Research Article

Participation of the Intron in the Reaction Catalyzed by the *Xenopus* tRNA Splicing Endonuclease

M. IRENE BALDI, EMILIO MATTOCCIA, EMANUELA BUFARDECI, STEFANIA FABBRI, GLAUCO P. TOCCHINI-VALENTINI*

Introns have generally been assumed to be passive in the transfer RNA splicing reaction. Experiments have now been done showing that the endonuclease is able to cut a precursor provided that a base in the single-stranded loop of the intron can pair with the base of the 5′ exon situated at the position that immediately follows the anticodon stem (position 33 in the yeast tRNA isoacceptor pre-$tRNA_3^{Leu}$, position 32 in yeast pre-$tRNA^{Phe}$). The elucidation of the role of the intron reveals that in addition to the conserved bases, there are positions in the mature domain which, although not necessarily occupied by the same base in all pre-tRNA's, nevertheless have a fundamental role in the splicing reaction. These positions are termed cardinal positions.

IN *XENOPUS*, A SINGLE ENDONUCLEASE CAN CLEAVE SEVERAL tRNA precursors to create substrates for subsequent ligation (*1*). The enzyme acts (i) by specific recognition and binding of the precursors and (ii) by cleaving of the two sites with consequent formation of the two halves and of the intron (*2*). The enzyme must therefore be dealing either directly or indirectly with features shared by all the precursors (*2, 3*).

Calculation of free energy minima (*4*) and the use of chemical and enzymatic structure-specific probes (*5*) suggest that all the precursors examined have a common tertiary structure (*6, 7*). In this structure, the tRNA portion of the precursor maintains the L-shaped conformation, stabilized by the interaction between the D and TψC loops. The 3′ splice site junction is always single-stranded. The introns, which are generally assumed to be passive in the splicing reaction, occupy the same relative location, interrupting the anticodon loop one base after the anticodon. There are no conserved sequences at the splice junctions.

The exact way in which endoribonuclease recognizes the precursors has not yet been determined. In general, however, the type of interaction of an RNA molecule with a protein bears greater resemblance to protein-protein binding than to protein-DNA binding. Analyses of simplified systems have shown that both the conformation and a few specific, highly conserved nucleotides are recognized (*8*). The binding of the *Xenopus* endonuclease to pre-tRNA presumably requires many interactions. We have previously shown that mutations in the mature domain affect precursor recognition. In particular, U^8 and C^{56} appear to be probable contact points between protein and RNA (*9*). We now show that the intron is an active participant in the splicing reaction.

Alongside the conserved bases, there are positions in the mature domain which, although not necessarily occupied by the same base in all pre-tRNA's, participate in the splicing reaction. We term these positions cardinal positions (CP).

The *Xenopus* endonuclease is, in fact, able to cut a precursor if a base in the single-stranded loop of the intron is allowed to pair with the base of the 5′ exon situated at the position that immediately follows the anticodon stem (position 33 in yeast pre-$tRNA_3^{Leu}$, position 32 in yeast pre-$tRNA^{Phe}$). This position is the first cardinal position (CP1). A pyrimidine, uracil or cytosine, is normally found at CP1, but we have found that a purine can accomplish the same function. A second cardinal position (CP2) is localized in the 3′ exon. CP2 corresponds to the 5′ end of the 3′ half; we now show that the base occurring at CP2 influences the determination of the 3′ splice site (Fig. 1).

Role of the intron in splicing. A number of studies of the effect of drastic alterations of the intron have been made. These alterations, in general, appeared to have little effect on the specificity or extent of splicing (*10, 11*). As a result, no general importance has yet been attributed to the sequence of the intron.

Another fact lending support to this view is that mutant pre-tRNA's with predominantly homopolymeric introns and anticodon loops are perfectly spliced (*12*). We constructed a variant of the yeast $tRNA^{Phe}$ gene. The mutant pre-tRNA, known as poly(C) (Fig. 2), is characterized by the predominance of C residues over the entire region comprising the anticodon loop and intron. Three G's (G^{42}, G^{49}, and G^{54}) were included in the intron to facilitate product characterization, while the C^{32} and A^{57} residues that form the 5′ and 3′ bases of the anticodon loop and the G^{37} at the position preceding the 5′ splice site were retained. This precursor, whose features parallel those of the poly(U) precursor (*12*), is accurately cut at both sites by the *Xenopus* endonuclease (Fig. 2).

