Question of commitment

Jim Manley

INTRONS, or non-coding sequences, are removed from messenger RNA precursors in a complex splicing process. Not only is the reaction pathway complicated, involving such baroque features as RNA lariats and branchpoints, but the number of factors required to catalyse the reaction is enormous, involving multiple small nuclear ribonucleoprotein particles (snRNPs) and an unknown (but large) number of accessory proteins. Despite this complexity, there has been considerable progress over the past decade in understanding the details of the reaction. But many crucial questions remain, one of which concerns the nature of the very first steps in the splicing reaction. How is a particular pre-mRNA initially recognized and committed to splicing and how are intron boundaries (splice sites) identified? This issue is especially important in multicellular organisms, where most mRNA precursors contain numerous introns, and many are subject to complicated patterns of alternative splicing. On page 82 of this issue[1], Xiang-Dong Fu provides part of the answer to the commitment question, and it turns out to be surprisingly simple. It seems that the presence of a single protein, one of the family of SR proteins, is sufficient to enable the splicing machinery to select a particular RNA in an *in vitro* splicing reaction.

Fu finds that, in an *in vitro* system, one SR protein — SC35 — is sufficient to allow human β-globin pre-mRNA to be preferentially spliced when an excess of a competitor RNA and splicing extract were added subsequently (the operative definition of commitment), whereas another human SR protein — ASF/SF2 — cannot. With an HIV tat RNA splicing substrate, the situation was reversed.

These findings are surprising in the light of earlier studies, carried out first in yeast[2,3] and then in mammalian[4,5] systems, which suggested that the U1 snRNP particle has a key early role in commitment of a pre-mRNA to splicing. It is well established that U1 snRNP recognizes the 5′ splice site of the intron, but it also appears that this interaction by itself is relatively weak, suggesting the need for an additional factor(s). Indeed, the experiments showing a requirement for U1 snRNP in the formation of the earliest detectable splicing complexes by necessity used unfractionated cellular extracts and relied on the ability to separate and identify snRNP-containing complexes. It is certainly possible that a simple protein–RNA complex could have gone undetected.

SR proteins enter the picture from a somewhat different perspective. They constitute a group of at least six proteins that are highly conserved from the nematode *Caenorhabditis elegans* to humans[6]. (They seem not to exist in yeast, however, perhaps indicating a mechanistic difference in an early step in splicing.) The prototypical member of this family is the human protein ASF/SF2, which was identified independently by two distinct *in vitro* assays. In one[7], it was purified as an activity able to switch the utilization of two 5′ splice sites in an alternatively spliced pre-mRNA, thereby recapitulating *in vitro* a cell-specific pattern of splicing observed *in vivo*. In another[8], the protein was purified by its ability to activate splicing of a simple pre-mRNA in a depleted extract, indicating that ASF/SF2 can function as an essential, as well as an alternative, splicing factor.

The sequence of the protein revealed interesting features[9,10], which indeed seem to be common to all SR proteins, including the well studied essential splicing factor, SC35 (ref. 11). At the amino terminus is an RNP-type RNA-binding domain (or RNA-recognition motif), undoubtedly involved in binding the pre-mRNA, and at the carboxy terminus is a region (called the RS domain) that consists of repeating arginine and serine residues (hence the name SR proteins). These features are intriguing because they are reminiscent of genetically defined regulators of splicing in *Drosophila*[12]. RS domains have now been found in a number of splicing factors, and, indeed, the presence of such a region can probably be considered diagnostic of a protein involved in splicing. Although the precise function of RS domains remains to be established, an attractive model, for which there is already some support[13], is that they constitute activating regions, perhaps analogous to the acidic domains found in certain transcription factors. Indeed, it may be that splicing and transcription activators share a similar modular organization, each containing a domain that binds nucleic acid and a discrete activating region.

