Perspective

Mechanisms of Alternative Pre-mRNA Splicing

TOM MANIATIS

MOST NUCLEAR MESSENGER RNA PRECURSORS (PRE-mRNA's) in higher eukaryotes contain multiple introns, which must be precisely excised by RNA splicing (*1*). Some pre-mRNA's are alternatively spliced in different cell types or at different times during development (*2*). Regulated alternative splicing can lead to the production of different proteins from a single pre-mRNA or can function as an on-off switch during development. In the latter case one mode of splicing generates an mRNA that lacks an open translational reading frame, while alternative splicing of the same pre-mRNA yields a functional protein (*3*). Alternative splicing is the primary mechanism involved in the regulatory hierarchy of sex determination in *Drosophila* and in many examples of tissue-specific gene expression (*2, 4*). An understanding of the mechanisms of RNA splicing is therefore of fundamental importance in developmental biology.

Virtually every imaginable pattern of alternative splicing has been reported (*2*), but relatively little is known about the regulatory mechanisms involved. However, recent advances on two different fronts have provided important clues. First, studies of basic splicing mechanisms suggest that differences in the activities or amounts of general splicing factors may participate in the regulation of alternative splicing. Second, studies of specific examples of regulated alternative splicing indicate that splice site selection can be negatively regulated by repressor proteins that bind specifically to pre-mRNA recognition sequences.

Pre-mRNA splicing occurs by a two step mechanism (*1*). In the first step, cleavage occurs at the 5′ splice site, and the guanosine residue at the 5′ end of the intron is covalently joined to an adenine residue near the 3′ splice site within a recognition element, designated the branch point sequence (BPS). The second step involves cleavage at the 3′ splice site and ligation of the exons. In addition to the conserved sequences at the 5′ and 3′ splice sites, the BPS (*5*) and flanking exon sequences (*6*) participate in splice site recognition. The splicing reaction takes place within spliceosomes, which are high molecular weight particles that contain pre-mRNA, small nuclear ribonucleoprotein particles (snRNP's) U1, U2, U4, U5, and U6 (*7*), and non-snRNP splicing factors (*8*). Spliceosomes assembled in vitro contain a large number of non-snRNP proteins (*9*), many of which may correspond to splicing activities identified by in vitro complementation or pre-mRNA binding studies [see (*1*) and (*8*) for reviews]. The functions of some spliceosome components have been determined. The U1 snRNP binds to the 5′ splice site, while the U2 snRNP binds to the branchpoint sequence (*7, 8*). In both cases, binding involves the formation of complementary base pairs between snRNA and pre-mRNA (*7, 8*). Binding of U2 snRNP requires a factor designated U2 auxiliary factor (U2AF) (*8*).

A central question in the study of both basic splicing mechanisms and regulated alternative splicing is how specific pairs of 5′ and 3′ splice sites are chosen for the splicing reaction (splice site selection). The majority of pre-mRNA's in higher eukaryotes contain multiple introns, and numerous cryptic 5′ and 3′ splice sites are scattered through both introns and exons. The splicing machinery must therefore discriminate between normal splice sites and cryptic sites. In addition, exon skipping must be avoided by selecting 5′ and 3′ splice sites that are within the same intron. The magnitude of this problem is illustrated by the existence of a pre-mRNA that contains 65 exons and 2 million nucleotides of RNA (*10*) and by an intron of 200,000 nucleotides (*11*).

On the basis of current information, similar mechanisms of splice site selection may be involved in general pre-mRNA splicing and in regulated alternative splicing. This possibility was demonstrated by two independent studies, one on the characterization of a general splicing factor called splicing factor 2 (SF2) (*12, 13*), and the other on a factor called alternative splicing factor (ASF), which participates in alternative splicing of a viral pre-mRNA (*14*).

