More than one way to splice an RNA: Branching without a bulge and splicing without branching in group II introns

VI T. CHU,¹ QIAOLIAN LIU,¹ MIRCEA PODAR,² PHILIP S. PERLMAN,²
and ANNA MARIE PYLE¹

¹Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10032, USA

²Department of Molecular Biology and Oncology, University of Texas Southwestern Medical Center,
Dallas, Texas 75235-9148, USA

ABSTRACT

Domain 6 (D6) of group II introns contains a bulged adenosine that serves as the branch-site during self-splicing. In addition to this adenosine, other structural features in D6 are likely to contribute to the efficiency of branching. To understand their role in promoting self-splicing, the branch-site and surrounding nucleotides were mutagenized. Detailed kinetic analysis on the self-splicing efficiency of the mutants revealed several interesting features. First, elimination of the branch-site does not preclude efficient splicing, which takes place instead through a hydrolytic first step. Second, pairing of the branch-site does not eliminate branching, particularly if the adenosine is involved in a mispair. Third, the G-U pairs that often surround group II intron branch-points contribute to the efficiency of branching. These results suggest that there is a strong driving force for promoting self-splicing by group II introns, which employ a versatile set of different mechanisms for ensuring that splicing is successful. In addition, the behavior of these mutants indicates that a bulged adenosine per se is not the important determinant for branch-site recognition in group II introns. Rather, the data suggest that the branch-site adenosine is recognized as a flipped base, a conformation that can be promoted by a variety of different substructures in RNA and DNA.

Keywords: base-flipping; catalysis; conformation; evolution; mismatch; ribozyme; self-splicing

INTRODUCTION

Group II intron self-splicing is generally thought to proceed through a series of two transesterification reactions (Michel & Ferat, 1995). Much like nuclear splicing (Padgett et al., 1984; Ruskin et al., 1984), the 2'-hydroxyl group on a bulged adenosine within the intron has been observed to act as the nucleophile during the first step of splicing. This adenosine is located in Domain 6 (D6) of group II introns, which is a short hairpin-loop structure located upstream of the 3'-splice site (Fig. 1). During the second step of splicing, the exons are ligated and the branched intron is released as a lariat-shaped molecule (Peebles et al., 1986; Schmelzer & Schweyen, 1986; van der Veen et al., 1986). Given this mechanistic pathway, the bulged adenosine of group II introns would seem to be of primary importance for splicing. However, certain group II introns lack

a bulged adenosine in Domain 6 (Michel et al., 1989). In addition, there is considerable evidence that splicing can proceed through an alternative pathway in which the 5'-splice site is cleaved by hydrolysis rather than transesterification during the first step (van der Veen et al., 1987; Jarrell et al., 1988; Daniels et al., 1996; Podar et al., 1998). Although a branch-point is not strictly required for self-splicing, lariat formation represents an important, and perhaps the major, pathway for self-splicing. Nonetheless, it is not clear how the branch-point adenosine is recognized, or if the bulged structure is important for transesterification.

To explore recognition of the branch-site adenosine, there has been extensive analysis of point mutations and single-atom changes at the branch-site in group II intron ai5 γ (Fig. 2). It has been shown that bulged G and U branch-points promote a low level of branching, with a rate that is 100-fold slower than that of the wild-type adenosine, whereas a bulged C at the branch-point does not result in any detectable transesterification (Liu et al., 1997). Because all three of these mutations severely damage the transesterification pathway, hy-

Reprint requests to: Anna Marie Pyle, Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10032, USA; e-mail: amp11@columbia.edu.

FIGURE 1. Schematic secondary structure of a group II intron and the reactions involved in self-splicing. There are two competing pathways in group II intron self-splicing. In the first step of splicing (labeled branching step 1), the 2'-hydroxyl group of a bulged adenosine in Domain 6 (shown in bold) attacks the 5' splice site (shown as an open circle). During the second step of splicing (labeled step 2), the free 3'-hydroxyl of the 5' exon attacks the 3'-splice site (shown as a closed circle), resulting in ligated exons and excised lariat intron (not shown). Alternatively, the first step of splicing can also occur through a hydrolytic pathway in which hydroxide or H_2O is the attacking nucleophile (labeled hydrolytic step 1). The 5' splice site is selected through base pairing interactions between the two exon binding sequences, EBS 1 and EBS 2 (shown as bold boxes), with two intron binding sequences, IBS 1 and IBS 2 (Jacquier & Michel, 1987), and to some extent, the $\epsilon-\epsilon'$ interaction (Michel & Ferat, 1995).

drolysis becomes the major pathway for the first step of splicing in these cases. To introduce less radical changes of the branch-point, single-atom changes were made to the adenosine, and their effects on branching rate indicated that a major site of recognition was the amino group (N6) of adenosine, whereas other positions, such as N7, played only a minor role in the reaction (Liu et al., 1997). In terms of relative reactivity of branch-point mutants, these results contrast sharply with those observed for the spliceosome, which preferentially tolerates bulged C mutations more than G or U (Hornig et al., 1986; Vijayraghavan et al., 1986; Query et al., 1996). Although the data on group II intron bulge mutants helped to show that adenosine is selected by a combination of steric selection and docking of N6, they did not reveal how the adenosine is presented by D6 so that it can be recognized by the core. To understand the conformational features that are important at the branch-site, it is important to consider the overall morphological features of D6.

The highly conserved bulged adenosine is only one of several structural features that are consistently observed in D6 (Fig. 2). For example, there is usually one, and there are often two, G-U pairs flanking the bulged adenosine (Michel et al., 1989). Given their conservation, it is possible that these flanking wobble pairs contribute to recognition and reaction by the branch-point adenosine. There are two other features that, although

FIGURE 2. Sequence and secondary structure of ai5 γ D6, indicating sites of mutation. The branch-point adenosine is labeled with an asterisk. On the right side of D6, bracketed nucleotides of the internal loop were mutated simultaneously to the indicated nucleotides to form the mutation D6 zip. Bracketed uridines on either side of the branch-point adenosine were changed simultaneously to C (GU $\rightarrow$ GC). The ΔA_{880} mutation indicates a variant in which the bulged nucleotide was deleted. On the left side of D6, prA-U and prA-G indicate two different mutations in which the branch-point adenosine was paired with a uridine or a guanosine, respectively, prepared by inserting U or G at the indicated position. Except for the mutations indicated, precursor RNAs used in this study were identical to the WT RNA transcribed from plasmid pJD20 (Koch et al., 1992; Daniels et al., 1996).

not phylogenetically conserved, are common structural elements found in D6. First of all, D6 is often capped by one of several different types of tetraloop at the terminus of the hairpin. This tetraloop is believed to play a role in conformational changes that occur between steps 1 and 2 and has been the subject of recent investigations on tertiary interactions with Domain 2 (Chanfreau & Jacquier, 1996; Costa et al., 1997). Second, the D6 of many introns contains an internal loop that is several base pairs away from the branch-site itself. This motif, like the wobble G-U pairs surrounding the branch-point, may be an important feature that plays a role in promoting branching by certain group II introns.

To explore the importance of branch-site structure, new mutations were made in several regions of D6 (Fig. 2). Three mutations directly disrupted the branch-point either by deletion of the adenosine (ΔA_{880}) or by pairing it with a uridine or a guanosine inserted on the other strand of D6 (prA-U, prA-G). The prA-G mutant is particularly interesting because it was initially obtained as a spontaneous revertant of the prA-U mutation in vivo (M. Podar & P.S. Perlman, unpubl. results). The mutation was reconstructed here for analysis of its in vitro kinetics. To explore the role of the flanking G-U

pairs, a mutant was constructed in which both uridines were changed to cytidines, generating G-C pairs on either side of the bulged adenosine (GU → GC, see also Podar et al., 1998). Finally, to study the function of the D6 internal loop, it was "zipped up" by the substitution of perfect A-U pairs in place of the mismatches normally found in this region (D6 zip). In all cases, kinetic effects and splicing reaction products were studied under two sets of conditions. (1) A permissive high-salt buffer was used that is known to stimulate and optimize the branching pathway (Peebles et al., 1987; Jarrell et al., 1988). In addition, (2) a low-salt buffer was used that renders the intron sensitive to subtle mutational effects that are otherwise masked under high salts (Peebles et al., 1987; Chanfreau & Jacquier, 1994) and that also correlates well with effects observed in vivo (Boulanger et al., 1995).

