Sequence requirements for branch formation in a group II self-splicing intron

Rachel Altura⁺, Brian Rymond[§], Bertrand Seraphin and Michael Rosbash

Department of Biology, Brandeis University, Waltham, MA 02254, USA

Received August 5, 1988; Revised and Accepted November 28, 1988

ABSTRACT

Evidence is presented for the existence of a specific intron-intron interaction, necessary for the formation of the branched product in the self-splicing reaction of a group II yeast mitochondrial intron. Trans-splicing reactions involving two RNA molecules (5'exon with covalently linked regions of intron and intron with covalently linked 3'exon) show that the presence of portions of intron domain I on the 5' molecule is necessary for the formation of branched products which are not seen with shorter 5' molecules. Modification/interference reactions show regions necessary for branch-formation and support a major role for specific regions of intron domain I. Further experiments, utilizing a truncated 3' molecule that is missing the conserved branchpoint nucleotide, indicate that domain VI may be required for a successful domain I interaction. A model for the formation of a proper branched structure includes implications for both cis and trans configurations.

INTRODUCTION

The self-splicing group II introns of mitochondria are good current models for the chemistry of nuclear pre-mRNA splicing (for review see ref. 1). In both cases, splicing occurs in two steps; the first consists of 5' cleavage and lariat formation and the second, 3' cleavage and exon ligation (see Fig. 1A and ref. 2,3). It is very likely that the two steps are coupled phosphate transfer transesterification reactions, as originally proposed for the group I self-splicing molecules (4,5).

A diagnostic and distinctive feature of the group II and nuclear pre-mRNA splicing reactions is the 2'-5' phosphodiester bond present at the lariat branchpoint. This bond presumably reflects the fact that the 2' OH of the branchpoint nucleotide acts as the nucleophile for 5' splice junction cleavage. Yet

there is no apparent reason why the 5' splice junction could not be cleaved by hydrolysis, i.e., by H_2O or OH^- . Although it appears that this is rarely if ever the case when the wild-type group II molecules (or nuclear pre-mRNA introns) are incubated under standard, splicing conditions (2,3), 5' splice junction hydrolysis does take place when group II molecules are incubated under altered reaction conditions (6). Also, 5' splice junction cleavage occurs exclusively by hydrolysis in the trans-splicing experiments of Jacquier and Rosbash (7). In these experiments, RNA containing exon I with a single nucleotide of intron (the Tag substrate) was incubated under self-splicing conditions with RNA containing two nucleotides of exon I, the entire intron and exon II. (We shall henceforth refer to this molecule as "the enzyme" because it presumably contains the catalytic activity for 5' splice junction cleavage. We note that the term is not strictly correct because the enzyme is cleaved in the course of the reaction.) Both RNAs were inert when incubated alone. Incubation of the two RNAs together led to efficient trans-splicing, in which the substrate 5' splice junction and the enzyme 3' splice junction were cleaved, and the substrate-derived exon I ligated to the enzyme-derived exon II (Fig. 1). Importantly, the substrate intron fragment was not covalently attached to the enzyme. This observation, in conjunction with other results, led to the conclusion that 5' splice junction cleavage took place via a hydrolysis mode rather than a branch mode (Fig. 1C rather than Fig. 1B). It is likely that the structure of the enzyme-substrate complex was aberrant so that, like under altered reaction conditions, activation of the 5' splice junction took place but without nucleophilic attack by the 2' OH of the branchpoint nucleotide. Presumably, the intron binding sites (IBS) present on exon I (8) are sufficient to promote a productive interaction with the enzyme but insufficient to allow branch formation. The presence of additional intron sequences on the substrate might therefore restore its ability to undergo branch formation. The experiments reported below verified this prediction and uncovered some unexpected and interesting consequences of extending the substrate intron.

FIGURE 1: Schematic representation of the group II self-splicing reaction. The two arrows indicate the two consecutive steps of the reaction: (1) 5' cleavage and the formation of the intron-3' exon intermediate; (2) 3' cleavage, exon ligation and release of the intron.

