In eukaryotic cells, RNA processing is physically separate from protein synthesis. As nuclear export of unspliced RNA is restricted, only mature messages are normally exposed to the translational machinery. In bacteria, however, splicing must be coordinated with the translation of nascent transcripts. These two processes make very different demands on the RNA substrate: Splicing of autocatalytic introns requires that the 5' and 3' splice sites be brought together as part of an elaborate tertiary structure, whereas translation requires that the mRNA be relatively free of secondary structure. Nonetheless, introns have been found in highly expressed genes in eubacteria, bacteriophages, mitochondria, and chloroplasts (for review, see Burke 1988). Clearly, there must be some means of balancing splicing of bacterial introns with cotranscriptional translation. In this issue, Semrad and Schroeder (1998) provide the surprising answer that splicing of a group I intron from phase T4 is facilitated by translation of the upstream open reading frame (ORF). This enhancement of splicing is achieved by modulating the long-range conformation of the pre-mRNA. Their results provide useful analogies for the coupling of eukaryotic pre-mRNA splicing with transcription.

Bacterial mRNAs exclusively contain group I or group II introns, and the three group I introns that are present in phase T4 are all able to self-splice in vitro (for review, see Belfort 1990). The introns are found in genes encoding thymidylate synthase (td), ribonucleotide reductase (nrdB) (Belfort 1990), and anaerobic ribonucleotide reductase (sunY, or nrdD) (Young et al. 1994). In addition to sequences that provide the necessary functions of self-splicing, td and sunY trans contain internal ORFs that encode double-stranded DNA endonucleases (Belfort 1989). The endonucleases trigger homing, or site-specific movement of the intron sequences to intronless alleles.

Self-splicing requires that the intron RNA fold into a unique secondary and tertiary structure (Cech and Herschlag 1996). The central core of this structure is highly conserved among group I introns and contains the active site where the transfer of phosphodiester bonds takes place. A helix containing the 5' splice site docks into the active site via hydrogen bonds with its ribose 2' hydroxyl groups (Pyle et al. 1992). Recognition of the 3' splice involves several weak interactions, including a 2-bp stem between the 3' end of the intron and nucleotides in the intron core (Burke et al. 1990). The folded structure of the RNA depends on coordination of magnesium ions, which are required for self-splicing (Cech and Herschlag 1996). Long-range interactions, such as base-pairing between hairpin loops, or tetraloop–helix receptor interactions, also stabilize the tertiary structure of the catalytic core (for review, see Brion and Westhof 1997).

Recent experiments on the folding kinetics of group I introns, as well as experiments carried out in the 1970s on tRNA, have begun to tease out the mechanisms by which RNAs reach a biologically active conformation (for review, see Draper 1996; Brion and Westhof 1997). Small hairpins can form in 10–100 µsec, and tRNAs can fold within milliseconds (Draper 1996). In contrast, larger RNAs fold in stages over much longer periods of time. The tertiary structure of the P4–P6 domain of the Tetrahymena group I intron, which can fold independently, appears in a few seconds (Sclavi et al. 1997). The core of the intron, however, takes several minutes to become completely folded (Zarrinkar and Williamson 1994; Banerjee and Turner 1995).

The very slow folding of longer RNAs arises, in part, from their tendency to form many alternative secondary structures. As RNA secondary structure is stable, incorrect base pairs have the potential to trap the molecule in inactive conformations that can persist for relatively long periods of time (for review, see Herschlag 1995; Thirumalai and Woodson 1996; Brion and Westhof 1997). These metastable states may be quite dissimilar to the final structure. Therefore, the folding process itself can result in RNA populations with different levels of biological activity. Misfolding of RNA has been shown to inhibit ribosome activity and spliceosome assembly (Goguel et al. 1993; Zavanelli et al. 1994; Uhlenbeck 1995). As discussed below, competition between metastable RNA conformations can also serve as a normal mechanism by which gene activity is regulated.