We also found that the poly(C) G54C precursor, in which the G at the antepenultimate position of the intron (−3 from the 3′ splice site) has been replaced by C, is not cut at both sites. This result suggests that G^{54} is important for the determination of the splice sites in the poly(C) precursor. Since, in both yeast and *Xenopus*, a single endonuclease activity is capable of recognizing and correctly

M. I. Baldi, E. Mattoccia, and G. P. Tocchini-Valentini are at the Institute of Cell Biology, Consiglio Nazionale delle Ricerche, Viale C. Marx 43, 00137 Roma, Italy. E. Bufardeci and S. Fabbri are with EniChem, Dipartimento di Biotecnologie 00015 Monterotondo, Roma, Italy. M. I. Baldi and G. P. Tocchini-Valentini are also affiliated with the Department of Molecular Genetics and Cell Biology, University of Chicago, Chicago, IL 60637.

*To whom correspondence should be addressed.

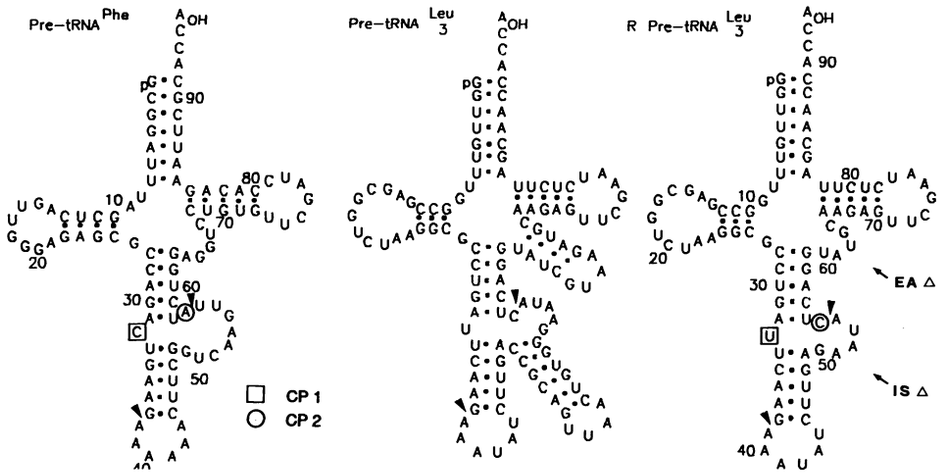

Fig. 1. Secondary structure of yeast pre-tRNA's. Nucleotide sequence of yeast pre-tRNA's. Arrows indicate the splice sites. In the secondary structure of reduced pre-$tRNA_3^{Leu}$ deletions in the extra arm and in the intervening sequence are indicated. CP, cardinal position.

cleaving the various precursors (*1, 13*), we would expect a recognition element for the endonuclease to be among the conserved bases. The sequence is not highly conserved at any of the intron positions, with one important exception. Guanosine appears at the antepenultimate position in approximately 80 percent of the precursors with known sequences. This consensus, although remarkable in itself, is not consistent with a specific requirement for guanosine (*14*). The frequency with which guanosine appears at the antepenultimate position, however, may reflect a structural requirement for pairing with another semiconserved position. If this is indeed the case, the existence of a structural requirement for a specific base pairing would explain why guanosine is not always present at the antepenultimate position. What all precursors must conserve is the possibility of forming the base pair. The question thus narrows down to which base undergoes pairing with the base at the antepenultimate position of the intron.

To gain further insight into the possible role of the base at the antepenultimate position, we focused our attention on two yeast pre-tRNA's: pre-$tRNA^{Phe}$ and pre-$tRNA_3^{Leu}$. Pre-$tRNA_3^{Leu}$ was chosen because, unlike the phenylalanine precursor, it has an adenosine instead of guanosine at the antepenultimate position. The $tRNA_3^{Leu}$ isoacceptor is a type 2 tRNA, and the structure of its precursor is considerably more complex than that of the phenylalanine precursor. To facilitate the comparison between the two precursors, we constructed a "reduced" form of pre-$tRNA_3^{Leu}$.

Construction of a reduced pre-$tRNA_3^{Leu}$ that is active in splicing. The most variable feature of tRNA is the so-called extra arm, which lies between the TψC and anticodon stems. Depending on the nature of the extra arm, tRNA's can be divided into two classes. Type 1 tRNA's, like yeast $tRNA^{Phe}$ have a small extra arm, consisting of only three to five bases; they represent about 75 percent of all tRNA's. Type 2 tRNA's, like yeast $tRNA_3^{Leu}$, have a large extra arm with 13 to 21 bases.