(A) The cis-splicing reaction, involving a full-length precursor RNA of 5' and 3' exons (solid lines) and intron (open box), results in an intron product which is a lariat.

(B,C) Alternate forms of the trans-splicing reaction, involving substrates with varying lengths of upstream intron covalently attached to the 5' exon (open box with star), result in the formation of a branched intron product (B) or a linear intron product with hydrolysis of the intron portion of the substrate(C).

MATERIALS AND METHODS

In Vitro Transcription

Recombinant plasmids spAE and spTE (see ref.7) were linearized by digestion with the following restriction endonucleases: TaqI,

SspI, Sau3A, HinfI (for spAE) and EcoRI, HpaII (for spTE).

Linear DNAs were then transcribed under conditions described by Melton et al. (9) at 40°C for 2 h. (20 to 50 μ Ci of 32 P-UTP (800 Ci/mmol) was added to those reactions resulting in internally-labelled RNAs). Purification was by DNase treatment, phenol extraction, and ethanol precipitation, followed by gel electrophoresis on 4% polyacrylamide-urea gels and elution overnight in PK buffer (0.1 M Tris/HCl, pH 7.5, 12.5 mM EDTA, 0.15 M NaCl, 1% SDS). Those transcripts derived from spAE plasmids were designated substrates and those transcripts derived from spTE plasmids were designated enzymes in the trans-splicing reactions.

3' End-Labeling of RNA

Unlabelled RNAs, generated from spAE plasmids, were 3' end-labelled with pCp, as previously described (10).

In Vitro Self-Splicing Reactions

Internally labelled or 3' end-labelled substrate (0.06 μ M, unless otherwise stated) was combined with unlabelled enzyme (0.5 μ M, unless otherwise stated) under standard splicing conditions, i.e., 5 mM EPPS (pH 8.0), 5 mM MgCl₂, 2 mM spermidine, and carrier yeast tRNA (0.2 μ g/ μ l). Reactions were performed at 45°C. (Reactions whose final volumes were 10 μ l or less were performed in sealed, siliconized capillary tubes.) Reactions were quenched at designated times (see Fig. legends) by mixing samples with one- to three-fold (vol/vol) excess loading buffer (98% formamide, 10 mM EDTA, 0.1% xylene cyanol, 0.1% bromphenol blue). Samples were heated for 5 min at 65°C prior to analysis by polyacrylamide-urea gel electrophoresis. In most cases, where resolution of RNA products of sizes greater than 100 nucleotides was necessary, 4% gels were employed. In cases where sizes of products were expected to be less than 30 nucleotides in length (i.e., for analysis of the Tag reaction), 20% gels were used.

Debranching Treatment

RNA products of trans-splicing reactions were debranched according to published procedures (11). HeLa cell nuclear extracts, used in place of cytoplasmic extracts, were a gift of Michael Green.

FIGURE 2: Restriction map of the RNA substrates and RNA enzymes used in the trans-splicing reaction. SP6 polymerase was used to transcribe the previously described templates spAE (denoted SUBSTRATE; ref.7) and spTE (denoted ENZYME; ref.7). The sizes of the exon regions (boxes) and intron portion (solid line) are indicated below the schematic of the plasmid insert. The various run-off transcripts used in this study are indicated below each fusion plasmid with the length of intron at the right.

Modification/interference assay

Because the Hinf substrate was too long for a convenient assay, we used a somewhat shorter Sau3A substrate (Fig. 2) which gave rise to essentially the same products as the Hinf substrate. Approximately 1 μ g of 3' end labelled substrate RNA (SP6 run-off transcripts of Sau3A-digested spAE DNA) was incubated with diethylpyrocarbonate (DEPC) or hydrazine under conditions

