Published in final edited form as:

Nat Struct Mol Biol. 2008 February ; 15(2): 213–215. doi:10.1038/nsmb.1378.

A protein-based EM label for RNA identifies the location of exons in spliceosomes

Eric A Alcid and Melissa S Jurica

Department of Molecular, Cell and Developmental Biology and Center for Molecular Biology of RNA, University of California, 1156 High Street, Santa Cruz, California 95064, USA

Abstract

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. This labeling strategy will serve as a general tool for analyzing the structures of RNA-containing macromolecular complexes.

The spliceosome is the large and dynamic macromolecular complex that carries out the process of splicing during eukaryotic gene expression. The structure of the spliceosome at several assembly states has been successfully tackled by EM and single-particle reconstruction^{1–6}. However, these structures lack easily identifiable features and do not have sufficient resolution to identify individual protein or RNA elements. Although progress in interpreting the spliceosome structure has been made by using gold cluster and antibody labeling⁴, these reagents remain challenging and expensive to use. The complexity and dynamics of the spliceosome, which is made up of more than 100 proteins and 5 RNAs, poses special challenges to structural analyses. Current purification protocols yield small amounts of material at low nanomolar concentrations. Thus, any potential label must have a high affinity for its target yet maintain a low background of binding. In this work, we aimed to develop a recombinant protein-based EM label that can identify key RNA features in the human spliceosome. In particular, we sought to identify the location of the two exons in the complex that are ligated together during splicing.

Results

Creation of a protein-based EM label directed to RNA

To generate the label, we fused the coliphage coat protein PP7 to the β -subunit of *Escherichia coli* DNA polymerase III, dnaN, to create a labeling protein we term Beta-PP7 (Fig. 1a). PP7 is a small protein that binds a 24-nucleotide (nt) RNA hairpin sequence with picomolar affinity⁷. dnaN, a so-called sliding clamp, dimerizes to form a donut-shaped structure of ~90 Å in diameter⁸. Beta-PP7 is highly expressed in *E. coli* and is well behaved following purification via a 6×His tag (data not shown). We examined the purified protein, which was negatively stained with uranyl acetate, by transmission EM and could easily

Correspondence should be addressed to M.S.J. (jurica@biology.ucsc.edu).

Author contributions

E.A.A. and M.S.J. designed the experiments; E.A.A. performed the experiments; E.A.A. and M.S.J. analyzed the data. E.A.A. prepared figures and M.S.J. wrote the paper.

identify the ring structure formed by dnaN (data not shown). Furthermore, on the basis of western analysis of pull-downs using a fusion of another coliphage coat protein, MS2, to maltose binding protein (MS2-MBP) and a pre-mRNA containing both MS2 and PP7 binding sites, we found that Beta-PP7 specifically binds RNA containing the PP7 target sequence (Fig. 1b). EM imaging of negatively stained eluate from the pull-downs revealed the presence of the donut-shaped label (Fig. 1c).

Labeling spliceosome complexes with Beta-PP7

To label the exons in the spliceosome, we inserted the PP7 target sequence into either the 5' or 3' exon of a pre-mRNA splicing substrate (Fig. 1d). The 5' exon PP7 site is located 59 nt upstream of the 5' splice site in a region where the pre-mRNA is accessible in assembled spliceosomes (J. Ilagan and M.S.J., unpublished observations). For 3' exon labeling, we placed the PP7 site at an accessible region 49 nt downstream of the 3' splice site. We purified C-complex spliceosomes assembled on this pre-mRNA in HeLa nuclear extract and arrested just before exon ligation as described previously⁹ with the following exception: after size exclusion and before affinity purification, we incubated the complexes with an ~100-fold excess of Beta-PP7. Excess unbound label was then washed away while the complexes were immobilized on the affinity resin.

In electron micrographs of negatively stained purified complexes, we repeatedly observe spliceosomes associated with a single donut of ~9 nm in diameter (Fig. 2a,b). Although the label presents two PP7 binding sites, we did not observe any dimerization of spliceosomes. It is likely that the low nanomolar concentration of spliceosomes combined with the large excess of Beta-PP7 results in the single labeling. A count of the particles that had a clear adjacent donut shape led us to estimate a labeling efficiency of ~25%. Because orientations of the spliceosome in which the label lies on top or underneath the complex or presents a side view would not be tallied, the labeling efficiency may actually be higher.

