Long-distance splicing

Alissa M. Anderson and Jonathan P. Staley*
Department of Molecular Genetics and Cell Biology, University of Chicago, 920 East 58th Street, Chicago, IL 60637

With the draft sequence of the human genome came the surprise that there were fewer genes than imagined. From where does complexity spring if not from the number of genes in an organism? RNA splicing provides at least part of the answer. Pre-mRNA splicing by alternative pathways is well known to expand an organism's protein diversity by generating distinct protein isoforms. Beyond *cis*-splicing at a single locus (Fig. 1*a*), there is evidence for specialized *cis*-splicing that results from read-through transcription of adjacent loci followed by splicing to generate transcription-induced chimeras (TICs) from two genes (Fig. 1*b*) (1, 2), as in the *TNSF12/TNSF13* chimera expressed in human T cells (3). In contrast to these *cis*-splicing events, *trans*-splicing joins exons from separate pre-mRNA transcripts. These transcripts can be encoded by different DNA strands at the same locus, as in *trans*-splicing of the *mod(mdg4)* gene in *Drosophila*, or by different alleles at the same locus, as for the *lola* gene, also in *Drosophila* (Fig. 1*c*) (4). All of these RNA splicing events involve transcripts from the same general region of the genome. In the work by Di Segni *et al.* in this issue of PNAS (5), the authors provide evidence suggesting yet another pathway to increase protein diversity, a pathway that involves *cis*-splicing of a single mRNA or *trans*-splicing of distinct mRNAs from distant genes by the tRNA-splicing machinery (Fig. 1 *d* and *e*).

Whereas splicing of nuclear pre-mRNA generally requires an RNA–protein machine, termed the spliceosome, splicing of eukaryotic tRNAs requires only proteins—an endonuclease, a ligase, and a phosphotransferase. These enzymes catalyze tRNA splicing in four steps (6, 7). First, the tRNA endonuclease cleaves both exon–intron junctions. This cleavage produces two half-tRNAs as well as a linear intron and at the cleavage sites, free hydroxyls at the cleaved 5′ ends and 2′,3′-cyclic phosphates at the cleaved 3′ ends. The next three steps are catalyzed by the tRNA ligase and involve the polynucleotide kinase, cyclic phosphodiesterase, adenylate synthetase, and ligase activities of this enzyme (7). First, the 5′-OH is phosphorylated and the 3′-cyclic phosphate is opened to form a 3′-OH and 2′-phosphate. Then, the two tRNA halves are ligated and, at least *in vitro*, the excised tRNA intron is also ligated, circularizing the intron (8). Finally, the 2′-phosphate is cleaved from the ribose. Just as pre-mRNA splicing generally occurs in the nucleus, pre-tRNA splicing in higher eukaryotes occurs in the nucleus, but pre-tRNA splicing in budding yeast occurs in the cytoplasm (9).

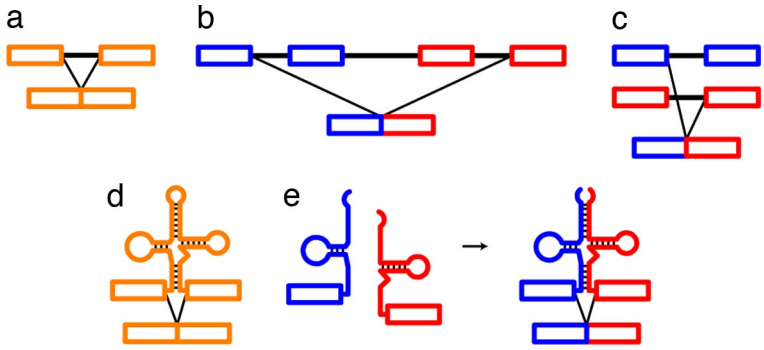

Fig. 1. RNA splicing pathways observed in eukaryotes (*a*–*c*) and engineered in yeast by Di Segni *et al.* (5) (*d* and *e*). (*a*) Canonical pre-mRNA *cis*-splicing yielding mRNA from pre-mRNA composed of two exons (orange boxes) and an intron (black line). (*b*) A transcription-induced chimera generated by read-through transcription from one gene (blue) into a downstream gene (red), followed by *cis*-splicing. (*c*) *Trans*-splicing of exons from two distinct pre-mRNAs (red and blue boxes). (*d*) *Cis*-splicing of a pre-mRNA embedded with a permuted tRNA. (*e*) *Trans*-splicing of two hybrid RNAs each composed of an mRNA fragment and a half tRNA. Note that in the tRNAs the aminoacyl acceptor stem is at the top and the anticodon loop is at the bottom (*d* and *e*).

How does the pre-tRNA splicing pathway recognize a substrate? In general, eukaryotic pre-tRNA is recognized for splicing, not by its intron sequences, as is largely the case in pre-mRNA splicing, but by the structure of the mature tRNA domain itself (10). In contrast, archaeal pre-tRNAs form a structure at the splice sites called a bulge-helix-bulge (BHB), which is required for recognition by the tRNA-splicing machinery. In either case, it is the tRNA-splicing endonuclease that recognizes the splice sites (11). Previous work from Tocchini-Valentini and colleagues (12, 13) has demonstrated that mature domain-independent splicing can occur in yeast or *Xenopus* if a sequence that will form a BHB structure is inserted into a model pre-tRNA or mRNA. Moreover, an exogenous archaeal endonuclease expressed endogenously in mammalian cells can splice a BHB structure inserted into a reporter mRNA (14).