Estimates of in vivo splicing rates are 10- to 50-fold faster than in vitro self-splicing (Brehm and Cech 1983; Zhang et al. 1995). This disjunction between in vitro and in vivo activity of catalytic RNAs implies that kinetic folding traps are normally overcome in the cell.
that facilitate splicing of group I and group II introns in fungal mitochondria were first identified as the result of splicing-deficient mutations (Gampel et al. 1989; Lambowitz and Perlman 1990). Biochemical experiments have established that these proteins promote splicing by stabilizing the correctly folded structure of the intron RNA. Neurospora crassa CYT-18 mitochondrial tRNA synthetase binds to a conserved region of the catalytic core and maintains the alignment of two double helices that form part of the active site (Caprara et al. 1996). The yeast protein CBP2 not only stabilizes the core of the intron, but also promotes docking of the splice site helix with the catalytic core (Weeks and Cech 1995).

Another class of RNA-binding proteins accelerates the resolution of misfolded RNA structures under certain conditions in vitro, in a manner analogous to protein chaperones (Herschlag 1995). These proteins, which include hnRNP A1 and HIV nucleocapsid protein, preferentially bind single-stranded RNA with low sequence specificity and promote nonspecific reannealing of RNA duplexes (Herschlag 1995). Ribosomal protein S12 enhances an intermolecular splicing reaction of a phage T4 intron and increases the turnover rate of a hammerhead ribozyme (Coetzee et al. 1994). It is not yet clear whether any of these proteins function as true RNA chaperones in vivo. However, a cold shock protein from Escherichia coli has also been shown to have RNA chaperone activity in vitro (Jiang et al. 1997), suggesting that its function is to compensate for overstabilization of RNA secondary structures at low temperatures.

The fact that introns are found in the highly streamlined genomes of bacteriophages is somewhat surprising, especially as intronless phage suffer no loss in viability (Belfort 1989). In part, their persistence in the T4 genome is ensured by the two homing endonucleases. However, the presence of related introns in other T-even phages, and coordination of splicing and expression of the homing endonucleases with the phage life cycle, suggest a long period of coadaptation (Belfort 1990). What factors, then, facilitate splicing of phage T4 pre-mRNA?

T4 td is expressed early in infection from a promoter upstream of the external ORF (Belfort 1990). Because the intron interrupts the reading frame, splicing must occur before translation to produce full-length td protein (Fig. 1). However, protein synthesis in E. coli begins shortly after transcription. As the td intron is 1016 bases long, ribosomes reach the 5' splice site long before the 3' splice site is transcribed. Movement of ribosomes into the interior of the intron inhibits splicing by disrupting the folded structure of the RNA (O¨hman-Hede ´n 1993). In td, this is prevented by a stop codon just after the 5' splice site (Belfort 1989). Termination of protein synthesis at this stop codon produces a truncated protein (NH2–TS) with no known function.

Clearly, expression of td requires coordination among transcription, splicing, and translation. To address this issue, Semrad and Schroeder began by setting out to investigate another phenomenon, namely, the inhibition of group I splicing by aminoglycoside antibiotics (von Ahsen et al. 1991). To uncouple translation of td pre-mRNA from splicing, they introduced stop codons upstream of the 5' splice site. Surprisingly, they found that splicing in vivo was reduced to nearly undetectable levels, even in the absence of antibiotics (Semrad and Schroeder 1998). Splicing was restored partially by a suppressor tRNA, as judged by both analysis of td RNA and the ability of strains to grow on minimal media lacking thymine. These observations suggested that translation itself was required for efficient splicing in vivo.

Reduced splicing of transcripts containing upstream stop codons was not a result of a change in mRNA stability, as stop codons had no effect on levels of td mRNA lacking the intron. Neither was the peptide product required for in vivo splicing, as tandem frameshift mutations that alter the peptide sequence while maintaining the continuity of the ORF did not inhibit splicing. What did matter, however, was the position of the stop codon. In general, splicing levels increased when stop codons were placed close to the 5' splice site, and decreased when they were introduced farther upstream.

