Cell, Vol. 87, 357–359, November 1, 1996, Copyright ©1996 by Cell Press

Mixed Mechanisms in Yeast Pre-mRNA Splicing?

Minireview

Michael Rosbash*
Howard Hughes Medical Institute
Department of Biology
Brandeis University
Waltham, Massachusetts 02254

Does anyone doubt that Life Sciences—and life—would be more dull without genetics? New information often comes from an expected direction—intellectually as well as geographically. But genetics breaks the rules. A mutant screen can take a laboratory as well as its field in a new and unexpected direction, making strange bedfellows of diverse disciplines in the process. This is one of several lessons from two exciting papers, containing the results of two genetic screens, in this issue of *Cell* (Cox and Walter, 1996; Sidrauski et al., 1996). Who imagined that protein folding and chaperones would hold new surprises for pre-mRNA splicing?

Cross-disciplining begins with trying on the acronyms. The UPR (unfolded protein response) is triggered by unfolded proteins accumulating in the ER (endoplasmic reticulum). This generates a signal that increases the transcription of genes encoding ER-localized chaperones. Therefore, the concentration increase of unfolded proteins must be initially sensed in the ER and ultimately delivered to the nucleus, to activate transcription of target genes. These include the known folding components *KAR2* (or *BIP*), *PDI1*, and *FKB2* (McMillan et al., 1994; Shamu et al., 1994).

There were two known components of the UPR: the Ire1p protein and the UPRE (UPR element), a 22 base pair region of DNA. The former is an inessential kinase that spans the ER membrane, senses the increase in unfolded protein concentration, and relays the signal to outside the ER. Signaling probably involves oligomerization and trans-autophosphorylation (Shamu and Walter, 1996). The UPRE is the key cis-acting upstream region of DNA that activates transcription of target genes. Its binding protein, presumed to be part of the UPR, had not been identified.

The first genetic screen resulted in three important findings. The first is the identification of the UPR-relevant transcription factor, Hac1p. The authors made a very sensitive UPR-reporter gene with four copies of the UPRE upstream of *lacZ*. In a background deleted of *IRE1*, they searched for genes that when overexpressed would constitutively turn on the UPR pathway. Three genes were identified: *IRE1* itself (as expected), the transcription factor *SWI4*, and *HAC1*. The two new genes were both inessential, making phenotypic characterization of gene knockouts straightforward. The Δ*swi4* strain still contained a robust UPR, indicating that Swi4p is probably not an important pathway component. In contrast, the Δ*hac1* strain had no UPR. From sequence analysis, Hac1p had been proposed to be a member of the bZIP family of transcription factors. Indeed, extracts from the Δ*hac1* strain have no UPRE-binding activity by gel shift analysis, and complexes from an epitope-tagged Hac1p strain are supershifted by anti-epitope antibody. Therefore, Hac1p binds directly to the UPRE, and another chapter in this regulatory tale appears to be wrapped up in a satisfying fashion.

The second finding is that regulation of Hac1p synthesis occurs through activation of pre-mRNA splicing. As expected, Hac1p is only detectable under inducing conditions, but *HAC1* RNA is present under non-inducing as well as inducing conditions. Size analysis, primer extension, cDNA cloning, and sequencing indicate that 252 internal nucleotides are missing from the uninduced RNA: a single splicing product, *HAC1*^{i} mRNA, is present in induced cells. Splicing is sufficient to activate the UPR pathway, because replacement of *HAC1* with an intron-less version results in constitutive activation. Therefore the signal activates splicing of the *HAC1* pre-mRNA to generate *HAC1*^{i} mRNA. Conceptual translation indicates that splicing removes the 10 C-terminal amino acids from the contiguous reading frame in the DNA sequence and replaces them with a different 18 amino acid sequence encoded at the beginning of the 3′ exon. An antibody against the new C-terminal sequence confirms that haC1p is the product of *HAC1*^{i} mRNA. This constitutes evidence for a novel regulatory mechanism: signal transduction–mediated splicing.

But uninduced *HAC1* pre-mRNA is not pre-mRNA in the traditional sense. It is an mRNA (*HAC1*^{u} mRNA) and it is found on polysomes in similar fashion to the spliced version, *HAC1*^{i} mRNA. Although polysome-associated pre-mRNA is not unprecedented (He et al., 1993), it is unusual for its abundance to be comparable to that of mRNA. This implies that the turnover properties of *HAC1*^{u} mRNA are similar to those of *HAC1*^{i} mRNA. Taken together with the 3′ position of the intron and the coding potential of the two mRNAs, it suggests that both mRNAs synthesize similar amounts of the two proteins.

