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

MARIA COSTA,^{1,3} ERIC L. CHRISTIAN,² and FRANÇOIS MICHEL¹

¹Centre de Génétique Moléculaire du C.N.R.S., 91190 Gif-sur-Yvette, France

²Department of Molecular Biology and Microbiology, Case Western Reserve University, Cleveland, Ohio 44106-4960, USA

ABSTRACT

Dimethyl sulfate modification was used to probe for tertiary structural elements in the group II intron PI.LSU/2 from the mitochondrial pre-ribosomal RNA of the brown alga *Pylaiella littoralis*. Modification of the lariat form of the intron under conditions that allow both native folding and conformational homogeneity is found to be generally consistent with secondary and tertiary structural features identified previously for group II ribozymes. A comparison of chemical probing at temperatures just below and above the first melting transition illustrates the cooperative unfolding of tertiary structure and identifies novel candidates for tertiary interactions in addition to defining elements of secondary structure. Substitution of the GAAA terminal loop of domain V is shown to be compatible with retention of conformational homogeneity (despite the loss of an important tertiary interaction), but produces a concise methylation footprint in domain I at the site previously shown to harbor the receptor for that loop. The analysis also identified two nucleotide positions in domain V with novel secondary and potential tertiary structural roles. The proposed refinement of domain V secondary structure is supported by an expanded comparative analysis of group II sequences and bears increased resemblance to U2:U6 snRNA pairing in the spliceosome.

Keywords: group II intron; methylation probing; ribozyme; RNA structure; spliceosomal RNAs

INTRODUCTION

The realization that group II self-splicing involves the same 2'-5' bonded lariat intermediate as spliceosome-catalyzed splicing sparked intense speculation about the extent of structural and mechanistic similarity of these two RNA processing systems (e.g., Cech, 1986; Sharp, 1991; Weiner, 1993). Since then, there has been considerable progress in understanding RNA–RNA interactions in the spliceosome (reviewed by Nilsen, 1998) and some 10 tertiary interactions and two crosslinks have been uncovered within and between the six secondary structure domains that make up group II introns (Jacquier & Michel, 1987, 1990; Michel & Jacquier, 1987; Costa & Michel, 1995; Podar et al., 1995; Chanfreau &

Jacquier, 1996; Costa et al., 1997a; Podar et al., 1998). Nevertheless, detailed structural information defining the mechanistic basis of these two splicing systems is only beginning to emerge.

Much of the progress in defining structure–function relationships in the group II system has been made with introns that are active in vitro (reviewed by Michel & Ferat, 1995). An important limitation in these studies, however, has been in obtaining homogenous populations of molecules with an active structure. Precursor transcripts from even the best-studied group II ribozyme—intron a5 γ from yeast mitochondria—react with complex kinetics, indicative of conformational heterogeneity (Daniels et al., 1996). Moreover, probing of these molecules with dimethyl sulfate (DMS) (Kwakman et al., 1990) failed to reveal evidence of tertiary structure, although recent DMS modification experiments of domain I were reported to be consistent with known tertiary contacts (Qin & Pyle, 1997).

In this context, the recent reporting by Costa et al. (1997b) of a group II intron that readily forms homog-

Reprint requests to: Maria Costa; Center for Molecular Biology of RNA, Sinsheimer Laboratories, University of California at Santa Cruz, Santa Cruz, California 95064, USA.

³Present address: Center for Molecular Biology of RNA, Sinsheimer Laboratories, University of California at Santa Cruz, Santa Cruz, California 95064, USA.

enous populations of molecules with a native conformation constituted welcome news. Specifically, the lariat form of intron PI.LSU/2, which is one of the mitochondrial ribosomal introns of the brown alga *Pyraliella littoralis*, was observed to migrate as a single band in nondenaturing gels. Furthermore, analysis of its melting profiles revealed an early, cooperative transition, irrespective of whether the molecules were (presumably) native or had been denatured (and renatured) prior to the experiment.

We now extend these findings, and confirm previous claims of homogeneity, by showing that molecules with a wild-type sequence react according to monophasic kinetics and, when treated with a chemically modifying agent under native and semi-denaturing conditions, yield clear-cut, contrasted patterns of local accessibility and protection that are generally consistent with known secondary and tertiary structure. Moreover, the same is true of mutant molecules that lack a critical interaction, but can be folded nevertheless under permissive conditions. These findings open the way to employ chemical footprinting for the direct identification of tertiary interactions as well as the refinement of current secondary structure models. The DMS modification studies in the current work, in fact, identify a new potential tertiary contact in domain V and refine the secondary structure of this essential component of group II introns at nucleotide positions known to be directly involved in catalysis. This altered domain V structure, which has an increased resemblance to the base pairings formed by snRNAs U2 and U6 in the spliceosome, is supported by an expanded comparative analysis of group II sequences.

RESULTS

Differential DMS modification of intron PI.LSU/2 identifies groups potentially involved in tertiary interactions

Comparison of chemical accessibility under native and semi-denaturing conditions provides a straightforward way to locate those groups that may be involved directly in tertiary contacts. Nevertheless, it is implicit in this approach that conditions under which only tertiary structure is destabilized not only exist, but (ideally) that they can be identified independently of chemical modification. In this respect, melting profiles can provide crucial information. One example is offered by the *Tetrahymena* and *sunY* group I introns, whose thermal denaturation involves a sharp initial transition. Chemical modification of these molecules at temperatures just below and above the first peak in melting profiles confirmed that the interactions being disrupted during that early transition were primarily tertiary contacts, even though a number of helical segments were also found to be affected (Banerjee et al., 1993; Jaeger et al., 1993).

Melting profiles of the lariat form of intron PI.LSU/2 likewise reveal a cooperative early transition, which extends from approximately 45 to 48 °C at 1 M NH₄Cl and 5 mM MgCl₂ (Costa et al., 1997b). We accordingly chose to probe this molecule at temperatures slightly below (42 °C) and above (50 °C) that range, using concentrations of DMS that ensure limited modification (see Materials and Methods). Modified bases were identified by primer extension with reverse transcriptase (Fig. 1), which detects methylation of the N1 position of adenines and of N3 of cytosines (Inoue & Cech, 1985; in our case, modification of cytosines was insignificant, except at position 232). Figure 2 makes it possible to compare extents of modification, as estimated from the intensity of reverse transcriptase stops (or pauses), with the proposed secondary structure of the PI.LSU/2 molecule, which was inferred by comparative sequence