In this issue, Di Segni *et al.* (5) designed an elegant system to ask whether tRNAs can be used to recruit the endogenous tRNA-splicing machinery in yeast to mediate mRNA splicing (5). The authors engineered constructs that would produce hybrid pre-tRNA/pre-mRNAs. The constructs contain *SUP4* $tRNA^{Tyr}$, a suppressor of nonsense ochre (UAA) mutants, within *STE2* or *STE3*, two genes involved in the mating signal transduction pathway in *Saccharomyces cerevisiae*. The authors tested for *cis*- or *trans*-splicing by using constructs either with the entire tRNA inserted into a single mRNA or with half-tRNAs fused to two mRNA fragments (Fig. 1 *d* and *e*). Functionality of the tRNA after splicing was assayed by the ability of the hybrid construct to confer prototrophy for methionine, lysine, or adenine in a strain harboring nonsense ochre mutations in genes involved in the biosynthesis of these nutrients. Additionally, a reporter for a functional mating signal transduction pathway indicated whether splicing of *STE2* or *STE3* into mature mRNAs was successful.

First, the authors tested whether the yeast cells could excise a pre-tRNA from the middle of a pre-mRNA. In this *cis*-splicing system, the mRNA essentially corresponds to the intron of a permuted pre-tRNA gene in which the 3′ half lies upstream of the 5′ half (15). Remarkably, the yeast cells not only process functional tRNA from the hybrid construct but also ligate the flank-

Author contributions: A.M.A. and J.P.S. wrote the paper.
The authors declare no conflict of interest.
See companion article on page 6864.
*To whom correspondence should be addressed. E-mail: jstaley@uchicago.edu.

ing "intron" to yield mature mRNA in a reaction that parallels the circularization of tRNA introns *in vitro*. Importantly, splicing of the hybrid transcript occurred precisely at the tRNA splicing sites. Production of the ligated mRNA was generally inefficient, which partially reflected the activity of a 5′-to-3′ exoribonuclease and may also reflect competition between the cytoplasmic tRNA endonuclease and the ribosome for the hybrid transcript, although cleavage and ligation of another transcript, *HAC1* (see below), has been shown to occur while that transcript is associated with polysomes (16). Despite inefficient formation of mRNA, these mRNAs were sufficient to produce functional Ste2p and Ste3p proteins. With this *cis*-splicing experiment, Di Segni *et al.* (5) demonstrate that splicing of an mRNA can be mediated by tRNA sequences within the transcript.

Next, Di Segni *et al*. (5) explored the possibility of long-distance *trans*-splicing mediated by tRNA halves. The authors cotransformed two kinds of hybrid constructs into one cell: one construct with the 5′ half of an mRNA fused upstream of the 3′ half of the tRNA, the other with the 5′ half of the tRNA fused upstream of the 3′ half of an mRNA. Complementarity between the two tRNA halves was predicted to juxtapose the two mRNA halves. Indeed, coexpression of hybrid constructs resulted in significant, albeit inefficient, *trans*-splicing and ligation of the mRNA halves. Given the pre-tRNA component of the substrate, splicing almost certainly involves the canonical tRNA-splicing pathway, although this remains to be proven. Nonetheless, these experiments establish the remarkable observation that an endogenous tRNA-like splicing pathway can mediate the formation of chimeric mRNAs in yeast.

In the future, it will be important to determine whether naturally expressed endogenous transcripts exploit this pathway for mRNA formation. Importantly, there is a precedent for splicing of a pre-mRNA by the tRNA-splicing machinery. This pre-mRNA is transcribed from the gene *HAC1*, which encodes a transcription factor in the unfolded protein response. Upon accumulation of unfolded proteins in the endoplasmic reticulum, *HAC1* pre-mRNA is inducibly spliced (17). Before induction, the 5′ UTR and intron of unspliced *HAC1* pre-mRNA base pair to form a secondary structure that stalls translation; after induction, excision of the intron by splicing removes this block to translation (16). Splicing of *HAC1* mRNA results from cytoplasmic cleavage by Ire1p and the subsequent ligation by the tRNA ligase (18); in this splicing reaction, Ire1p substitutes for the tRNA endonuclease. Importantly, in archaea there is also a precedent for *trans*-splicing of tRNA halves (19). Thus, the two key requirements for pre-mRNA splicing in *trans* by the tRNA-splicing machinery, proven in principle by Di Segni *et al*. (5), have already been observed in nature.

tRNA-mediated splicing of mRNAs could provide an interesting tool for synthetic biology.