RESULTS

Survey of self-splicing reactions

When the splicing reactions are compared under high-salt conditions, the most striking result is that the ΔA_{880}

mutant is completely unable to branch and form lariat intron (Fig. 3, lanes 5–7). Nonetheless, splicing proceeds relatively efficiently through hydrolysis in which linear intron and spliced exons are generated. Similarly, the prA-U mutant splices almost exclusively through the hydrolytic pathway (Fig. 3, lanes 8–10), although, in this case, a tiny amount of branched lariat product can be observed after extended exposure of the autoradiogram (Fig. 3, lane 18). Behavior of the ΔA_{880} and prA-U mutants is analogous to that reported in other studies (Schmelzer & Müller, 1987; van der Veen et al., 1987). A novel mutation was studied in which the A is paired with a G, forming a noncanonical base pair interaction (prA-G). In contrast to the ΔA_{880} and prA-U mutants, the prA-G mutant displays normal splicing behavior. In fact, the apparent efficiency and extent of branching in this mutant is comparable to that of the WT intron (Fig. 3, lanes 19–21; Table 1). Therefore, when both the bulge and the adenosine are physically removed, as in the case of ΔA_{880} , branching is completely eliminated. However, when the A is present, but paired in a strong Watson-Crick interaction, branching occurs at a dramatically reduced level. When the A is present but paired in a weaker interaction (prA-G), branching is

FIGURE 3. Short time courses for in vitro splicing of WT and mutant precursor RNAs. Internally ^{32}P -labeled precursor RNAs were incubated in high-salt splicing buffer [500 mM $(\text{NH}_4)_2\text{SO}_4$, 100 mM MgCl_2 , 40 mM MOPS, pH 7.5]. Aliquots of each reaction were taken at times indicated and products were separated on a 5% denaturing polyacrylamide gel. The composition of reaction precursors and products are diagrammed on the left. Under high-salt splicing buffer, the wild-type precursor intron (band B) is converted to lariat intron (band A), linear intron (band C), and spliced exons (band D). Because the first step is rate limiting and the second step is fast, lariat 3'-exon intermediate (*) from the branching pathway and linear 3'-exon intermediate from the hydrolysis pathway do not accumulate in abundance during any of the reactions. After extended autoradiographic exposures of Figure 3 and other time courses, only WT and D6zip mutants contained traces of the lariat 3'-exon at the 2-min time point, which disappeared at later time points. Although it is possible that the prA-G mutant might reveal second step defects under other conditions, under these conditions (Materials and Methods) there was no greater propensity for the second step of prA-G to be inhibited. Figure 3 was cropped to allow visualization of the wells and any lariat 3'-exon that might be apparent. When the branch-point adenosine is deleted or paired with uridine, there is a noticeable absence of lariat intron product and an abundance of linear intron and spliced exons (lanes 5–10). Extended exposure of lanes 7 and 10 reveals a tiny amount of lariat intron produced by the prA-U mutant (lane 18).

TABLE 1. Apparent rate constants for self-splicing by mutant introns under two conditions.^a

Intron Mutant	A. High salt ^b		B. Low salt ^b	
	k_{br}^c (10^{-2} min^{-1})	k_{hy} (10^{-2} min^{-1})	k_{br} (10^{-3} min^{-1})	k_{hy} (10^{-2} min^{-1})
WT	14	0.43	4.1	0.63
ΔA_{880}	ND ^d	1.8	ND	0.99
prA-U	0.012	1.7	ND	0.35
prA-G	9.4	0.63	0.72	0.48
D6 zip	9.9	0.70	1.2	1.1
GU $\rightarrow$ GC	4.4	0.93	0.37	1.1

^aObserved rate constants for defined sets of reaction conditions are reported here. To date, they have not been shown to represent rate constants for specific chemical or conformational events. Standard deviations in all values are provided in the figure captions.

^bHigh-salt conditions: 500 mM $(\text{NH}_4)_2\text{SO}_4$, 100 mM MgCl_2 , 40 mM MOPS, pH 7.5, with NaOH. Low-salt conditions: 100 mM MgCl_2 , 40 mM MOPS, pH 7.5, with KOH.

^cVariables k_{br} and k_{hy} are the rate of branching and the rate of hydrolysis, respectively, derived from the plots in Figures 4, 5, 7, and 8.

^dND indicates that no reaction products could be detected, even after many weeks of exposure on film or by extensive phosphor-imager scanning.

unaffected. These three sets of observations suggest that branching involves more than the recognition of a bulged structure per se. Additionally, the data indicate that the lack of a functional branch-point does not prevent the intron from splicing readily through an alternative pathway in which water or hydroxide is the nucleophile during the first step. For all D6 mutants, the hydrolytic rates are the same or slightly faster than wild-type hydrolysis (Fig. 3; Table 1), indicating that the intron active site is not perturbed by mutations of the branch-point.

Under high-salt conditions, mutation of the other structural elements in D6 does not severely affect the rate of branching. Behavior of the D6 zip RNA is very similar to that of WT, suggesting that the internal loop within D6 does not have a direct role in the mechanism of branching (Fig. 3, lanes 11–13). Another mutation of D6 that does not severely affect branching is the GU $\rightarrow$ GC mutation. Under high-salt conditions, the rate of branching by this mutant appears to be only slightly impaired (Fig. 3, lanes 14–16). This cursory analysis suggests that lariat formation is not catastrophically curtailed when structural elements surrounding the branch-point are altered. Although this type of analysis is useful for rapid evaluation of splicing reaction products by the mutant introns, a more detailed kinetic examination of the splicing reactions is required to interpret the effects of mutations and to understand the basis for the observed behavior of the mutants. To this end, extensive time courses were performed under high- and low-salt conditions.

Splicing kinetics of ΔA_{880} and prA-U

Mutations that delete the branch-point adenosine (ΔA_{880}) or pair it with a uridine (prA-U) cause splicing to proceed rapidly through the hydrolytic pathway under high-salt conditions (Fig. 4A,B). One of the most notable features of these reactions is that hydrolytic splicing appears to proceed faster in these mutants (0.017 min^{-1}) than in the WT intron (0.0043 min^{-1} , Table 1). Despite appearances, the deletion of one pathway (branching) does not cause the other pathway (hydrolysis) to become inherently more efficient. Instead, the fourfold differences in rate can be attributed to the fact that there are two reactive conformations of the ai5 γ intron under the high-salt condition (Podar et al., 1995; Daniels et al., 1996). In the wild-type intron, the faster of the two populations (A) represents $\sim 60\%$ of all the molecules, and the slow population (B) represents $\sim 40\%$ of the molecules. When the WT intron reacts under high-salt conditions in the presence of $(\text{NH}_4)_2\text{SO}_4$, the fast population A reacts exclusively through branching and the slow population B reacts through hydrolysis (Daniels et al., 1996). In the mutant introns, where branching is made impossible by deletion of the bulged adenosine, population A is still more reactive than population B, but it is limited to reaction through hydrolysis. Therefore, hydrolytic splicing by the mutant population A occurs $\sim$ fourfold faster than hydrolytic splicing by WT population B. Interestingly, in wild-type reactions containing high concentrations of KCl, the fast population reacts through both hydrolysis and branching mechanisms (Jarrell et al., 1988; Daniels et al., 1996). In that case, the rate of hydrolytic splicing by the fast population is comparable to the rates of hydrolytic splicing by mutants containing a deleted or paired A-U branch-point. Therefore, the faster hydrolytic splicing by the mutant introns compared with WT is due to higher overall reactivity of population A, which, in these cases, cannot react efficiently through branching.

Another notable feature of these mutants is that a tiny amount of branching can be observed for the prA-U mutant under high-salt conditions (Fig. 3, lane 18; Fig. 4B, inset) (van der Veen et al., 1987). This indicates that, even though it is involved in a stable base pair, the branch-point A has some propensity to participate in transesterification. Only by deletion of the A, replacement with C or dA as observed previously (Liu et al., 1997), can one eliminate the branching pathway entirely. In the case of the prA-U mutant, branching occurs in parallel with hydrolysis (they both have the same reaction half-time and k_T) (Fig. 4B, inset), but normalization for the extent of reaction (Equations 4 and 5, Materials and Methods) results in a rate of branching (k_{br}) that is extremely slow ($1.24 \times 10^{-4} \text{ min}^{-1}$, consuming only 0.6% of precursor RNA, Table 1).