FIGURE 1. Example of mapping of DMS modifications by primer extension with reverse transcriptase. Modification reactions were performed in a buffer containing 5 mM MgCl₂ (see Materials and Methods). Native conditions: Modification was performed at 42 °C for 5 min (+), 10 min (++), or 30 min (+++); lane (–), mock-treated RNA. Semi-denaturing conditions: Modification was performed at 50 °C for 3 min (+), 5 min (++), or 20 min (+++); lane (–), mock-treated RNA. C, T, A, and G are sequencing lanes generated by reverse transcription of the wild-type PI.LSU/2 lariat intron molecule in the presence of each one of the four ddNTPs. Arrows point to positions that are significantly more accessible to modification at 50 °C than at 42 °C. Note that primer extension stops one nucleotide before a modified base (Inoue & Cech, 1985).

analysis (Fontaine et al., 1995; Michel et al., 1989). At 42 °C (Fig. 2A), the adenines supposedly involved in secondary structure pairings are essentially inaccessible, except for A307, A315, and A2389, all three of which were found to be highly reactive. These bases were accordingly left unpaired in the secondary structure models of Figure 2, which are otherwise identical to the one in Figure 5 of Fontaine et al. (1995). In addition, all adenines currently known to participate in tertiary interactions (see below) are observed to be protected (note that because the molecule probed was the intron lariat, the EBS1 and EBS2 sequences involved in intron–exon pairing were not expected to be protected, and are indeed accessible).

Raising the temperature to 50 °C (Fig. 2B) does not change significantly the accessibility of helical segments, except that there is no longer any indication of base pairing in the basal helix of domain IV and some destabilization is also seen in domain III. In contrast, there is no evidence that any of the known tertiary interactions still exists at 50 °C. All four A's that participate in the α - α' classical base pairs (Michel & Jacquier, 1987; Harris-Kerr et al., 1993) are reactive to DMS (even though to varying extents), whereas they were protected at 42 °C. A similar change in protection pattern is observed for the adenines of the GUAA and GAAA terminal loops that are part of the θ - θ' and ζ - ζ' interactions (Fig. 5; Costa & Michel, 1995; Costa et al., 1997a). Disruption of the latter contact is also manifested in the altered pattern of chemical reactivity of the proposed receptor motif for the ζ' loop (Costa & Michel, 1995), at the junction of stems ID⁽ⁱ⁾ and ID⁽ⁱⁱ⁾. In fact, virtually all A's in so-called "single-stranded" secondary structure components are found to be accessible to some degree, suggesting that the conditions used allow accurate discrimination between elements of secondary and tertiary structure.

The current experimental conditions, in fact, identify two nucleotide positions within domain V of particular interest where differential DMS probing indicates novel roles in secondary or tertiary structure. Interestingly, extensive modification of A2389 is observed under native conditions. Modification at this position was not expected, because this nucleotide, like the rest of domain V (dV), is well conserved in group II introns and, until now, had been assumed to base pair with U2374 in secondary structure models. The high reactivity of the N1 position of A2389 is all the more striking because its immediate neighbor on the 3' side, A2390, is fully protected in molecules with a native structure. The latter nucleotide nevertheless becomes accessible under semi-denaturing conditions or in separate dV or dV-dVI transcripts, so that the two bases are then modified to about the same extent (Figs. 4A, 5). These data suggest a potential tertiary contact involving position A2390 and imply a refinement of the dV secondary structure (see below).

Substitution of the dV terminal loop results in a concise footprint in domain I, but no change in the rest of dV

Whether or not N1 of A2390 is itself engaged in a tertiary contact, its reactivity to DMS thus constitutes a sensitive indicator of the state of the middle part of dV. These findings are of interest, inasmuch as the small domain V is a central piece of the group II ribozyme structure, with an inordinate number of components that are already known to be involved in interactions with the rest of the molecule and/or to be essential for function (Chanfreau & Jacquier, 1994; Boulanger et al., 1995; Peebles et al., 1995; Abramovitz et al., 1996). In the case of the dV terminal loop, the interacting partner has, in fact, been identified and, as mentioned above, consists of the ζ "receptor" motif in domain I (Costa & Michel, 1995). Because a single, rather short helix connects the GAAA dV terminal loop of intron PI.LSU/2 to its two-nucleotide bulge, it was interesting to determine whether changing the loop sequence would have repercussions on the accessibility of A2389 and A2390.

The dV GAAA terminal loop was changed to UUCG, a sequence that markedly stabilizes the helices it caps (Tuerk et al., 1988), but does not interact fruitfully with receptors for GNRA loops (Jaeger et al., 1994; Murphy & Cech, 1994). The resulting "L5:UUCG" precursor molecule retains little activity under conditions found to be optimal for wild-type transcripts (10 mM MgCl₂, 45 °C; Costa et al., 1997b). Reactivity of the "L5:UUCG" precursor, however, is largely restored by raising the magnesium concentration to 20 mM (data not shown), making it possible to generate large amounts of the lariat intron form for further characterization.

A major advantage of the lariat intron over precursor molecules is that it is inactive by itself, so that its folding and reactivity can be investigated separately. When wild-type lariat intron molecules that have been pre-incubated in magnesium-containing buffers are mixed with 5' exon molecules (see Materials and Methods), a partial debranching reaction ensues. This reaction converts a fraction of the intron lariat into intron–5' exon linear molecules, as already described for several other group II introns (Augustin et al., 1990; Chin & Pyle, 1995; Costa et al., 1997a). In the case of intron PI.LSU/2, we found that debranching with 5' exon in excess of intron is too rapid at pH 7 to be followed, but the reaction can be slowed down by decreasing the pH. Examination of the time course of debranching at pH 5.4 (Fig. 3A) reveals essentially single-phase kinetics and neither the apparent first-order reaction rate, nor the final fraction of debranched molecules changes markedly between 10 and 50 mM MgCl₂.

The L5:UUCG lariat intron can be debranched as well, and the final fraction of reaction product, which again remains essentially constant within the range of magnesium concentrations investigated, is, in fact, some-

A

FIGURE 2. Sites of DMS modification on the lariat form of intron PI.LSU/2. Arrowheads point to 5' and 3' splice sites. For secondary structure modeling and nomenclature, see Michel et al. (1989) and Fontaine et al. (1995) [the stem-loop carrying the EBS1 sequence is labeled ID3 in order to preserve consistency with the yeast $\alpha 5\gamma$ intron (Jacquier & Michel, 1987)]. EBS1-IBS1, EBS2-IBS2, greek letters, curved arrows, and boxing indicate sequences that participate in tertiary interactions (Michel & Ferat, 1995; Chanfreau & Jacquier, 1996; Costa et al., 1997a). Note that the terminal loop of domain IV was shortened to 56 nt in the molecule used throughout this work (see Materials and Methods). **A:** Native conditions (42 °C, see Fig. 1 and Materials and Methods for detailed conditions). Arrows indicate sites and extents of modification. The latter were estimated for each position by comparing data from two to six separate experiments (all but the first 5 and last 14 nt of the intron could be read by reverse transcription). **B:** Semi-denaturing conditions (50 °C). Arrows as in A, + (or -) signs indicate bases that are significantly more (or less) accessible at 50 °C than at 42 °C. (Figure continues on facing page.)