previously determined to generate less than one lesion per molecule (12) and purified as described above. Splicing was initiated by mixing equimolar amounts of unlabelled enzyme RNA (run-off transcripts from either HpaII- or EcoRI-digested SpTE) and modified, radioactive substrate RNA in a 20 μ l reaction containing high salt (500 mM $(\text{NH}_4)_2\text{OAc}$, 100 mM MgCl_2 , 0.1% SDS, 40 mM Tris-HCl, pH 7.7) to improve the product yield. As a control, a portion of the substrate RNA was incubated without added enzyme. After 3 hours at 45°C the reaction was stopped by the addition of 20 μ l of 500 mM EDTA, 100 μ g/ml tRNA and the RNA precipitated by adding 20 μ l 7.5 M $(\text{NH}_4)_2\text{OAc}$ and 150 μ l of 100% ethanol. The RNA was recovered by spinning for 10 min in a microfuge, and the pellet was washed twice with 80% ethanol and dried. The two labelled products of splicing (the branched molecules and hydrolysis product) were separated from the unreacted substrate RNA on a denaturing 5% gel, the positions of each species located by autoradiography, and RNA recovered, as described above. Chemical cleavage at the sites of DEPC modification was induced by aniline treatment (12) and the resulting cleavage patterns analyzed on a 6% sequencing gel. To precisely identify the positions of modification an RNA sequence ladder (13) of the substrate RNA was run in adjacent lanes of each experiment. Similar results were obtained under low salt conditions (the same as were used for the other experiments in this paper) except that branch formation was further reduced by modification in the 5' splice site (intron nt 1-6) and the extent of enhancement was less pronounced. In general, the bands containing modified nucleotides at the sites of 5' cleavage were difficult to assay in these experiments because their migration rates are similar to those of full-length, unmodified molecules.

RESULTS

To understand why the Tag substrate undergoes splicing with little or no branch formation, we decided to synthesize other substrates with more intron sequence at their 3' ends. Cleavage of the template DNA with HinfI and transcription with SP6 polymerase generated a substrate RNA (Hinf substrate) which differed from the Tag substrate by the presence of an additional

215 nucleotides of intron (intron nucleotides 2-216) at its 3' end (Fig. 2). ^{32}P 3' end-labelled Hinf substrate was added to non-radioactive, complete enzyme (Fig. 2) and incubated under standard splicing conditions. (We note that in this experiment a substantial amount of intron sequence, nucleotides 1-216, was present both on the substrate molecule and the enzyme molecule, in contrast to the Tag substrate reaction in which only intron nucleotide 1 was duplicated on both molecules.) A small fraction of the radioactivity was released as a linear 217 nucleotide RNA in a time and enzyme-dependent fashion, consistent with the notion that some hydrolysis occurred at the proper 5' splice junction. A much larger fraction of the substrate radioactivity was transferred to a series of three closely migrating species of RNA (Fig. 3A and data not shown). The size of these RNAs was approximately that expected for the branched products of a lariat-like trans-splicing reaction, i.e., for the transfer of 217 nucleotides of intron from the Hinf substrate to the normal branch site on the enzyme (8 nucleotides from the 3' end of the intron (14)). Under identical conditions, incubation of enzyme and equivalent amounts of end-labelled Tag substrate resulted in hydrolysis and release of the pGp ^{32}Cp intron with little or no branch formation, as previously described (7 and Fig. 3B).

To verify that these radioactive products (labelled B in Fig. 3A) were branched molecules with a proper 2'-5' phosphodiester bond, the three different RNA species were separately purified and individually incubated with debranching activity. Debranching of species B3 (the fastest migrating of the three bands) resulted in the appearance of a labelled RNA of the correct size for the end-labelled intron portion of the Hinf substrate (Fig. 4). Similar results were obtained with species B1 and B2 (data not shown). The generation of a 217 nucleotide species with debranching activity suggested that 5' cleavage took place at (approximately) the correct location and that the intron portion of the Hinf substrate was branched during splicing to another molecule of RNA (presumably the intron portion of the enzyme) via a 2'-5' phosphodiester bond. In conjunction with the size of this branched molecule and the small fraction of the Hinf substrate cleaved by hydrolysis, the debranching data suggest

FIGURE 3: Time course of the trans-splicing reaction involving the *HinfI* and *TagI* substrates and the complete enzyme. Substrates, synthesized by *in vitro* transcription of *HinfI*- or *TagI*-digested spAE plasmids, were 3' end-labelled with ^{32}P prior to the splicing reaction. Enzyme RNA was nonradioactive.

