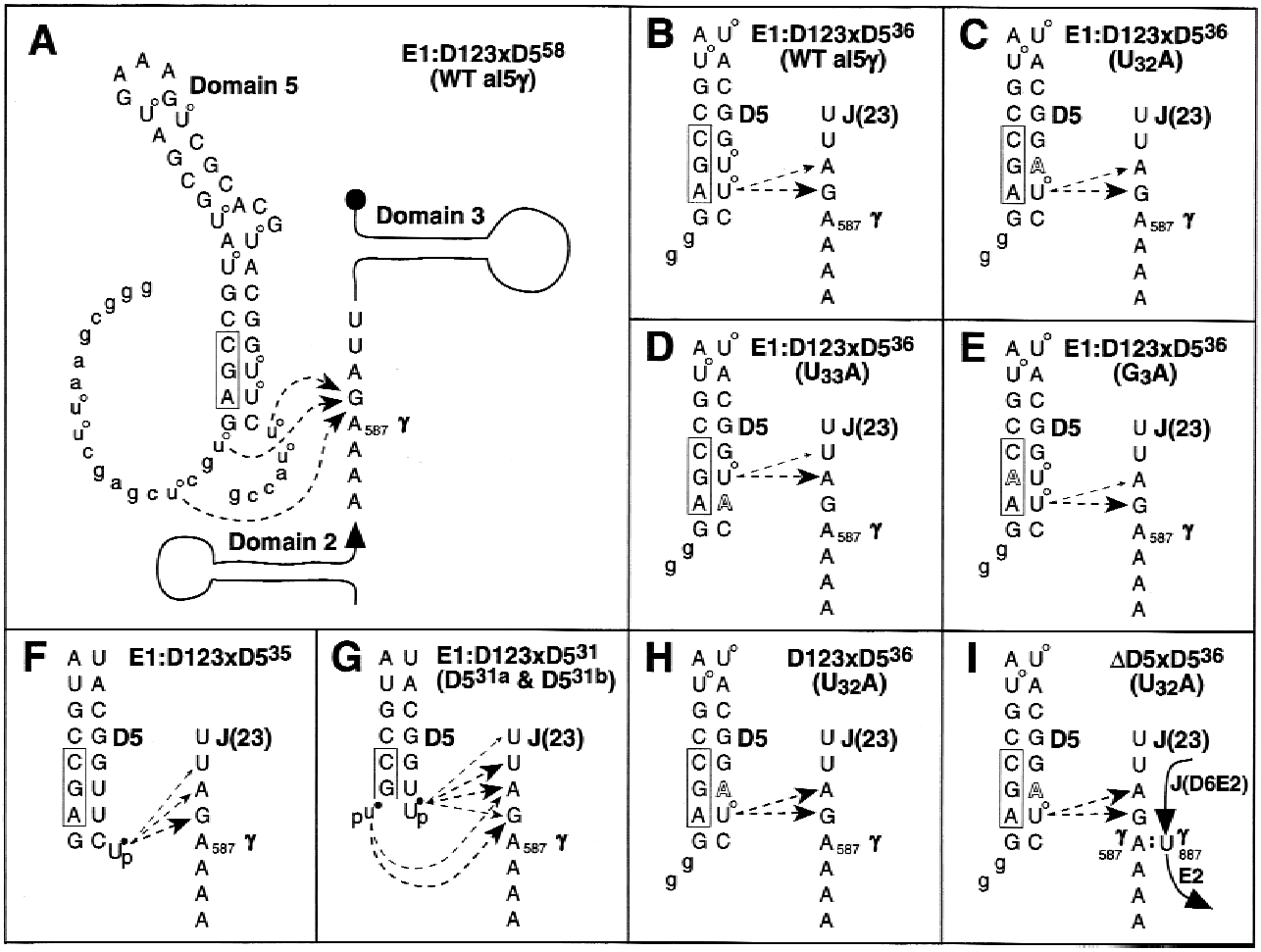

FIGURE 5. Photocrosslinking by D5 to conserved nucleotide residues in J(23). This diagram illustrates the photocrosslinks from D5 to J(23) defined by this study. In each panel, the most prominent photocrosslinks identified from that particular *D5 donor to the J(23) region of the specific acceptor RNA are shown as dashed arrows. **A:** Prominent photocrosslinks from the randomly labeled wild-type *D5^{58} donor to J(23) of the E1:D123 acceptor. The primary sequence of *D5^{58} from al5γ is shown arranged in its secondary structure, with uppercase characters emphasizing the D5 substructure itself and lowercase letters specifying the leader and trailer regions. The primary sequence of the J(23) region is shown attached to a hairpin sketched for D2 with an arrowhead to indicate the 3′ direction or a hairpin for D3 with a bullet to mark the 3′ end. The γ nucleotide A_{587} is numbered and labeled explicitly. **B:** Prominent photocrosslinks from the randomly labeled wild-type *D5^{36} donor to J(23) of the E1:D123 acceptor. **C:** Prominent photocrosslinks from the randomly labeled $U_{32} \rightarrow$ A-*D5^{36} donor to J(23) of the E1:D123 acceptor. **D:** Prominent photocrosslinks from the randomly labeled $U_{33} \rightarrow$ A-*D5^{36} donor to J(23) of the E1:D123 acceptor. **E:** Prominent photocrosslinks from the randomly labeled $G_3 \rightarrow$ A-*D5^{36} donor to J(23) of the E1:D123 acceptor. **F:** Prominent photocrosslinks from the mono-substituted 3′-p4sUp-labeled *D5^{35} donor to J(23) of the E1:D123 acceptor. **G:** Prominent photocrosslinks from the mono-substituted 5′-p4sU-labeled *D5^{31a} and 3′-p4sUp-labeled *D5^{31b} donors to J(23) of the E1:D123 acceptor. **H:** Prominent photocrosslinks from the randomly labeled $U_{32} \rightarrow$ A-*D5^{36} donor to J(23) of the D123 acceptor. **I:** Prominent photocrosslinks from the randomly labeled $U_{32} \rightarrow$ A-*D5^{36} donor to J(23) of the ΔD5 acceptor. This panel emphasizes that the ΔD5 RNA includes J(D6E2), E2, and the 3′SJ with the γ' pairing partner U_{887}. Part of the RNA strand is shown as a curved line, with the arrowheads indicating the 3′ direction. For panel A, the sequences of the entire *D5^{58} donor and for J(23) are shown, whereas for panels B–I, the primary sequences of the lower helix of *D5^{36} and for J(23) from al5γ are shown. For all panels, uppercase characters indicate residues in the D5 substructure itself or in J(23) and lowercase letters specify the leader and trailer regions of D5; letters in outline font in the D5 sequence denote specific substitutions introduced to limit the potential sites for photocrosslinking 4sU donor residues, for example. For all panels, a heavier line indicates a more intense crosslink, and a thinner line indicates a less intense crosslink; open circles mark the U positions that are partly substituted with 4sU; solid circles mark the fully occupied 4sU positions; the rectangle on D5 highlights the catalytically crucial trinucleotide of D5.