what higher than for wild-type molecules (Fig. 3B). On the other hand, the apparent rate of reaction is strongly dependent on magnesium, for it is 14-fold lower than for the wild-type lariat at 10 mM MgCl_2 , but almost reaches the wild-type value at 50 mM MgCl_2 ($k_{\text{obs}} = 0.81 \pm 0.08 \text{ min}^{-1}$ for L5:UUCG lariat molecules, compared to $1.03 \pm 0.10 \text{ min}^{-1}$ for wild-type). When the final fraction of debranched molecules is plotted as a function of the concentration of 5' exon, typical saturation behavior is observed for both the wild-type and L5:UUCG mutant (Fig. 3C). By fitting these curves, apparent " K_m " values are obtained and these can be converted in turn into K_d (i.e., $k_{\text{off}}/k_{\text{on}}$) estimates (see

legend of Fig. 3) by assuming the following scheme (L, lariat intron; E5, 5' exon; P, linear intron-5' exon):

The resulting K_d values— $0.75 \pm 0.06 \mu\text{M}$ for the wild-type and $0.81 \pm 0.06 \mu\text{M}$ for the L5:UUCG mutant—do not differ significantly, which confirms that when L5:UUCG lariat molecules are reacted under permissive conditions, their behavior is essentially indistinguishable from that of wild-type.

FIGURE 2. *Continued.*

In Figure 3D, melting profiles of the wild-type and L5:UUCG lariat introns are compared. As reported previously (Costa et al., 1997b), denaturation of wild-type molecules begins with a markedly cooperative transition with a T_m of ca. 52 °C at 10 mM MgCl₂. Melting of the L5:UUCG lariat also shows evidence of an initial cooperative transition, but with a lower T_m (by about 6 °C) and a somewhat broader temperature range. Nevertheless, comparison of the two melting profiles reveals no major change in hyperchromicity or in the baseline at temperatures below 40 °C, so that conformational homogeneity does not appear to be compromised by the L5:UUCG mutation.

Further confirmation that the L5:UUCG mutation is compatible with the formation of homogenous populations of molecules with a correct three-dimensional architecture was obtained by treatment with DMS under permissive conditions (50 mM MgCl₂ and 37 °C), followed by reverse transcription of the intron with a full set of primers. When patterns of chemical modification

were compared for L5:UUCG and wild-type samples that had been treated in parallel (e.g., Fig. 4), no significant change was observed, except in the receptor for the dV terminal loop at the junction of stems ID⁽ⁱ⁾ and ID⁽ⁱⁱ⁾ (and also of course at the site of the substitution itself). As seen in Figure 4B, A305, which is essentially protected in wild-type molecules, is modified by DMS in mutant ones, whereas the accessibility of A307 is reduced significantly. On the other side of the receptor, some slight modification of A165 was observed in mutant but not in wild-type introns molecules (not shown). Interestingly, these changes parallel the effects of raising the temperature to 50 °C (see Fig. 2B). On the other hand, although semi-denaturing conditions result in equal accessibility of A2389 and A2390 in dV (Fig. 5C), the L5:UUCG substitution has no detectable effect on the reactivity to DMS of these two bases (Fig. 5D). In conclusion, there is no evidence that disruption of the ζ - ζ' interaction is of any consequence for the structure of the neighboring two-nucleotide bulge and the interactions of

FIGURE 3. Kinetic and thermodynamic characterization of wild-type and L5:UUCG lariat introns. **A:** Time course of debranching of the wild-type intron at 40°C and a quasi-saturating (5 μ M) 5' exon concentration (see Materials and Methods for detailed conditions; reactions were followed for 70 min, with no further increase in the extent of debranching). Data were fitted with the program KaleidaGraph (Abelbeck Software) by assuming a monophasic reversible reaction (e.g., Fersht, 1977), i.e., fraction reacted = $[k_1/(k_1 + k_{-1})]\{1 - \exp[-(k_{-1} + k_1)t]\}$, with k_1 and k_{-1} being rates of debranching and branching, respectively. Circles and solid line, 10 mM MgCl_2 ; squares and dashed line, 20 mM MgCl_2 ; diamonds and dotted line, 50 mM MgCl_2 . **B:** Time course of debranching of L5:UUCG mutant lariat molecules. Conditions, symbols, and fits as in A. **C:** 5' Exon saturation curves. Abscissa, 5' exon concentration. Ordinates, observed fraction of debranched molecules at equilibrium (f_{eq}). Data were fitted to the following equation (see text), $f_{eq} = [k_1/(k_1 + k_{-1})][[E5]/([E5] + K_m)]$, with $K_m = [k_{-1}/(k_1 + k_{-1})] \cdot K_D$. Circles and dashed line, wild-type lariat ($K_m = 0.59 \pm 0.05 \mu\text{M}$, $K_D = 0.75 \mu\text{M}$). Squares and solid line, L5:UUCG mutant lariat ($K_m = 0.60 \pm 0.05 \mu\text{M}$, $K_D = 0.81 \mu\text{M}$). Reaction conditions were as in A and B with 10 mM (wild-type) or 50 mM (L5:UUCG mutant) MgCl_2 . **D:** Derivate melting profiles of wild-type (solid line) and L5:UUCG (dashed line) lariat molecules at 40 mM potassium cacodylate, pH 7.5, 1 M NH_4Cl , 10 mM MgCl_2 , and 0.02% SDS. Ordinates are normalized to an absorbance of 1 at 30°C.

that bulge with the rest of the intron as long as the overall fold of the molecule is maintained.

An alternative dV structure suggested by comparative sequence analysis is consistent with DMS modification

Until now, the consensus secondary structure of domain V had been assumed to be the one shown in

Figure 7A, with two helices of eight and six base pairs separated by a dinucleotide bulge. Although this structure was first proposed at a time when only four group II introns had been sequenced (Michel et al., 1982), it was to prove compatible with the vast majority of the group II sequences that have become available since then and was not questioned when the secondary structure of group II introns was last reassessed (Michel et al., 1989). Still, there are several base pairs

FIGURE 4. Mapping by primer extension of changes in DMS modification of domain V and the domain I tetraloop receptor between wild-type and mutant constructs. All modification reactions were performed at 37 °C. Lane (-), mock-treated RNA. C, T, A, and G lanes as in Figure 1. **A:** Domain V. L5:UUCG intron lariat molecules were probed in the presence of 50 mM MgCl₂ (see Materials and Methods) for 5 min (+) or 10 min (++). Wild-type (w.t.) intron lariat molecules and a transcript consisting of domains V and VI (dV-dVI) were modified in the presence of 10 mM MgCl₂ (see also Materials and Methods). w.t. lanes: (+), 7 min; (++), 15 min. dV-dVI construct: (+), 15 min. Numbering of the dV-dVI transcript starts from the first nucleotide of domain V (position 2366 in Fig. 2). **B:** Domain I tetraloop receptor. Same reaction conditions as in A. Nucleotide numbering is according to Figure 2.