SF2 was originally identified by in vitro complementation studies as an activity required for the first step in the splicing reaction (*15*). When SF2 was purified to homogeneity, it was found to be a protein doublet with a molecular size of approximately 33 kilodaltons (kD) (*13*). Purified SF2, which binds nonspecifically to RNA, is required for the assembly and stabilization of spliceosomes and promotes the annealing of complementary RNA (*13*).

The most remarkable characteristic of SF2 is that it has a profound effect on 5′ splice site selection (*12*). This effect was illustrated by analysis of a model pre-mRNA substrate that contained two competing 5′ splice sites and a single 3′ splice site (*6, 12*). In the presence of low concentrations of SF2, the outside (strong) site is preferentially used. However, as the amount of SF2 is increased, splicing switches to the inside (weak) 5′ splice site. One interpretation of this observation is that in the presence of limiting amounts of SF2 only the strong site is recognized by the splicing machinery. However, when the amount of SF2 is no longer limiting, both 5′ splice sites are recognized, but the weak site, which is in closest proximity to the 3′ splice site, is preferentially used. This possibility is consistent with the observation that the proximal 5′ splice site is used when the strength of competing splice sites is identical (*6*). Its effect on 5′ splice site selection and its RNA annealing activity suggest that SF2 may facilitate base pair interactions between U1 snRNA and the 5′ splice site (*12, 13*).

Studies of the alternative splicing of a pre-mRNA encoded by the early region of simian virus 40 (SV40) suggest that general splicing factors participate in regulated alternative splicing (*14*). SV40 early pre-mRNA contains two alternative 5′ splice sites and a single 3′ splice site. Use of the distal 5′ splice site generates an mRNA that encodes large T antigen (T), while use of the proximal 5′ splice site generates an mRNA that encodes small t antigen (t). Although both splicing products are produced in all cell types, the ratio of small t to large T mRNA was shown to be 10- to 20-fold higher in adenovirus-transformed human embryonic kidney cells than in other mammalian cells (*16*). Recently, this cell specific difference in SV40 splicing was reproduced in vitro, and a complementation assay was used to purify ASF (*14*). Like SF2, ASF promotes the use of the proximal 5′ splice site. In fact, comparison of ASF and SF2 strongly suggests that they are the same protein or proteins. The two factors display similar chromatographic properties, and they have similar molecular sizes. In addition, ASF can complement SF2-deficient extracts, and SF2 can increase the t-T ratio (*17*).

These observations suggest that cell-specific differences in the concentrations or activities of general splicing factors could regulate alternative splicing. If this is the case, pre-mRNA's that are not alternatively spliced must withstand cell-specific fluctuations in the concentrations of splicing factors like SF2. Consistent with this

Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, MA 02138.

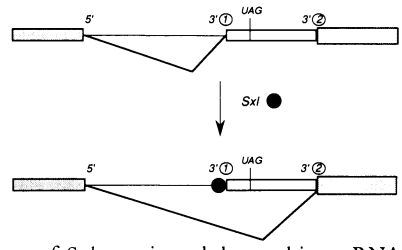

Fig. 1. Model for the negative control of *tra* pre-mRNA splicing by *Sxl* protein. The diagram shows the alternatively spliced regions of *tra* pre-mRNA. The pre-mRNA contains a single 5′ splice site and two competing 3′ splice sites. Splice site 1 is used in the absence of *Sxl* protein and the resulting mRNA contains a translational termination signal near the NH_2-terminus of the open reading frame. According to the model, *Sxl* protein (indicated by the black ball) binds to splice site 1 of the pre-mRNA, thereby switching splice site selection from site 1 to site 2. The use of site 2 produces an mRNA that encodes functional protein. The shaded boxes correspond to exons included in the functional mRNA, while the white box indicates an exon that is included in the nonfunctional mRNA.

possibility, high concentrations of SF2 stimulate the use of cryptic 5′ splice sites in vitro only when the normal 5′ splice site is weakened by mutation (*12*).