nied by a shift of the donor site by one base. By reversing this same argument, it also becomes quite clear that $U_{32} \rightarrow$ 4sU of $U_{33} \rightarrow$ A-*D5^{36} crosslinked to A_{589}. In addition, these data suggest that the 3′ half of J(23) near D5 may be arranged with a definite polarity, so that G_{588} is closer to $U_{33} \rightarrow$ 4sU of D5, whereas A_{589} is closer to $U_{32} \rightarrow$ 4sU (Fig. 5).

Next, the sequence variant $G_3 \rightarrow$ A-*D5^{36} was prepared as a crosslinking donor. This mutant is catalytically inactive for D5 function in *trans*-splicing, but it

nevertheless occupies the normal intron binding site for D5, because $G_3 \rightarrow$ A-*D5[36] is known to be a competitive inhibitor (Peebles et al., 1995). The extent of photocrosslinking by $G_3 \rightarrow$ A-*D5[36] to E1:D123 was similar to that of wild-type *D5[36], but only one adduct was made, and it migrated like wild-type E1:D123x*D5[36] (data not shown). As expected, this E1:D123x$G_3 \rightarrow$ A-*D5[36] was inactive during incubation in splicing reaction buffers (data not shown). To map the crosslinking sites from $G_3 \rightarrow$ A-*D5[36], partial alkaline hydrolysis and RTase primer extension analysis of the crosslinking adduct were done, with the same results as for wild-type *D5[36] (data not shown). These findings verify that the inactive variant $G_3 \rightarrow$ A-D5[36] binds to the intron transcript in the same location as wild-type D5[36] (diagrammed in Fig. 5E). This result also shows explicitly that X1[36] most likely is E1:D123x*D5[36] and that X2[36] is probably D123x*D5[36], confirming the identifications inferred above from the condition-of-reaction and precursor-product observations.

Mono-substituted 4sU-D5 RNAs also target J(23) of E1:D123

In part, the pattern of crosslinking from the 4sU-labeled *D5[58] RNA was complex because a few different 4sU residues were active as crosslinking donors. To reduce that complexity while retaining the potent photocrosslinking-donor activity intrinsic to single-stranded 4sU donors, we constructed a mono-substituted crosslinking donor from an active 34-mer D5 transcript by site-specifically attaching 4sU to its 3′ end with T4 RNA ligase. This 4sU-labeled crosslinking donor of 35 nt (*D5[35]) is fully active as D5 and contains the entire helical D5 substructure (Fig. 1E). The crosslinking activity of *D5[35] was tested, and the extent of crosslinking by *D5[35] was greater than for the 4sU-labeled *D5[36] types. The acceptor sites for *D5[35] were in J(23), with the main crosslinking site at G_{588}, but some crosslinking to A_{589} and U_{590} was detectable (Fig. 5F). The mobility and cleavage activity of the *D5[35] adducts otherwise resembled the *D5[36] adducts rather closely (data not shown).

To place photocrosslinking donor 4sU residues as close as possible to the catalytically crucial G_3 residue of D5 (Fig. 1E), two additional crosslinking reagents were made from a truncated version of D5. This 30-mer D5 (D5[30]) lacked two base pairs, specifically $G_1 \cdot C_{34}$ and $A_2 \cdot U_{33}$, and failed to promote 5′SJ hydrolysis, but competed effectively against an active D5 when both were added to the same *trans*-splicing reaction (data not shown). This activity as a competitive inhibitor shows that D5[30] binds to the same site as active D5[36] or at least to an overlapping site.