in domain V that could never be formally established by phylogenetic analysis due to an insufficient number of compensatory changes in natural sequences. For instance, all but three of the currently known group II introns have A and C at positions 2 and 4 of dV, so that proof of base pairing between nt 2–33 and 4–31 had to

FIGURE 5. Quantitation of DMS methylation of domain V as a function of conditions and constructs. **A:** dV alone transcript (see Materials and Methods). A gel similar to the ones shown in Figure 4 was scanned using a PhosphorImager. Thick curve, no DMS control; thin curve, DMS modification (15 min, 37 °C). Bottom, sequence of domain V; bases were assigned to peaks by taking into account that reverse transcriptase stops or pauses at the position immediately 3' of a modified base. **B:** dV-dVI transcript (see Materials and Methods), same as in A. **C:** Wild-type lariat as a function of temperature. Thick curve, no DMS control; thin curve, DMS modification (3 min) at 50 °C; dashed curve, DMS modification (5 min) at 42 °C. **D:** Wild-type versus L5:UUCG lariat molecules. Thick curve, no DMS control; thin curve, wild-type (7 min, 37 °C); dashed curve, L5:UUCG mutant (same conditions).

await the demonstration that some double substitutions that restore base pairing have compensatory effects with regard to the ability to self-splice in vitro (Peebles et al., 1995) or can be isolated by looking for second-site suppressors in vivo (Boulanger et al., 1995). Similarly, the existence of the G3:U32 wobble pair is supported by the fact that, although none of the single substitutions at G3 are tolerated, the A:C combination (which can adopt the same geometry as G:U upon protonation of A-N1) retains detectable activity in vitro (Peebles et al., 1995) and was generated by second site suppression in vivo (Boulanger et al., 1995).

Yet another base pair for which comparative evidence was lacking is the one between U9 and A24. As was already apparent back in 1989, and is made even clearer by an expanded analysis of 140 sequences (Fig. 6; Table 1), a significant number of introns (22 of 140) have a base other than A at position 24, but only one of them carries a compensatory change at position 9 (resulting in a potential A-U base pair). On the other hand, 6 of 10 group II sequences that lack a U at

TABLE 1. Base combinations observed at positions 9, 24, and 26 of domain V in 140 group II sequences.

24					26						
9		A	G	C	U			A	G	C	U
	U	<u>112</u>	<u>1</u>	5	12	9	U	<u>1</u>	<u>128</u>	1	0
	C	<u>5</u>	0	2	1		C	<u>5</u>	<u>3</u>	0	0
	G	0	0	0	0		G	0	0	0	0
	A	1	0	0	<u>1</u>		A	0	1	0	<u>1</u>

^aUnderlined numbers indicate Watson-Crick (solid line) and wobble (dashed line) pairs.

position 9 have a second change at position 26, which is one of the other sites facing U9, and is usually a G (132/140 introns). One of these sequences has A9:U26 and the remaining five (which belong to at least three separate subdomains of the group II phylogenetic tree) have C9:A26. Therefore, comparative sequence analysis provides no evidence of base pairing between A24

FIGURE 6. Variations observed at positions 9, 24, and 26 of domain V in 112 group II introns with distinct insertion sites and 28 sequence variants with base changes at one or several of the nucleotides that form and surround the bulge (positions 7–11 and 22–28). Center: Consensus structure and sequence of domain V. Uppercase letters reflect greater than 90% conservation; lowercase letters, 80–90%. R, purine; Y, pyrimidine; M, A, or C; K, G, or U; H, any base but G; N-N', a canonical base pair. Dashed lines indicate potential base pairs between 9 and 24 and 9 and 26. ^(a) The following sequences (mt, mitochondrial-encoded; cp, chloroplast-encoded; accession numbers between brackets) show noncanonical variations at the positions indicated: *Lycopodium squarrosum* mtcox3 (X76273), 27UU; *Marchantia* mtcox1-2 (M68929) and angiosperm mtcox2-1 (X01108, X57694, X17394, X63625), 8U; angiosperm mtnad2-3 (M81726, X96536), 8C; *Oryza sativa* cprps12-1 (X15901), cprps12-2 of seed plants (Z00044, M17841, X89563, X15901) and *Marchantia* cprps12-2 (M68929), 10U (the latter also has 23G); *Pinus thunbergii* petD (D17510), 22G. Left: Sequences that have mutations (circled bases) at positions 24 and 26, but retain U9. ^(b) *Marchantia* mtcox3-1 (M68929) and *Pellia epiphylla* mtcox3 (X92740) have 8U. Right: Sequences with mutations at position 9. Estimates of the number of sequences of independent evolutionary origin are indicated above arrows for each group of observed variations.

and U9, and instead supports the existence of a wobble pair between U9 and G26.

As shown in Figure 7B, replacement of the U9:A24 Watson–Crick pair by a U9:G26 wobble pair results in shifting the two-nucleotide bulge by one base pair, so that the two helices that make up domain V now comprise nine and five base pairs, rather than eight and six base pairs. As for A24, which corresponds to A2389 in Figure 2, it is now left without a partner, which neatly explains its accessibility to DMS in isolated dV transcripts. What could not have been anticipated from phylogenetic analyses, though, is that A2389 remains accessible in lariat molecules with a native structure, whereas the other bulged nucleotide, A2390 (=A25), is protected.

DISCUSSION

Uniform and correct folding of both wild-type and mutant molecules

Our probing of intron PI.LSU/2 by DMS fully supports the claims of Costa et al. (1997b) regarding conformational homogeneity of the lariat form of that molecule and the basis for its cooperative unfolding. Such clear-cut patterns of high accessibility alternating with nearly complete protection as are observed throughout the

intron (e.g., Figs. 1, 4, 5) imply that the molecules that were probed had an essentially uniform structure. Furthermore, positions inaccessible to DMS modification at temperatures below the melting range include all but three of the bases predicted to be protected by the current structural model of group II introns, suggesting that the structure probed was the native one. These data provide direct experimental evidence for a number of structural features that have been inferred indirectly by phylogenetic and genetic analysis and indicate that the α - α' , θ - θ' , and ζ - ζ' tertiary interactions actually coexist in what must be the ground state of the lariat intron.

Linking of the early, magnesium-dependent transition seen in melting profiles of intron PI.LSU/2 to the disruption of tertiary structure (Costa et al., 1997b) was based on analogy with similar observations on group I introns (Banerjee et al., 1993; Jaeger et al., 1993), tRNA, and a ribosomal RNA subdomain (reviewed by Draper, 1996). In probing the PI.LSU/2 intron structure at a temperature just beyond the initial magnesium-dependent transition, we have verified that known tertiary interactions are no longer present, although much of the secondary structure remains intact. The current work builds upon previous DMS studies that were shown to be consistent with the proposed secondary structure of intron a5 γ , but could not discern elements of tertiary structure (Kwakman et al., 1990).