The third finding is the important role of protein degradation, which follows from the observations described above. Although the synthesis of Hac1p^{u} is implied rather than shown, the notion is that Hac1p^{u} is very unstable; this accounts for the absence of detectable protein in uninducing conditions. The splicing-dependent switch in C-termini would then stabilize the protein. It is the C-terminal tail of the Hac1p^{u} that is destabilizing, as a tailless version of Hac1p—missing the new 18 amino acids but also missing the old 10 amino acids—is fully functional. A role for protein degradation is also supported by two additional facts: the C-terminal half of the protein is rich in Pro, Glu, Asp, Ser, and Thr, the so-called PEST sequences that target proteins for ubiquitin-mediated proteolysis. More significant is experimental evidence for the involvement of the ubiquitin pathway in the degradation of the Hac1p^{u} product. The yeast *UBC* genes encode members of the ubiquitin conjugating enzymes. Mutants in different *UBC* genes activate the UPR pathway in the absence of inducing conditions and without bypassing *HAC1* function.

As a consequence of this work, the UPR pathway will attract the attention of three subdisciplines of biochemistry: transcription, RNA splicing, and protein degradation. But it is for splicing that there are real surprises with the promise of more to come.

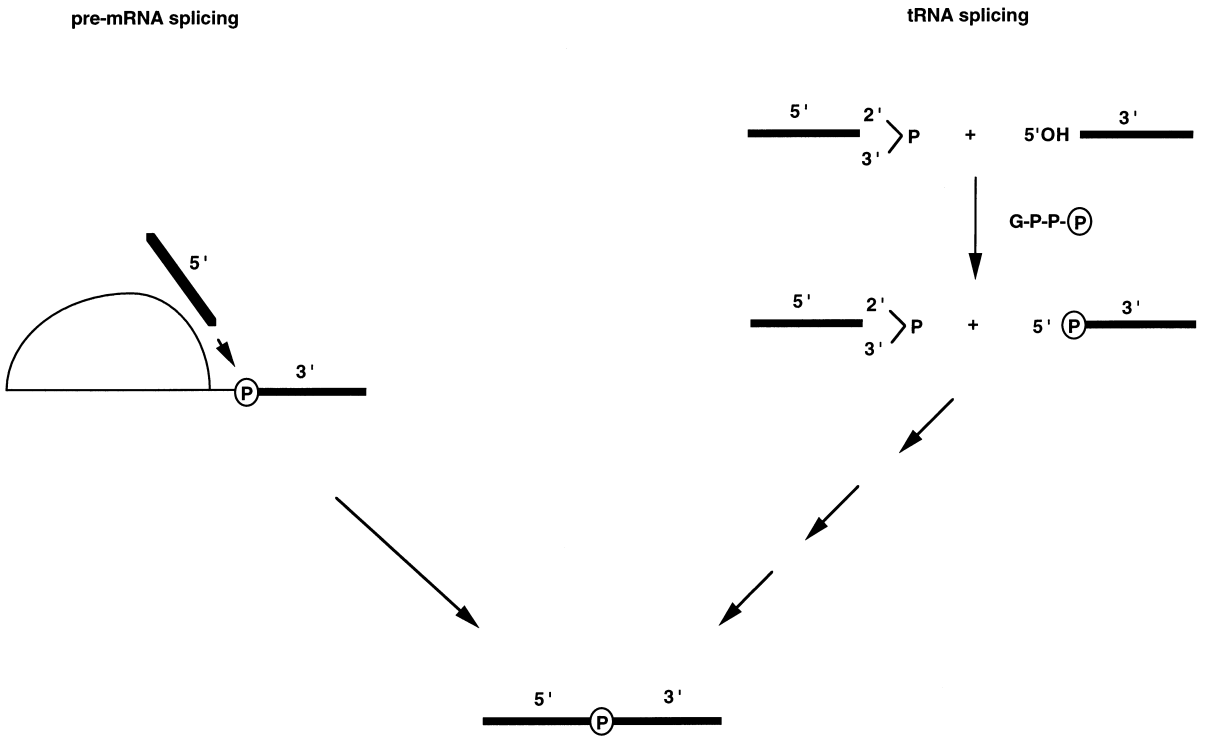

Figure 1. The Source of the Phosphate at the mRNA Exon–Exon Junction

This phosphate is circled. The 5′ and 3′ exons are labeled. On the left is depicted the second step of pre-mRNA splicing. The key phosphate is between the last nucleotide of the intron and the first nucleotide of the 3′ exon. The intron is in the form of a lariat-intermediate after the first step of splicing. The phosphodiester bond containing this phosphate is attacked by the 3′ end of the 5′ exon and retained in the mRNA product at the exon-exon junction. On the right (top) are depicted the two substrates of the multiple tRNA splicing kinase and ligase reactions (indicated by multiple arrows). The key phosphate is derived from an exogenous nucleotide, which serves as the source of phosphate added to the 5′ end of the 3′ exon. It is retained and becomes the phosphate at the mRNA exon–exon junction.