Although it is yet unknown whether the *trans*-splicing pathway can occur in mammals, bioinformatics will likely provide clues as to whether the tRNA-splicing machinery is utilized to increase protein diversity in higher eukaryotes by ligating pre-tRNA/pre-mRNA hybrids transcribed from distant regions of the genome. Eukaryotic genomes can include thousands of tRNA genes and pseudogenes (20), so it is tempting to speculate that these loci facilitate the formation of chimeric mRNAs. However, such tRNA-mediated splicing would likely leave behind a byproduct embedded in the spliced message, a byproduct derived from a pre-tRNA motif, such as the BHB, that recruits the tRNA endonuclease, unlike in pre-mRNA splicing, in which no motif remains in the mature mRNA. Although the structure, rather than the sequence of this motif, can serve as the major determinant for recruitment, sequence constraints could be circumvented by recursive splicing in which tRNA-mediated *trans*-splicing would join two different transcripts together in a chimera and subsequent pre-mRNA splicing would excise the tRNA-associated sequences. This mechanism might be utilized by higher eukaryotes, in which both pre-mRNA and pre-tRNA splicing occur in the nucleus (9).

Aside from the biological implications of this work, tRNA-mediated splicing of mRNAs could provide an interesting tool for synthetic biology. *Trans*-splicing mediated by tRNAs could provide a way to easily mix and match protein domains to screen for combinations with novel functions. For example, a large-scale screen could be enabled by simply coexpressing pairs of different hybrid fragments in yeast. In any case, the potential of long-distance tRNA-mediated splicing of mRNAs promises that the human genome has many more surprises in store.

1. Akiva P, *et al.* (2006) Transcription-mediated gene fusion in the human genome. *Genome Res.* 16:30–36.
2. Parra G, *et al.* (2006) Tandem chimerism as a means to increase protein complexity in the human genome. *Genome Res.* 16:37–44.
3. Pradet-Balade B, *et al.* (2002) An endogenous hybrid mRNA encodes TWE-PRIL, a functional cell surface TWEAK-APRIL fusion protein. *EMBO J.* 21:5711–5720.
4. Horiuchi T, Aigaki T (2006) Alternative trans-splicing: A novel mode of pre-mRNA processing. *Biol Cell* 98:135–140.
5. Di Segni G, Gastaldi S, Tocchini-Valentini GP (2008) *Cis*- and *trans*-splicing of mRNAs mediated by tRNA sequences in eukaryotic cells. *Proc Natl Acad Sci USA* 105:6864–6869.
6. Abelson J, Trotta CR, Li H (1998) tRNA splicing. *J Biol Chem* 273:12685–12688.
7. Phizicky EM, Greer CL (1993) Pre-tRNA splicing: Variation on a theme or exception to the rule? *Trends Biochem Sci* 18:31–34.
8. Gegenheimer P, Gabius HJ, Peebles CL, Abelson J (1983) An RNA ligase from wheat germ which participates in transfer RNA splicing in vitro. *J Biol Chem* 258:8365–8373.
9. Hopper AK, Shaheen HH (2008) A decade of surprises for tRNA nuclear-cytoplasmic dynamics. *Trends Cell Biol* 18:98–104.
10. Reyes VM, Abelson J (1988) Substrate recognition and splice site determination in yeast tRNA splicing. *Cell* 55:719–730.
11. Li H, Trotta CR, Abelson J (1998) Crystal structure and evolution of a transfer RNA splicing enzyme. *Science* 280:279–284.
12. Di Segni G, Borghese L, Sebastiani S, Tocchini-Valentini GP (2005) A pre-tRNA carrying intron features typical of Archaea is spliced in yeast. *RNA* 11:70–76.
13. Fruscoloni P, Baldi MI, Tocchini-Valentini GP (2001) Cleavage of non-tRNA substrates by eukaryal tRNA splicing endonucleases. *EMBO Rep* 2:217–221.
14. Deidda G, Rossi N, Tocchini-Valentini GP (2003) An archaeal endoribonuclease catalyzes cis- and trans-nonspliceosomal splicing in mouse cells. *Nat Biotechnol* 21:1499–1504.
15. Soma A, *et al.* (2007) Permuted tRNA genes expressed via a circular RNA intermediate in *Cyanidioschyzon merolae*. *Science* 318:450–453.
16. Ruegsegger U, Leber JH, Walter P (2001) Block of HAC1 mRNA translation by long-range base pairing is released by cytoplasmic splicing upon induction of the unfolded protein response. *Cell* 107:103–114.
17. Rosbash M (1996) Mixed mechanisms in yeast pre-mRNA splicing? *Cell* 87:357–359.
18. Gonzalez TN, Sidrauski C, Dorfler S, Walter P (1999) Mechanism of non-spliceosomal mRNA splicing in the unfolded protein response pathway. *EMBO J* 18:3119–3132.
19. Randau L, Münch R, Hohn MJ, Jahn D, Söll D (2005) *Nanoarchaeum equitans* creates functional tRNAs from separate genes for their 5′- and 3′-halves. *Nature* 433:537–541.
20. Goodenbour JM, Pan T (2006) Diversity of tRNA genes in eukaryotes. *Nucleic Acids Res* 34:6137–6146.