To make D5[30] as a mono-substituted photocrosslinking donor, primed transcription was done with 5′-4sUpG to yield a 31-mer (4sU-D5[31a]) with a single unpaired 4sU residue replacing A_2 of D5 (Fig. 1E). This 4sU-D5[31a] was 5′-end-labeled on the 4sU itself with ^{32}P to yield radioactive *D5[31a]. A second mono-substituted crosslinking reagent (*D5[31b]) was made from D5[30] by attaching 5′-^{32}P-p4sUp to the 3′ end of D5[30] by T4 RNA ligase. This *D5[31b] has a single, unpaired 4sU at its 3′ end in place of U_{33} (Fig. 1E). Both *D5[31a] and *D5[31b] were catalytically inactive, but both competed as well as their parent compound D5[30]. Both *D5[31a] and *D5[31b] made one crosslinked adduct with E1:D123 at the same mobility as E1:D123x*D5[36]. Both *D5[31a] and *D5[31b] were catalytically inactive, so it was not surprising that both E1:D123x*D5[31a] and E1:D123x*D5[31b] failed to react when incubated in splicing buffers (data not shown). The major crosslinking acceptor sites for E1:D123x*D5[31a] and E1:D123x*D5[31b] were mapped by primer extension; *D5[31a] crosslinked to G_{588} and A_{589}, whereas *D5[31b] crosslinked to A_{589} and U_{590} (Fig. 5G). The finding that both *D5[31a] and *D5[31b] crosslink to J(23) with patterns so similar to D5[58] and D5[35] strongly supports our interpretation that the base of D5 is situated near J(23) in the active structure prior to step 1. This places the three highly conserved base pairs in the lower helix of D5 close to the conserved trinucleotide 5′-$A_{587}G_{588}A_{589}$ in the 3′ half of J(23).

D5 crosslinking donors target J(23) of D123 and ΔD5 transcripts

In order to exclude the possibility that photocrosslinking by D5 donors to J(23) was limited to E1:D123 acceptor transcripts, ΔD5 transcripts were compared with D123 RNA in crosslinking experiments. The $U_{32} \rightarrow$ A-*D5[36] crosslinking donor was used because the apparent efficiency of crosslinking is somewhat higher than for the wild-type *D5[36] crosslinking donor. After UV-light irradiation, the main crosslinked adducts were purified and tested for activity. The adduct D123xU_{32}A-*D5[36] was quite unreactive, as expected for an end product of splicing. The adduct ΔD5xU_{32}A-*D5[36] showed some self-cleavage activity, comparable to E1:D123xU_{32}A-*D5[36] (see Fig. 4A). Next, the crosslinks present in D123xU_{32}A-*D5[36] and ΔD5xU_{32}A-*D5[36] were mapped as described above for E1:D123x*D5[36] by alkaline hydrolysis and RTase primer extension (results diagrammed in Fig. 5H,I). The pattern of crosslinking acceptor sites was virtually identical for D123xU32A-*D5[36], ΔD5xU_{32}A-*D5[36], and E1:D123xU_{32}A-*D5[36], although it appeared that G_{588} and A_{589} were targeted almost equally in D123xU_{32}A-*D5[36] and ΔD5xU_{32}A-*D5[36], whereas G_{588} was targeted more strongly in E1:D123xU_{32}A-*D5[36] than A_{589}. Otherwise, the mobility and other properties of D123xU_{32}A-*D5[36] correspond closely to the properties of the X2[36] species identified by the crosslinking of wild-type *D5[36] to E1:D123 (see Fig. 3); this confirms our designation of X2[36] as D123x*D5[36]. The similarity of the crosslinking patterns identified in D123xU_{32}A-*D5[36] and

E1:D123xU_{32}A-*D5[36] shows that this J(23)xD5 crosslink can form before or after the first step of the reaction, and thus it does not differentiate between the proposed alternative conformations of the ribozyme active site (Chanfreau & Jacquier, 1996). The similarity of the crosslinking locations in E1:D123 and ΔD5 indicates that J(23) is positioned in virtually the same place with respect to D5, whether or not the $\gamma \cdot \gamma'$ tertiary interaction can form. This shows that the interpretations of the E1:D123 crosslinking studies presented here can be generalized to the more complete ribozyme complex represented by the ΔD5·D5 system, and probably to full-length self-splicing transcripts and splicing in vivo as well.

DISCUSSION

We have implemented a photocrosslinking strategy with 4sU derivatives of several D5 sequence variants and structural isomers. This has allowed us to identify for the first time a close approach between D5 and J(23), the joining segment between D2 and D3 in group II intron transcripts (Fig. 5). These crosslinked adducts probably represent an active conformation of the ribozyme, because many of the adducts between D5 and E1:D123 retain detectable site-specific hydrolysis activity. In addition to the single observations shown (Fig. 2B, 3B, and 4A), reaction time courses were done for certain adducts, most notably the wild-type D5 forms of E1:D123x*D5[58] and E1:D123x*D5[36] (data not shown). These kinetic assays show that these photocrosslinked adducts undergo a progressive albeit slow site-specific cleavage reaction that competes with nonspecific RNA hydrolysis. We infer that the specific hydrolysis activity of the photocrosslinked adducts represents a slow reaction by a homogeneous population, rather than a rapid reaction by an active subset of the population. However, the specific hydrolysis reaction by these photocrosslinked adducts does appear to proceed more slowly than a similar reaction by a self-cleaving transcript with a comparable sequence content [for example, an E1:D12345 transcript (see Franzen et al., 1994)].