The suggestion of novel splicing features is immediately apparent from the intron sequence: the 5′ ss (splice site) and 3′ ss are unprecedented for yeast pre-mRNAs (Rymond and Rosbash, 1992). Although there is an ambiguity in the precise intron location, neither possibility has canonical GT and AG sequences at the 5′ and 3′ ends of the intron. But there is also a branchpoint-like sequence in a reasonable location close to the 3′ ss; although the sequence is very similar to the highly conserved yeast consensus (TACTAA*GA rather than TAC TAA*CA; the * follows the A residue that forms the branch), the G is unusual if not unique. These cis-acting sequence features are complemented by an unexpected splicing factor uncovered in the second genetic screen (Sidrauski et al., 1996).

By constructing a strain conditionally dependent on the UPR for survival, Sidrauski et al. (1996) were able to carry out a synthetic lethal screen for mutants unable to induce the pathway. This resulted in a curious finding: the identification of a missense mutation in the well studied tRNA ligase gene (*RLG1*; Apostol et al., 1991). Further characterization made it certain that this allele (*rlg1–100*) is sufficient to prevent induction of the pathway and that this is due to the absence of spliced mRNA. The RNA analysis makes the story even more intriguing, as activation of the pathway in the mutant background not only inhibits splicing but also causes the disappearance of pre-mRNA, i.e., the *HAC1*^{u} mRNA species. The suggestion is that the tRNA ligase gene functions to ligate the two half molecules subsequent to endonucleolytic cleavage (see Figure 1, right panel); in the absence of functional or adequate ligase activity, the half molecule substrates are rapidly degraded. The provocative speculation is that a tRNA-processing-like mechanism splices a pre-mRNA.

Control experiments address two related issues. First, as the tRNA ligase mutant allele has no effect on RNA levels from the intronless gene, the absence of intron sequences per se does not lead to rapid degradation of *HAC1*^{i} mRNA. Second, the *rlg1–100* allele has no effect on tRNA splicing. Because the *RLG1* gene and tRNA splicing are essential for viability, this suggests that the missense allele affects *HAC1*^{u} splicing specifically: either *rlg1–100* is a weak loss of function allele and the UPR pathway is more sensitive, or the mutation affects the protein in a pathway-specific fashion.

There is certainly something unusual going on here, because *HAC1*^{u} splicing is also insensitive to two classical PRP (pre-mRNA processing) temperature-sensitive (ts) mutants. *PRP2* encodes an RNA-dependent ATPase that is an essential pre-mRNA splicing factor. *PRP8* encodes a component of U5 snRNP, which is also essential for splicing. Both proteins are necessary prior to the first catalytic step, 5′ splice site cleavage and lariat formation. Heating the two ts strains inhibits the splicing of all known pre-mRNA splicing substrates (Rymond

and Rosbash, 1992). But the levels of *HAC1* mRNA are unaffected, consistent with the notion that it is produced through a novel splicing pathway. The half-life of the spliced mRNA is about 20 minutes, indicating that the wildtype-like mRNA levels are not due to preexisting mRNA that persists stably during the 2–3 hour incubation at the non-permissive temperature. So it would seem that the genetics is also providing us with strange biochemical bedfellows: no effect from inactivation of the conventional pre-mRNA splicing factors Prp8p and Prp2p; a pre-mRNA with an unusual but semi-conventional branchpoint sequence; and a role for tRNA ligase.

It is possible that the effect of the single tRNA ligase allele *rlg1–100* is indirect and reflects an epiphenomenon unrelated to *HAC1* RNA ligation. Based on this reasoning, the CC-CG (or the alternative, GC-CC) 5′ and 3′ splice site sequences and the branchpoint sequence might presage a situation not unlike the recently described AT-AC intron family in higher eukaryotes (Hall and Padgett, 1994; Hall and Padgett, 1996; Tarn and Steitz, 1996a,b). The analogy predicts a classical pre-mRNA splicing mechanism and a dedicated set of splicing factors for *HAC1* expression. It also predicts the existence of *PRP8*- and *PRP2*-like genes, which would explain the absence of a splicing effect in the classic ts mutant strains. In addition to ignoring tRNA ligase, the hypothesis has an additional liability: the sequenced yeast genome does not have obvious *PRP8*- and *PRP2*-like genes. They could of course be functionally related to *PRP8* and *PRP2* without looking very similar at the primary sequence level.

