

## Structural Analysis of the Human Spliceosome by Electron Microscopy

M.S. Jurica,\*† J. Ilagan\* and E.A. Alcid\*

\*Department of Molecular, Cell & Developmental Biology, University of California, Santa Cruz, CA 95064

† Center for Molecular Biology of RNA

A common feature of eukaryotic genes is the presence of intervening sequences that interrupt the functional coding sequences. The intervening sequences, referred to as introns, are removed from RNA transcripts of genes, while the functional sequences, called exons, are joined in a process called splicing. Splicing is catalyzed by the spliceosome, a large cellular complex composed of RNA and protein including the U- rich small nuclear ribonucleoproteins (U snRNPs). The spliceosome is a highly dynamic macromolecular machine and undergoes several conformational changes during splicing. As with many macromolecular machines, the major challenges to structural studies of the spliceosome are its dynamic nature and limited amounts of material available for study. Our research group has pioneered a method to purify spliceosomes arrested at the second chemical step of the splicing chemistry at an intermediate known as C complex [1]. At this stage, the spliceosome consists of U2, U5, and U6 snRNPs plus ~40 additional proteins. Using cryo-electron microscopy and single particle reconstruction techniques we generated a 3D model of the complex at 30 Å resolution (Fig. 1) [2]. This model represents an important step in understanding the structure of the spliceosome, but mechanistic interpretation of the model has been limited. To further explore the structure of the spliceosome, we are investigating the effect of different biochemical manipulations of the spliceosome on its structure. We are also developing labels to identify the location of key spliceosome components. Our progress on these two fronts will be presented.

We are employing several biochemical approaches to alter the structure of our purified splicing complexes. These include directed cleavage of the substrate RNA [3], partial disassociation by mild denaturant treatment, and capturing the complex at other assembly intermediates. By 2D image analysis and difference mapping of the different complexes (Fig. 2), we are able to identify different regions of the spliceosome, which we then relate back to the different manipulations. These studies are revealing the core complex and providing information regarding the nature of more flexible regions.

Good EM labels for components of the spliceosome have proven difficult to generate. To locate key RNA features in the structure of the spliceosome by EM, we fused a sequence-specific RNA binding protein to a protein with a distinct donut-shaped structure. We used this fusion to label spliceosomes assembled on a pre-mRNA that contained the target sequence in the exons. The label is clearly visible in EM images of the spliceosome, and subsequent image processing with averaging shows that the exons sit close to each other in the complex (Fig. 3) [4]. This approach has been extended by several other groups to label both proteins and nucleic acids [5, 6]. Currently, we are creating additional RNA directed labels to use EM to investigate the global structure of alternatively spliced introns and their regulation by splicing factors.

### References

- [1] M.S. Jurica, et al., *RNA*, 8 (2002) 426.
- [2] M.S. Jurica, et al., *Nat Struct Mol Biol*, 11 (2004) 265.
- [3] J. Ilagan, et al., *J Mol Biol*, 394 (2009) 363.

- [4] E.A. Alcid, et al., *Nat Struct Mol Biol*, 15 (2008) 213.  
 [5] D. Flemming, et al., *Nat Struct Mol Biol*, 17 (2010) 775.  
 [6] M.E. Stroupe, et al., *RNA*, 15 (2009) 244.  
 [7] This work was funded by National Institutes of Health grant 5R01GM72649 to M.S.J.

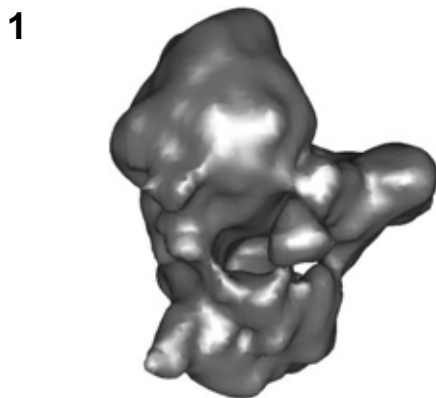


FIG. 1. 3D model of human C complex spliceosome reconstructed from ~1000 single particle cryo-EM images.

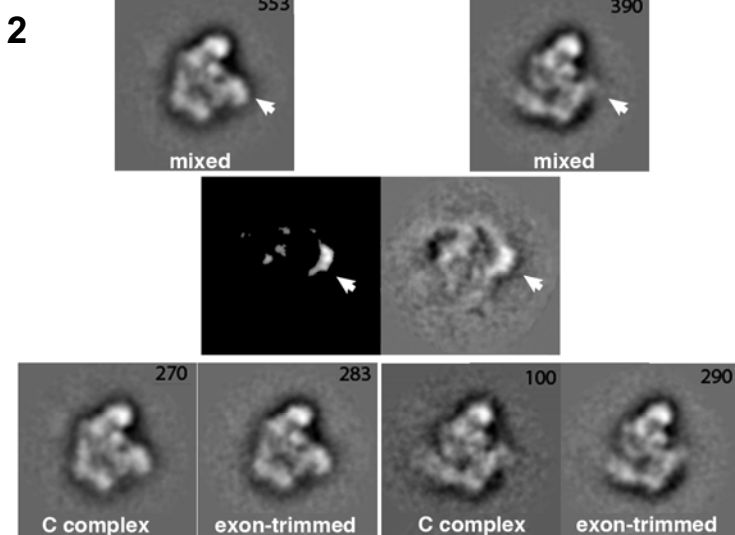


FIG. 2. 2D image analysis of negative stained images of C complex spliceosome. We compared intact complexes with complexes from which exon sequences were trimmed. The top row shows two related class average views from a mixed data set with a difference maps shown below (The difference map on the left is shown at a  $5.0 \sigma$  threshold). The last row separates C complex and exon-trimmed images that make up the mixed average shown above. The numbers in the top right indicate the number of individual images in each averaged view. The arrow points to a density feature that is differentially represented in the C complex vs. exon-trimmed samples.

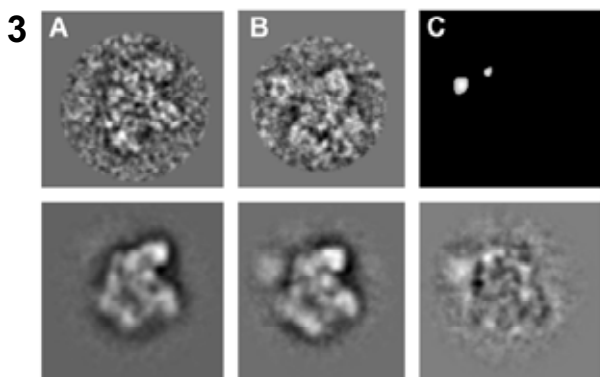


FIG. 3. 2D image processing of C complex spliceosomes with a protein label directed to the 5' exon. A. Typical raw image (top) and class average view (bottom) of the unlabeled complex. B. Same as (A) but with labeled complexes. C. Difference maps of the class average views in (A) and (B). The top difference map is shown at a  $3.0 \sigma$  threshold.