In addition to these characteristics of the mature domain, another feature that distinguishes pre-$tRNA_3^{Leu}$ from pre-$tRNA^{Phe}$ is the length of the intervening sequence (IVS) which is 32 bases long in the former and 19 bases long in the latter. To convert pre-$tRNA_3^{Leu}$, we used the EA deletion (which reduces the size of the extra arm) and the IS deletion (which reduces the length of the IVS). The double mutant EAΔ, ISΔ (Fig. 1), also known as reduced pre-$tRNA_3^{Leu}$ (R pre-$tRNA_3^{Leu}$), is predicted to have a cloverleaf structure quite similar to that of a type 1 tRNA.

The double mutant is perfectly spliced by the *Xenopus* splicing endonuclease (Fig. 3); the length of the 5′ half is unchanged but, as we expected, the length of the 3′ half and that of the IVS are reduced in each case. The sequence of the IVS was verified by RNA sequence

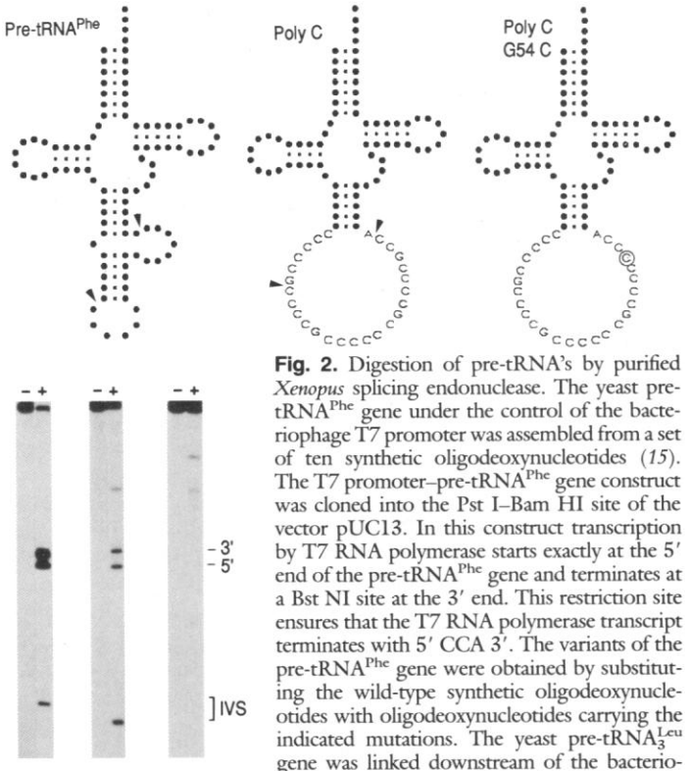

Fig. 2. Digestion of pre-tRNA's by purified *Xenopus* splicing endonuclease. The yeast pre-$tRNA^{Phe}$ gene under the control of the bacteriophage T7 promoter was assembled from a set of ten synthetic oligodeoxynucleotides (*15*). The T7 promoter–pre-$tRNA^{Phe}$ gene construct was cloned into the Pst I–Bam HI site of the vector pUC13. In this construct transcription by T7 RNA polymerase starts exactly at the 5′ end of the pre-$tRNA^{Phe}$ gene and terminates at a Bst NI site at the 3′ end. This restriction site ensures that the T7 RNA polymerase transcript terminates with 5′ CCA 3′. The variants of the pre-$tRNA^{Phe}$ gene were obtained by substituting the wild-type synthetic oligodeoxynucleotides with oligodeoxynucleotides carrying the indicated mutations. The yeast pre-$tRNA_3^{Leu}$ gene was linked downstream of the bacteriophage T7 promoter by site-directed mutagenesis (*16*) of plasmid pTGEM2 (*9*). By the same method a Bst NI restriction endonuclease site was introduced precisely at the 3′ end of the gene ensuring that the T7 RNA polymerase transcript terminates with 5′ CCA 3′. Variants of the pre-$tRNA_3^{Leu}$ gene were obtained by site-directed mutagenesis with the oligodeoxynucleotides as primers on appropriate single-stranded DNA templates. To synthesize tRNA precursors, plasmid DNA's were cleaved with Bst NI restriction enzyme and transcribed with T7 RNA polymerase (*17*). When $tRNA_3^{Leu}$ DNA was used in the transcription reaction the final concentration of the labeled triphosphate was raised to 100 μM. Transcripts were purified by polyacrylamide gel electrophoresis. The tRNA-splicing endonuclease was prepared from *Xenopus laevis* germinal vesicles extract (*3*) and further purified on a tRNA-agarose column. Yeast tRNA (Sigma) was oxidized with 4.6 mM sodium periodate (*18*). The oxidized tRNA was coupled to the adipic acid-dihydroxide agarose (Sigma) (*19*). The enzyme fraction used in the assays was purified more than 1000-fold over the GV extract. Pre-tRNA's labeled with α-^{32}P were incubated with the purified endonuclease in a 30-μl reaction mixture for 60 minutes at 22°C (*20*); the final concentration of precursor in the assay was approximately 1 nM. Identities are shown as cloverleaf structures, in corresponding order. Bases relevant to the study are labeled; the circled base indicates the mutation; arrows indicate the splice sites. Lanes on the left (marked −) show the electrophoretic separation of the precursor incubated with buffer only. Lanes on the right (marked +) show the electrophoretic separation of the products of endonuclease reaction.