A more radical mixed mechanism is required to integrate a catalytic role for tRNA ligase with the TACTAAC-like sequence. The classic transesterification mechanism that governs pre-mRNA splicing makes 5′ splice site cleavage without coupled branchpoint formation, and 3′ splice site cleavage without coupled ligation, an impossibility; the 3′ exon half-molecule cannot be a reaction intermediate (Moore et al., 1993). But there are RNA endonucleases that, like the tRNA endonuclease, generate a 2′-3′ cyclic phosphate and a 5′OH ends (Symons, 1989). There are also mutants of group I and group II self-splicing introns in which the nucleophile is water or hydroxide ion rather than a nucleotide 2′OH or 3′OH (e.g., Van der Veen et al., 1987). A 3′ splice site cleavage product generated in this manner could be further processed to become a suitable substrate for tRNA ligase. It is therefore not impossible that *HAC1* RNA cleavage takes place by some novel hybrid mechanism that utilizes RNA-based cleavage at least for the 3′ splice site and a protein enzyme for exon ligation (Phizicky and Greer, 1993). Characterization of the rapidly degraded molecules might provide evidence for branch formation, which would clarify the mechanism of 5′ splice site cleavage.

The third possibility, the most straightforward suggestion from the discovery of the tRNA ligase mutant, is that *HAC1* RNA splicing occurs via a simple two enzyme endonuclease-ligase tRNA-processing-like mechanism (Phizicky and Greer, 1993). There may still be some contribution from constitutive pre-mRNA splicing factors—a protein that recognizes the branchpoint sequence for example—but this would not involve RNA catalysis. For example, a PRP protein could contribute to nuclear retention of the *HAC1* primary transcript, thus preventing cytoplasmic transport prior to nuclear splicing. Testing known pre-mRNA splicing components in addition to Prp2p and Prp8p might identify relevant protein factors. Alternatively, the apparent branchpoint conservation may be fortuitous.

One of these mechanistic possibilities might emerge with further application of the holy trinity of modern biology: molecular biology, biochemistry, and genetics. Mutagenesis of the intron should indicate the importance of the splice sites and putative branchpoint sequences. It might also be informative to examine the *HAC1* intron sequence in closely related yeast species to examine the phylogenetic conservation of intron sequence and structure. Is there something that resembles tRNA? Does the enzyme recognize exon or intron elements? Biochemistry should provide a definitive distinction between a pre-mRNA-like mechanism and a tRNA-like mechanism. For example, an in vitro system that can accurately splice *HAC1* pre-mRNA should identify the source of the phosphate at the exon–exon junction of the final mRNA product. In tRNA splicing, the phosphate is derived from a nucleotide cofactor; in pre-mRNA splicing, it comes from the 3′ splice site (see Figure 1). Finally, further genetic screens should identify additional splicing components if there are more to be found. Of course, we should not be surprised if genetics continues to take this story in unexpected directions.

Selected Reading

Apostol, B.L., Westaway, S.K., Abelson, J., and Greer, C.L. (1991). J. Biol. Chem. *266*, 7445–7455.

Cox, J.S., and Walter, P. (1996). Cell *87*, this issue.

Hall, J.C., and Padgett, R.A. (1994). J. Mol. Biol. *239*, 357–365.

Hall, S.L., and Padgett, R.A. (1996). Science *271*, 1716–1718.

He, F., Peltz, S.W., Donahue, J.L., Rosbash, M., and Jacobson, A. (1993). Proc. Natl. Acad. Sci. USA *90*, 7034–7038.

McMillan, D.R., Gething, M.-J., and Sambrook, J. (1994). Curr. Opin. Biotechnol. *5*, 540–545.

Moore, M.J., Query, C.C., and Sharp, P.A. (1993). The RNA World. R.F. Gesteland and J. F. Atkins, eds. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press), pp. 303–357.

Phizicky, E.M., and Greer, C.L. (1993). Trends Biol. Sci. *18*, 31–34.

Rymond, B.C., and Rosbash, M. (1992). The Molecular and Cellular Biology of the Yeast Saccharomyces: Gene Expression. E. W. Jones, J.R. Pringle, and J. R. Broach, eds. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press), pp. 143–192.

Shamu, C.E., Cox, J.S., and Walter, P. (1994). Trends Cell Biol. *4*, 56–60.

Shamu, C.E., and Walter, P. (1996). EMBO J. *15*, 3028–3039.

Sidrauski, C., Cox, J.S., and Walter, P. (1996). Cell *87*, this issue.

Symons, R.H. (1989). Trends Biochem. Sci. *14*, 445–450.

Tarn, W.-Y., and Steitz, J.A. (1996a). Cell *84*, 801–811.

Tarn, W.-Y., and Steitz, J.A. (1996b). Science *273*, 1824–1832.

Van der Veen, R., Kwakman, J.H., and Grivell, L.A. (1987). EMBO J. *6*, 3827–3831.