FIGURE 7. Comparison of current (A) and proposed (B) consensus secondary structures for domain V to consensus secondary structure for spliceosomal U2/U6 helix 1b and the U6 intramolecular stem (C). Uppercase letters reflect greater than 90% conservation; lowercase letters, 80–90%. R (or r), Y (or y), M, and K as in Figure 6. Circles: Positions involved in chemistry (Chanfreau & Jacquier, 1994; Peebles et al., 1995; Abramovitz et al., 1996). Squares: Positions involved in both chemistry and binding. Shaded letters: Phosphorothioate effects 5' to the indicated nucleotide. Arrows: Helical distance between the conserved AGC trinucleotide and the dinucleotide bulge.

The buffer conditions used in the current work for DMS probing (5 mM MgCl_2 and a high monovalent cation concentration) were chosen to maximize the difference in secondary and tertiary structural stability in order to more clearly interpret chemical protection patterns in the context of group II structure. Raising the concentration of MgCl_2 should result in an increasing number of secondary structure pairings melting together with the three-dimensional structure of the intron, whereas lowering magnesium conditions might make it possible to identify separate domains of tertiary folding (see discussions in Draper, 1996 and Costa et al., 1997b). It must also be realized that some of the adenines that underwent significant changes in accessibility between 42 and 50 °C may have done so due to some indirect steric effects of the disruption of three-dimensional structure, rather than because of their direct involvement in a tertiary contact. Nevertheless, the propensity for terminal loops, especially of the GNRA family, to be involved in RNA–RNA interactions (reviewed by Westhof et al., 1996; Michel & Costa, 1998) indicates that additional tertiary interactions are most likely to exist for the GCAA and GAAA loops that cap stems IIB and IIIC, and seem probable as well for the terminal loops that cap stems IA and IIIA. In addition, numerous adenines in RNA internal loops have been found to be involved directly in tertiary contacts (e.g., Cate et al., 1996a). We expect many of the adenosines that become modified by DMS above the first temperature transition to be involved similarly in noncanonical interactions.

The PI.LSU/2 intron has unusually low magnesium requirements for folding and function, which led Costa et al. (1997b) to anticipate that even seriously destabilizing alterations could be compensated by raising the concentration of magnesium. We now show that this is the case for a molecule in which an important tertiary interaction (ζ - ζ'), which is a major contributor to recognition of dV by the rest of the intron (Franzen et al., 1993; Chanfreau & Jacquier, 1994; Costa & Michel, 1995; Abramovitz et al., 1996), has been disrupted. As expected, the L5:UUCG lariat intron is destabilized markedly (Fig. 3D), and it also reacts poorly compared to the wild-type intron at 10 mM MgCl_2 (Fig. 3A,B), possibly because the temperature at which debranching was assayed (40 °C) is too close to the melting range of the mutant molecule. Raising the concentration of magnesium to 20 mM is nevertheless sufficient for the two lariats to become essentially indistinguishable at that temperature by kinetic criteria. DMS modification under these permissive conditions indicates that the population of L5:UUCG molecules is just as uniformly and correctly folded as its wild-type counterpart—except, of course, at the expected ζ receptor site in domain I, where a well-delimited footprint is observed. It is quite evident that extending this approach to components suspected to be involved in ter-

tiary interactions, but with unknown partner nucleotides, should lead, in most cases, to the rapid identification of the missing contacts.

DMS probing of the ζ - ζ' motif

The concise footprint we observe in L5:UUCG mutant molecules at the junction of helices ID⁽ⁱ⁾ and ID⁽ⁱⁱ⁾ confirms the location of the ζ receptor for the well-conserved dV terminal loop, ζ' , which reads GAAA in most group II introns. This receptor was identified initially through comparative analysis, followed by the demonstration in intron a5 γ of compensatory effects as a result of appropriate double substitutions at the two supposedly interacting sites (Costa & Michel, 1995; see also Jestin et al., 1997, for confirmation by modification interference). The identity of the ζ receptor was questioned by Abramovitz et al. (1996), because of the presence in intron a5 γ at the proposed docking site of two G:C pairs with opposite orientations rather than the expected tandem repeats. As already suggested, however, by Costa and Michel (1995), and soon afterward established by the crystal structure of a GAAA receptor (Cate et al., 1996a), only the G-C pair that is correctly oriented in intron a5 γ actually contacts the loop. Moreover, in vitro selection for RNA receptors of GAAA loops (Costa & Michel, 1997) has now made it clear that the other G:C pair (the one “wrongly oriented” in a5 γ) is not an essential feature of the 11-nt canonical GAAA receptor [even though a hypothetical model according to which that pair would play a part in the early stages of docking has been proposed recently (Butcher et al., 1997)].

Another feature of the 11-nt receptor for GAAA loops that is not strictly conserved is its G:U closing pair. Even though this is the pair on which the A-A platform of the receptor is stacked (Cate et al., 1996b), it is replaced in the ζ receptor of quite a few group II introns, including PI.LSU/2, by an A:U combination (A307:U164 in Fig. 2). In fact, one of the three A-A platforms that were uncovered by Cate et al. (1996b) in the P4-P6 domain of the *Tetrahymena* intron also happens to be stacked on an A:U pair, although not of the standard Watson–Crick configuration. Based on this observation, Cate et al. (1996b) and, subsequently, Costa and Michel (1997) have argued that base pairs on which an A-A platform is stacked would most likely turn out to have a noncanonical geometry. The fact is that A307 is most likely not bound to U164 by Watson–Crick pairing in the native PI.LSU/2 molecule given that its N1 position is highly accessible to DMS modification. Interestingly, this base is significantly less reactive when the receptor becomes unbound either as a consequence of substituting the dV terminal loop (Fig. 4B) or raising the temperature (Fig. 2B; this is, in fact, the only example throughout the molecule of DMS reactiv-

ity being significantly weaker at 50 °C than at 42 °C). Extensive rearrangement of the 11-nt receptor upon binding of its loop partner has been proposed by Cate et al. (1996a), based on the same changes in DMS methylation as we see here, and was confirmed by an NMR structure of the unbound receptor (Butcher et al., 1997). It remains to be seen whether—as suggested by our own DMS modifications—the structure of the unbound ζ receptor is indifferent to its interaction with the rest of the intron.