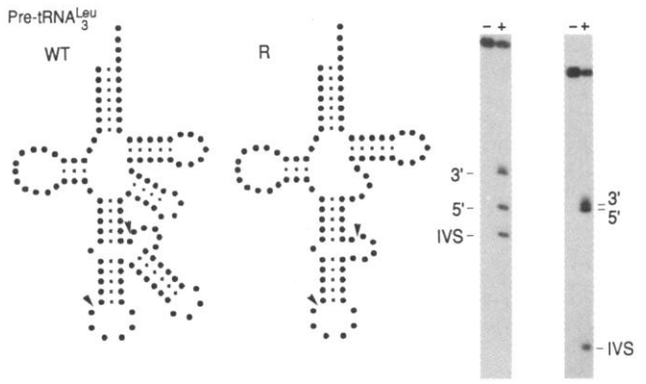

Fig. 3. Digestion of pre-tRNA's by purified *Xenopus* splicing endonuclease. Identities are shown as cloverleaf structures, in corresponding order. Arrows indicate the splice sites. Lanes on the left (marked −) show the electrophoretic separation of the precursor incubated with buffer only. Lanes on the right (marked +) show the electrophoretic separation of the products of endonuclease reaction.

analysis. Not only do the changes introduced not alter the position of the splice sites, but the K_m and V_{max} of the mutant precursor are similar to those of the wild-type leucine and phenylalanine tRNA precursors. The K_m values (10^{-10} *M*) were for pre-tRNA^{Phe}, 7.0; for pre-tRNA$_3^{Leu}$ and the mutant, 6.0 and 7.0. The V_{max} values (10^{-16} mol min^{-1}) were for pre-tRNA^{Phe}, 26; for pre-tRNA$_3^{Leu}$ and the mutant, 18 and 28.

The first cardinal position is in the 5′ exon. If a structural requirement for a specific base pairing does in fact exist, there should be a base that interacts with A^{52}, the antepenultimate base of the reduced pre-tRNA$_3^{Leu}$.

A likely candidate could be the base at position 33 (the first base after the 5-bp anticodon stem), where there is always a pyrimidine, most often cytidine, at this position in the various precursors. The pyrimidine character ensures the possibility of pairing with the two purines normally found at the antepenultimate position, according to the generally accepted pair types U · A, U · G, C · G. Another factor that makes the pyrimidine at position 33 an attractive candidate is that it lies just opposite the 3′ site, and is therefore in the vicinity of the antepenultimate base in the intron of normal pre-tRNA's.

To test our hypothesis, we produced the mutant of the reduced pre-tRNA$_3^{Leu}$ U33C (Fig. 4), in which the base at position 33 is C and cannot pair with A^{52} (the antepenultimate base of the intron in the wild-type precursor). There is only one G in the single-stranded loop of the mutant pre-tRNA; if C^{33} could pair with G^{50}, we should still have, according to our hypothesis, an active precursor. In Fig. 4, we show that the U33C precursor is indeed an excellent precursor for the *Xenopus* endonuclease. There is a single intron product however, which is two bases shorter than the wild-type intron. RNA sequence analysis shows that the 3′ cleavage site has shifted two bases in the 5′ direction. The 3′ half is, as a result, two bases longer. This finding indicates that, at least in the case of the U33C precursor, the base at position 33 not only satisfies the requirement for an active precursor by pairing with a base in the single-stranded loop of the intron, but it also has a role in the determination of the 3′ splice site. The base in the single-stranded loop of the intron that pairs with the base at position 33 becomes automatically the antepenultimate base. We refer to position 33 as the first cardinal position (CP1) because, as we shall see, there is also a second such key position, in the 3′ exon. The question then arises whether the phenotype of the U33C mutation be suppressed with a compensatory second mutation.