(A) ΔA / High Salt **(B) prA-U / High Salt**

FIGURE 4. Kinetic analysis of the ΔA_{880} and prA-U mutants under high-salt conditions. Experimental data for the decrease in precursor (filled squares) and the evolution of linear (filled triangles) and lariat (closed circles) intron products are plotted and fit using Equations 1 and 2, respectively. **A:** Time course for the self-splicing reaction of ΔA_{880} . ΔA_{880} precursor is converted exclusively to linear intron product over time. Fraction of ΔA_{880} precursor molecules that reacts over time (a_1) = 0.880 (± 0.010) and the overall rate of reaction (k_T) = 0.0182 (± 0.00063) min^{-1} . Evolution of linear intron product was fit using Equation 2, where f_{hy} = 0.878 (± 0.010) and k_T = 0.0185 (± 0.00063) min^{-1} . Because $a_1 = f_{hy}$ in this case and $k_T(\text{precursor}) = k_T(\text{linear intron})$, the overall rate of the reaction, $k_T = k_{hy}$ = 0.0181 (± 0.00063) min^{-1} . **B:** Time course for the self-splicing reaction of mutant prA-U. The prA-U precursor (closed squares) is converted predominantly to linear intron product (closed triangles) and a trace of lariat intron product (closed circles, inset). The fractional coefficients (f_{hy} and f_{br}) for the relative accumulation of linear and lariat intron product, represented by the asymptotes of the product curves, are 0.856 (± 0.016) and 0.00632 (± 0.00057), respectively. The k_T value determined from precursor fit is 0.0166 (± 0.00085) min^{-1} . The k_T values determined from the fit of linear and lariat intron product evolution are 0.0166 (± 0.00084) min^{-1} and 0.0174 (± 0.0044) min^{-1} , respectively. The k_T value used in subsequent calculations was the average of k_T values obtained from lariat and linear intron fits. Applying Equations 4 and 5, $k_{br} = 1.24 \times 10^{-4}$ min^{-1} and $k_{hy} = 0.0169$ min^{-1} .

The data shown thus far were obtained under highly permissive salt conditions (100 mM MgCl_2 , 500 mM $(\text{NH}_4)_2\text{SO}_4$, pH 7.5), in which mutations generally have small effects and the branching reaction is greatly favored (Peebles et al., 1987; Jarrell et al., 1988; Koch et al., 1992). To examine more subtle effects on intron function, the kinetics of self-splicing were also studied under a low-salt condition (100 mM MgCl_2 , pH 7.5) in which the intron is less stabilized and branching and hydrolysis occur with lower efficiency (Peebles et al.,

1987; Chanfreau & Jacquier, 1994). Under this less permissive condition, splicing by the wild-type intron is relatively inefficient and the rate of hydrolysis dominates over that of branching (Table 1).

As expected, under low-salt conditions, the branch-point mutants ΔA_{880} and prA-U splice only through the hydrolytic pathway, with rates and kinetic profiles similar to that of WT (Table 1) (Daniels et al., 1996). Mutations of the branch-point A to C, G, or U also result in a complete elimination of branching under these low-

salt conditions, with splicing rate constants comparable to the ΔA_{880} and prA-U mutants (Q. Liu & A.M. Pyle, data not shown).

Under both low- and high-salt conditions, there is a lag in the splicing reaction at short times. This lag is much smaller under high-salt conditions than low-salt conditions. The source of this lag is not clear at this time, although it is believed to represent a delay in the folding of intron molecules immediately after mixing with reaction salts (Qin & Pyle, 1997). It is unlikely that this lag has been introduced by mutations of the intron, because the WT intron has a similar lag under low-salt conditions where hydrolysis is the dominant reaction pathway.

Splicing kinetics of prA-G mutant

Quantitative analysis of reactivity by the prA-G mutant under high-salt conditions confirms that it is very similar to that of WT (Fig. 5A; Table 1). However, the prA-G intron is less reactive under low-salt conditions. Although the rate of hydrolytic splicing remains the same as WT, the rate of transesterification is sixfold slower (Fig. 5B; Table 1). Interestingly, the prA-G mutation is less disruptive to branching than the GU $\rightarrow$ GC mutation (see below), in which the branch-point remains bulged and intact but the neighboring base pair interactions have been strengthened (Table 1). These trends observed at lower salt parallel the behavior of the mutant introns *in vivo*. In that case, the yeast strain containing the prA-G mutant intron splices efficiently (M. Podar & P.S. Perlman, unpubl. results), whereas the strain transformed with the GU $\rightarrow$ GC mutant is nearly splicing deficient (Podar et al., 1998).

Mapping the branch-point of the prA-G mutant

One explanation for efficient branching by the prA-G mutant is the possibility of cryptic splicing at an adjacent site. To explore this possibility, a general strategy was developed for high-resolution mapping of lariat branch-points in group II introns (or any other RNA containing a modified site). Although a branch-point can be mapped effectively by inclusion of a single radioactive phosphate at the branch-site followed by nuclease digestion, it requires the creation of new synthetic constructs for any mutant tested (Gaur et al., 1997; Liu et al., 1997). Mapping has also been reported by 3' end-labeling of the lariat RNA followed by alkaline cleavage (Gaur et al., 1997); however, we observe that the large size of the *ai5 γ* intron (887 nt) complicates this analysis. A hydrolysis ladder of six nucleotides (the distance from nt 881 to the labeled 3' end) is difficult to detect on this large RNA molecule unless it is over-cleaved, necessitating strong conditions that can destabilize the 2'-5' link.

To map the prA-G branch-point, DNAzymes were used to clip the branched region out of the parent lariat, allowing subsequent end-labeling and mapping to high resolution. The DNAzymes act essentially like RNA restriction enzymes to cleave specifically on either side of the branch, yielding a small Y-shaped molecule (Fig. 6A). The free 5'-OH and 2'3'-PO₃ ends generated by the cleavage allows the small branched RNA to be kinased at a single 5' end with γ -³²P-ATP. The labeled RNA is then subjected to partial alkaline hydrolysis and the reaction products are separated on a 20% denaturing polyacrylamide gel. Because the 2'-OH on the branch nucleotide is engaged in a 2'-5' linkage, the phosphodiester linkage immediately 3' to it is resistant to hydrolysis. This results in an eight-nucleotide gap in the hydrolysis ladder (the length of intron sequence downstream of the branch-site, including the branch-point A), allowing one to count from the 5' end of the fragment to the branch-site, thereby revealing its precise location (Fig. 6A). When the WT and prA-G mutant are both treated with the DNAzymes, followed by 5' labeling and mapping of the branched products, it is clear that they both result in a gap at the same site (Fig. 6B), which is 17 nt from the 5' end (DNAzyme 1 was targeted 17 nt upstream of the branch-point in each case). It is possible that certain populations of molecules may be branching at nucleotides 5' upstream of the adenosine, in which case, those cryptic sites would be obscured by an alkaline hydrolysis ladder of 5' labeled molecules. To preclude this possibility, DNAzyme-cleaved branched molecules were also mapped from the 3' end (data not shown). As mapped from both the 5' and 3' ends, the prA-G mutant branches primarily at the correct nucleotide, despite the fact that it is no longer distinguished as a bulged residue (Fig. 6B). That branching is accurate *in vivo* is supported by primer extension studies performed on isolated lariat-3'-exon molecules resulting from splicing of the prA-G mutant in yeast. In that case, the lariat species were all branched from the correct A residue (M. Podar & P.S. Perlman, unpubl. results), as observed here *in vitro*.

Splicing kinetics of the GU $\rightarrow$ GC and D6zip mutants

Replacement of the G-U pairs that flank the branch point with G-C pairs has noticeable, although not catastrophic effects on the rate of branching. Under high-salt conditions, k_{br} is reduced by threefold, although k_{hy} is similar to that of the WT intron (Fig. 7A; Table 1). The kinetic behavior of the GU $\rightarrow$ GC mutant (Fig. 7A) suggests that, like the WT intron, this mutant has two conformational populations and that the fast population reacts primarily through branching. The inhibition of branching becomes especially apparent at lower salt concentrations (Fig. 7B). Under those conditions, k_{br} is