The observation that splicing activity varied with the position of the stop codon indicated that a folding defect in the pre-mRNA could be responsible for inhibition. Moreover, the requirement for translation suggested that ribosomes altered the RNA folding pattern. To test this idea, proteins that are known to enhance splicing of the td intron were overexpressed in strains carrying the td gene constructs. Splicing of pre-mRNAs containing upstream stop codons were improved by overexpression of both S12, which enhances splicing in a nonspecific manner (Coetzee et al. 1994), and CYT-18, which specifically stabilizes the folded structure of the core (Mohr et al. 1994).
These results are consistent with a structural defect in the pre-mRNA and not with active site inhibition. Comparison of td mutations revealed that stop codons >80 nucleotides upstream of the 5′ splice site had the most deleterious effects on splicing. Strikingly, positions −81 to −73 in the 5′ exon are complementary to 9 bases at the 3′ end of the intron (Fig. 1). These nucleotides normally form helices within the intron (P9.0a and P9.0b) that stabilize the catalytic core and permit 3′ splice site recognition (Jaeger et al. 1993). As a result, premature base-pairing of the intron terminus with the 5′ exon would inhibit self-splicing. Consistent with this model, splicing of pre-mRNA containing a stop codon at −82 is increased from <1% to 60% of wild-type levels by mutations in the 5′ exon that destroy complementarity with the intron (Semrad and Schroeder 1998).

Semrad and Schroeder propose a model in which ribosomes unfold the 5′ exon RNA as they move down the nascent transcript, terminating at the stop codon just after the 5′ splice site (Fig. 1). As transcription continues, the intron begins to fold into the tertiary structure required for self-splicing. Ribosomes upstream prevent inappropriate base-pairing between the 3′ end of the intron and the 5′ exon that could compete with productive interactions in the intron. Splicing removes the stop at the 5′ end of the intron, permitting translation of full-length thymidylate synthase to begin.

Several observations raise the question of whether translation and splicing are even more tightly coordinated than this model would suggest (Semrad and Schroeder 1998). It is not known whether ribosomes immediately dissociate upon reaching the stop codon at the 5′ end of the intron, or whether they remain temporarily associated with the pre-mRNA. However, the failure of mutations between −81 and −73 to restore splicing to wild-type levels, and the fact that stop codons close to the 5′ splice site also affect splicing, lead Semrad and Schroeder to suggest that the ribosome also stabilizes the intron structure. This could occur as the result of direct interactions between the ribosome and the folded intron RNA.

There are several reasons to think that group I introns can bind to ribosomes. First, one face of the intron resembles the L-shape of tRNA and the td intron competes with tRNA for binding to CYT-18 tRNA synthetase (Caprara et al. 1996, and references therein). Second, the Tetrahymena group I intron becomes tightly associated with 50S ribosomes when expressed in E. coli (Nikolcheva and Woodson 1997) and is able to integrate into its splice junction in 23S rRNA by a reversal of the self-splicing reaction (Roman and Woodson 1995). The integration sites are concentrated in regions of the 23S rRNA that interact with tRNA and elongation factors. Preliminary results show that the td intron is also bound by E. coli ribosomes (R. Schroeder, pers. comm.). Finally, E. coli 70S ribosomes can accommodate a variety of RNA structures, including tSarA RNA, which has the properties of both tRNA and mRNA (Keiler et al. 1996). Within T4, gene 60 contains an intervening sequence that is not removed from the message but is instead precisely skipped during translation (Atkins et al. 1990). This ribosome “hop” requires a specific structure in the mRNA. Thus, ribosomes parallel the actions of RNA binding proteins. They not only resolve misfolded structures but may also stabilize certain folded conformations, perhaps by bringing them into tRNA sites.

The work of Semrad and Schroeder provides an example in which in vivo splicing depends on the exon context. In td, splicing is inhibited by improper base-pairing between the 5′ exon and the 3′ end of the intron. Self-splicing of Tetrahymena pre-rRNA is attenuated by a conserved RNA hairpin in the 5′ exon that competes with the 5′ splice site helix (Woodson and Cech 1991). The equilibrium between these alternative conformations must be balanced to ensure splicing of the pre-rRNA, and refolding of the rRNA into its conserved secondary structure after the intron is removed. This balance is maintained partly by a second alternative hairpin in the 5′ exon that stabilizes the active form of the pre-rRNA (Woodson and Emerick 1993) in a manner similar to translational attenuation (see below). In other cases, exon structure promotes self-splicing. The structure of the tRNA exons enhances self-splicing of group I pre-rRNA in Anabaena (Zaug et al. 1993). Exon secondary structure has also been found to inhibit splicing of nuclear pre-mRNAs and a mitochondrial group II intron in yeast (Séraphin et al. 1988; Libri et al. 1995).