It turns out that the *Xenopus* endonuclease cleaves the two double-mutant reduced pre-tRNA$_3^{Leu}$ precursors (i) U33C, A52G and (ii) U33C, C55A and produces introns and halves of exactly the same size as those derived from the wild-type precursor (Fig. 4).

We discuss below the U33C, C55A mutant, where we also define and characterize the second cardinal position (CP2); first, however, we must describe the phenotype of the U33C, A52G variant.

In the single-stranded loop of this precursor, there are two guanosines, G^{50} and G^{52}. We assume that the latter guanosine's position at the center of the loop favors the interaction of C^{33} with G^{52} over that of C^{33} with G^{50}. This would explain why the intron and the 3′ half derived from the U33C, A52G precursor are identical in size to the products derived from the wild type; the antipenultimate position of the intron corresponds to position 52 in both cases.

The location of second cardinal position in the 3′ exon. In an effort to generalize the conclusions reached in experiments with the reduced pre-tRNA$_3^{Leu}$ precursor and its variants, we now turn to yeast pre-tRNA^{Phe}. Let us first point out a few features that distinguish the phenylalanine precursor from that of leucine. In pre-tRNA^{Phe} we find a C instead of a U at CP1 (now position 32, since the D loop region of tRNA$_3^{Leu}$ contains an extra base). At the antepenultimate position there is a G which, according to our hypothesis, pairs with C32 to yield the precursor which is spliced by the *Xenopus* endonuclease. Furthermore, the base at the 5′ end of the 3′ half is an A, whereas at the homologous position in reduced pre-tRNA$_3^{Leu}$ there is a C.

In the double mutant pre-tRNA^{Phe} A53G, G54A, the 3′ splice site has not shifted (Fig. 5). The intron and 3′ half are of the same

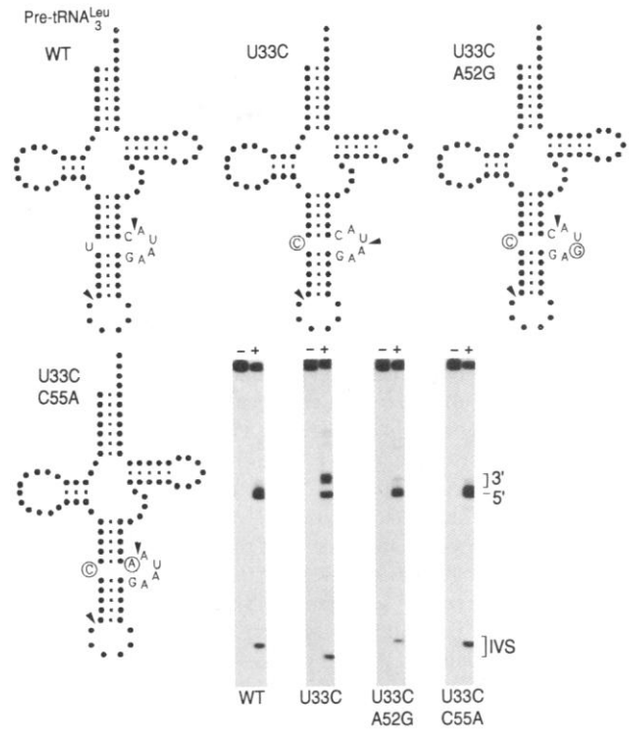

Fig. 4. Digestion of pre-tRNA's by purified *Xenopus* splicing endonuclease. Identities are shown as cloverleaf structures, in corresponding order. Relevant bases are labeled; circled bases indicate mutations; arrows indicate the splice sites. Lanes on the left (marked −) show the electrophoretic separation of the precursor incubated with buffer only. Lanes on the right (marked +) show the electrophoretic separation of the products of endonuclease reaction.

size as their wild-type counterparts. On the basis of what had been observed in the variants of reduced pre-tRNA$_3^{Leu}$, we might expect to see an intron that is one base shorter and a 3′ half that is one base longer; in other words we might expect the 3′ splice site to have shifted one base in the 5′ direction. The C at CP1 cannot have paired with A^{54}, but only with G^{53}, which should have become the antepenultimate base. However, the shift of the 3′ splice site occurs instead in the triple mutant pre-tRNA^{Phe} A53G, G54A, A57C. It is the A57C substitution, therefore, that is the determinant. If there is a C at the 5′ end of the 3′ exon, the phenylalanine precursor behaves like the leucine one. Therefore we refer to the base at the 5′ end of the 3′ exon as the base at the second cardinal position (CP2).