FIGURE 5. Kinetic analysis of the prA-G mutant. **A:** Under high-salt conditions, the predominant product is lariat intron, which is generated at a faster rate than linear intron. The reaction of precursor was biphasic and, like WT (Daniels et al., 1996; Liu et al., 1997), fits biexponential Equation 3, where $a_1 = 0.616 (\pm 0.026)$, $a_2 = 0.399 (\pm 0.022)$, $k_T^A = 0.140 (\pm 0.014) \text{ min}^{-1}$, and $k_T^B = 0.00479 (\pm 0.00059) \text{ min}^{-1}$. Evolution of linear and lariat intron fit a single exponential Equation 2. The fractional coefficients (f_{hy} and f_{br}) for evolution of linear and lariat intron products are $0.282 (\pm 0.029)$ and $0.605 (\pm 0.0097)$, respectively. k_T values derived from product curves of linear and lariat evolution are $0.0063 (\pm 0.0012) \text{ min}^{-1}$ and $0.0944 (\pm 0.0045) \text{ min}^{-1}$. Under these conditions, reactive precursor molecules are partitioned into two distinct populations (A and B) that react exclusively through branching or hydrolysis splicing pathways, respectively (a_1 from the precursor fit = f_{br} from lariat intron fit; a_2 from precursor fit = f_{hy} from linear intron fit). Therefore, the individual k_T values for lariat and linear intron evolution were equal to k_{br} and k_{hy} , respectively. Thus, $k_{hy} = 0.00630 (\pm 0.0012) \text{ min}^{-1}$ and $k_{br} = 0.0944 (\pm 0.0045) \text{ min}^{-1}$. **B:** Splicing time course under low-salt conditions. The prA-G precursor is converted predominantly to linear intron, although a small amount of lariat evolution is shown (inset). After correcting for a 30-min lag observed under this condition, the decrease in precursor (closed squares) fit single exponential Equation 1, where $a_1 = 0.820 (\pm 0.017)$ and $k_T = 0.0089 (\pm 0.0004) \text{ min}^{-1}$. The evolution of lariat (closed circles, inset) and linear (closed triangles) intron products were fit to Equation 2, where fractional coefficients, $f_{br} = 0.119 (\pm 0.014)$ and $f_{hy} = 0.701 (\pm 0.021)$, respectively. The k_T values derived from these curves were $0.00820 (\pm 0.0021) \text{ min}^{-1}$ and $0.00904 (\pm 0.00061) \text{ min}^{-1}$, respectively. Applying Equations 4 and 5, $k_{br} = 0.00125 \text{ min}^{-1}$ and $k_{hy} = 0.0074 \text{ min}^{-1}$.

FIGURE 6. Using DNAzymes to map the branch-point of the prA-G mutant. **A:** Scheme for mapping the branch-point of lariat RNAs. A purified lariat molecule is treated with two DNAzymes flanking the branch-site. DNAzyme 1 targets the sequence 17 nt upstream of the putative branch-point (closed circle). DNAzyme 2 targets the sequence 103 nt from the intron 5' end (now part of the branched linkage). The resulting Y-shaped molecule (127 nt) is gel purified, resulting in a product that can be 5' end-labeled (*) with $\gamma^{32}\text{P}$ -ATP, and subjected to alkaline hydrolysis. Separation on a 20% polyacrylamide gel produces a ladder in which nucleotides preceding the branched nucleotide are separated by an 8-nt gap from higher molecular weight species (HMW) that contain the branch. **B:** Mapping the branch-points of WT and prA-G lariats. Shown are the partial alkaline hydrolysis ladders of 5'-labeled DNAzyme cleavage products of prA-G (lane 1) and WT (lane 2) lariats. An RNA oligonucleotide (17nt-BBP) corresponding to the 17 nt preceding the WT-A branchpoint (lanes 3 and 4) was used as a sequencing and size marker after treatment with alkali and nuclease T1. RNAse T1 cleaves after G residues, thus indicating the size and sequence of the mapped fragment. The pattern of band intensity within a given lane is very similar for prA-G and WT introns, indicating that they branch to the same site. Phosphorimager quantitation reveals that there are slightly more counts in the prA-G lane than WT, which may account for the darker bands in the gap of prA-G. This fact, together with the proximity to the 3' end of the intron, suggests that the faint bands in the gap do not represent products of cryptic splicing. However, quantitative analysis of background bands relative to the branch-site revealed that this assay will be unable to detect cryptic branching that is $\leq 10\%$ of the total branched product.

reduced 11-fold relative to WT under the same conditions (Table 1). These data suggest that the highly conserved wobble pairs flanking the branch-point may serve an important role in transesterification. However, mutations of these wobble pairs are not sufficient to eliminate the branching pathway. The strong effects on branching parallel the significant splicing defects observed for the GU $\rightarrow$ GC mutant in vivo, in which the mutant strain has a low level of splicing that is partitioned through both branching and hydrolytic pathways.

Superficially, the splicing of the D6zip mutant appears to be almost indistinguishable from that of WT (Fig. 3). However, there are salient differences that are important to note. Under high-salt conditions, the behavior is very much like WT except that the rate of branching is slightly (30%) lower (Fig. 8A; Table 1). Differences between the two versions of the intron are more dramatic at low salt, in which hydrolysis dominates over branching (Fig. 8B; Table 1). Branching is significantly slower in the mutant than in the WT intron,

(A) GU→GC / High Salt **(B) GU→GC / Low Salt**

FIGURE 7. Kinetic analysis of the GU → GC mutant. **A:** Splicing time course under high-salt conditions. The decrease in precursor was fit with Equation 3, where $a_1 = 0.566$ (± 0.032), $a_2 = 0.433$ (± 0.033), $k_T^A = 0.0457$ (± 0.0045) min^{-1} , and $k_T^B = 0.00266$ (± 0.00053) min^{-1} . Evolution of linear and lariat intron was fit with Equation 2. The resulting fractional coefficients (f_{hy} and f_{br}) were 0.274 (± 0.030) and 0.498 (± 0.0065), respectively, and the k_T values were 0.0093 (± 0.0023) min^{-1} and 0.0438 (± 0.0017) min^{-1} , respectively. Because precursor partitioned distinctly between two populations going through either hydrolysis or branching pathways, k_T values of linear and lariat intron evolution were equivalent to k_{hy} and k_{br} , respectively. **B:** Splicing time course under low-salt conditions. The decrease in GU → GC precursor (closed squares) was fit with Equation 1, resulting in $a_1 = 0.900$ (± 0.035) and $k_T = 0.00951$ (± 0.00094) min^{-1} . Evolution of lariat (closed circles, inset) and linear (closed triangles) intron products fit Equation 2, resulting in fractional coefficients $f_{br} = 0.0281$ (± 0.0028) and $f_{hy} = 0.875$ (± 0.035), respectively, and k_T values of 0.0142 (± 0.0041) min^{-1} and 0.00933 (± 0.00092) min^{-1} , respectively. The similarity in k_T values indicates a parallel kinetic scheme, so, by applying Equations 4 and 5, one obtains $k_{br} = 0.00037$ min^{-1} and $k_{hy} = 0.0114$ min^{-1} .

showing a relative rate diminution of $\sim$ fourfold. These data suggest that, although the hydrolytic pathway is relatively unaffected by this particular mutation, there are small effects on branching that become especially apparent under less permissive reaction conditions (low salt).

DISCUSSION

Mutations in the morphology of the branch-site reveal that the group II intron is a highly versatile molecule. It can catalyze lariat formation when the branch-point A is involved in a paired (or mispaired) configuration, and,

(A) D6 zip / High Salt **(B) D6 zip / Low Salt**

FIGURE 8. Kinetic analysis of the D6zip mutant. **A:** Splicing time course under high-salt conditions. The decrease in precursor was fit with Equation 3, where $a_1 = 0.599 (\pm 0.011)$, $a_2 = 0.405 (\pm 0.011)$, $k_T^A = 0.1257 (\pm 0.0051) \text{ min}^{-1}$, and $k_T^B = 0.00711 (\pm 0.00032) \text{ min}^{-1}$. Evolution of linear and lariat intron products was fit to Equation 2, resulting in fractional coefficients (f_{hy} and f_{br}) of $0.344 (\pm 0.037)$ and $0.599 (\pm 0.009)$, respectively, and k_T values of $0.0070 (\pm 0.0015) \text{ min}^{-1}$ and $0.0987 (\pm 0.0037) \text{ min}^{-1}$. Precursor molecules partitioned into two distinct populations that reacted exclusively through hydrolysis or branching pathways and therefore the k_T values for evolution of linear and lariat intron were equivalent to k_{hy} and k_{br} , respectively. **B:** Splicing time course under low-salt conditions. The decrease in D6zip precursor (closed squares) fit Equation 1, resulting in $a_1 = 1.030 (\pm 0.057)$ and $k_T = 0.0103 (\pm 0.0013) \text{ min}^{-1}$. The evolution of lariat (closed circles, inset) and linear (closed triangles) intron products were fit with Equation 2, resulting in $f_{br} = 0.0960 (\pm 0.0062)$ and $f_{hy} = 0.925 (\pm 0.057)$, respectively, and k_T values $0.0150 (\pm 0.0025) \text{ min}^{-1}$ and $0.0100 (\pm 0.0015) \text{ min}^{-1}$, respectively, suggesting a parallel kinetic regime. By applying Equations 4 and 5, $k_{br} = 0.0012 \text{ min}^{-1}$ and $k_{hy} = 0.0114 \text{ min}^{-1}$.

even when branching fails to occur, the intron has a default pathway that allows splicing to proceed through hydrolysis during the first step. These findings are significant because they reveal a strong driving force for group II intron splicing and they provide insight into the mechanism of reaction.

