Regulation of alternative splicing in cell nuclei

Department of
Organic Chemistry

Prof. Joseph Sperling

Yael Hacham, Eyal Kamhi

☎ 972 8 934 2509, 2182
FAX 972 8 934 4142
@ j.sperling@weizmann.ac.il
www.weizmann.ac.il/Organic_Chemistry/sperling.shtml

Most eukaryotic pre-mRNAs contain intervening sequences (introns) that must be removed in order to place the coding sequences (exons) in a protein-reading frame. The precise removal of introns from pre-mRNAs by the pre-mRNA splicing machine is, therefore, an essential step in the regulation of gene expression. The human genome project revealed that alternative splicing and pre-mRNA processing are major contributors to the functional diversity of products encoded in the human genome. It has been estimated that 40-75% of the ~25,000 human genes undergo alternative splicing, thereby increasing the coding potential of the human genome by more than one order of magnitude. It is therefore not surprising that changes in splice-site selection and alternative splicing are frequently associated with human pathologies and cancer.

The mechanism of the chemical transformations involved in pre-mRNA splicing, has been extensively worked out based on studies in vitro, mainly of pre-mRNAs composed of two exons and an intervening intron. However, the regulation of splicing and alternative splicing of the multi-intronic pre-mRNAs transcribed in vivo is still not well understood, particularly from the standpoint of the structure of the supraspliceosome – the nuclear macromolecular machine that catalyzes the splicing reactions. To better understand the complex regulation of alternative splicing we employed, in collaboration with the group of Prof. Ruth Sperling at the Hebrew University, a combination of functional, biochemical and structural analyses of the supra-spliceosome complex isolated from living mammalian cells.

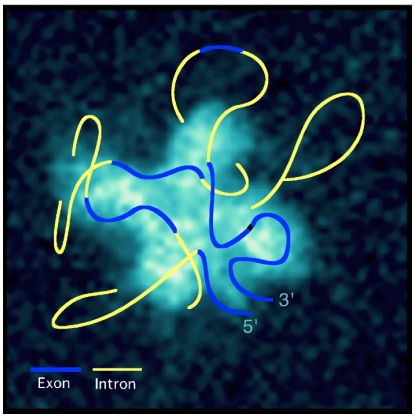

Fig. 1 *The supraspliceosome model. The model is based on our EM structural studies (Medalia et al., 1997; Sperling et al., 1997; Müller et al., 1998). Lines representing a pre-mRNA undergoing processing (exons in blue; introns in yellow) have been superimposed on the STEM image of an unstained supraspliceosome to illustrate the proposed model. The image itself clearly reveals the four subcomplexes, which are proposed to form a frame allowing the juxtaposition of exon ends about to be spliced, while introns are looped around and out of each of the respective spliceosome cores. This configuration can account for alternative splicing (e.g., a skipped exon is flanked by two introns looped out from the top subunit); and for the processive, but not sequential (5' to 3'), removal of introns from multi-intron pre-mRNAs (e.g., splicing of the intron nearest to the 3' end occurs on the far right subunit while upstream unspliced introns are still bound to the particle). (From Müller et al., 1998).*

The supraspliceosome structure

The supraspliceosome is one of the largest protein-RNA complexes of which structural analysis has thus far been attempted. To study this huge splicing machine we took a top-down approach. Namely, we isolated the splicing machine from nuclei of living cells and initiated the analyses of its composition, structure and function.

To develop a protocol for the isolation of supraspliceosomes from mammalian cell nuclei under physiological conditions, we optimized conditions whereby over 85% of ^{3}H-labeled RNA polymerase II (pol II) transcripts could be released to the nuclear supernatant after mild sonication of the nuclei and removing the chromatin. Subsequent fractionation of such nuclear supernatants in sucrose gradients revealed that the general population of labeled pol II transcripts, including several specific pre-mRNA transcripts, sedimented at the 200S region of the gradient (Sperling et al., 1985; Spann et al., 1989; Sperling and Sperling, 1998). Visualization by EM of aliquots from fractions across the gradient revealed a peak of large tetrameric structures, having overall dimensions of 50 x 50 x 35 nm (Spann et al., 1989). Immunoprecipitation experiments, using antibodies against splicing factors, showed that all five spliceosomal U snRNPs are integral components of the supraspliceosome (Sperling and Sperling, 1990; Miriami et al., 1995; Yitzhaki et al., 1996; Sperling and Sperling, 1998). Similar experiments showed that all known protein splicing factors, including the non-snRNP protein splicing factor U2AF and all members of the SR protein family, are also integral components of supraspliceosomes (Sperling and Sperling, 1990; Miriami et al., 1995; Yitzhaki et al., 1996; Sperling and Sperling, 1998). A remarkable feature of supraspliceosomes is that they package pre-mRNA transcripts of different sizes and of different number of introns into complexes of a unique size and hydrodynamic properties, indicating their universal nature.