We can now go back to the mutant-reduced pre-tRNA$_3^{Leu}$ U33C, C55A (Fig. 4). There the 3′ splice site does not shift because there is an A at the CP2. When there is an A at CP2, the *Xenopus* enzyme cleaves the phosphodiester bond at the 5′ position of the sugar moiety of adenosine; the 3′ cleavage site is completely determined by the A at the second cardinal position. The base at CP1, however, must pair with a base of the intron for the structural requirement described above to be satisfied. When there is an A at CP2 the base of the intron that interacts with the base at the first cardinal position does not necessarily become "antepenultimate." In Fig. 6, we show that the mutant pre-tRNA^{Phe} G54C is not cut at both sites. There is an A at CP2 as in wild-type tRNA^{Phe}, but the structural requirement is not satisfied. The C at CP1 cannot find an intron base to pair within the single-stranded loop.

In natural pre-tRNA's, a pyrimidine, U or C, is always at CP1. The base of the intron which, according to our view, pairs to satisfy the structural requirement must necessarily be a purine. The question then arises whether the structural requirement can be satisfied by a purine at the cardinal site and a pyrimidine in the intron?

Both the single mutant pre-tRNA^{Phe} C32G and the double mutant pre-tRNA^{Phe} C32G, G54C are spliced (Fig. 6). In both cases, intron and halves are the same size as their wild-type counterparts. An A is present at CP2, while the structural requirement is satisfied in one case by the pair G^{32}-C^{51}, in the other by the pair G^{32}-C^{54}.

The effects of a U or a G at the second cardinal position. Reduced pre-tRNA$_3^{Leu}$ has a C, and pre-tRNA^{Phe} has an A at CP2. If it is C, the base of the intron that pairs with the base at CP1 automatically becomes antepenultimate. If it is A, and the structural

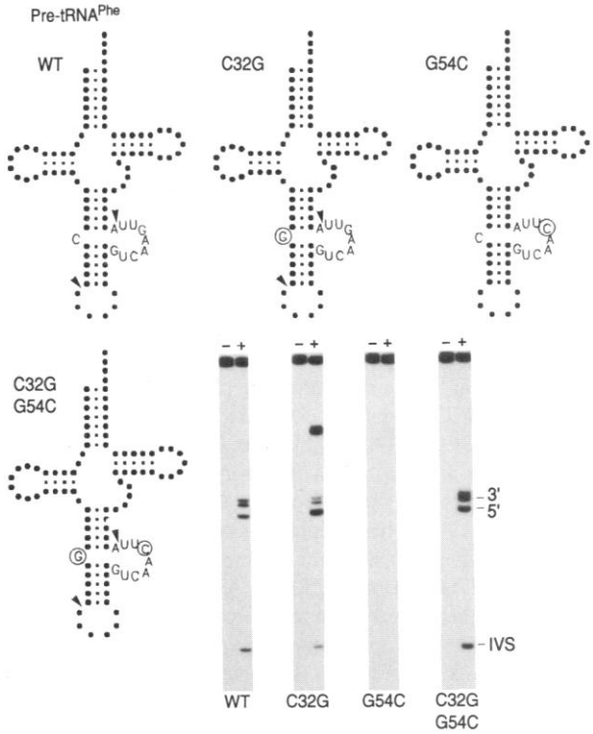

Fig. 6. Digestion of pre-tRNA's by purified *Xenopus* splicing endonuclease. Identities are shown as cloverleaf structures, in corresponding order. Bases relevant to the study are labeled; circled bases indicate mutations; arrows indicate the splice sites. Lanes on the left (marked −) show the electrophoretic separation of the precursor incubated with buffer only. Lanes on the right (marked +) show the electrophoretic separation of the products of endonuclease reaction.