Mechanistic implications of branching from a paired A: Recognition of a flipped base?

Although A₈₈₀ is base paired in the prA-U mutant, it has a small propensity to branch, as shown by the evolution of lariat species at very long reaction times. When

this base pairing is weakened, as in the prA-G mutant, the ability to branch is dramatically increased. In fact, the prA-G mutant branches with an efficiency and precision that is comparable to WT under most conditions. We have also tested the effects of having either an adenosine (prA-A) or a cytosine (prA-C) lying opposite the branchpoint A. All three mismatched mutants (prA-A, prA-C, and prA-G) result in several hundred-fold increases in the rate and extent of branching by comparison to the prA-U mutant (data not shown). Interestingly, the purine–purine mismatches, which would be expected to result in the greatest helical distortion, have slightly larger effects (~threefold greater than prA-C). Taken together, the data are consistent with the notion that dynamic flexibility, rather than a bulged structure per se, is required at the branch-site, and that this is important for promoting an extrahelical conformation of the branch-site adenosine. The likelihood that a dynamic extrahelical (flipped) adenosine is involved in branch point recognition is supported by several pieces of evidence. First, crystal structures of RNA helices containing single-adenosine bulges indicate that the base is flipped out of the helix (Valegard et al., 1994; Portmann et al., 1996). Solution NMR studies have shown that bulged adenosines exist in a dynamic equilibrium, continually sampling both the flipped and helically stacked states (Borer et al., 1995). Second, a previous study on the importance of specific branch-site functional groups showed that N6 was a particularly important determinant for recognition by the rest of the intron (Liu et al., 1997). Because the major groove substituent N6 has poor steric accessibility in the stacked state, it is likely that adenosine becomes extrahelical before docking into the ribozyme active site.

Here we suggest that structural features that are likely to increase the population of flipped versus stacked adenosine residues tend to increase the amount of branching that is observed during self-splicing. Interestingly, it has been observed that adenosine becomes extrahelical most readily when a bulged A is flanked by two pyrimidine residues (Portmann et al., 1996). In group II introns, flanking pyrimidine nucleosides are a highly conserved feature at the branch-site (Michel et al., 1989). The important determinant in successful branching seems to be that the equilibrium between stacked and flipped adenosine conformations must be shifted to the right (Fig. 9, top). When the branch-point A exists in the context of a bulge or mismatched pair, the equilibrium appears to be strongly shifted toward base flipping and branching. However, strengthening the flanking base pairs (which increases local conformational rigidity), or involving the adenosine in a strong Watson–Crick interaction shifts the equilibrium to the left, making branching less favorable (Fig. 9, bottom). Taken together, the trends reported here have two implications of major importance. First, they provide strong evidence that the branch-site is in conformational equilibrium

between stacked and flipped states, with the flipped state being the active form for transesterification. Second, the results demonstrate that the most important feature of a good branch site is not that it contains a bulged nucleotide, but that it promotes a flipped adenosine conformation, through one of several possible structural strategies that include mispairing.

It is important to note that extrahelical adenosines have been observed to adopt a variety of conformational forms, which differ in the position and accessibility of the adenosine 2'-hydroxyl group (Portmann et al., 1996). In one crystal form, the 2'-OH is actually aligned to attack the adjacent 3'-phosphate, rather than a phosphodiester linkage on another molecule. Other conformational forms, resulting from a different distribution of possible backbone torsion angles, position the 2'-OH differently, potentially in a configuration capable of serving as a nucleophile during splicing. This underscores the importance of the flanking G-U pairs commonly seen in group II introns. They are likely to give a unique conformational identity to the flipped adenosine nucleoside at the branch-point.

Base-flipping has now been shown to be important for many different processes, including DNA methylation and repair (Roberts, 1995; Cheng & Blumenthal, 1996). Perhaps in a parallel with the branching reaction, DNA methylation occurs more readily when the target (flipped) base is involved in a mispair (Klimasauskas & Roberts, 1995; Yang et al., 1995). This is consistent with the fact that there are considerable thermodynamic and kinetic barriers to base-flipping. Although a certain amount of rapid proton exchange and helical breathing is expected for nucleotides within a duplex (Leroy et al., 1985a, 1985b), the estimated rate for spontaneous flipping completely out of an RNA duplex is considerably slower than the rate of branching observed by the prA-U mutant. In DNA, base-flipping is accelerated by enzymes that lower the energetic barriers by providing stabilizing interactions that compensate for lost stacking and hydrogen bonding. Because even the prA-U mutant is capable of branching, it is likely that the group II intron active site binds the branch-point in a manner that makes base-flipping, or at least local melting of D6, particularly favorable. This effect might also contribute to efficiency of branching in the WT intron because, in solution, bulged adenosines are in conformational equilibrium with a stacked (unflipped) state, which can be the preferred conformation in certain instances (Borer et al., 1995).

Although base-flipping is an energetically costly process in a helix, the energetic barrier for flipping of a nucleotide within a single base bulge is expected to be much lower. Behavior of the prA-G mutant suggests that there are other energetic strategies for facilitating flipping of a bulged A residue. This was first indicated by *in vivo* studies of spontaneous revertants of the prA-U mutant, in which the U was transformed into a G,

FIGURE 9. Model for base-flipping at the branch-point. Top: Even when a base is paired and stacked in a duplex, it is in dynamic equilibrium with a flipped form of the base. For a Watson–Crick pair, the equilibrium is displaced to the left. When this occurs at the branch-point of a group II intron, splicing will proceed predominantly through the hydrolytic pathway. For a mispair such as A–G, the equilibrium shifts to the right. When this occurs at the branch site, transesterification (branching) should predominate over hydrolysis during the first step of splicing. Bottom (from left to right): From the results described herein, the intron active site appears to promote especially efficient flipping of adenines in a bulge or in mismatch configurations such as an A–G pair. Both of the latter are more facile than flipping a bulged A that is surrounded by G–C pairs rather than G–U pairs, and the least facile is flipping of an A involved in an A–U pair.

thus creating the prA–G configuration. This suppressor permits efficient splicing with branching *in vivo*. Biochemical analysis of whole-cell RNA, including northern blots and primer extension studies, showed that the abundant free introns are branched from the correct adenosine on D6 (M. Podar & P.S. Perlman, unpubl. results). In order to explore this effect, the prA–G mutant was constructed and analyzed *in vitro* as shown here. In correspondence with the *in vivo* results, the prA–G mutant branches with an efficiency and precision that is close to that of the WT intron *in vitro*. This is remarkable in light of previous work on the stability and structure of G–A pairs in RNA helices. Structural studies have shown that single purine–purine mismatches do not necessarily have drastic effects on the nucleic acid structure (Maskos et al., 1993). In certain contexts, G–A pairs are stabilized by hydrogen bonds between them and stacking interactions (either inter-strand or intra-strand) with neighboring base pairs. It is therefore likely that the D6 stem containing the prA–G mutation has the average appearance of a continuous duplex. However, when bound in the core of the intron, the G–A pair may become dynamic, resulting in a low energetic barrier for flipping of the A.

Although the data presented herein are consistent with a base-flipping event that is necessary for branching, direct evidence is needed to support the model. Experimental approaches to address this issue may

include chemical modification studies for probing the accessibility of base functionalities, solution or solid-state NMR experiments on the dynamic flexibility of the branch-site in model D6 RNA helices, and even computational approaches to monitoring the dynamic behavior of the branch-point in different sequence contexts.