Structural studies of the supraspliceosome

The complexity and large size of the supraspliceosome makes EM the method of choice for its structural analysis. Three-dimensional (3-D) image reconstruction of individual supraspliceosomes by automated electron tomography of negatively stained (Medalia et al., 1997; Sperling et al., 1997), and of frozen hydrated complexes (Medalia et al., 2002) showed the supraspliceosome as forming a closed structure, composed mainly of four similar subcomplexes. Mass measurements by scanning transmission EM (STEM) showed that the supraspliceosome has a mass of 21 MDa, whereas each of its main

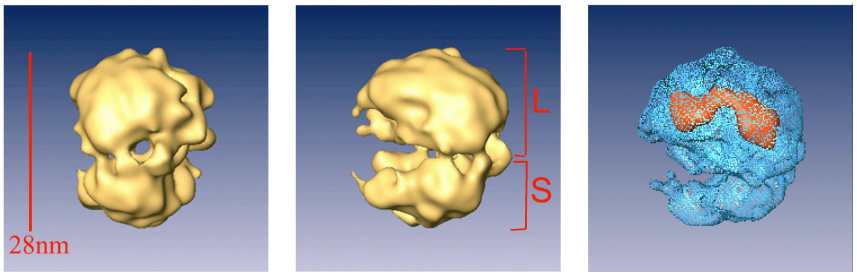

Fig. 2 *Surface representation of the 3-D structure of the native spliceosome at 20 Å resolution. Left and middle panels: Two views of the native spliceosome, determined by cryo-EM single particle technique, are shown separated by counterclockwise rotation of 60º about the central axis. The large (L) and small (S) subunits are indicated. Right panel: High-threshold rendering (red surface) shows the high-density mass region, which represents the stable RNAs (presumably the five snRNPs—penta snRNP) within the large subunit of the native spliceosome (from Azubel et al., 2004).*

subcomplexes has a mass of 4.8 MDa (Müller et al., 1998). Using a positive staining protocol, which allowed us the visualization of nucleic acids, we could show strands and loops of RNA emanating from positively stained supraspliceosomes (Müller et al., 1998), indicating that the RNA is loosely bound and therefore accessible for probing. Two-dimensional image restoration of ice-embedded particles revealed new high-resolution features, such as holes within the subcomplexes and fibers, presumably the pre-mRNA covered with proteins, through which the subcomplexes are interconnected (Medalia et al., 2002). These observations support our working hypothesis that each supraspliceosome packs one pre-mRNA and can therefore be responsible for its post-transcriptional processing. In this model (Figure 1), the supraspliceosome provides a platform to juxtapose exons about to be spliced, and each of the four major subcomplexes represents a spliceosome that can splice the intron wound around it. The supraspliceosome complex provides simultaneous multiple interactions for the assembly of the splicing machine with an intron, likely resulting in better efficiency and specificity of recognition of large introns. Furthermore, it should provide better specificity and efficiency of recognition of multi-intronic pre-mRNAs, which constitute the majority of the pre-mRNA population (Sperling et al., 1997; Müller et al., 1998). Thus, the supraspliceosome configuration is compatible with the requirements for alternative splicing and with the fact that the splicing of multi-intronic pre-mRNAs does not occur in a sequential manner.

Structural studies of native spliceosomes

The four substructures of the supraspliceosome are interconnected in a flexible way and may thus adopt different angular settings, which impose a significant restriction on reaching high resolution in EM image analyses. We have therefore developed a methodology to prepare and isolate the monomeric spliceosomal subcomplexes (termed native spliceosomes) from supraspliceosomes, and carried out image reconstruction of frozen hydrated particles by the single particle technique.

For the isolating of native spliceosomes we specifically cleaved the general population of pre-mRNAs within supraspliceosomes, while keeping the snRNAs within these substructures intact. The resulting monomeric native spliceosomes were purified by centrifugation in a glycerol gradient, where they sedimented at the 60-70S region. The biochemical composition of the 60-70S native spliceosomes is very similar to that of 200S supraspliceosomes and they exhibit RNA splicing *in vitro* – hence the term native spliceosomes (Azubel et al., 2004; Azubel et al., 2006). The isolation of native spliceosomes and their relative stability enabled us to perform 3-D cryo-EM structural analysis by the single particle technique (Azubel et al., 2004). The 3-D structure was reconstructed from 9297 raw single-particle images at a resolution of 20 Å, which is the highest resolution obtained thus far for any splicing-active complex.