Regulation of alternative splicing can also be mediated by specialized proteins that control alternative splicing of specific pre-mRNA's (*2*). The first evidence for the existence of such proteins was provided by studies of the hierarchy of alternative splicing in the sex determination pathway of *Drosophila* (*4*). Mutations that knock out genes required for alternative splicing are not lethal, so they must not encode general splicing factors. Although these regulatory proteins could act by modifying the amounts or activities of general splicing factors, at least in one case, the protein appears to act by directly binding to the affected pre-mRNA (see below).

The *Drosophila* sex determination pathway is a cascade of alternative RNA processing events (*2, 4*). *Sex lethal* (*Sxl*), one of the earliest genes in the pathway, encodes a protein that regulates the alternative splicing of its own pre-mRNA, and of *transformer* (*tra*) pre-mRNA, the next gene in the pathway (*18*). Alternative splicing of both *Sxl* and *tra* pre-mRNAs involves competition between two 3′ splice sites for interaction with a common 5′ splice site (Fig. 1). Recent studies of *tra* pre-mRNA splicing in transgenic flies (*19*) and in transfected cells in culture (*20*) support the conclusion that the *Sxl* protein regulates 3′ splice site selection by binding specifically to one of the two competing 3′ splice sites. In males, no *Sxl* protein is produced, and splice site 1 is preferentially used. This leads to the production of mRNA that lacks a long open reading frame. In contrast, *Sxl* is produced in females and binds specifically to splice site 1. This RNA-protein interaction suppresses the use of splice site 1 and leads to the preferential use of splice site 2, generating an mRNA that encodes a functional protein. The ability of the *Sxl* protein to recognize a specific RNA sequence is most likely due to the presence of a consensus sequence motif shared by many RNA-binding proteins (*21*).

Drosophila P-transposase pre-mRNA contains three introns, two of which are removed in somatic cells to encode a transposition repressor (*22*). In germ cells, all three introns are removed to encode transposase. The germ cell–specific splicing of the third intron could be due to the presence of a specific splicing repressor in somatic cells, a splicing activator in germ cells, or both. Recent in vitro splicing studies provide evidence for the existence of an intron 3–specific splicing repressor in somatic cells that binds to the exon adjacent to the 5′ splice site (*23*).

Unlike introns 1 and 2 of *P*-transposase pre-mRNA, intron 3 is not spliced in extracts prepared from *Drosophila* somatic cells. However, all three introns are spliced in human HeLa cell nuclear extracts (*23*). Evidence that *Drosophila* somatic cells contain a repressor of intron 3 splicing was provided by the observation that splicing of this intron in HeLa cell extracts could be specifically blocked by the addition of *Drosophila* somatic cell extracts. The target of the putative repressor appears to be located within the exon, upstream from the 5′ splice site (*23*). A small RNA that contains the putative target sequence relieves the inhibition of splicing when added to in vitro extracts. This region of the exon is also required for regulated splicing in vivo (*24*).

Because the inhibitory activity is observed with extracts from cells that do not contain the *P*-element, the putative repressor must be encoded by a *Drosophila* gene. Therefore, the negative mechanism involved in transposase splicing is likely to be involved in regulating germline-specific genes in somatic cells (*23*).

Another example of negative regulation of splicing is provided by the suppressor of *white apricot* locus of *Drosphila* [$su(w^a)$] (*3*). In this case, the $su(w^a)$ gene product blocks the splicing of the first two introns of its own pre-mRNA. In the absence of this splicing event a nonfunctional mRNA is produced. The target of this splicing repressor is located near the branchpoint sequences of the affected introns (*25*).

Negative control of pre-mRNA splicing does not appear to be restricted to *Drosophila*. In yeast, splicing of ribosomal protein pre-mRNA's that contain a single intron is negatively autoregulated (*26*). In mammals, a sequence near a 3′ splice site in the calcitonin–calcitonin gene-related peptide (CGRP) pre-mRNA appears to inhibit the production of calcitonin transcripts in CGRP-producing cells (*27*). Similarly, a negative regulatory sequence located between the branchpoint sequence and the 3′ splice site of introns in the rat (*28*) and chicken (*29*) β-tropomyosin pre-mRNA may prevent the use of the adjacent exon in nonmuscle cells. Remarkably, there is as yet no example in which splice site selection is under positive control.