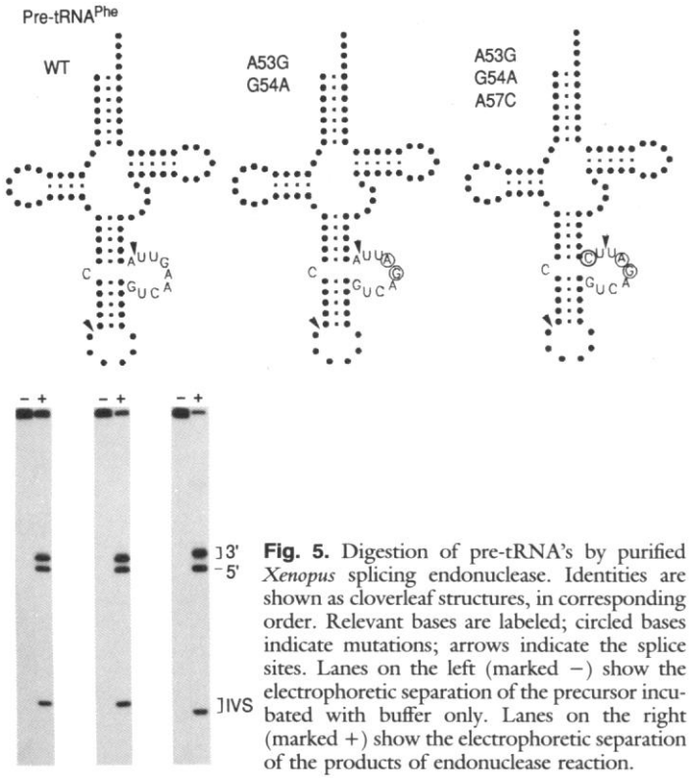

Fig. 5. Digestion of pre-tRNA's by purified *Xenopus* splicing endonuclease. Identities are shown as cloverleaf structures, in corresponding order. Relevant bases are labeled; circled bases indicate mutations; arrows indicate the splice sites. Lanes on the left (marked −) show the electrophoretic separation of the precursor incubated with buffer only. Lanes on the right (marked +) show the electrophoretic separation of the products of endonuclease reaction.

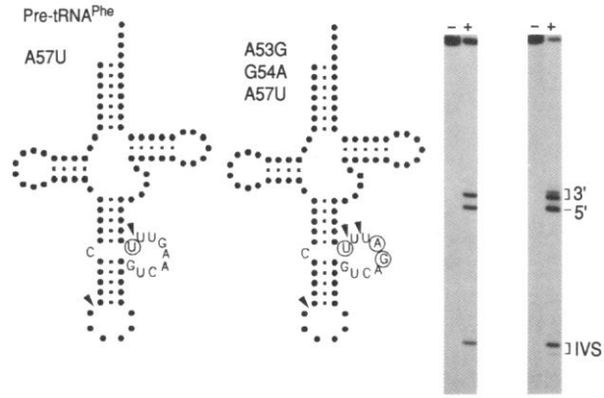

Fig. 7. Digestion of pre-tRNA's by purified *Xenopus* splicing endonuclease. Identities are shown above, as cloverleaf structures, in corresponding order. Relevant bases are labeled; circled bases indicate mutations; arrows indicate the splice sites. Lanes on the left (marked −) show the electrophoretic separation of the precursor incubated with buffer only. Lanes on the right (marked +) show the electrophoretic separation of the products of endonuclease reaction.

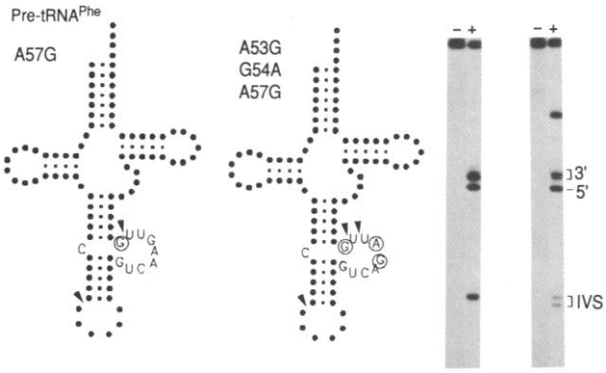

Fig. 8. Digestion of pre-tRNA's by purified *Xenopus* splicing endonuclease. Identities are shown as cloverleaf structures, in corresponding order. Bases relevant to the study are labeled; circled bases indicate mutations; arrows indicate the splice sites. Lanes on the left (marked −) show the electrophoretic separation of the precursor incubated with buffer only. Lanes on the right (marked +) show the electrophoretic separation of the products of endonuclease reaction.

requirement is satisfied, the location of the 3′ site is such that the A is always found at the 5′ end of the 3′ half.