Image processing of individual particles to locate exons

To determine the location of the label in the complex, we selected a large number of individual labeled spliceosome images from labeled samples and carried out several rounds of classification and alignment to obtain class-averaged views. Previously, we observed that the spliceosome exhibits a few preferred orientations in negative stain⁹, and with the Beta-PP7 labeled samples we again obtain class-averaged views reflecting these orientations. Furthermore, in a subset of class-averaged views, we observe an additional feature that has dimensions appropriate for Beta-PP7 (Fig. 2a,b). To verify the feature as the label, we compared spliceosome images from unlabeled and labeled samples in the following way. For a particular class-averaged view showing the feature, we extracted the constituent images (typically on the order of 400 images) and mixed these with a similar number of images derived from a comparable class-averaged view obtained from unlabeled splicing complexes. We carried out several rounds of alignment with these images and then separated them into two classes. As expected, the classes represent an identical view, but one class contained the feature and was composed entirely of images from the labeled image set (slightly fewer than the input 400), whereas the other class lacked the feature (Fig. 2c, right). We then subtracted the two averages from each other to generate a difference image and observed that the strongest difference feature is consistent with Beta-PP7 in both shape and dimensions (Fig. 2d,e). We carried out this analysis with the label directed to both the 5' and 3' exons independently. In both cases, the label was located on the same face of the spliceosome structure, as may be expected for this stage of spliceosome assembly in which complexes are arrested before exon ligation. Given that the label is bound to RNA, which could serve a flexible tether, we were somewhat surprised at the relatively tight distribution of the averaged label density. It is conceivable that a wider distribution of label positions

exists and was not selected for during classification. However, we did not observe any evidence for this possibility in the other class-averaged views.

Discussion

Labeling features of macromolecular complexes is an important, yet challenging, undertaking for interpreting structures derived from EM reconstruction. Our goal in creating the Beta-PP7 label was to identify pre-mRNA features in the structure of C-complex spliceosomes. One important attribute for any exogenous label is efficiency of binding. Although not perfectly efficient, the 25% labeling we achieved with Beta-PP7 is a significant achievement for labeling single particles in EM images. There are at least two possible sources for the substoichiometric labeling observed. First, as discussed previously, our method for estimating labeling efficiency is conservative. Second, our purification yields complexes at low concentration and, although we use an excess of label, binding kinetics suggest that we still cannot achieve 100% labeling given the time available for binding. Even so, the labeling efficiency we achieved clearly allowed us to identify the position of the label by routine image processing and statistical analysis. Furthermore, the labeling is considerably higher than that achieved either by our previous strategy of using biotinylated antisense oligonucleotides targeting exon sequences with a streptavidin-gold conjugate⁹ or by attempts at antibody labeling. Because labeled particles can be rare, most single-particle labeling experiments are reported with a handful of individually labeled complexes and are interpreted by eye, not by averaging as we demonstrate in this study. We attribute the relatively high efficiency of labeling to the picomolar affinity of PP7 for its binding site and our ability to economically produce and use a large excess of the protein during a onestep labeling procedure.

Another consideration for an EM label is the precision with which it can be used to locate its target. Given the large size of Beta-PP7 and the fact that its target RNA may provide a flexible tether, its precision may be limited. However, relative to the large spliceosome, we can still make some predictions based on the position of the Beta-PP7 label directed to the exons in the complex. First, as noted previously, the 5' and 3' exons are positioned close to each other, possibly extending from the same face of the complex, and reflecting their positions before exon ligation. Comparison to cryo-negative stain views used to calculate the three-dimensional structure of C-complex spliceosomes² reveals that this area correlates to the region where the 'top' and 'bottom' domains of that structure meet. Recent work by the Stark and Luhrmann groups indicates that the bottom domain is the U5 small nuclear ribonucleoprotein (snRNP)⁴. One U5 snRNP protein, pre-mRNA processing factor 8 homolog (PRP8), has been shown to interact with both exons near the splice sites in the C complex, suggesting that this protein may be located near the domain interfaces^{10–12}.

In summary, we have established a new protein-based label against RNA for EM studies of macromolecular complexes. The label is straightforward to produce, is highly specific and has allowed us to identify key RNA features in the human spliceosome. Furthermore, Beta-PP7 has the benefit of being easily identified in raw EM images and will thereby serve as a tool for analyzing the structures of the ever-growing number of large macromolecular complexes that contain or process RNA.

