"Five Easy Pieces"

PHILLIP A. SHARP

THE DISCOVERY OF SELF-SPLICING OR CATALYTIC INTRONS sparked interest in the role of RNA catalysis in cellular processes. There are several cellular processes in which highly conserved RNAs are critical components. For example, protein translation by ribosomes requires the participation of tRNAs and ribosomal RNAs. Could the amino acid transferase activity of this process be catalyzed by RNA? Analogously, the splicing of nuclear precursor to mRNAs (pre-mRNAs) requires the activities of five different small nuclear RNAs (snRNAs) in formation of a spliceosome. Could the cleavage and joining reactions in the spliceosome be RNA-catalyzed? There have been strong indications for several years that the answer to the second question is yes, but concrete evidence is still not available. The only two known cases where branched RNAs are formed in nature are in the splicing of nuclear pre-mRNAs in the spliceosome process and the splicing of group II catalytic introns (*1*). Recent studies of the structure of genes in chloroplast may now have provided a glimpse of the type of intermediate steps that could have occurred in the evolutionary transition between group II introns and spliceosomal snRNAs (*2*).

In *Chlamydomonas reinhardii* the three exons encoding the psaA protein are scattered on both strands of the 140-kb genome at widely spaced intervals. The introns flanking these exons contain sequences reminiscent of group II introns. Almost all group II introns can be folded into a common structure with six spokes of duplex segments, I, II, III, IV, V, and VI, radiating from a circle of conserved sequences (*3*). The necessity of joining such widely dispersed exons led to the suggestion that each set of two exons was joined by formation of a competent catalytic intron for trans-splicing of the exons through duplex pairing of the two parts of the intron. In fact, previous studies in vitro have shown a similar reaction (*4*). Consistent with this model, the intron sequences flanking exons 2 and 3 of the *psa*A gene can pair via an intermolecular spoke IV to generate an apparently complete six-spoke group II–type structure.

If a functional group II intron can be formed from two separate RNAs, could it be formed by coalescing more than two parts? Apparently, yes. Goldschmidt-Clermont *et al.* (*2*) found that splicing of *psa*A exons 1 and 2, but not 2 and 3, could be destroyed by mutations in a second chloroplast gene *tsc*A. Sequence analysis of *tsc*A revealed no open reading frame for protein, but RNA analysis detected a small 430-nt species. A tripartite combination of the small RNA (*tsc*A-RNA) and sequences in the two introns flanking *psa*A exons 1 and 2 form a prototype group II intron. The three RNAs were paired through formation of intermolecular spokes I and IV of the consensus intron structure (Fig. 1). This example of generation of a small RNA that probably functions in trans to promote the formation of a structure competent for splicing encourages speculation that group II introns could be the evolutionary precursors of snRNAs.

Genes whose expression depends upon trans-splicing are not uncommon in the genomes of cytoplasmic organelles in plants (*5, 6*). The *nad*1 gene of wheat, which encodes subunit 1 of the respiratory chain NADH dehydrogenase, is scattered over the mitochondrial genome in five segments (*7*). Two of the exons are separated by a prototype cis-splicing group II intron. The other three exons are flanked by sequences common to group II introns, and trans-splicing, like that proposed for *psa*A, is thought to be responsible for formation of the mature mRNA. The trans-spliced introns were probably created by DNA rearrangements with end points in the middle of the two cis-spliced group II introns flanking an exon. The equivalent *nad*1 gene from the mitochondria of another plant, *Oenothera* (*8*), is also distributed in five exons, but cis-spliced group II introns separate two pairs of exons in this case, with trans-splicing probably being responsible only for joining one pair. This suggests that the highly structured group II introns can be easily converted by DNA rearrangements into trans-spliced segments and also perhaps into trans-acting parts of a catalytic intron.

The original structure of genes was almost certainly RNA, and it has been suggested that the joining of short RNA exons by trans- and cis-splicing could have produced complex functional proteins. The presence of group II catalytic introns in this early stage of evolution could certainly fulfill the hypothetical role of promoting efficient cis- and trans-splicing. During evolution of the progenitor cell for eukaryotes, group II introns could have been fragmented into small trans-acting RNAs as described above. Division of a group II intron into "five easy pieces" could have generated the precursors for the five snRNAs that form the spliceosome in the nuclei of eukaryotic cells. A note of caution is warranted. A simple prediction of this hypothesis is that the consensus secondary and tertiary structures of the group II intron family should be present in structures formed by snRNAs. As yet, the equivalent sequences or structures have not yet been identified in the combination of snRNAs found in the spliceosome.

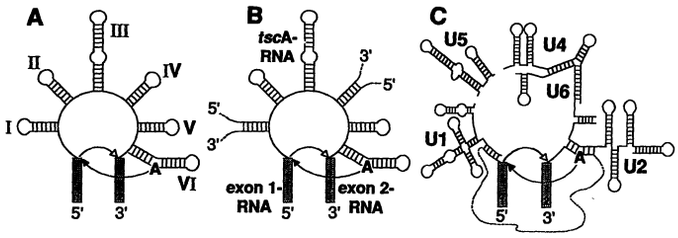

Fig. 1. Intermediate between group II catalytic intron and snRNAs? (**A**) The consensus structure for group II introns is illustrated as six helices (spokes) (*3*). Sequences that form part of these helices as well as sequences that connect the helices are conserved among group II introns. The loops at the tops of the hairpins can be of variable length. The adenosine residue in spoke VI is the branch site that is used in the first step during cleavage at the 5′ splice site (arrow with filled head). The second step in splicing is a reaction of the 5′ exon with the 3′ splice site producing a lariat intron and spliced exon products (arrow with open head). (**B**) The three RNAs (exon 1–RNA, *tsc*A-RNA, and exon 2–RNA) coalesce to form a prototype group II catalytic intron (*2*). The three RNAs are bound together by base pairing in formation of spokes I and IV. (**C**) The hypothetical splicing of nuclear pre-mRNAs by the five snRNAs. The adenosine paired in duplex with U2 snRNA is the site of branch formation. Base-pairing interactions between U1 snRNA and the 5′ splice site, U4 and U6 snRNAs, U6 and U2 snRNAs, and U2 snRNA and the branch site are illustrated. The steps in splicing are the same as in (A).

Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139.

REFERENCES

1. C. Guthrie, *Science* **253**, 147 (1991).
2. M. Goldschmidt-Clermont *et al.*, *Cell* **65**, 135 (1991).
3. F. Michel, K. Unesano, H. Ozeki, *Gene* **82**, 5 (1989).
4. K. A. Jarrell, R. C. Dietrich, P. S. Perlman, *Mol. Cell. Biol.* **8**, 2361 (1988).
5. H. Fukuzawa *et al.*, *FEBS Lett.* **198**, 11 (1986).
6. N. Zaita, K. Torazawa, K. Shinozaki, M. Sugiura, *ibid.* **210**, 153 (1987).
7. Y. Chapdelaine and L. Bonen, *Cell* **65**, 465 (1991).
8. B. Wissinger *et al.*, *ibid.*, p. 473.