(A) 2.0 μM of *HinfI* substrate RNA and 0.9 μM of enzyme RNA were incubated under standard conditions. Time points were analyzed as described in Materials and Methods. Arrow B indicates the position of the putative branched intron products (B1-3). Arrow H indicates the 217 nucleotide hydrolysis product.

(B) 3 μM of end-labelled *TagI* substrate RNA and 0.3 μM of unlabelled enzyme RNA were incubated together. Reactions were terminated as indicated above. Arrow H indicates the 2-nucleotide hydrolysis product. No radioactivity was transferred to a higher molecular weight RNA.

that the *HinfI* substrate interacted with the enzyme such that 5' splice junction cleavage occurred predominantly via a branch (lariat-like) mechanism. We presume that the three different branched molecules were due to the use of three different branch sites on the enzyme (see Discussion). Thus, the presence of intron sequences on the *HinfI* substrate that are absent from the *TagI* substrate must be necessary for branch formation.

FIGURE 4: Characterization of the *Hinf* product of the trans-reaction. Three high-molecular weight products of the trans-splicing reaction, involving the *Hinf* substrate, were prepared from a preparative splicing reaction for treatment with debranching enzymes. The fastest migrating of the three species (B3), shown here, was dried and suspended in 20 μ l of debranching buffer before the addition of 5 μ l of HeLa cell extract. The reaction was incubated for 60 min at 30°C before the addition of stop buffer. Products were analyzed on a 4% polyacrylamide gel. Lane 1 shows the putative 217 nucleotide debranched product (position indicated by the lower arrow) and a small amount of branch remaining after debranching (position indicated by the upper arrow). Lane 2 is the 123 basepair marker lane. Mock incubation without the addition of debranching activity did not give rise to a 217 nt band (data not shown).

To gain more insight into what fraction of the first 216 nucleotides of the intron are involved in branch formation and what specific role they may be playing, we applied modification/interference procedures to the substrate:enzyme reaction (12). In brief, 3' end-labelled *Sau*3A substrate (see Fig. 2 legend) was treated gently with modification reagents (DEPC or hydrazine) such that less than one modification event per molecule took place (see Fig. 5 legend). The modified substrate was incubated with enzyme and the two radioactive products (193 nt linear RNA and branch) were purified. These were subjected to

FIGURE 5: Modification/interference of branch formation and hydrolysis. Shown is the chemical cleavage pattern of the DEPC-modified *Sau* substrate (input), the branched intron product (branch), the hydrolysis product (hydrolysis) and the substrate remaining unspliced after incubation in the *trans*-splicing reaction (unspliced). Brackets indicate regions in which the yield of modified nucleotides is substantially reduced (downward pointing arrows) in the branched products. Also, the position labelled 1 (downward pointing arrow) refers to nucleotide 1 of the intron, also found in reduced yields in the branched products and resolved on longer runs but not visible in this figure. The other regions (upward pointing arrows) indicate positions of nucleotides enhanced in the branch or hydrolysis products. The numbers at the right indicate the positions of intron nucleotides 1 and 116.

the chemical reactions which cleave specifically at the locations of modified nucleotides. The resulting cleavage ladders, when compared with the control lanes, indicated those modified nucleotides that influence hydrolysis or branch formation (e.g., Fig. 5).

In general, modification of the substrate within the intron had little or no effect on hydrolysis as the cleavage ladder derived from the hydrolysis product resembled closely that of the input RNA (compare lanes 1 and 3, Fig. 5). In contrast, modification of the substrate intron sequences had a pronounced effect on branch formation (compare lanes 1 and 2, Fig. 5). Modified residues could be grouped into three categories on the basis of the nature of the effect on branch formation. Some modified residues had little or no effect on branch formation (which may indicate that the sequences at these locations are not critical for branch formation - but see (12), Materials and Methods, and Fig. 5 legend), most modified residues inhibited branch formation, and some enhanced branch formation (Fig. 5). Modification of some of the residues in this latter category also had a mild, positive effect on hydrolysis (compare lanes 1 and 3, Fig. 5).