It is interesting to consider what happens if instead there is a U or a G at CP2. In Fig. 7, we show that the products derived from the pre-tRNA^{Phe} A57U mutant are normal in size. In the case of pre-tRNA^{Phe} A53G, G54A, A57U, there is a single 5′ half, but two kinds of 3′ halves and two kinds of introns. In 70 percent of cases, a U at CP2 determines the 3′ cleavage site; in the remaining 30 percent of cases, G^{53} is the antepenultimate base.

To illustrate what happens when there is a G at CP2, we used pre-tRNA^{Phe} A57G and pre-tRNA^{Phe} A53G, G54A, A57G (Fig. 8). In the case of the triple mutant there are two kinds of 3′ halves and two kinds of introns, in approximately equal proportions. We conclude, therefore, that the relative 3′ site determining power of the four bases when they occupy CP2 is A(100), U(70), G(50), C(0).

We have come to two conclusions regarding the mechanisms of pre-tRNA splicing catalyzed by the *Xenopus* endonuclease.

First, the intron is not passive in the splicing reaction. Although the sequence is not highly conserved, one base of the single-stranded loop must pair with the base at CP1. Nuclear magnetic resonance studies should clarify whether the base pair preexists in the pre-tRNA molecule or whether it is formed as a result of the interaction with the enzyme.

Our second conclusion is that adjacent to the conserved bases that act as probable contact points between protein and RNA there are positions in the mature domain which, although not necessarily occupied by the same base in all pre-tRNA's, play a fundamental role in the splicing reaction. We have, in fact, identified two such cardinal positions: CP1 and CP2. The existence of cardinal positions is entirely in line with previous results indicating that there is a measuring mechanism governing splice sites determination (*9, 12*).

REFERENCES AND NOTES

1. A. Otsuka, A. De Paolis, G. P. Tocchini-Valentini, *Mol. Cell Biol.* **1**, 269 (1981).
2. M. I. Baldi, E. Mattoccia, S. Ciafré, D. Gandini-Attardi, G. P. Tocchini-Valentini, *Cell* **47**, 965 (1986).
3. D. Gandini-Attardi, M. I. Baldi, E. Mattoccia, G. P. Tocchini-Valentini, *Methods Enzymol.* **181**, 510 (1989).
4. I. Tinoco, Jr., *et al.*, *Nature New Biol.* **246**, 40 (1973).
5. P. Wrede, R. Wurst, J. Vournakis, A. Rich, *J. Biol. Chem.* **254**, 9608 (1979).
6. H. Swerdlow and C. Guthrie, *ibid.* **259**, 5197 (1984).
7. M. C. Lee and G. Knapp, *ibid.* **260**, 3108 (1985).
8. M. Wickens and J. E. Dahlberg, *Cell* **51**, 339 (1987).
9. E. Mattoccia, M. I. Baldi, D. Gandini-Attardi, S. Ciafrè, G. P. Tocchini-Valentini, *ibid.* **55**, 731 (1988).
10. J. D. Johnson *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 2564 (1980).
11. G. L. Raymond and J. D. Johnson, *Nucleic Acids Res.* **11**, 5969 (1983).
12. V. M. Reyes and J. Abelson, *Cell* **55**, 719 (1988).
13. C. L. Peebles, D. Gegenheimer, J. Abelson, *ibid.* **32**, 525 (1983).
14. E. Szekely, H. G. Belford, C. L. Greer, *J. Biol. Chem.* **263**, 13839 (1988).
15. V. M. Reyes and J. Abelson, *Anal. Biochem.* **166**, 90 (1987).
16. E. Mattoccia, M. I. Baldi, G. Pande, R. Ogden, G. P. Tocchini-Valentini, *Cell* **32**, 67 (1983).
17. D. A. Melton *et al.*, *Nucleic Acids Res.* **12**, 7035 (1984).
18. S. R. Fahnestock and N. Nomura, *Proc. Natl. Acad. Sci. U.S.A.* **69**, 363 (1972).
19. N. Ulbrich and I. G. Wool, *J. Biol. Chem.* **253**, 9049 (1978).
20. D. Gandini-Attardi, I. Margarit, G. P. Tocchini-Valentini, *EMBO J.* **4**, 3289 (1985).
21. We thank G. Di Franco for technical assistance and A. Sebastiano for typing the manuscript. Supported in part by Progetto Finalizzato Ingegneria Genetica, by Progetto Finalizzato Biotecnologie e Biostrumentazione, and by HFSPO (Human Frontiers Science Program Organization).

9 December 1991; accepted 21 January 1992