Because J(23) is partly conserved in primary sequence (Michel et al., 1989), the proximity of J(23) in active complexes to the catalytically critical residues of D5 is quite intriguing. Maybe the J(23) residues G_{588} and A_{589} of aI5γ are highly conserved because they interact directly with highly conserved residues in D5. Because our study relied on separately folded D5 RNAs, this close approach is definitively demonstrated only for the *trans*-reaction. However, it seems likely that D5 is near J(23) in the full-length transcripts used for self splicing, and that this J(23) interaction probably exists when D5 acts within the context of pre-mRNA transcripts in vivo. J(23) is conserved in both length and primary sequence among subgroup IIA introns; J(23) contains four residues that can be summarized as 5′-D2-(A/C)(A/G)GA-D3-3′ (Michel et al., 1989). In contrast, J(23) varies significantly in structure among subgroup IIB introns, ranging from five to nine residues in length with a consensus sequence that can be summarized as 5′-D2-$X_{0\text{-}4}$(A/G)(A/G)GA(C/U)$X_{0\text{-}1}$-D3-3′ (Michel et al., 1989).

The most highly conserved region of J(23) among all group II introns is its 3′ half, beginning with γ, a purine that has been shown to co-vary with γ', the pyrimidine that is the last base of group II introns (Jacquier & Michel, 1990; Michel et al., 1989). Curiously, the functional basis of the preference for purines at γ and for pyrimidines at γ' is not understood. Analysis of site-directed mutations within aI5γ showed that γ (A_{587}) and γ' (U_{887}) interact functionally by Watson–Crick base pairing such that those mutations influence the rate of step 2 of the splicing reaction (Jacquier & Michel, 1990). Mutations at γ or γ' also affect the accuracy of 3′SJ definition (Jacquier & Jacquesson-Breuleux, 1991). Therefore, the γ base within J(23) must be near the 3′SJ in the folded intron to allow base pairing with the γ' base. Next to γ on its 3′ side is the dinucleotide 5′-$G_{588}A_{589}$ in aI5γ, and 5′-GA occurs in this location in virtually all group II introns, suggesting that this highly conserved dinucleotide plays an important role in splicing. Similarly, the nucleotide next to γ on its 5′ side is A_{586} in aI5γ, and this location prefers an A or a C among the subgroup IIA introns, although it is semi-conserved as a purine among subgroup IIB introns. Virtually no other conserved features of length or sequence are discernible in J(23), either 5′ of position 586 or 3′ of position 589 as numbered in aI5γ. So, the important residues in J(23), whatever their functional role(s), are most likely within the tetranucleotide 5′-AAGA spanning positions 586–589 in aI5γ that specifically include the γ position and the main sites of D5 crosslinking.

Apparently, J(23) is dispensable for minimal first-step ribozyme function in vitro, and this conclusion implies that J(23) does not contribute any critical part of the active center of the ribozyme nor any essential tertiary contact within the intron, as shown by activity and binding assays (Michels & Pyle, 1995). Later it was reported that D5 binds at least 50-fold less strongly to E1:D1 than to E1:D123 (Jestin et al., 1997); however, this conclusion is at variance with the previous report (Michels & Pyle, 1995). Thus, it remains unclear whether J(23) plays an important role for D5 binding, in the branching reaction, in step 2 of splicing in vitro, or for splicing in vivo, because Michels and Pyle (1995) examined the hydrolytic step 1 model reaction. In another study (Chanfreau & Jacquier, 1996), modification-interference analysis found no critical residues in J(23) specific for the first step, reporting less-than-twofold step 1 effects for DEPC-modifications to A_{586}, A_{587}, G_{588}, and A_{589}. However, the step 2 interference effects were reported to be as much as four- to sixfold for A_{587} and A_{589}, and two- to fourfold for A_{586}, although the effect was still less than twofold for G_{588} (Chanfreau & Jac-

quier, 1996). These facts suggest that the conserved residues in J(23) may play their main role in step 2 of the splicing reaction. However, Chanfreau and Jacquier (1996) also reported briefly that point mutations at A_{589} affected both steps of splicing, so the functional specificity of the J(23) region for step 2 remains uncertain.