The 887 nucleotide AI5 intron has six major secondary structural domains (domain I-VI), originally proposed based on phylogenetic comparisons (15, 16). The 193 nucleotide intron portion of the Sau3A substrate contains a substantial fraction of domain I (Figure 6). Correlating the modification pattern (Fig. 5 and data not shown for pyrimidine cleavage patterns and other purine cleavage patterns) with the proposed structure of domain I led to the conclusion that modified nucleotides relatively neutral for branch formation were predominantly clustered in putative non-paired regions (Fig. 6). This is perhaps best illustrated by two neutral regions situated close to the 5' splice site (intron residues 11-14 and 35-44; residues 35-44 lie between the two bracketed regions in Fig. 5 and are indicated in Fig. 6 by double arrows). These regions are in unpaired loops and are surrounded by inhibitory regions that reside in putative base-paired stems (intron residues 7-10 and 15-34). In the case

FIGURE 6: Distribution of modification pattern on domain I of the AI5 secondary structure. Modified nucleotides that are neutral (solid circles) or enhance (solid triangles) branch formation are indicated on a predicted secondary structure map of the first 193 nucleotides of the AI5 intron (14,15). A dashed line connects intron nucleotides 185-186. D helices are not shown and are indicated by a D flanked by arrows. Modified nucleotides that are undesignated in the structure inhibit the formation of the branch. Also shown are enzyme intron nucleotides 388-415 that have been predicted to base pair with the first 193 nucleotides of domain I. Intron nucleotides 1 and 193 are indicated, as well as the neutral nucleotides 35-44, by the double arrows.

of one of these neutral regions (the AAAG sequence at intron position 11-14), the two flanking inhibitory regions (intron positions 7-10 and 15-34, respectively) are complementary to domain I sequences absent from the substrate-intron and present only on the enzyme. This pattern suggested that portions of domain I were being formed in *trans*, by pairing between substrate-intron sequences and enzyme-intron sequences. It

suggested further that these interactions are required for branch formation.

More generally, we interpret these observations in the following way. The addition of the 200-225 nucleotides of intron to the Tag substrate (thereby generating the longer substrates) shifted the splicing mode from predominantly hydrolysis to predominantly branch formation. Branch formation required that exon I be covalently attached to sufficient intron sequences so that appropriate interactions (in large part intron-intron interactions) take place in trans with the enzyme. Positions of modification that inhibited branch formation therefore identified important intron regions that function in this way. The small fraction of hydrolysis that was observed with the longer substrates was insensitive to modification of these regions because it occurred without the participation of the substrate intron region, as in the case of the Tag substrate. Positions of modification that enhanced branch formation (and in some cases hydrolysis) may reflect regions of the intron which function better when in cis with the rest of the enzyme. Since these sequences are duplicated, their presence in the substrate may compete for the formation of an important cis interaction in the enzyme, thereby producing a less active trans structure.

To further explore the relationship between hydrolysis and branch formation, we generated an enzyme which could only cleave by hydrolysis and not by branch formation. By digesting the enzyme template DNA with HpaII, we eliminated domain VI, which contains the presumed sites of branch formation. We expected that this truncated enzyme might be unable to undergo branch formation with the Hinf substrate (with the low level of hydrolysis remaining unaffected) yet be indistinguishable from the complete enzyme in cleaving the Tag substrate via hydrolysis, i.e., branch formation would be eliminated and the amount of hydrolysis with the Tag substrate would be considerably greater than the amount of hydrolysis with the Hinf substrate. (We note that Jarrell *et al.* also tested splicing under conditions in which domain VI was absent and found no branched products (17).) As expected, the truncated enzyme was only able to cleave via hydrolysis, with both the Tag and the Hinf substrates as well as

FIGURE 7: Time course of reactions with the two different enzymes. The substrates, indicated on the figure, were internally labelled. The same number of moles of substrate were added in all six cases so that the rate of mRNA formation (labelled only in exon I) in the presence of the complete enzyme could be compared directly with the rate of exon I formation in the presence of the truncated enzyme. Products of both (A) and (B) were assayed on the same gel.