The role of flanking G–U pairs

Wobble pairs on either side of the bulged adenosine are likely to affect the local structure and the conformational dynamics at the branch site. It is conceivable that flanking G–U pairs would result in a greater propensity for flipping of the branch point A out of the duplex. This is consistent with NMR studies on DNA oligonucleotides, in which G–T pairs were found to increase the opening rate of neighboring base pairs (Pardi et al., 1982). The G–U pairs may also help to position the extrahelical A such that the 2′-OH is in a particularly reactive conformation. Consistent with this, mutation of the G–U pairs to G–C pairs has a strong effect on the rate of branching, which is evident even under high-salt conditions (Fig. 7). Overall, the kinetic effects observed in this study are not catastrophic (4-fold at high-salt, 11-fold at low salt). However, under the conditions used in this study, branching is not limited by the chemical rate of catalysis. Instead, the rate is limited by a conformational change that may, or may not, involve

positioning of the branch-point adenosine. It is therefore possible that the wobble pairs may be inducing a considerably larger effect than that observed here, because a slow conformational change may obscure fundamental differences in branching rates. This would be consistent with the fact that the GU $\rightarrow$ GC mutant is very disruptive to branching in vivo (Podar et al., 1998). Although branched product is observed in vivo, it is produced inefficiently with strong defects in both the first and second steps of splicing. These mutational studies therefore indicate that an important role may exist for the semi-conserved G-U pairs around the branch-site of self-splicing group II introns. They may provide flexibility at the branch-site, thereby promoting a flipped adenosine conformation that is particularly conducive to branching.

An excess of conformational stabilization: Pairing of the D6 internal loop

Previous studies of the D6 stem structure have suggested that high thermodynamic stability of D6 duplexes is important for proper branch-point function (Schmelzer & Schweyen, 1986). This is reasonable, given that a very loose structure would induce too much conformational variability in and around the branch-point A. However, the D6 stem of the *ai5 γ* intron is particularly stable, and may have evolved a specific conformational state for promoting optimal branching efficiency. Consistent with this notion, mutating the D6 internal loop into a paired helical region results in small, but significant, defects in branching efficiency. Rates are reduced by 30% at high-salt and by fourfold under low-salt conditions (Fig. 8). As in the case of the GU $\rightarrow$ GC mutant, these effects are likely to represent lower limits on the magnitude of branching defects because the identity of the rate-limiting step is not known in either case. Thus, the actual mutational effects could be larger, and this may be reflected in reduced function in vivo. Therefore, although stability of the D6 stem undoubtedly plays an important role up to a point (Schmelzer & Schweyen, 1986), too much stability may inhibit branching.

Implications of an efficient hydrolytic splicing pathway

Lariat formation is the defining reaction of group II introns and the mechanistic link between group II introns and nuclear splicing by the eukaryotic spliceosome. However, the data provided here show that group II introns do not need a branch-point in order to splice efficiently in vitro. When the branch-point is deleted or mutated, splicing readily proceeds through an alternative pathway in which water or hydroxide is the nucleophile during the first step of splicing. This reaction was

reported previously to compete with branching in self-splicing reactions of the WT intron, giving rise to linear rather than lariat forms of the intron (Jarrell et al., 1988; Daniels et al., 1996). Here we show that, upon deletion of the branch-point A, hydrolysis functions as the primary pathway during the first step of self-splicing even under conditions that favor the branching pathway in the WT intron. This is similar to the behavior observed when the branch-point adenosine was mutated to a cytosine (Liu et al., 1997). In that case, branching was undetectable and splicing proceeded efficiently through a hydrolytic first step.

An efficient hydrolytic pathway for mutants with disrupted branch-points is significant because it shows that, unlike the spliceosome, the group II intron uses an alternative mechanistic pathway to promote splicing when branching is disrupted, rather than selecting cryptic branch-sites for lariat formation. This is consistent with results from previous work indicating that branching from bulged G or U mutants occurs with high fidelity and that incorporation of a bulged C or a deoxy A at the branch-site eliminated lariat formation (Liu et al., 1997). Taken together, results from these two studies show that group II introns have a mechanism for carefully specifying the proper branch site with high fidelity. When that branch site is unavailable, and even to some extent when it is available, splicing proceeds through an alternative hydrolytic pathway.

It is interesting to consider why H₂O or OH⁻ are good nucleophiles during cleavage of the 5'-splice site. One might expect that the nucleophile was required to be a large, precisely positioned, nucleotide residue in order to participate in the expected transition state for phosphodiester cleavage. The fact that H₂O can function as the nucleophile suggests that there is significant strain exerted on the scissile bond. That is to say, the ribozyme active site is arranged to create a "hot phosphate," which is precisely situated to react with a specific nucleophile such as the branch-point 2'-hydroxyl group. When that nucleophile is not present, the phosphodiester bond is so reactive that it is attacked by other nucleophiles in the vicinity, such as molecules of the solvent.

The findings reported here are consistent with the behavior of group II introns in vivo. When the mutant introns ΔA_{880} , prA-U, GU $\rightarrow$ GC, A880C, A880U, or A880G were transformed into the mitochondria of yeast strains, it was recently found that the resultant mutant strains grew in glycerol medium (Podar et al., 1998). This respiratory function, which depends on the formation of accurately spliced mRNA, together with the isolation and mapping of linear intron product RNAs, showed that a hydrolytic first step of splicing occurs in vivo. Because only lariat introns are likely to be infectious genetic elements, a hydrolytic splicing reaction in vivo may be important for regulating splicing versus mobility by group II introns. Alternatively, the pathway

may represent a vestigial reaction derived from more ancient forms of the intron. Finally, the findings presented herein explain why there are group II introns that appear to lack branch-points entirely, and they suggest that alternative processes of branch-site recognition or hydrolysis may occur, to some extent, during the process of nuclear splicing.

MATERIALS AND METHODS

Plasmids and transcripts

Plasmids pVC00, pVC01, pVC02, pVC03, and pVC04 encoding the precursor RNAs GU → GC, D6 zip, ΔA₈₈₀, prA-U, and prA-G, respectively, were created by template-directed mutagenesis of plasmid pJD20 (Jarrell et al., 1988) using published procedures (Kunkel et al., 1991). The WT ai5γ RNA and the mutant pre-mRNA molecules, containing the modified intron (887 nt for WT) flanked by a 293-nt 5'-exon and a 318-nt 3'-exon, were transcribed from plasmids linearized with *Hind* III, using T7 RNA polymerase that was expressed and purified according to published procedures (Davanloo et al., 1984). Transcription was performed under standard conditions, in which ³²P-UTP was added to internally label the transcripts (Daniels et al., 1996). The RNA transcripts were isolated, purified, and stored as described previously (Daniels et al., 1996).

Reaction conditions and kinetic analyses

Precursor RNA molecules (in a storage buffer of 10 mM MOPS/1 mM EDTA, pH 6.5) were taken from the freezer and aliquots were removed for reaction. Prior to reaction, the RNA was heated at 95 °C for 1 min and then allowed to assume a final temperature of 42 °C. After this preincubation step, precursor RNA (~1 nM final concentration) was combined with reaction buffer in a total volume of 20 μL. Two types of reaction buffers were used in this study. One set of reactions contained 100 mM MgCl₂ and 40 mM MOPS brought to pH 7.5 with KOH, final concentrations (low-salt buffer). Another set of reactions contained 100 mM MgCl₂, 500 mM (NH₄)₂SO₄, and 40 mM MOPS, brought to pH 7.5 with NaOH, final concentrations (high-salt buffer). Reactions were incubated at 42 °C, aliquots (1–2 μL) were removed at the indicated time points, added to quench buffer, and then the samples were subjected to denaturing electrophoresis. After separation of reaction products, polyacrylamide gels were dried and quantitated using a Packard Instantimager. The radioactivity in each band was counted and background subtracted as described previously (Daniels et al., 1996). The resulting data were fit in a manner similar to that reported previously (Daniels et al., 1996; Liu et al., 1997), after incorporating the following expressions into a graphing program (Kaleidagraph, Abelbeck Software).

For reactions in which the decrease in precursor follows a single exponential decay (WT and mutants under low-salt conditions, ΔA₈₈₀ and prA-U under all salt conditions), reflecting a single population of reactive molecules, the data were fit to the following equation:

$$\frac{P(t)}{P(0)} = n_1 + a(e^{-k_T(t)}) \quad (1)$$

where $P(t)$ represents the amount of precursor RNA at any time (t), and $P(0)$ represents the amount of precursor present at the beginning of reaction ($t = 0$). The fitting was not constrained by the necessity for the reaction to reach completion, so the value a represents the reactive population of molecules and n_1 represents the unreactive fraction. The term k_T represents the overall rate of the splicing reaction (by all pathways).

For reactions in which the evolution of a specific product follows a single exponential (WT and mutants under all reaction conditions), the data were fit as follows:

$$\frac{I(t)}{I(0)} = m_1 + f_{br,hy}(1 - e^{-k_T(t)}) \quad (2)$$

where $I(t)$ represents the amount of free intron product RNA (either lariat or linear) at any time (t), and $I(0)$ represents the amount of intron present at the beginning of the reaction ($t = 0$). To prevent the fit from being constrained at its limits, m_1 represents the fraction of intron product at $t = 0$ and $f_{br,hy}$ represents the fraction of a particular intron product (relative to precursor and all other intron products) observed at the end of reaction. Specifically, f_{br} represents the fraction of lariat intron and f_{hy} represents the fraction of linear intron. The simplifying assumption that all precursor molecules (P) are stoichiometrically converted to intron products (I) is reasonable given that the second step of splicing in all cases observed here is very rapid. Under the conditions stated, there were insignificant levels of reaction intermediates (intron–3'-exon species) detected that would complicate the analysis.