Understanding the mechanisms involved in the control of alternative splicing will require the development of in vitro systems that accurately reflect the situation in vivo. Although negative regulation of *P*-transposase mRNA can be detected in a mixed HeLa-*Drosophila* cell extract system (*22*), cell-specific regulation of splicing in vitro has yet to be achieved.

REFERENCES

1. A. R. Krainer and T. Maniatis, in *Frontiers in Molecular Biology: Transcription and Splicing*, B. D. Hames and D. M. Glover, Eds. (IRL Press, Oxford and Washington, DC, 1988), p. 131.
2. C. W. Smith, J. G. Patton, B. Nadal-Ginard, *Annu. Rev. Genet.* **23**, 527 (1989); M. McKeown, *Genet. Eng.* **12**, 139 (1990).
3. P. M. Bingham *et al.*, *Trends Genet.* **4**, 134 (1988).
4. B. S. Baker, *Nature* **340**, 521 (1989).
5. R. Reed and T. Maniatis, *Genes Dev.* **2**, 1268 (1989); Y. Zhuang, A. M. Goldstein, A. M. Weiner, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 2752 (1989).
6. R. Reed and T. Maniatis, *Cell* **46**, 681 (1986).
7. T. Maniatis and R. Reed, *Nature* **325**, 673 (1987); J. Steitz *et al.*, in *Structure and Function of Major and Minor Small Nuclear Ribonucleoprotein Particles*, M. L. Birnsteil, Ed. (Springer-Verlag, Berlin and Heidelberg, 1988), p. 115.
8. A. Bindereif and M. R. Green, *Genet. Eng.* **12**, 201 (1990).
9. R. Reed, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 8031 (1990).
10. C. A. Feener *et al.*, *Nature* **338**, 509 (1989).
11. A. Bernards *et al.*, *Mol. Cell. Biol.* **7**, 3231 (1987).
12. A. R. Krainer, G. C. Conway, D. Kozak, *Cell* **62**, 35 (1990).
13. ———, *Genes Dev.* **4**, 1158 (1990).
14. H. Ge and J. L. Manley, *Cell* **62**, 25 (1990).
15. A. R. Krainer and T. Maniatis, *ibid.* **42**, 725 (1985).
16. X. Y. Fu and J. L. Manley, *Mol. Cell. Biol.* **7**, 738 (1987).
17. A. Krainer and J. Manley, personal communication.
18. L. R. Bell *et al.*, *Cell* **55**, 1037 (1988).
19. B. A. Sosnowski, J. M. Belote, M. C. McKeown, *ibid.* **58**, 449 (1989).
20. K. Inoue *et al.*, *Nature* **344**, 461 (1990).
21. R. J. Bandziulis, M. S. Swanson, G. Dreyfuss, *Genes Dev.* **3**, 431 (1989).
22. F. A. Laski, D. C. Rio, G. M. Rubin, *Cell* **44**, 7 (1986); D. C. Rio, F. A. Laski, G. M. Rubin, *ibid.*, p. 21; H. M. Robertson and W. R. Engels, *Genetics* **123**, 815 (1989); S. Misra and D. C. Rio, *Cell* **62**, 269 (1990).
23. C. W. Siebel and D. C. Rio, *Science* **248**, 1200 (1990).
24. F. A. Laski and G. M. Rubin, *Genes Dev.* **3**, 720 (1989).
25. P. Bingham, personal communication.
26. M. D. Dabeva *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 5854 (1986).
27. R. B. Emeson *et al.*, *Nature* **341**, 76 (1989).
28. D. M. Helfman *et al.*, *Genes Dev.* **4**, 98 (1990)
29. M. Goux-Pelletan, *EMBO J.* **9**, 241 (1990).