Figure 2 summarizes these structural studies. The structure (Figure 2, left and middle images) depicts the native spliceosome as an elongated globular particle made up of two distinct subunits. This finding is consistent with our previous STEM mass measurements, which revealed two major, equally populated, distinct groups of small particles with masses of 1.5 MDa and 3.1 MDa, which together add up to close to the 4.8 MDa mass of the native spliceosome (Müller et al., 1998). The two subunits are interconnected to each other leaving a tunnel in between, which is large enough to allow the pre-mRNA to pass through. The other side of the native spliceosome exposes a cavity that could provide a place to transiently store the pre-mRNA. Because RNA molecules are quite susceptible to degradation, and the number of bases that are involved in the splicing reaction represents only a small fraction of the entire pre-mRNA, the cavity might protect the part of the pre-mRNA that is not directly involved in the splicing reaction from non-specific degradation.

The two subunits vary also with respect to the distribution of high densities within the native spliceosome. Because RNA is denser than protein, the localization of regions of high density can provide some information about internal organization of RNA and protein components. We have shown that the large subunit is a suitable candidate to accommodate the penta-snRNP, as the high density regions were found on the large subunit (Figure 2, right image), and the mass and volume of the large subunit are similar to the estimated mass and volume of the penta-snRNP (Azubel et al., 2004). Further structural analysis of native spliceosomes in the context of intact supraspliceosomes, using electron microscopy combined with image processing, revealed classes

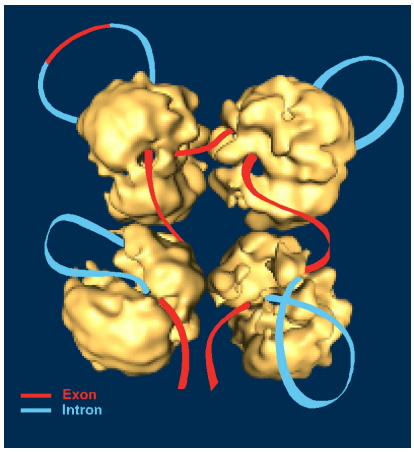

Fig. 3 *The supraspliceosome model. A schematic model of the supraspliceosome in which an imaginative pre-mRNA (introns in blue, exons in red) is connecting four native spliceosomes. The supraspliceosome presents a platform onto which the exons can be aligned and splice junctions can be checked before splicing occurs (note an alternative exon depicted in the upper left corner) (From Azubel et al., 2006).*

in which the small subunits, which are in the center of the supraspliceosome, are related to one another by a right angle forming a four-fold pattern (Cohen-Krausz et al., 2007). These studies culminated in an integrated working model (Figure 3) that can account for alternative splicing in general and for alternative 5′ splice site selection in particular (see below).

Alternative 5′ splice site selection

The remarkable accuracy and specificity of the splicing reaction stems from an array of cis and trans-acting factors. The cis acting elements include the 5′ and 3′ splice sites (SSs), a branch point, a poly-pyrimidine track, and splicing enhancer and silencer sequences. The trans-acting factors include the five spliceosomal uridine-rich small nuclear ribonucleoprotein particles (snRNPs); several non-snRNP proteins, such as the SR-protein family; and hnRNP proteins.

A bioinformatic survey of EST data bases revealed that alternative 5′SS selection accounts for ~8% of the alternative splicing events that are conserved between the human and the mouse genomes. This estimate is by far lower than expected because a survey carried out by Miriami et al. (2002) on 446 annotated genes revealed that sequences that conform to the 5′SS consensus are highly abundant in introns, yet, there is no evidence that such latent 5′SSs are selected for splicing. In this study, a total of 10490 latent 5′ SSs were found within 1601 introns. It was also found that more than 95% of these intronic latent 5′SSs have an in frame stop codon between the authentic and the latent 5′SSs. Thus, utilization of such a latent 5′SS could potentially introduce a premature termination codon (PTC) in the mature mRNA. Such PTC-carrying mRNAs could be harmful to the cell by producing a truncated protein, which could function in a deleterious way (e.g. loss-of-function; dominant-negative) . Following this observation we hypothesized that selection of latent 5′SSs is suppressed such that the introduction of PTCs into mRNAs is avoided. Hence, the phenomenon was termed stop codon mediated Suppression Of Splicing (SOS).