Inspection of our D5 crosslinking data provides one view of the arrangement of the J(23) RNA strand with respect to the RNA chains that make up the lower helix of D5 (Fig. 5). The $D5^{58}$ crosslinking results show that the 4sU residues in the single-stranded leader and trailer segments of the D5 transcript and closest to the base of the D5 helix are the most active for crosslinking, and that their preferred target is G_{588} and its nearest neighbors (Fig. 5A). This result places J(23) very close to the base of D5, but the inherent flexibility of the single strands of this donor precludes a more detailed structural interpretation. The close approach of J(23) to the lower helix of D5 is confirmed by the mono-substituted $D5^{35}$ derivative with a single 4sU at its 3′ end, because that location is equivalent to one of the preferred 4sU donor sites in $D5^{58}$. $D5^{35}$ crosslinks most strongly to G_{588}, but detectably to A_{589} and U_{590} as well (Fig. 5F). The results from crosslinking by the $D5^{31a}$ and $D5^{31b}$ derivatives suggest that J(23) crosses D5 in a path that is approximately perpendicular to the helix axis of D5, so that the 5′ to 3′ direction of J(23) encounters the 5′ strand of D5 before the 3′ strand (Fig. 5G). The crosslinking from the $D5^{31}$ derivatives suggests this geometry because the mono-substituted 5′-end crosslinker $D5^{31a}$ targeted G_{588} and A_{589}, whereas the 3′-end crosslinker *$D5^{31b}$ targeted A_{589} and U_{590}. However, in both of these $D5^{31}$ derivatives, as in $D5^{35}$, the 4sU was introduced without any base pairing partner (Fig. 5G), so these 4sU residues have a significant degree of conformational mobility. Moreover, the $D5^{31}$ crosslinking donors are both inactive D5 RNAs that are missing two base pairs from the lower helix of D5, so we cannot be certain that they bind to J(23) appropriately. That objection does not apply to a straightforward interpretation of the $D5^{58}$ or $D5^{35}$ results, and it is clear that J(23) passes quite close to the lower helix of D5 in those active ribozyme complexes.

A more detailed conformational picture can be drawn from the crosslinking results with the active $D5^{36}$ derivatives, where the donor 4sU residues at position 33 or 32 are within the helix of D5. The D5 helix probably is a modified A-form structure with a deep, narrow major groove and a shallow minor groove. The base paired 4sU residues, offering much less conformational flexibility than the unpaired 4sU donors summarized above, would have their sulfur atom at the 4 position pointing into the major groove. The $G_3 \cdot U_{32}$ wobble base pair might widen the major groove, as seen in the duplex structure of an oligonucleotide (Holbrook et al., 1991), presumably allowing access to the functional groups on U_{33} or U_{32} of D5 in the major groove. We suggest, on the basis of our crosslinking results, that G_{588} and A_{589} of J(23) lie along this widened major groove of D5, at least over the region where the conserved trimer 5′-$A_2G_3C_4$ is paired to 5′-$G_{31}U_{32}U_{33}$. The results of crosslinking with the $D5^{36}$ variants suggest that J(23) may be approximately antiparallel to the 5′-$G_{31}U_{32}U_{33}$ strand of D5; $U_{33} \rightarrow$ 4sU targeted G_{588}, whereas $U_{32} \rightarrow$ 4sU targeted A_{589} (Fig. 5C,D). We favor this latter model based mainly on the $D5^{36}$ crosslinking donors primarily because the $D5^{36}$ donors include functional D5 RNAs and also because some of their crosslinked adducts retain hydrolysis activity.

The same antiparallel direction of J(23) is supported independently by comparing the $D5^{35}$ and $D5^{31b}$ results (Fig. 5F,G). In these D5 donors, the 4sU is attached to the 3′ end, but the 4sU is positioned lower by two base pairs in $D5^{35}$ compared to $D5^{31b}$. Crosslinking is primarily to G_{588} from $D5^{35}$ and to A_{589} from $D5^{31b}$, with this shift again suggesting that J(23) may be oriented antiparallel to the 5′-$G_{31}U_{32}U_{33}$ strand of D5. If J(23) were antiparallel to the 5′-$G_{31}U_{32}U_{33}$ strand of D5, J(23) would then be parallel to the D5 strand with the conserved trimer 5′-$A_2G_3C_4$. Further studies are needed to examine this proposed model for the structural relationship between D5 and J(23).

The crosslinking acceptor sites in the 3′ half of J(23) were all observed to be in or near the moderately conserved trinucleotide sequence 5′-$A_{587}G_{588}A_{589}$ that includes the γ nucleotide A_{587}. Inside the complete folded transcript, J(23) must be close to the 3′SJ by virtue of the $\gamma \cdot \gamma'$ interaction. Because we have shown definitely that D5 is close to J(23), a straightforward extension of this reasoning says that D5 must be close to the 3′SJ as well. Comparable crosslinks were formed between D5 and E1:D123 or between D5 and ΔD5, a pre-mRNA transcript that includes the $\gamma \cdot \gamma'$ base pair, thus verifying that D5 binds near the 3′SJ of the complete ribozyme (Fig. 5C,I). Furthermore, it seems likely that the 3′SJ itself must be very close to the 5′SJ, because the 3′-OH on that end of E1 is destined to attack the phosphodiester bond at the 3′SJ to complete step 2 of splicing. Likewise, the branchpoint A_{880} in D6 must be close to D5 as well, because the 2′-OH on A_{880} must attack the 5′SJ to achieve step 1 of the branching pathway of self splicing. It seems unlikely to us that the splicing substrates or intermediate products within the active site undergo significant translational movements through space during the time that elapses between the two steps of the splicing reaction. Instead, it seems much more likely that all of these components are brought close together prior to the first chemical step by the folding of the intron. By demonstrating that J(23) makes intimate contact with D5, we can infer that D5 is near the 3′SJ, the 5′SJ, and the branchpoint as well. Because we and others have suggested previously that D5 participates directly in the catalytic center of the

group II intron ribozyme (Jarrell et al., 1988a; Koch et al., 1992; Franzen et al., 1993; Chanfreau & Jacquier, 1994; Pyle & Green, 1994; Boulanger et al., 1995; Peebles et al., 1995; Abramovitz et al., 1996), and because it seems necessary for the active center to be close to its substrates in the folded structure, our finding that D5 is near all three substrates of splicing is a most welcome and rewarding outcome.