For reactions in which the decrease in precursor follows a double-exponential decay (WT, D6 zip, GU → GC, and prA-G mutants in high-salt buffer), representing two distinct populations or conformations of reacting molecules, the data were fit as follows:

$$\frac{P(t)}{P(0)} = a(e^{-k_T^A(t)}) + b(e^{-k_T^B(t)}) \quad (3)$$

where a and b represent the fraction of molecules reacting to completion for fast population A and slow population B, respectively. The values k_T^A and k_T^B represent the total rates of reaction by populations A and B, respectively, through either reaction pathway (hydrolysis or transesterification).

Under low-salt conditions, the WT and mutant introns all reacted as a single population of precursor molecules. For the mutant introns ΔA₈₈₀ and prA-U, k_{hy} was equivalent to k_T . For introns that were capable of lariat formation (WT, D6 zip, GU → GC, and prA-G mutants), branching and hydrolysis arose from the same uniformly reactive population of precursor. As a result, their reaction progress curves followed the paradigm of parallel reaction kinetics (Fersht, 1985; Jencks, 1987). In this kinetic regime, the apparent rate of reaction (k_T) is the same upon fitting the evolution of either reaction product, lariat or linear intron. The k_T value is deconvoluted into k_{hy} and k_{br} by correcting for the fraction of each type of product (f_{br} or f_{hy}) at the end of reaction. Parallel kinetic

behavior in group II intron self-splicing has been discussed at length elsewhere (Daniels et al., 1996). Under high-salt conditions, slightly more complex behavior was observed that will be discussed as indicated.

From the above equations, values of k_T and the fractional coefficients f_{br} or f_{hy} for the two different types of intron product (lariat or linear), were used to calculate the actual rates of branching (k_{br}) and hydrolysis (k_{hy}) using the equations below:

$$k_{br} = k_T \left(\frac{f_{br}}{f_{br} + f_{hy}} \right) \quad (4)$$

$$k_{hy} = k_T \left(\frac{f_{hy}}{f_{br} + f_{hy}} \right) \quad (5)$$

Coefficients f_{br} and f_{hy} represent the relative amount of lariat and linear intron products, respectively, that accumulate by the end of reaction. These values are determined from the horizontal asymptotes of the product evolution curves shown in Figures 4, 5, 7, and 8. In most cases, the k_T values used to solve Equations 4 and 5 were obtained from the fit of precursor reaction (Equation 3).

Isolation of lariat RNAs for branch-point mapping studies

The WT and prA-G mutants were transcribed from *Hind* III-digested plasmids pJD20 and pVC04, respectively, in an 8-mL reaction volume under standard conditions for large-scale transcription (Pyle & Green, 1994), except that 8 μ L of α - 32 P-UTP was added to dope the transcripts with a low level of radioactivity for autoradiographic visualization. After transcription, the precursor RNAs were separated from unincorporated NTPs by spinning the reactions in a Centricon 50 (Amicon) filter, and then washing the isolated transcripts by spinning the filter with buffer (10 mM MOPS, 1 mM EDTA, pH 6.5) several times. After transferring to an Eppendorff tube, purified precursor RNAs were then incubated at 42 °C for 3 h in a high-salt buffer [500 mM (NH₄)₂SO₄, 100 mM MgCl₂, and 40 mM MOPS, pH 7.5, in a 1-mL volume] to promote the branching reaction. After splicing, the samples were spun again in a Centricon 50 filter to separate the RNA precursor and products from the salt. The purified RNAs were then suspended again in buffer (10 mM MOPS/1 mM EDTA, pH 6.5, in 40 μ L). After addition of an equal volume of urea dye buffer, the sample was loaded onto a 6% denaturing polyacrylamide gel and subjected to electrophoresis to separate reaction products. Gel slices containing lariat RNAs were excised, crushed, and eluted overnight at 4 °C on a shaking platform (Fisher) in a MOPS storage buffer of 10 mM MOPS/1 mM EDTA, pH 6.5. Lariat RNA from the eluate was precipitated and the concentration was determined spectrophotometrically.

DNAzyme reaction conditions for mapping the branch-site

Two oligodeoxyribonucleotide DNAzymes were used simultaneously to cleave lariat molecules between positions

A862 and C863 (at the 5'-region of D6, DNAzyme 1), and between positions A103 and U104 (at the C1 loop of D1, DNAzyme 2) of intron ai5 γ : 5'-agaccttacaagGGCTAGCTACAACGAtttccccc-3' and 5'-gcttagcatatttaGGCTAGCTAACGAgttaccattttaa-3', respectively, where uppercase nucleotides denote the catalytic domain and lowercase nucleotides represent binding arms for recognition of specific sites on the lariat substrate. Binding arms were designed such that the $-\Delta G^\circ$ of helix formation was expected to be ≥ -11 kcal/mol based on nearest-neighbor calculations (G. Joyce, pers. comm.) (Freier et al., 1986). DNAzymes of this core sequence cleave RNA molecules after a single unpaired purine that is followed by a paired pyrimidine, resulting in 5'-OH and 2'-3' cyclic phosphate termini (Santoro & Joyce, 1997).

Prior to reaction, lariat RNA and the two DNAzymes were preincubated separately at 90 °C for 2 min, spun briefly, and placed immediately on ice. Then the two DNAzymes and buffer salts (see below) were added simultaneously to lariat RNA and the whole reaction mixture was incubated at 37 °C for 5 h. Reactions were performed under single-turnover conditions, with a large excess of DNAzyme relative to lariat RNA substrate. DNAzyme reaction volume was 150 μ L, containing 300 nM lariat RNA, 60 μ M each of DNAzyme 1 and DNAzyme 2, 150 mM NaCl, 100 mM MgCl₂, and 40 mM Tris, pH 8.0. This set of conditions differs from published procedures, in which DNAzymes were used at lower concentrations and function with multiple-turnover. The conditions described herein are designed to maximize specific RNA cleavage without random degradation. After 5 h of DNAzyme incubation, the reaction was halted by an equal volume of urea/EDTA quench buffer, loaded onto a 6% denaturing polyacrylamide gel, and subjected to electrophoresis. The cleavage product (127 nt in length) that results from lariat cleavage by the two DNAzymes was visualized by radioanalytic imaging, excised, eluted, precipitated, and dissolved in 7 μ L 10 mM MOPS, pH 6.5/1 mM EDTA.

Using the above procedure, only a fraction (1/1,000) of the 127-mer branched molecules contains an internal 32 P label, which was used solely for purposes of fragment isolation. To map the branch-site in these molecules, it was necessary to label them quantitatively on the 5' terminus with γ - 32 P ATP and polynucleotide kinase (NEB) under standard reaction conditions. These 5'-labeled products were excised, eluted, precipitated, and dissolved in MOPS storage buffer.

Conditions for mapping the branched oligonucleotide

Generating the alkaline-hydrolysis ladder

In a 10- μ L reaction volume, approximately 250,000 cpm of 5' end-labeled 127-mer RNA was incubated at 90 °C for 5 min in a buffer containing 50 mM NaHCO₃, pH 9.0, and 5 μ g tRNA (Kuchino & Nishimura, 1989). After incubation, the reaction was immediately placed on ice and an equal volume of urea/EDTA quench buffer was added. This mixture was then loaded on a 20% gel, subjected to electrophoresis, and visualized by radioanalytic imaging.

RNAse T1 digestion conditions

RNAse T1 cleaves after G residues, leaving the same 5'-OH and 3'-PO₃ ends as those arising from an alkaline hydrolysis (Donis-Keller et al., 1977). For this reaction, 250,000 cpm of 5' end-labeled marker RNA (see legend, Fig. 6B) was denatured at 95°C for 2 min before the addition of RNAse T1 buffer (20 mM sodium citrate, pH 5.0, in a volume of 20 µL), that had been pre-heated for 1 min at 50°C. The reaction mixture was incubated at 50°C for 5 min, then put on ice. Then 50 units of RNAse T1 (Boehringer) were added and the mixture was heated at 50°C for 30 min, before adding urea/dye buffer and loading alongside branch-point hydrolysis ladders on a 20% polyacrylamide gel.