(A) Reaction with the complete enzyme. Arrows B show the position of the three branched products formed as a result of the reaction with the *HinfI* and *SspI* substrates and the complete enzyme (the latter are visible on the autoradiogram but not on the figure). Arrow H shows the position of the 216 nucleotide hydrolysis product of the *HinfI* reaction. (Analogous products of the *SspI* and *TaqI* reactions are not resolved on this low percentage gel.) Arrow M shows the position of the mature mRNA (exons 1 and 2). As is clear, the levels of message formed with each substrate were, within a factor of two or so, identical.

(B) Reaction with truncated enzyme, which consists of the first step of the splicing reaction (5' cleavage), via hydrolysis. Arrow H (above) shows the position of the 216 nucleotide hydrolysis product formed as a result of the reaction with the

Hinf substrate and the truncated enzyme. Arrow H (below) shows the position of the 30 nucleotide hydrolysis product formed as a result of the reaction with the Ssp substrate and the truncated enzyme. Arrow E1 shows the position of the 5' exon. (It is noted that, due to the nature of transcription, some RNA substrates were heterogeneous at their 3' ends and therefore could give rise to heterogeneous, mature hydrolysis products during the course of the splicing reaction.)

with a third substrate of intermediate intron length (Fig. 7). However, two features of this experiment were surprising. First, 5' splice site cleavage was five- to ten-fold faster with the truncated enzyme than with the complete enzyme for all three substrates. We surmise that the rate of 5' cleavage is normally limited by some interaction which was eliminated by the removal of domain VI, hence the higher rate of 5' cleavage, even for the Tag substrate. Second, the time course and extent of hydrolysis with the Hinf substrate (and, within a factor of two, the Ssp substrate) were the same as those observed with the Tag substrate. This latter observation indicated that the splicing which occurred by branch formation with the complete enzyme was shifted to the hydrolysis mode with the truncated enzyme. When we applied the modification/interference procedures described above, but with the truncated enzyme instead of the complete enzyme, 5' splice site cleavage, as predicted, was insensitive to the inhibitory effects of modification of the intron portion of the Sau3A substrate (although a residual level of enhancement was still detectable) (Fig. 8). Presumably, in the absence of domain VI, the Hinf intron-enzyme interaction necessary for branch formation fails to take place or no longer has a strong effect on the splicing reaction.

DISCUSSION

Our observations suggest that trans-splicing through a branched intermediate requires that the intron sequences of the Hinf substrate interact with intron sequences on the enzyme, resulting in the formation of a structure resembling that of the normal, cis-group II molecule. Apparently, there is a minimum length of substrate intron required for a productive interaction to take place. Although incubation of 3' end-labelled Tag substrate with complete enzyme sometimes caused the transfer of a small amount

FIGURE 8: The effect of removing domain VI on modification/interference. Shown are the cleavage patterns of modified purines in the substrate (input), in the hydrolysis product, obtained using the *Hpa*II truncated enzyme (hydrolysis*), and in the hydrolysis product (hydrolysis) and branched molecules (branch) isolated from *trans*-splicing reactions using the full-length enzyme.

of radioactivity to a higher molecular weight RNA, we have been unable to confirm that these nucleotides were attached via a 2'-5' phosphodiester band. We conclude that one nucleotide of intron is not sufficient for branch formation. However, intermediate lengths of intron (between 1 and 216 nucleotides) resulted in real branch formation (cf. Figs. 2-4). Even a substrate intron length of 13-28 nucleotides (generated by oligonucleotide-directed RNase H cleavage) was sufficient for some branch formation to take place, as evidenced by the release

from the putative branch of a ladder of appropriately sized radioactive oligonucleotides with debranching enzyme activity. In general, the fraction of 5' cleavage which took place via branch formation increased with increasing length of substrate intron (data not shown). Presumably, this indicates that the larger the substrate intron the better it is able to interact with the enzyme in a manner which mimics the structure of the intact cis molecule.