ζ - ζ' is the only established contact between domain V and the rest of the intron, but certainly not the only one to exist. In fact, many chemical groups on dV have already been suggested to be involved in interactions with the rest of the molecule (Chanfreau & Jacquier, 1994; Abramovitz et al., 1996; Jestin et al., 1997; Konforti et al., 1998). Thus, the small domain V appears to be held tightly by a network of contacts with other intron components, which may explain why substitution of its terminal loop had no detectable effect on the conformation of its middle part (Figs. 4, 5), as well as the retention of conformational homogeneity of this mutant intron as a whole.

An alternative structure for domain V

The alternative structure we propose for domain V (Fig. 7B) reflects a refinement of base pairing in which position 9 interacts with position 26 rather than position 24, and was derived from a combination of comparative sequence analysis and DMS probing. The proposed 9-26 pairing changes the position and composition of the dinucleotide bulge and extends the proximal helix by one base pair, altering the relative positions of nucleotides essential for splicing (compare Fig. 7A and B, see below). In support of this proposed pairing, Peebles and coworkers recently have found that a potential U-G wobble pair at positions 9 and 26 is strictly maintained during *in vitro* selection experiments, whereas the potential to form canonical or wobble pairs between positions 9 and 24 is not (C.L. Peebles, pers. comm.). Moreover, a kinetic analysis of all 16 pairwise combinations of positions 9 and 26 shows that the U-G wobble pair is kinetically favored over canonical base pairs at these positions and that the relative rate constants for the different mutant pairs roughly parallels the observed frequency of nucleotide pairs in phylogeny (C.L. Peebles, pers. comm.).

An extensive mutagenic analysis of the dinucleotide bulge that emphasized the importance of maintaining a correct geometry in the middle part of domain V has also been performed by Perlman and coworkers (Schmidt et al., 1996). In this study, there is one case where a mismatch between positions 9 and position 26 (U9 C26) is active in a *cis*-splicing assay *in vitro*. However, G26 appears to be strongly preferred because the only revertants seen at this position return to the

original G (Schmidt et al., 1996). Similarly, it may also be important for the 9-26 pair to be stacked on the end of a helix. Mutations that potentially disrupt the pairing of nucleotides adjacent to 9-26 in helix I are strongly inhibited for splicing both *in vitro* and *in vivo* (Schmidt et al., 1996). Again, the only revertants observed restore the equivalent of the 8-27 interaction adjacent to the 9-26 pair.

The reorganization of the two-nucleotide bulge is of particular interest because the middle part of dV contains an inordinate number of bases and backbone elements that already have been ascribed a structural or chemical role. In intron a5 γ , alteration of the bases at positions 26 (G) and 27 (U), and of the phosphate 5' of position 24 (and, to a lesser degree, of position 25) has been shown to interfere with binding of another intron component, most likely domain III (Chanfreau & Jacquier, 1994; Jestin et al., 1997). To this list of components directly or indirectly involved in the binding of dV, we can now add the base at position 25 (A2390 in Fig. 2), which undergoes a spectacular change in DMS reactivity upon folding of the entire PI.LSU/2 molecule, but remains accessible in transcripts of domain V, or domains V and VI alone (position 25 is a C in intron a5 γ , and was not assayed by interference). Our comparative sequence analyses suggest that G26 forms a wobble pair with U9 [even though removal of uridine at position 9 by hydrazine was not found to markedly affect group II splicing *in vitro* (Chanfreau & Jacquier, 1994)]. In this respect, the strong tendency for a wobble pair adjacent to the dinucleotide bulge (132/140 cases) may indicate a specific geometry or flexibility within this active site element. It is also interesting to note that, in group I introns, the G:U pair at the 5' splice site is bound by consecutive sheared A:A pairs in the J4/5 internal loop of these molecules (Strobel et al., 1998) and that in a5 γ , PI.LSU/2, and other subgroup IIB introns, the internal loop that interrupts the basal stem of domain III happens to have the same consensus sequence as J4/5.

The middle part of dV also includes several chemical groups—the phosphate 5' of position 25 (Chanfreau & Jacquier, 1994) and the 2'-hydroxyls of nt 23, 24, and 26 (Abramovitz et al., 1996)—that are thought to be important in catalysis (and not so much in the ground state). Although direct chemical probing reflects the ground state structure (and then, in our case, only of the lariat intron form), a strength (or weakness) of comparative sequence analysis is that it is essentially indifferent to the possible dynamic nature of the interactions and constraints it uncovers (see, e.g., discussions in Michel & Costa, 1998). Although the new structure we propose for domain V is fully consistent with both the comparative and probing data available at the moment, we cannot rule out the possibility of the dV two-nucleotide bulge being a dynamic element in group II function.

Implications for group II-spliceosome comparisons

The 9-26 pairing in dV may also be relevant to structural parallels that have been proposed between dV and conserved elements within U6 snRNA (Madhani & Guthrie, 1992; Chanfreau & Jacquier, 1994; Boulanger et al., 1995; Peebles et al., 1995; Yu et al., 1995). These comparisons center on similarities observed in the overall secondary structure, the effects of mutations in the conserved AGC trinucleotide in each of these molecules, and in the relative positions of phosphate oxygens important to group II catalysis and pre-mRNA splicing (Fabrizio & Abelson, 1992; Madhani & Guthrie, 1992; Chanfreau & Jacquier, 1994; Peebles et al., 1995; Yu et al., 1995). The spliceosomal counterpart to dV is proposed to be formed by the stacking of U2/U6 helix Ib and the U6 intramolecular stem-loop whose consensus secondary structure is shown in Figure 7C (Chanfreau & Jacquier, 1994; Yu et al., 1995). The proposed base pair between positions 9 and 26 in dV makes this structural analogy between dV and the U6/U2 stem loop even more compelling (Fig. 7A,B,C). The proposed pair not only creates a more equivalent helical distance between the highly conserved AGC trinucleotide and the dinucleotide bulge, but also shifts the location of the phosphorothioate effect in the bulge to an analogous position to that observed in U6. The proposed secondary structure thus lends further support to the use of group II introns as potential analogues to the complex mechanism of pre-mRNA processing within the spliceosome.

MATERIALS AND METHODS

DNA constructs

The "wild-type" PI.LSU/2 intron construct used in this work, which lacks all but 56 nt of the terminal loop of domain IV and is flanked by 5' and 3' exons of 126 and 106 nt, is described in Costa et al. (1997b). The domain V mutant version of this molecule (L5:UUCG, see Results) was obtained by subcloning of PCR fragments and was verified by sequencing the entire length of the insert.

RNA synthesis and purification

All transcripts were internally labeled by transcription in the presence of [α - 32 P]UTP. RNA synthesis and purification were performed as described by Costa et al. (1997b). Templates for synthesis of the wild-type and L5:UUCG precursor transcripts were generated by *Aco65* I digestion of the corresponding plasmids. DNA templates for synthesis of domain V (dV) and domain V and VI (dV-dVI) transcripts were generated by PCR on wild-type plasmid DNA with primer 5'-GGAGGACTGCAGTAATACGACTCACTATAGGAAAGCAG AATGCGTGAG (T7 promoter underlined) on the one hand, and primers 5'-TTGGGAGGTACCGGATGGGTCTCATCAGA ACCGTGCTTGCGAC (dV) and 5'-ATGGGAGGTACCGGTC

TCAGTTGGATAGGTAGACGA (dV-dVI), on the other hand. In the case of dV-dVI, the amplified DNA matrix was digested with *BsmA* I prior to transcription.