In order to validate this hypothesis, Li et al. (2002) carried out an exhaustive mutational analysis on two model genes: CAD (abbreviated for the multifunctional enzyme carbamoyl-phosphate syntethase, aspartate trans-carbamylase, and dihydro-orotase) and IDUA (abbreviated for α-L-iduronidase). In these analyses latent splicing was elicited from mini-gene constructs only when all in frame PTCs were eliminated either by direct mutations or by frame shifting. These experiments have specifically implicated the PTCs as the causal factor for the suppression of splicing from the latent 5′SS. Because a stop codon is inherently defined by an open reading frame (ORF), the SOS machinery needs a start point to register and establish the reading frame for a proper identification of a PTC. To this end we have whether the translation start codons (AUG sequences), and specifically the first AUG, affect the SOS phenomenon. Transfections with CAD constructs wherein start codons were mutated, elicited latent splicing, thus showing that AUG sequences are required to sustain the SOS mechanism (Kamhi et al., 2006). Importantly, although the SOS mechanism is dependent on the establishment of an ORF, this mechanism is fully active when protein translation, or even the pioneer round of translation, is inhibited (Kamhi et al., 2006). These and other published data (Li et al., 2002; Wachtel et al., 2004), portray the SOS mechanism as an RNA surveillance mechanism that is distinct from the nonsense mediated mRNA decay (NMD) pathway.

New evidences from our group now show that the initiator tRNA (ini-tRNA) may act, in a manner that is independent on its role in protein translation, as an

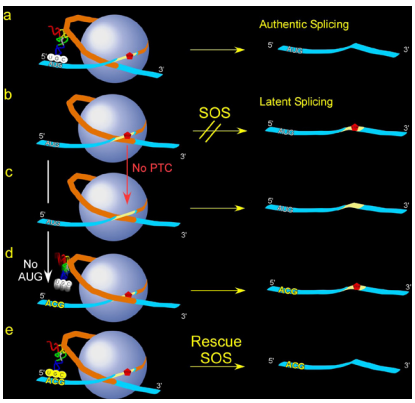

Fig. 4 *Cartoon depiction of the working hypothesis for the SOS mechanism. The balloons represent native spliceosomes around which the pre-mRNA (exons in light blue; introns in orange; latent exon in yellow) is wrapped, so that the 5′ and 3′ splice sites are juxtaposed. When the authentic 5′SS is juxtaposed to the 3′SS (a), an ORF is established between the first AUG, which is recognized by Ini-tRNA, and the normal stop codon, enabling splicing to commence. In an altered conformation, where the latent 5′SS is aligned with the 3′SS (b), a pre-mature stop codon does not allow the recognition of the ORF, resulting in the suppression of latent splicing. When the Premature intronic stop codon is eliminated (c), SOS is relieved and latent splicing can take place. When the AUG is mutated (d), the Ini-tRNA no longer recognizes an AUG and the stop codon cannot be recognized in an ORF thus leading to latent splicing. The recognition of the mutated AUG by a mutated Ini-tRNA (e) re-establishes an ORF and rescues SOS.*

SOS trans-acting-factor that marks the AUG to establish a reading frame. We tested this possibility using gene constructs harboring a mutation in their first ATG. As expected, SOS for pre-mRNAs expressed from these constructs was abrogated and they exhibited latent splicing. However, SOS could be rescued when each of the ATG mutants was cotransfected with a mutated ini-tRNA construct in which the anti-codon was mutated to complement the respective ATG mutation. This was demonstrated on two minigene constructs, CAD and IDUA, each carrying either ATG to ACG mutation or a ATG to AAA mutation. Importantly, cotransfection of the ATG mutant with an ini-tRNA mutant harboring a non-complementary anti-codon mutation, did not affect the level of latent RNA. It should be noted that the effect of the ini-tRNA on SOS is not dependent of translation, because the rescue of the AUG phenotype by the complementary ini-tRNA mutant was maintained in cells where translation initiation had been inhibited. Moreover, we were able to show that the mutated forms of the ini-tRNA are not charged with amino acid, suggesting that only the RNA moiety, of the tRNA is required for SOS. Consistent with this finding, we also show that ini-tRNA, but neither the elongator methionyl-tRNA nor lysyl-tRNA, fractionated with supraspliceosomes and native spliceosomes, suggesting a possible involvement of ini-tRNA in splicing regulation.

Our study present evidence for the recognition of ORFs in the cell nucleus prior to RNA splicing, and points at a possible connection between the maintenances of ORF and splice site selection. A working model that accounts for this connection is shown in Figure 4.