Even with 216 nucleotides of intron present in the Hinf substrate, however, the structure of the enzyme-substrate complex was almost certainly abnormal as compared to that of the cis molecule. This is evidenced by the fact that a small amount of 5' cleavage still occurred by hydrolysis and, more importantly, by the fact that at least three different branched molecules were detectable with the Hinf substrate. (Multiple branched species were also generated with other, shorter substrates (data not shown).) Because each branched species liberated a 217 nucleotide RNA upon incubation with debranching enzyme, we suspect that, in all cases, 5' cleavage had occurred at the correct site but that for at least two of these molecules an incorrect nucleotide was used as the nucleophile, hence the aberrant electrophoretic mobility. Although in our view this interpretation is likely to be correct, we have not proven that 5' cleavage occurred at precisely the correct phosphodiester bond and that an incorrect branch site was the only unusual feature of the molecule. In any case, the Hinf substrate favors branch formation (even incorrect branch formation) over hydrolysis. This recalls similar observations with mammalian, nuclear splicing systems in which introns utilize alternate branchpoints when the normal branchpoint is mutated, apparently with little difficulty (18, 19).

The inhibitory features of the modification pattern indicate that branch formation is dependent upon a substrate intron-enzyme interaction which forms a portion of the domain I-like structure in trans. Presumably the substrate intron sequence displaces or otherwise competes with the identical sequence present on the enzyme. An implication is that branch formation requires that some fraction of domain I be covalently linked to

exon I. As the rate of 5' splice site cleavage was quite similar for the Tag and Hinf substrates and as the modifications that inhibit branch formation failed to enhance hydrolysis, it is possible that a covalent association of exon I and a subset of domain I aids in promoting branch formation rather than suppressing hydrolysis. For example, the correct (cis) structure may contain features which prevent 5' splice junction activation unless an appropriate branchpoint 2' OH nucleophile is within striking distance, thereby precluding hydrolysis. This is in contrast to the more conventional explanation, namely, that the correct cis structure may exclude competing nucleophiles, such as H₂O or OH⁻, from the active site (7, 20).

It is somewhat surprising that the rate of 5' splice site cleavage was relatively insensitive to the nature of the attacking nucleophile (a 2' OH in the case of the branched products and presumably H₂O or OH⁻ products in the case of the hydrolysis products). This suggests that some other step, perhaps a conformational change or intron-intron interaction, was rate limiting. As the rate of 5' splice junction cleavage was quite similar for the Tag and Hinf substrates, the formation of the domain I-like structure between the Hinf substrate and the enzyme was probably not involved in this putative rate-limiting step; it was however, affected by the removal of domain VI (which contains the normal branchpoint), as the time course with this truncated enzyme is five- to ten-fold faster than with the normal enzyme. As incubation of the longer substrates with this truncated enzyme failed to exhibit branch formation and the associated sensitivity of 5' cleavage to the inhibitory effects of chemical modification of the substrate intron, it would appear that the absence of domain VI prevented substrate-enzyme interactions necessary for an effective trans formation of domain I. An attractive but by no means exclusive hypothesis to account for the observed rate enhancement is that other, perhaps related domain I interactions also occur in cis, involve domain VI, and are rate limiting for 5' splice junction cleavage. The enhancement regions may identify some of the domain I regions that participate in interactions of this nature. Possibly, the enhanced rate observed in the absence of domain VI reflects

FIGURE 9: Model of substrate/enzyme interactions in the first step of the splicing reaction. Shown are the Tag substrate (Tag S) and Enzyme (E), which interact at the IBS-EBS regions, (solid circle) and the Hinf substrate (Hinf S) and Enzyme (E), which interact at the IBS-EBS regions, as well as at additional regions (open circle) involving the intron portion domain I of the substrate and a portion of the intron of the enzyme. The additional intermolecular interactions of the Hinf reaction allow the substrate to proceed through a commitment step for branch formation (B), absent in the Tag reaction ([]). K is the rate-determining step of the splicing reaction, preceding the formation of exon 1 and the released guanosine nucleotide of the Tag reaction or the formation of exon 1 and the branched intermediate of the Hinf reaction. K remains the same for both reactions with the complete enzyme shown. Removal of domain VI affects both B and K, B being eliminated under these conditions and K being increased and thus causing an increase in the overall rate of the splicing reaction.

interactions which may provide some (as yet undetected) measure of fidelity or control at the expense of rate. A summary of these ideas is depicted in the model presented in Figure 9.