Samples of wild-type and L5:UUCG lariat intron molecules destined for kinetic analyses, chemical modification, and melting experiments were prepared from preparative splicing reactions performed under optimal conditions. Gel-purified precursor RNA was incubated for 80 min (wild-type) or 120 min (L5:UUCG mutant) at 45 °C in 40 mM Tris-HCl, pH 7.5 (25 °C), 1 M NH₄Cl, 0.02% (w/v) SDS, and 10 mM (wild-type) or 20 mM (L5:UUCG mutant) MgCl₂. Splicing products were then loaded onto 8 M urea/4% polyacrylamide gels and lariat intron RNA was recovered from the gel as described by Costa and Michel (1995). All RNA concentrations were determined from UV absorbance measurements.

Kinetic analyses

For debranching of lariat molecules, a ribooligonucleotide (rE5) with the same sequence as the last 13 residues of the natural 5' exon (5'-UGUUUAUUAAAAA) was synthesized (Genset, Paris). Debranching reactions were performed in 40 mM Na-MES, pH 5.4, 1 M NH₄Cl, 0.02% (w/v) SDS, with diverse concentrations of MgCl₂ (see figure legends). For renaturation of purified intron lariat RNA, samples were preincubated in water at 55 °C (a temperature slightly above the melting range for tertiary structure), an equal volume of 2 $\times$ -concentrated splicing buffer was added, and the temperature was slowly (-0.8 °C/min) decreased to 40 °C. Debranching was initiated by addition of an equal volume of rE5 solution in 1 $\times$ -concentrated buffer. For all reactions, the initial lariat intron concentration was set at 10 nM. Reactions were stopped by addition of an equal volume of a solution of 98% (v/v) formamide, 0.025% (w/v) bromophenol blue, 0.025% (w/v) xylene cyanol, with a concentration of Na₂EDTA, pH 8.0, sufficient to exceed that of MgCl₂. Reaction products were analyzed by electrophoresis on 8 M urea/4% polyacrylamide gels and radioactivity was quantitated on fixed, dried gels using a PhosphorImager (Molecular Dynamics). The linear intron-5' exon resulting from debranching is only 13 nt longer than the broken lariat form, which, in the case of group II introns, migrates as a linear molecule. Nevertheless, broken lariat molecules constituted such a small fraction of our preparations (typically between 3 and 5% of total intron) that their presence did not interfere with quantitation of debranched molecules.

DMS modification

In addition to MgCl₂ (see figure legends), buffers used for modification experiments contained 40 mM Na-HEPES, pH 7.6 at 37 °C, 1 M NH₄Cl, and 0.02% (w/v) SDS. Typical modification reactions contained either 13 nM (lariat intron) or 53 nM (dV and dV-dVI) RNA in 50 μ L. In order to ensure that populations of molecules being probed were folded uniformly and correctly, RNA samples were renatured (see Kinetic analyses) by slow-cooling from 55 °C to either 37 or 42 °C (see figure legends). After a further 15-min incubation at the final temperature, 1 μ L of DMS diluted in ethanol was added where appropriate to a final dilution of 1:400 to 1:600 (see figure legends for duration of the modification step).

Reactions were stopped by precipitating the RNA at -70°C after addition of 10 μg of bulk yeast tRNA (Boehringer), 60 μL of 0.5 M sodium acetate, and 300 μL of ethanol. RNA pellets were washed twice with 70% ethanol prior to drying. Modified RNAs were resuspended in water at a final concentration of 0.08 μM prior to reverse transcription.

Reverse transcription

Reverse transcription was performed as described by Costa et al. (1997b) with two minor modifications: annealing of 0.08 pmol of DMS- (or mock-) treated RNA and 0.24 pmol of (5'- ^{32}P)-labeled complementary DNA primer was performed in 5 μL and all elongation and sequencing reactions were performed at 46°C . Reactions were stopped by addition of 10 μL of a solution containing 60 mM Na_2EDTA and 0.04% SDS. Degradation of the RNA template was ensured by addition of 2 μL of NaOH 3 M, followed by incubation for 3 min at 90°C , and a further 40 min at 40°C . The solution was neutralized by addition of 2 μL of 3 M HCl. Samples were subsequently precipitated in the presence of 0.3 M sodium acetate and three volumes of ethanol. Pellets were rinsed with 75% ethanol before being resuspended in formamide loading solution (see Kinetic analyses) with 10 mM Na_2EDTA , pH 8.0, added. Samples were heated to 80°C for 10 min prior to loading onto denaturing polyacrylamide gels. Fixed and dried gels were quantitated with a PhosphorImager (Molecular Dynamics).

Optical melting curves

Melting curves were generated from 5 μg of wild-type or L5:UUCG mutant lariat intron RNA in 600 μL of solution containing 40 mM KCacodylate, pH 7.5 at 25°C , 1 M NH_4Cl , 10 mM MgCl_2 , and 0.02% SDS. Lariat intron RNA was renatured as indicated in Kinetic analyses, except that samples were slow-cooled ($0.5^{\circ}\text{C}/\text{min}$) down to 25°C . Melting was monitored by measuring absorbance at 260 nm on a Uvikon 941 Plus spectrophotometer (Kontron) equipped with two cells that were thermostated by a circulating water bath coupled to a thermoprogrammer. Data were collected every 0.05°C from 30 to 80°C while samples were being heated at a rate of $0.2^{\circ}\text{C}/\text{min}$.

ACKNOWLEDGMENTS

We thank D. McPheeters for providing sequence alignments of U2 and U6, Luc Jaeger for help in melting experiments, and J. Gott, M. Harris, A. Jacquier, V. Knoop, D. McPheeters, T. Nilsen, C. Peebles, P. Perlman, A.M. Pyle, and J.A. Wise for helpful discussions. This work was supported by a fellowship from the Fondation pour la Recherche Médicale to M.C. and a postdoctoral fellowship AI09293 from the National Institutes of Health to E.L.C.

Manuscript accepted without revision June 8, 1998

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.