Initial work suggested that SR proteins might be functionally interchangeable[6,14], raising the possibility that their activities could be redundant *in vivo*. However, several recent studies, including the present work by Fu, detect differences in the behaviour of certain SR proteins. A *Drosophila* SR protein (Rbpl) and ASF/SF2 display significant differences in their ability to function as essential as well as alternative splicing factors[15], and several mammalian SR proteins also differ from each other in their effects on alternative splicing[16]. The activation of the female-specific intron in the *Drosophila doublesex* pre-mRNA by Transformer and Transformer-2 proteins also requires SR proteins, and this requirement can be met only by a subset of these proteins[17]. So the different SR proteins may have distinct roles *in vivo*, perhaps modulating the splicing of specific classes of transcripts. The idea that SR proteins may be involved in regulation is also supported by differences in their tissue distribution[16], the existence of alternatively spliced forms[7], and by the ability of an snRNP-associated protein kinase to phosphorylate serines in the ASF/SF2 RS domain[18].

The findings of Fu[1] are important not only because they suggest an early determinative role for SR proteins in the assembly of splicing complexes on pre-mRNA, but also because they add to the view that SR proteins perform distinct functions that may be involved in splicing control. The simplest explanation for the differential action of SC35 and ASF/SF2 is that the two proteins, and, by extension, other SR proteins, have distinct sequence preferences in their interaction with pre-mRNA.

Like many unexpected findings, Fu's results raise at least as many questions as they answer. For example, which RNA sequences are recognized by SR proteins, and do different proteins recognize different sequences? What do the proteins do after binding to the pre-mRNA? And do all SR proteins do the same thing? The answers to these questions will undoubtedly provide exciting insights into the mechanism and control of metazoan pre-mRNA splicing, but the important role of SR proteins in the earliest steps of the process now seems firmly established. □

Jim Manley is in the Department of Biological Sciences, Columbia University, New York, New York 10027, USA.

1. Fu, X.-D. *Nature* **365**, 82–85 (1993).
2. Seraphin, B. & Rosbash, M. *Cell* **59**, 349–358 (1989).
3. Ruby, S. R. & Abelson, J. *Science* **242**, 1028–1035 (1988).
4. Michaud, S. & Reed, R. *Genes Dev.* **5**, 2534–2546 (1991).
5. Jamison, S. F., Crow, A. & Garcia-Blanco, M. *Molec. cell. Biol.* **12**, 4279–4287 (1992).
6. Zahler, A. M., Lane, W. S., Stolk, L. A. & Roth, M. B. *Genes Dev.* **6**, 837–847 (1992).
7. Ge, H. & Manley, J. L. *Cell* **62**, 25–34 (1990).
8. Krainer, A. R., Conway, G. C. & Kozak, D. *Cell* **62**, 35–42 (1990).
9. Ge, H., Zuo, P. & Manley, J. L. *Cell* **66**, 372–382 (1991).
10. Krainer, A. R., Mayeda, A., Kozak, D. & Binns, G. *Cell* **66**, 383–394 (1991).
11. Fu, X.-D. & Maniatis, T. *Science* **256**, 535–538 (1992).
12. Baker, B. S. *Nature* **340**, 521–524 (1989).
13. Valcárcel, J., Singh, R., Zamore, P. D. & Green, M. R. *Nature* **362**, 171–175 (1993).
14. Fu, X.-D., Mayeda, A., Maniatis, T. & Krainer, A. R. *Proc. natn. Acad. Sci. U.S.A.* **89**, 11224–11228 (1992).
15. Kim, Y.-J., Zou, P., Manley, J. L. & Baker, B. S. *Genes Dev.* **6**, 2569–2579 (1992).
16. Zahler, A. M., Neugebaur, K. M., Lane, W. S. & Roth, M. B. *Science* **260**, 219–222 (1993).
17. Tian, M. & Maniatis, T. *Cell* **74**, 105–114 (1993).
18. Woppmann, A. *et al. Nucl. Acids Res.* **21**, 2815–2822 (1993).