An interesting question is whether the long-range pairing in td pre-mRNA has any regulatory function in phase T4. One possibility is that this pairing is merely fortuitous. It has been tolerated during T4 evolution simply because upstream ribosomes normally prevent it from interfering with splicing. The other possibility is that it provides an escape mechanism for expression of the homing endonuclease. Under normal conditions, translation of the internal ORF is repressed by a stem–loop in the td pre-mRNA (Gott et al. 1988). The ORF is expressed from an internal late promoter within the intron; the stem–loop cannot form in transcripts from the late promoter (Gott et al. 1988). If T4 infects a cell in which the translational capacity is reduced, unspliced td RNA is likely to accumulate. Perhaps under these conditions the endonuclease can be translated from the misfolded pre-mRNA, stimulating recombination among T4 genes (Belfort 1990).

The idea that ribosomes alter the equilibrium between alternative states of the mRNA reaches back to the classic discovery of translational attenuation and antitermination of the Trp operon in E. coli (Yanofsky 1981). When tryptophan is abundant, ribosomes read through Trp codons in the 5′ leader and prevent folding of an antiterminator stem–loop, leading to transcription termination. In low tryptophan, the antiterminator stem prevents termination by sequestering nucleotides that form part of the downstream terminator stem–loop. In Bacillus subtilis, an unusual RNA-binding protein, TRAP, takes the place of the ribosome (Gollnick 1994). TRAP binds the RNA and disrupts the antiterminator stem in a tryptophan-dependent manner. There are now
many other examples, such as the 54 operon in E. coli (Spedding and Draper 1993), in which translation initiation is controlled by specific structures in the mRNA.

Many examples of translational and antisense regulation in bacteria have made it clear that not only the equilibrium between two structures, but also the kinetics of RNA folding can be effectively exploited to control gene expression (e.g., Ma et al. 1994; Gultyaev et al. 1995; Poot et al. 1997). Translation of the M52 coat protein is repressed by a cloverleaf secondary structure in the 5′ leader of the mRNA that renders the ribosome binding site inaccessible (Poot et al. 1997). Translation can occur on nascent transcripts but not full-length mRNA. Delayed folding of the cloverleaf structure after transcription provides a window of time in which ribosomes can bind (Poot et al. 1997). Deletions that reduce the cloverleaf to a small hairpin completely suppress translation, whereas insertions that increase the size of the cloverleaf remove all inhibition.

Antisense regulation frequently relies on the transition between a complex set of alternative conformations (Gultyaev et al. 1995 and references therein). Again, the final structure of the target RNA is often determined by the pathway of folding during transcription (Gultyaev et al. 1995). In the hok killer gene of plasmid R1, recently it has been shown that the appearance of mRNA that is capable of binding either antisense sok RNA or ribosomes is delayed by slow equilibration of metastable intermediates (Franch et al. 1997; Gultyaev et al. 1997). These intermediate structures are well supported by phylogenetic comparisons and thus appear to be functionally important for regulation of toxin genes.

In eukaryotes, moderately stable RNA structures provide a timing mechanism for the binding of ribosomes or gene regulatory factors to nascent transcripts. In eukaryotes, a similar principle may fine tune nuclear splice-site selection. First, there are several examples in which pre-mRNA structure regulates the usage of splice sites (Clouet-D’Orval et al. 1991; Eng and Warner 1991; Libri et al. 1991). Second, splicing and 3′-end processing occur cotranscriptionally and are coordinated by interactions between splicing and polyadenylation factors and the polymerase II elongation complex (Corden and Patturajan 1997; McCracken et al. 1997). Third, nearly every stage of splicing involves conformational rearrangements in the pre-mRNA and small nuclear RNAs that make up the splicing complex (Staley and Guthrie 1998). These rearrangements are catalyzed by proteins from a superfamily of ATPases, which include RNA and DNA helicases (Staley and Guthrie 1998). Our lessons from bacteria suggest that interconversion among metastable RNA conformations provides a simple mechanism for framing the temporal response to available splicing factors in eukaryotes.

Acknowledgments
I thank Renee Schroeder, Marlene Belfort, and Stephen Mount for comments on the manuscript, and Katharina Semrad for help with Figure 1.

References


Ribosomes as group I splicing factors


Ironing out the kinks: splicing and translation in bacteria

Sarah A. Woodson

Genes Dev. 1998, 12:

References

This article cites 53 articles, 20 of which can be accessed free at: http://genesdev.cshlp.org/content/12/9/1243.full.html#ref-list-1

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.