- Augustin S, Müller MW, Schweyen RJ. 1990. Reverse self-splicing of group II intron RNAs in vitro. *Nature* 343:383–386.
- Banerjee AR, Jaeger JA, Turner DH. 1993. Thermal unfolding of a group I ribozyme: The low-temperature transition is primarily disruption of tertiary structure. *Biochemistry* 32:153–163.
- 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.
- Butcher SE, Dieckmann T, Feigon J. 1997. Solution structure of a GAAA tetraloop receptor RNA. *EMBO J* 16:7490–7499.
- Cate JH, Gooding AR, Podell E, Zhou KH, Golden BL, Kundrot CE, Cech TR, Doudna JA. 1996a. Crystal structure of a group I ribozyme domain—Principles of RNA packing. *Science* 273:1678–1685.
- Cate JH, Gooding AR, Podell E, Zhou KH, Golden BL, Szewczak AA, Kundrot CE, Cech TR, Doudna JA. 1996b. RNA tertiary structure mediation by adenosine platforms. *Science* 273:1696–1699.
- Cech TR. 1986. The generality of self-splicing RNA: Relationship to nuclear mRNA splicing. *Cell* 44:207–210.
- 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, Dème E, Jacquier A, Michel F. 1997a. Multiple tertiary interactions involving domain II of group II self-splicing introns. *J Mol Biol* 267:520–536.
- Costa M, Fontaine JM, Loiseaux-de Goër S, Michel F. 1997b. A group II self-splicing intron from the brown alga *Pylaiella littoralis* is active at unusually low magnesium concentrations and forms populations of molecules with a uniform conformation. *J Mol Biol* 274:353–364.
- Costa M, Michel F. 1995. Frequent use of the same tertiary motif by self-splicing RNAs. *EMBO J* 14:1276–1285.
- Costa M, Michel F. 1997. Rules for RNA recognition of GNRA tetraloops deduced by in vitro selection: Comparison with in vivo evolution. *EMBO J* 11:3289–3302.
- Daniels DL, Michels WJ Jr, Pyle AM. 1996. Two competing pathways for self-splicing by group II introns: A quantitative analysis of in vitro reaction rates and products. *J Mol Biol* 256:31–49.
- Draper DE. 1996. Strategies for RNA folding. *Trends Biochem Sci* 21:145–149.
- Fabrizio P, Abelson J. 1992. Thiophosphates in yeast U6 snRNA specifically affect pre-mRNA splicing in vitro. *Nucleic Acids Res* 20:3659–3664.
- Fersht A. 1977. *Enzyme structure and mechanism*. New York: W.H. Freeman and Company.
- Fontaine JM, Rousvoal S, Leblanc C, Kloareg B, Loiseaux-de Goër S. 1995. The mitochondrial LSU rDNA of the brown alga *Pylaiella littoralis* reveals α -proteobacterial features and is split by four group IIB introns with an atypical phylogeny. *J Mol Biol* 251:378–389.
- Franzen JS, Zhang MC, 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.
- Harris-Kerr CL, Zhang M, Peebles CL. 1993. The phylogenetically predicted base-pairing interaction between α and α' is required for group II splicing in vitro. *Proc Natl Acad Sci USA* 90:10658–10662.
- Inoue T, Cech TR. 1985. Secondary structure of the circular form of the *Tetrahymena* rRNA intervening sequence: A technique for RNA structure analysis using chemical probes and reverse transcriptase. *Proc Natl Acad Sci USA* 82:648–652.
- Jacquier A, Michel F. 1987. Multiple exon-binding sites in class II self-splicing introns. *Cell* 50:17–29.
- Jacquier A, Michel F. 1990. Base-pairing interactions involving the 5' and 3'-terminal nucleotides of group II introns. *J Mol Biol* 213:437–447.
- Jaeger L, Michel F, Westhof E. 1994. Involvement of GNRA tetraloop in long-range RNA tertiary interactions. *J Mol Biol* 236:1271–1276.
- Jaeger L, Westhof E, Michel F. 1993. Monitoring of the cooperative unfolding of the *sunY* group I intron of bacteriophage T4. The

- active form of the *sunY* ribozyme is stabilized by multiple interactions with 3' terminal intron components. *J Mol Biol* 234:331–346.
- Jestin JL, Dème 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.
- 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.
- Kwakman JHJM, Konings DAM, Hogeweg P, Pel HJ, Grivell LA. 1990. Structural analysis of a group II intron by chemical modifications and minimal energy calculations. *J Biomol Str and Dyn* 8:413–430.
- Madhani H, Guthrie C. 1992. A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome. *Cell* 71:803–817.
- Michel F, Costa M. 1998. Inferring RNA structure by phylogenetic and genetic analyses. In: Simons RW, Grunberg-Manago M, eds. *RNA structure and function*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press, pp 175–202.
- Michel F, Ferat JL. 1995. Structure and activities of group II introns. *Annu Rev Biochem* 64:435–461.
- Michel F, Jacquier A. 1987. Long-range intron-exon and intron-intron pairings involved in self-splicing of class II catalytic introns. *Cold Spring Harbor Symp Quant Biol* 52:201–212.
- Michel F, Jacquier A, Dujon B. 1982. Comparison of fungal mitochondrial introns reveals extensive homologies in RNA secondary structure. *Biochimie* 64:867–881.
- Michel F, Umesono K, Ozeki H. 1989. Comparative and functional anatomy of group II catalytic introns—A review. *Gene* 82:5–30.
- Murphy FL, Cech TR. 1994. GAAA tetraloop and conserved bulge stabilize tertiary structure of a group I intron domain. *J Mol Biol* 236:49–63.
- Nilsen TW. 1998. RNA–RNA interactions in nuclear pre-mRNA splicing. In: Simons RW, Grunberg-Manago M, eds. *RNA structure and function*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press, pp 279–307.
- 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.
- 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, 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.
- 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^{++} ions in the tertiary structure. *Biochemistry* 36:4718–4730.
- 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.
- Sharp PA. 1991. "Five easy pieces". *Science* 254:663.
- Strobel SA, Ortoleva-Donnelly L, Ryder SP, Cate JH, Moncoeur E. 1998. Complementary sets of noncanonical base pairs mediate RNA helix packing in the group I intron active site. *Nature Struct Biol* 5:60–66.
- Tuerk C, Gauss P, Thermes C, Groebe DR, Gayle M, Guild N, Stormo G, d'Aubenton-Carafa Y, Uhlenbeck OC, Tinoco I Jr, Brody EN, Gold L. 1988. CUUCGG hairpins: Extraordinarily stable RNA secondary structures associated with various biochemical processes. *Proc Natl Acad Sci USA* 85:1364–1368.
- Weiner AM. 1993. mRNA splicing and autocatalytic introns: Distant cousins or the products of chemical determinism? *Cell* 72:161–164.
- Westhof A, Masquida B, Jaeger L. RNA tectonics: Towards RNA design. *Folding & Design* 1:R78–R88.
- Yu YT, Maroney P, Darzynkiewicz E, Nilsen T. 1995. U6 snRNA function in nuclear pre-mRNA splicing: A phosphorothioate interference analysis of the U6 phosphate backbone. *RNA* 1:46–54.