Whereas many of these interpretations are speculative, some of them are easily testable. First, a similar rate effect of deleting domain VI may be detectable in cis, i.e., by deleting domain VI and exon II from the normal group II molecule. Second, mutations in the enhancement portions of domain I of the normal group II molecule should inhibit splicing substantially. Finally, the strong effect of modification within much of the Hinf I intron on branch formation suggests that a single modified residue can have a pronounced effect on the trans-formation (folding?) of some of domain I. This may reflect an impediment to branch formation which may be testable by more direct means.

ACKNOWLEDGEMENTS

We thank Michael Green for his generous gift of debranching extract and Claudio Pikielny for his comments on the manuscript. We especially thank all members of the Rosbash laboratory for their thoughtful suggestions during the course of this study and, in the preparation of the manuscript, specifically, Hildur V. Colot for her writing suggestions, and Tobie Tishman for her assistance in compiling this manuscript. Most of this work was submitted by R.A. to the Brandeis University Department of Biology in partial fulfillment of the requirements for an undergraduate honors degree. This work was supported by National Institutes of Health grant GM23549.

Present addresses: ⁺Washington University, School of Medicine, St Louis, MO 63110 and [§]University of Kentucky, School of Biological Sciences, Lexington, KY 40506-0225, USA

REFERENCES

1. Cech, T.R. (1986) *Cell* **44**, 207-210.
2. Van der Veen, R., Arnberg, A.C., Van der Horst, G., Boren, L., Tabak, H.F. and Grivell, L.A. (1986) *Cell* **44**, 225-234.
3. Peebles, C.L., Perlman, P.S., Mecklenburg, K.L., Petrillo, M.L., Tabor, J.H., Jarrell, K.A. and Cheng, H.L. (1986) *Cell* **44**, 213-223.
4. Zaug, A.J., Grabowski, P.J. and Cech, T.R. (1983) *Nature* **301**, 578-583.
5. Cech, T.R., Zaug, A.J. and Grabowski, P.J. (1981) *Cell* **27**, 487-496.
6. Jarrell, K.A., Peebles, C.L., Dietrich, R.C., Romiti, S.L. and Perlman, P.S. (1988) *J. Biol. Chem.* **263**, 3432-3439.
7. Jacquier, A. and Rosbash, M. (1986) *Science* **234**, 1099-1104.
8. Jacquier, A. and Michel, F. (1987) *Cell* **50**, 17-29.
9. Melton, D.A., Krieg, P.A., Rebagliati, M.R., Maniatis, T., Zinn, K. and Green, M.R. (1984) *NAR* **12**, 7035-7056.
10. Pikielny, C.W. and Rosbash, M. (1986) *Cell* **45**, 869-877.
11. Ruskin, B. and Green, M.R. (1985) *Science* **229**, 135-140.
12. Rymond, B.C. and Rosbash, M. (1988) *Genes and Development* **2**, 428-439.
13. Peattie, D.A. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 1760-1764.
14. Schmelzer, C. and Schweyen, R.J. (1986) *Cell* **46**, 557-565.
15. Michel, F. and Dujon, B. (1982) *Biochimie* **64**, 867-881.
16. Michel, F. and Dujon, B. (1983) *EMBO J.* **2**, 33-38.
17. Jarrell, K.A., Dietrich, R.C. and Perlman, P.S. (1988) *Mol. Cell. Biol.* **8**, 2361-2366.
18. Reed, R. and Maniatis, T. (1985) *Cell* **41**, 95-105.
19. Ruskin, B., Greene, J.M. and Green, M.R. (1985) *Cell* **41**, 833-844.
20. van der Veen, R., Kwakman, J.H.J.M. and Grivell, L.A. (1987) *EMBO J.* **6**, 3827-3831.