Finally, it must be considered a possibility that J(23) itself contributes important residues to the intron active site. If J(23) participates in the ribozyme active center, then the best candidates are the two nucleotides in the highly conserved 5′-$G_{588}A_{589}$ element, because this dinucleotide is both conserved and positioned close to the splicing substrates and the conserved residues of D5 in the folded intron. We have evidence that point mutations affecting G_{588} or A_{589} impair self-splicing activity, *trans*-ribozyme assembly, and splicing in vivo (P. Faix, M. Podar, D. Perera, J. Franzen, P. Perlman, & C. Peebles, unpubl. data), establishing the importance of J(23) residues for normal group II intron function. A comparison of the J(23) conserved 5′-AGA trimer with the conserved sequences in the spliceosomal snRNAs finds several matches, but one intriguing match is to the U6 snRNA sequence 5′-ACAGAGA, which is conserved from yeast to human (Brow & Guthrie, 1988), positioned close to the 5′SJ as detected by photocrosslinking (Sawa & Shimura, 1992; Sontheimer & Steitz, 1993), adjacent to helix I (Wolff et al., 1994), and important for splicing as shown by the deleterious effects of several point mutations in this sequence both in vitro and in vivo (Fabrizio & Abelson, 1990; Madhani et al., 1990; Datta & Wiener, 1993). Additional work will be required to determine precisely what role the conserved residues of J(23) actually play in the function and assembly of the group II intron ribozyme active site.

MATERIALS AND METHODS

Plasmid and oligonucleotide transcription templates

The template for the D5 58-mer ($D5^{58}$) was plasmid pJD-I5′-75 cleaved by *Hpa* II; and the template for long E1:D123 (1,008 nt) was pJD-I3′-673 cleaved by *Hin*d III (Fig. 1) (Jarrell et al., 1988b); this E1 is 296 nt, and D123 is 712 nt. A shorter form of E1:D123 (827 nt) was made from plasmid pMP827 cleaved by *Hin*d III; this E1 is 115 nt, and D123 is 712 nt. Plasmid pMP827 was derived from pJD-I3′-673 by cleavage at the unique *Mun* I and *Eco*R I sites to delete 181 bp from the intron-distal E1 portion of the transcription unit. The template for the shortest E1:D123 RNA used here (758 nt) was made by PCR with primers "T7 E1 forward" (5′-TAATA CGACTCACTATAGGTGTTGCCTTAGCTAAC) and "D3 reverse" (5′-GATAGGCATATCACCTATAGTATAAG); this E1 is 70 nt, and D123 is 688 nt. The template for the D123 (688 nt) RNA was made by PCR with primers "T7 D1 forward" (5′-TAATACGACTCACTATAGAGCGGTCTGAAAGTTATC) and "D3 reverse." The template for the 988-nt ΔD5 transcript was a PCR product amplified with primers "T7 E1 forward" and "CP90-305" (5′-GAGGACTTCAATAGTAGTATCCTGC) from the plasmid pJD20-delta5 (Dib-Hajj et al., 1993); this E1 is 70 nt, this ΔD5 intron is 855 nt, and this E2 is 63 nt. For the D5 36-mer ($D5^{36}$), 34-mer ($D5^{34}$), and 30-mer ($D5^{30}$) transcripts, synthetic DNA templates were used as described (Milligan & Uhlenbeck, 1989; Franzen et al., 1993; Peebles et al., 1995).

Transcription

For photocrosslinking, $D5^{58}$ and various $D5^{36}$ RNAs were transcribed with 4sUTP and labeled randomly at a low substitution ratio essentially as described (Lemaigre-Dubreuil et al., 1991). Addition of 4sUTP was adjusted to achieve an average of slightly less than one 4sUMP per D5 transcript by empirically determining the incorporation of 4sU by spectrophotometry. The nominal concentration of 4sUTP ranged from 10% to 25% of the total UTP. The 4sU-$D5^{58}$ or 4sU-$D5^{36}$ mixtures promoted group II intron *trans*-splicing reactions with no significant change in the measured kinetic coefficients (data not shown; Franzen et al., 1993). Highly substituted 4sU-$D5^{36}$ made by transcription with no added UTP had an average of 5.5 4sU residues, a slightly reduced T_m (64 °C), the same K_m, and a fivefold lower k_{cat} than unsubstituted $D5^{36}$ (data not shown; Peebles et al., 1995). $D5^{30}$, $D5^{34}$, and the acceptor transcripts like E1:D123 were made without 4sUTP or radioactive substrate addition. All transcription reactions, with or without 4sUTP, were done essentially as described (Jarrell et al., 1988b; Franzen et al., 1993; Peebles et al., 1995). Dinucleotide-primed transcription of $D5^{30}$ used UpG or 4sUpG (Sigma Chemicals) and reduced GTP as described (Sontheimer & Steitz, 1993).