ACKNOWLEDGMENTS

We thank Kevin Chin and Patrick Zarrinkar for reading and commenting on the manuscript. In addition, we thank Eckhard Jankowsky and Marc Boudvillain for helpful discussions. This work was supported by NIH grant RO150313 and an NYI award from the NSF to A.M.P.

Received May 6, 1998; returned for revision June 8, 1998;
revised manuscript received July 6, 1998

REFERENCES

- Borer PN, Lin Y, Wang S, Roggenbuck MW, Gott JM, Uhlenbeck OC, Pelczar I. 1995. Proton NMR and structural features of a 24-nucleotide RNA hairpin. *Biochemistry* 34:6488–6503.
- 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.
- 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.
- Cheng X, Blumenthal RM. 1996. Finding a basis for flipping bases. *Structure* 4:639–645.
- Costa M, Deme E, Jacquier A, Michel F. 1997. Multiple tertiary interactions involving domain II of group II self-splicing introns. *J Mol Biol* 267:520–536.
- Daniels D, Michels WJ, 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.
- Davanloo P, Rosenberg AH, Dunn JJ, Studier FW. 1984. Cloning and expression of the gene for bacteriophage T7 DNA. *Proc Natl Acad Sci USA* 81:2035–2039.
- Donis-Keller H, Maxam AM, Gilbert W. 1977. Sequencing RNA with endonucleases (SIC). *Nucleic Acids Res* 4:2527–2538.
- Fersht A. 1985. *Enzyme structure and mechanism*. New York: W.H. Freeman.
- Freier SM, Kierzek R, Jaeger JA, Sugimoto N, Caruthers MA, Neilson T, Turner DH. 1986. Improved free energy parameters for predictions of RNA duplex stability. *Proc Natl Acad Sci USA* 83:9373–9377.
- Gaur RK, McLaughlin LW, Green MR. 1997. Functional group substitutions of the branchpoint adenosine in a nuclear pre-mRNA and a group II intron. *RNA* 3:861–869.
- Hornig H, Aebi M, Weissman C. 1986. Effect of mutations at the lariat branch acceptor site on β -globin pre-mRNA splicing in vitro. *Nature* 324:589–591.
- Jacquier A, Michel F. 1987. Multiple exon-binding sites in class II self-splicing introns. *Cell* 50:17–29.
- Jarrell KA, Peebles CL, Dietrich RC, Romiti SL, Perlman PS. 1988. Group II intron self-splicing: Alternative reaction conditions yield novel products. *J Biol Chem* 263:3432–3439.
- Jencks WP. 1987. *Catalysis in chemistry and enzymology*. New York: Dover. pp 569–574.
- Klimasauskas S, Roberts RJ. 1995. M HhaI binds tightly to substrates containing mismatches at the target base. *Nucleic Acids Res* 23:1388–1395.
- Koch JL, Boulanger SC, Dib-Hajj SD, Hebbard SK, Perlman PS. 1992. Group II Introns deleted for multiple substructures retain self-splicing activity. *Mol Cell Biol* 12:1950–1958.
- Kuchino Y, Nishimura S. 1989. Enzymatic RNA sequencing. *Methods Enzymol* 180:154–163.
- Kunkel TA, Bebenek K, McClary J. 1991. Efficient site-directed mutagenesis using uracil-containing DNA. *Methods Enzymol* 204:125–139.
- Leroy JL, Bolo N, Figueroa N, Plateau P, Gueron M. 1985a. Internal motions of transfer RNA: A study of exchanging protons by magnetic resonance. *J Biomol Struct Dyn* 2:915–939.
- Leroy JL, Broseta D, Gueron M. 1985b. Proton exchange and base-pair kinetics of poly(rA).poly(rU) and poly(rI).poly(rC). *J Mol Biol* 184:165–178.
- Liu Q, Green JB, Khodadadi A, Haeberli P, Beigelman L, Pyle AM. 1997. Branch-site selection in a group II intron mediated by active recognition of the adenine amino group and steric exclusion of non-adenine functionalities. *J Mol Biol* 267:163–171.
- Maskos K, Gunn BM, LeBlanc DA, Morden KM. 1993. NMR study of G-A and A-A pairing in (dGCGAATAAGCG)₂. *Biochemistry* 32:3583–3595.
- Michel F, Ferat JL. 1995. Structure and activities of group II introns. *Annu Rev Biochem* 64:435–461.
- Michel F, Umeson K, Ozeki H. 1989. Comparative and functional anatomy of group II catalytic introns—A review. *Gene* 82:5–30.
- Padgett RA, Konarska MM, Grabowski PJ, Hardy SF, Sharp PA. 1984. Lariat RNAs as intermediates and products in the splicing of messenger RNA precursors. *Science* 225:898–903.
- Pardi A, Morden KM, Patel DJ, Tinoco I. 1982. Kinetics for exchange of imino protons in the d(CGCAATTCGCG) double helix and in two similar helices that contain a G-T base pair and an extra adenine. *Biochemistry* 21:6567–6574.
- Peebles CL, Benatan EJ, Jarrell KA, Perlman PS. 1987. Group II intron self-splicing: Development of alternative reaction conditions and identification of a predicted intermediate. *Cold Spring Harbor Symp Quant Biol* 52:223–232.
- Peebles CL, Perlman PS, Mecklenburg KL, Petrillo ML, Tabor JH, Jarrell KA, Cheng HL. 1986. A self-splicing RNA excises an intron lariat. *Cell* 44:213–223.
- Podar M, Chu VT, Pyle AM, Perlman PS. 1998. Group II intron splicing in-vivo by first step hydrolysis. *Nature* 391:915–918.
- Podar M, Dib-Hajj S, Perlman PS. 1995. A UV-induced Mg²⁺-dependent crosslink traps an active form of domain 3 of a self-splicing group II intron. *RNA* 1:828–840.
- Portmann S, Grimm S, Workman C, Usman N, Egli M. 1996. Crystal structures of an A-form duplex with single-adenosine bulges: a conformational basis for site-specific RNA self-cleavage. *Chemistry and Biology* 3:173–184.
- Pyle AM, Green JB. 1994. Building a kinetic framework for group II intron ribozyme activity: Quantitation of interdomain binding and reaction rate. *Biochemistry* 33:2716–2725.
- Qin PZ, Pyle AM. 1997. Stopped-flow fluorescence spectroscopy reveals that domain 1 of a group II intron is an independent folding unit. *Biochemistry* 36:4718–4730.
- Query CC, Strobel SA, Sharp PA. 1996. Three recognition events at the branch-site adenine. *EMBO J* 15:1392–1402.
- Roberts RJ. 1995. On base flipping. *Cell* 82:9–12.
- Ruskin B, Krainer AR, Maniatis T, Green MR. 1984. Excision of an intact intron as a novel lariat structure during pre-mRNA splicing in vitro. *Cell* 38:317–331.
- Santoro SW, Joyce GF. 1997. A general purpose RNA-cleaving DNA enzyme. *Proc Natl Acad Sci USA* 94:4262–4266.
- Schmelzer C, Müller MW. 1987. Self-splicing of group II introns in vitro: Lariat formation and 3'-splice site selection in mutant RNAs. *Cell* 51:753–762.

- Schmelzer C, Schweyen RJ. 1986. Self-splicing of group II introns in vitro: Mapping of the branch point and mutational inhibition of lariat formation. *Cell* 46:557–565.
- Valegard K, Murray JB, Stockley PG, Stonehouse NJ, Liljas L. 1994. Crystal structure of an RNA bacteriophage protein–operator complex. *Nature* 371:623–626.
- van der Veen R, Arnberg AC, van der Horst G, Bonen L, Tabak HF, Grivell LA. 1986. Excised group II introns in yeast mitochondria are lariats and can be formed by self-splicing in vitro. *Cell* 44:225–234.
- van der Veen R, Kwakman JHJM, Grivell LA. 1987. Mutations at the lariat acceptor site allow self-splicing of a group II intron without lariat formation. *EMBO J* 6:3827–3831.
- Vijayraghavan U, Parker R, Tamm J, Iimura Y, Rossi J, Abelson J, Guthrie C. 1986. Mutations in conserved intron sequences affect multiple steps in the yeast splicing pathway, particularly assembly of the spliceosome. *EMBO J* 5:1683–1695.
- Yang AS, Shen JC, Zingg JM, Mi S, Jones PA. 1995. Hha1 and HpaII DNA methyltransferases bind DNA mismatches, methylate uracil and block DNA repair. *Nucleic Acids Res* 23:1380–1387.