Selected publications

Sperling, R., Sperling, J., Levine, A.D., Spann, P., Stark, G.R. and Kornberg, R.D. (1985) Abundant nuclear ribonucleoprotein form of CAD RNA. Mol. Cell Biol., 5, 569-575.

Spann, P., Feinerman, M., Sperling, J. and Sperling, R. (1989) Isolation and visualization of large compact ribonucleoprotein particles of specific nuclear RNAs. Proc. Natl. Acad. Sci. USA, 86, 466-470.

Sperling, R. and Sperling, J. (1990) Large nuclear ribonucleoprotein particles of specific RNA polymerase II transcripts. In Strauss, P.R. and Wilson, S.H. (eds.), The Eukaryotic Nucleus, Molecular Biochemistry and Macromolecular Assemblies. Telford Press, Caldwell, NJ, Vol. 2, pp. 453-476.

Miriami, E., Angenitzki, M., Sperling, R. and Sperling, J. (1995) Magnesium cations are required for the association of U small nuclear ribonucleoproteins and SR proteins with pre-mRNA in 200 S large nuclear ribonucleoprotein particles. jmb, 246, 254-263.

Yitzhaki, S., Miriami, E., Sperling, J. and Sperling, R. (1996) Phosphorylated Ser/Arg-rich proteins: Limiting factors in the assembly of 200S large nuclear ribonucleoprotein particles. Proc. Natl. Acad. Sci. USA, 93, 8830-8835.

Medalia, O., Koster, A.J., Tocilj, A., Angenitzki, M., Sperling, J., Berkovitch, Y.Z. and Sperling, R. (1997) Automated electron tomography of large nuclear RNP (lnRNP) particles- the naturally assembled complexes of precursor messenger RNA and splicing factors. J. Struct. Biol., 120, 228-236.

Sperling, R., Koster, A.J., Melamed-Bessudo, C., Rubinstein, A., Angenitzki, M., Berkovitch-Yellin, Z. and Sperling, J. (1997) Three-dimensional image reconstruction of large nuclear RNP (lnRNP) particles by automated electron tomography. J. Mol.Biol., 267, 570-583.

Müller, S., Wolpensinger, B., Angenitzki, M., Engel, A., Sperling, J. and Sperling, R. (1998) A supraspliceosome model for large nuclear ribonucleoprotein particles based on mass determinations by scanning transmission electron microscopy. jmb, 283, 383-394.

Sperling, R. and Sperling, J. (1998) The lnRNP particle - A naturally assembled complex of pre-mRNA and splicing factors. In Schenkel, J. (ed.), RNP Particles, Splicing and Autoimmune Diseases. Springer, pp. 29-47.

Li, B., Wachtel, C., Miriami, E., Yahalom, G., Friedlander, G., Sharon, G., Sperling, R. and Sperling, J. (2002) Stop codons affect 5′ splice site selection by surveillance of splicing. Proc. Natl. Acad. Sci. USA, 99, 5277-5282.

Medalia, O., Typke, D., Hegerl, R., Angenitzki, M., Sperling, J. and Sperling, R. (2002) Cryoelectron microscopy and cryoelectron tomography of the nuclear pre-mRNA processing machine. J. Struct. Biol., 138, 74-84.

Miriami, E., Motro, U., Sperling, J. and Sperling, R. (2002) Conservation of an open-reading frame as an element affecting 5′ splice site selection. J. Struct. Biol., 140, 116-122.

Azubel, M., Wolf, S.G., Sperling, J. and Sperling, R. (2004) Three-dimensional structure of the native spliceosome by cryo-electron microscopy. Mol. Cell, 15, 833-839.

Wachtel, C., Li, B., Sperling, J. and Sperling, R. (2004) Stop codon-mediated suppression of splicing is a novel nuclear scanning mechanism not affected by elements of protein synthesis and NMD. RNA, 10, 1740-1750.

Azubel, M., Habib, N., Sperling, J. and Sperling, R. (2006) Native spliceosomes assemble with pre-mRNA to form supraspliceosomes. J. Mol. Biol., 356, 955-966.

Kamhi, E., Yahalom, G., Kass, G., Hacham, Y., Sperling, R. and Sperling, J. (2006) AUG sequences are required to sustain nonsense-codon-mediated suppression of splicing. Nucl. Acids Res., 34, 3421-3433.

Cohen-Krausz, S., Sperling, R. and Sperling, J. (2007) Exploring the Architecture of the Intact Supraspliceosome Using Electron Microscopy. Journal of Molecular Biology, 368, 319-327.