Radioactive end-labeling and D5 photocrosslinker end-labeling

T4 polynucleotide kinase and RNA ligase were used according to the commercial supplier's instructions to make the various *D5 RNAs with radioactive 5′- or 3′-end labels. 4sU-D5 was 5′-end-labeled by T4 polynucleotide kinase and γ-^{32}P-ATP to make 5′-^{32}P-D5 (*D5). Radioactive 5′-, 3′-nucleoside-bis-phosphates ^{32}P-pUp, ^{32}P-p4sUp, and ^{32}P-pCp were prepared from the 3′ monophosphates with T4 polynucleotide kinase and γ-^{32}P-ATP or were obtained commercially. Cp and Up were from Sigma Chemicals. 4sUp was made by thiolating Cp (3,000 μmol) with H_2S in aqueous pyridine (Miura et al., 1973). The dried residue with the modified nucleotide was purified by anion-exchange chromatography on DEAE-cellulose, and 4sUp was identified by its absorption spectrum with a λ_{max} at 320 nm. The peak fractions were dried, and the product 4sUp (2,000 μmol) was checked by its spectrum and mobility on thin-layer chromatography.

Photocrosslinking and crosslinked-conjugate purification

Purified D5 RNAs labeled with 4sU and ^{32}P were mixed with acceptor RNAs (E1:D123, D123, or ΔD5) in splicing buffer. Irradiation with UV light was done for 10–20 min at 45 °C with

a single Black-Ray lamp (UV Products) emitting UV light with a maximum around 360 nm and with a short-wave cut-off filter interposed. The 5′SJ hydrolysis reaction is active under some of these conditions (see Figs. 1, 2). Samples were separated on 4% polyacrylamide gels, the radioactive zones were excised after electrophoresis and autoradiography, and the RNAs were eluted and concentrated. Photocrosslinked adducts were checked routinely for 5′SJ hydrolysis activity.

Splicing-related hydrolysis reactions

Reactants for *trans*-hydrolysis controls or photocrosslinked adducts were incubated for 5–45 min at 45 °C in a high-salt splicing buffer [with either 1 M NH_4Cl, 1 M KCl, or 1 M $(NH_4)_2SO_4$) plus 0.1 M $MgCl_2$, 20 mM HEPES-NaOH, pH 7.5]. Control *trans*-hydrolysis assays were done for 10–30 min at 45 °C with 20 nM radioactively labeled E1:D123 RNA (Jarrell et al., 1988a) plus 1 μM D5[36] (Franzen et al., 1993). Competitive inhibition was assayed as described (Peebles et al., 1995). Samples from splicing or photocrosslinking assays were readied for gel electrophoresis either by adding excess Na_3EDTA before heating at 100 °C for 2 min, prior to adding the urea loading dye mix (8 M urea plus 20% sucrose plus 0.025% xylene cyanol and 0.025% bromphenol blue), or by adding two volumes of formamide loading dye mix (90% formamide plus both dyes) and heating at 65 °C.

RNA sequencing and alkaline hydrolysis

To generate a calibrating sequence ladder or to digest the crosslinked adducts for mapping the *D5 donor sites, RNA was dissolved in alkaline hydrolysis buffer (50 mM Na_2CO_3-$NaHCO_3$, pH 9.0, 1 mM Na_3EDTA) and heated at 100 °C for 5 min before chilling and adding one volume of a loading dye (either 8 M urea or 90% formamide) and loading the sample on a 12% polyacrylamide gel (Podar et al., 1995). Other samples of *D5 or crosslinked adducts were treated with RNase T1 for 10 min at 50 °C in 4 M urea, 50 mM sodium acetate, pH 5.0 (Podar et al., 1995).

Primer extension and DNA sequencing

Primer extension with RTase was performed as described (Podar et al., 1995). DNA sequencing was done with Sequenase (Amersham) on double- or single-stranded DNA as appropriate. The primer used for the J(23) region was "D3 reverse;" other defined primers were used to survey the entire intron for crosslinks in control experiments (data not shown). Nucleotide positions within the intron acceptor transcripts are numbered from G_1, the first residue at the 5′ end of aI5γ, through U_{887}, the last nucleotide at the 3′ end of aI5γ, whereas positions in the separate D5 RNA start from the first paired residue at the 5′ end of the substructure, so that intron position G_{815} corresponds to D5 position G_1 and intron position C_{848} corresponds to D5 position C_{34}.

ACKNOWLEDGMENTS

We thank our colleagues in Pittsburgh for reading this manuscript and providing valuable comments. This work was supported by grants from the National Science Foundation (9206525) and the American Cancer Society (NP-894) to C.L.P. and from the National Institutes of Health (GM31480) and the Robert A. Welch Foundation (I-1211) to P.S.P. We gratefully acknowledge the material assistance provided by Professor P. Wipf and Dr. G. Srikanth from the Department of Chemistry at the University of Pittsburgh for thiolating cytidine nucleotides to 4sU derivatives for us. The University of Pittsburgh Department of Biological Sciences has provided research support for our analytical instrumentation and to the DNA Synthesis Facility to prepare defined DNA.

Received September 15, 1997; returned for revision October 20, 1997; revised manuscript received November 19, 1997

REFERENCES

Abramovitz DL, Friedman RA, Pyle AM. 1996. Catalytic role of 2′-hydroxyl groups within a group II intron active site. *Science 271*:1410–1413.