Methods

Construction and purification of the Beta-PP7 fusion protein and PP7-tagged pre-mRNA

The full-length β -subunit of the *E. coli* DNA polymerase, *dnaN*, was subcloned into the NcoI and EcoRI sites of pET28ZZTPP7H to generate the plasmid pETBetaPP7H, which encodes Beta (residues 1–366) fused via a 27-amino-acid linker to the N terminus of PP7

(residues 1–137) with a C-terminal 6×His tag. The protein was expressed in *E. coli* and purified sequentially by nickel affinity, heparin cation exchange and size-exclusion chromatographies.

PP7-tagged pre-mRNA was generated by subcloning a primer extension product of the 24-nt PP7 binding site into either the KpnI-SacI sites (5' exon tag) or the BamHI site (3' exon tag) of an AdML gene construct containing three MS2 binding sites in the intron[**Ed: Only 1 intron.**] and an AG to GG 3' splice-site mutation⁹.

Pull-down experiments

10 pmol PP7-tagged pre-mRNA was incubated with 1 nmol MS2-MBP, 1 nmol Beta-PP7 and 50 µl of amylose resin (New England Biolabs) in 20 mM Tris, pH 7.9, 150 mM KCl and 2 mM MgCl₂ for 2 h at 4 °C. The resin was washed and complexes were eluted with maltose. Identical control experiments were carried out with pre-mRNA substrate lacking the PP7 tag. For western blots, eluted fractions from the amylose resin were separated on a 10% (v/v) SDS-PAGE gel, transferred to a nitrocellulose membrane and probed with an antibody to the 6×His tag [**ED: OK**] of Beta-PP7 (Santa Cruz Biotechnology).

Assembly and labeling of C-complex spliceosomes

C-complex spliceosomes were assembled and purified as described⁹ with the following exception: before affinity purification, we added 1 nmol Beta-PP7 to the spliceosome-containing fractions and incubated them for 1 h at 4 °C.

Electron microscopy and image processing of labeled complexes

5 µl of freshly prepared C-complex spliceosome solution[**ED: OK**] was deposited onto glow-discharged carbon-coated grids and stained with 2% (w/v) uranyl acetate solution. Samples were analyzed in a JEOL 1230 electron microscope operating at 120 kV. Micrographs were exposed under low-dose conditions at 60,000× magnification on a 4K × 4K Gatan CCD camera, resulting in images being sampled at 4.9 Å per pixel. Roughly 6,500 5' exon–labeled spliceosomes, 10,000 3' exon–labeled spliceosomes and 6,500 unlabeled spliceosomes were selected with the EMAN program Boxer¹⁴. Following 2×2 pixel averaging in SPIDER¹³ the IMAGIC software package¹⁵ was used to align images and create class sums and difference images.

Acknowledgments

J. Ilagan provided support for EM data collection and image processing. M. O' Donnell (Rockefeller University, New York, USA) provided a plasmid encoding full-length dnaN and K. Collins (University of California, Berkeley, USA) gave us pET28ZZTTP7H including the PP7 sequence. We thank N. Grigorieff (Brandeis University, Waltham, Massachusetts, USA) and members of the Jurica laboratory for advice and discussion. This work was funded by US National Institutes of Health grant 5R01GM72649 to M.S.J., which included a Diversity Supplement for E.A.A.

References

1. Boehringer D, et al. Three-dimensional structure of a pre-catalytic human spliceosomal complex B. *Nat Struct Mol Biol* 2004;11:463–468. [PubMed: 15098019]
2. Jurica MS, Sousa D, Moore MJ, Grigorieff N. Three-dimensional structure of C complex spliceosomes by electron microscopy. *Nat Struct Mol Biol* 2004;11:265–269. [PubMed: 14981503]
3. Deckert J, et al. Protein composition and electron microscopy structure of affinity-purified human spliceosomal B complexes isolated under physiological conditions. *Mol Cell Biol* 2006;26:5528–5543. [PubMed: 16809785]

4. Sander B, et al. Organization of core spliceosomal components U5 snRNA loop I and U4/U6 Di-snRNP within U4/U6. U5 Tri-snRNP as revealed by electron cryomicroscopy. *Mol Cell* 2006;24:267–278. [PubMed: 17052460]
5. Behzadnia N, et al. Composition and three-dimensional EM structure of double affinity-purified, human prespliceosomal A complexes. *EMBO J* 2007;26:1737–1748. [PubMed: 17332742]
6. Ohi MD, Ren L, Wall JS, Gould KL, Walz T. Structural characterization of the fission yeast U5. U2/U6 spliceosome complex. *Proc Natl Acad Sci USA* 2007;104:3195–3200. [PubMed: 17360628]
7. Lim F, Peabody DS. RNA recognition site of PP7 coat protein. *Nucleic Acids Res* 2002;30:4138–4144. [PubMed: 12364592]
8. Kong XP, Onrust R, O'Donnell M, Kuriyan J. Three-dimensional structure of the beta subunit of *E. coli* DNA polymerase III holoenzyme: a sliding DNA clamp. *Cell* 1992;69:425–437. [PubMed: 1349852]
9. Jurica MS, Licklider LJ, Gygi SR, Grigorieff N, Moore MJ. Purification and characterization of native spliceosomes suitable for three-dimensional structural analysis. *RNA* 2002;8:426–439. [PubMed: 11991638]
10. Chiara MD, et al. Identification of proteins that interact with exon sequences, splice sites, and the branchpoint sequence during each stage of spliceosome assembly. *Mol Cell Biol* 1996;16:3317–3326. [PubMed: 8668147]
11. Chiara MD, Palandjian L, Feld Kramer R, Reed R. Evidence that U5 snRNP recognizes the 3' splice site for catalytic step II in mammals. *EMBO J* 1997;16:4746–4759. [PubMed: 9303319]
12. Reichert VL, Le Hir H, Jurica MS, Moore MJ. 5' exon interactions within the human spliceosome establish a framework for exon junction complex structure and assembly. *Genes Dev* 2002;16:2778–2791. [PubMed: 12414731]
13. Frank J, et al. SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J Struct Biol* 1996;116:190–199. [PubMed: 8742743]
14. Ludtke SJ, Baldwin PR, Chiu W. EMAN: semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol* 1999;128:82–97. [PubMed: 10600563]
15. van Heel M, Harauz G, Orlova EV, Schmidt R, Schatz M. A new generation of IMAGIC image processing system. *J Struct Biol* 1996;116:17–24. [PubMed: 8742718]
16. Tars K, Fridborg K, Bundule M, Liljas L. The three-dimensional structure of bacteriophage PP7 from *Pseudomonas aeruginosa* at 3.7-Å resolution. *Virology* 2000;272:331–337. [PubMed: 10873776]
17. Frank, J. Three-Dimensional Electron Microscopy of Macromolecular Assemblies. Academic; New York: 1996.

Figure 1.

Beta-PP7 label design, RNA binding and imaging. **(a)** Schematic of Beta-PP7 modeled from the separate X-ray crystallographic structures of the coliphage coat protein PP7 and the β subunit of *E. coli* DNA polymerase III, dnaN[ED: OK]^{8,16}. dnaN dimerizes to form a 'donut' structure (shaded darker gray), and two copies of PP7 (lighter-gray crescents) serve to bind RNA. **(b)** Western analysis of Beta-PP7 pull-downs. Using a fusion of the coliphage coat protein MS2 and maltose binding protein (MS2-MBP) pull-downs of Beta PP7 were carried out with MS2-tagged pre-mRNA substrates [ED: Changed as requested] either possessing a PP7 site (+) or lacking one (–). The antibody was specific to the 6×His tag on Beta-PP7. **(c)** EM micrograph of the + PP7 elution stained with 2% (w/v) uranyl acetate. Arrows indicate the Beta-PP7 rings. **(d)** Schematic of the pre-mRNA substrate for spliceosome purification and labeling. A single PP7 site (large black hairpin) is inserted at indicated sites of a pre-mRNA splicing substrate containing 3 MS2 sites (small gray hairpins) used for affinity purification of C-complex spliceosomes.

Figure 2.

Difference maps of two-dimensional averages of labeled and unlabeled C complexes. **(a)** Filtered single image of C complex labeled at the 5' exon and the corresponding two-dimensional average. **(b)** As in **a**, with the C complex labeled at the 3' exon instead of the 5' exon. **(c)** As in **a**, with unlabeled C complex. **(d)** From left to right: the difference map between the two-dimensional averages of 5' exon – labeled C complex and unlabeled C complex; the statistically significant region of the difference map at the 99% confidence level¹⁷ and the difference map (in white) superimposed on a two-dimensional average of unlabeled C complex. **(e)** As in **d**, with the C complex labeled at the 3' exon instead of the 5' exon.