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

Brow DA, Guthrie C. 1988. Spliceosomal RNA U6 is remarkably conserved from yeast to mammals. *Nature 334*:213–218.

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.

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

Datta B, Weiner AM. 1993. The phylogenetically invariant ACAGAGA and AGC sequences of U6 small nuclear RNA are more tolerant of mutation in human cells than in *Saccharomyces cerevisiae*. *Mol Cell Biol 13*:5377–5382.

Dib-Hajj SD, Boulanger SC, Hebbar SK, Peebles CL, Franzen JS, 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.

Fabrizio P, Abelson J. 1990. Two domains of yeast U6 small nuclear RNA required for both steps of nuclear precursor messenger RNA splicing. *Science 250*:404–409.

Favre A. 1990. 4-Thiouridine as an intrinsic photoaffinity probe of nucleic acid structure and interactions. In: Morrison H, ed. *Bioorganic photochemistry: Photochemistry and the nucleic acids, vol 1.* New York: John Wiley & Sons, Inc. pp 379–425.

Favre A, Fourrey JL. 1995. Structural probing of small endonucleolytic ribozymes in solution using thio-substituted nucleobases as intrinsic photolabels. *Acc Chem Res 28*:375–382.

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.

Franzen JS, Zhang M, Peebles CL. 1993. Kinetic analysis of the 5′ splice junction hydrolysis of a group II intron promoted by domain 5. *Nucleic Acids Res 21*:627–634.

Holbrook SR, Cheong C, Tinoco I Jr, Kim SH. 1991. Crystal structure of an RNA double helix incorporating a track of non-Watson–Crick base pairs. *Nature 353*:579–581.

Jacquier A, Jacquesson-Breuleux N. 1991. Splice site selection and role of the lariat in a group II intron. *J Mol Biol 219*:415–428.

Jacquier A, Michel F. 1990. Base-pairing interactions involving the 5′ and 3′-terminal nucleotides of group II self-splicing introns. *J Mol Biol 213*:437–447.

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

Jestin JL, Déme E, Jacquier A. 1997. Identification of structural elements critical for inter-domain 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.
Lemaigre-Dubreuil Y, Expert-Bezançon A, Favre A. 1991. Conformation and structural fluctuations of a 218 nucleotides long rRNA fragment: 4-Thiouridine as an intrinsic photolabeling probe. *Nucleic Acids Res 19*:3653–3660.
Madhani HD, Bordonne R, Guthrie C. 1990. Multiple roles for U6 snRNA in the splicing pathway. *Genes & Dev 4*:2264–2277.
Michel F, Ferat JL. 1995. Structure and activities of group II introns. *Annu 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, 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.
Miura K, Shiga M, Ueda T. 1973. Nucleosides and nucleotides: VI. Preparation of diribonucleoside monophosphates containing 4-thiouridine. *J Biochem 73*:1279–1284.
Padgett RA, Podar M, Boulanger SC, Perlman PS. 1994. The stereochemical course of group II intron self-splicing. *Science 266*:1685–1688.
Peebles CL, Perlman PS, Mecklenburg KL, Petrillo ML, Tabor JH, Jarrell KA, Chen HL. 1986. A self-splicing RNA excises an intron lariat. *Cell 44*:213–223.
Peebles CL, Zhang M, Perlman PS, Franzen JS. 1995. Catalytically critical nucleotides in domain 5 of a group II intron. *Proc Natl Acad Sci USA 92*:4422–4426.
Podar M, Chu VT, Pyle AM, Perlman PS. 1998. Group II intron splicing in vivo by first step hydrolysis. *Nature*. Forthcoming.
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.
Pyle AM, Green JB. 1994. Building a kinetic framework for group II ribozyme activity: Quantitation of interdomain binding and reaction rate. *Biochemistry 33*:2716–2725.
Sawa H, Shimura Y. 1992. Association of U6 snRNA with the 5′-splice site region of pre-mRNA in the spliceosome. *Genes & Dev 6*:244–254.
Sontheimer EJ, Steitz JA. 1993. The U5 and U6 small nuclear RNAs as active site components of the spliceosome. *Science 262*:1989–1996.
Tanner NK, Hanna MM, Abelson J. 1988. Binding interactions between yeast tRNA ligase and a precursor transfer ribonucleic acid containing two photoreactive uridine analogues. *Biochemistry 27*:8852–8861.
van der Veen R, Arnberg AC, van der Horst L, 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.
Wyatt JR, Sontheimer EJ, Steitz JA. 1992. Site-specific cross-linking of mammalian U5 snRNP to the 5′ splice site before the first step of pre-mRNA splicing. *Genes & Dev 6*:2542–2553.
Woissard A, Fourrey JL, Favre A. 1994. Multiple folded conformations of a hammerhead ribozyme domain under cleavage conditions. *J Mol Biol 239*:366–370.
Wolff T, Menssen R, Hammel J, Bindereif A. 1994. Splicing function of mammalian U6 small nuclear RNA: Conserved positions in central domain and helix I are essential during the first and second step of pre-mRNA splicing. *Proc Natl Acad Sci USA 91*:903–907.
Zimmerly S, Guo H, Eskes R, Yang J, Perlman PS, Lambowitz AM. 1995. A group II intron RNA is a catalytic component of a DNA endonuclease involved in intron mobility. *Cell 83*:529–538.