Eric J. Montemayor, TP1 application for NCCAT, Dec. 2018

Supplementary Information: I have spent the last 10 years investigating the structure and dynamics of protein/RNA complexes (Montemayor *et al*., 2014, *NSMB*; Montemayor *et al*., 2018, *Nature Comm*.). We recently reported the crystal structure of the U6 small-nuclear ribonucleoprotein from *S. cerevisiae* at 2.7 Å resolution, thereby showing that the RNA remodeling "active site" of Prp24 is positioned close to the Lsm2-8 ring in three-dimensional space. However, the U6 snRNP from *S. cerevisiae* is an evolutionary outlier that lacks the conserved N-terminal domain of Prp24 that is present in the human and *S. pombe* orthologs of the protein. This N-terminal domain contains many Half-A-Tetracopeptide (HAT) repeats and is a protein-protein interaction hub that expands the function of Prp24 and U6 snRNPs. For example, the N-terminal domain interacts with deubiquitinases and components for the RNA transcriptional machinery. Notably, human Prp24 has also been shown to play a role in expression of mature HIV-1.

We have reconstituted U6 snRNPs from *S. pombe*, which include the N-terminal protein-protein interaction domain that was lacking in our prior structures from *S. cerevisiae* (Fig. A). The overall mass of this complex is ~450 kDa, and the purified samples are homogenous via non-denaturing polyacrylamide gel electrophoresis and gel filtration chromatography (Fig B). The purified snRNPs have already been used to acquire small-angle X-ray scattering data showing that the particle is elongated with maximum dimension of ~ 300 Å and has a calculated molecular mass that comports with the expected size of an intact complex. We have prepared negative stain samples and have visualized ring-like structures that likely correspond to the associated Lsm2-8 complex (~80 kDa) (Fig. C). However, the negative stain images lack densities corresponding to the N-terminal HAT domain of Prp24 (~100 kDa), suggesting denaturation of Prp24 on grids. We are currently pursuing single particle analysis on spot-it-on grids via a collaboration with NRAMM, but have yet to successfully visualize intact particles in ice, suggesting optimization of buffer conditions or addition of protein cofactors (i.e. one of the above mentioned deubiquitinases) will be needed to sufficiently stabilize the complex.

We already have a fallback project in motion, in case the *S. pombe* U6 snRNP project is found to be intractable during the early phases of the TP1 training interval. We have recently discovered that the homologous Lsm1-7 complex (~ 80 kDa) has unusually tight binding affinity for polyuridine tracts that are followed by a single adenosine residue (Fig D). This binding activity is completely unexpected and has implications for the mechanism by which messenger RNA degradation is regulated in eukaryotes. Despite its smaller size, this complex has proven refractory to crystallization. Since the Lsm2-8 complex appears to be the (currently) best behaved part of the *S. pombe* U6 snRNP, we propose to determine the structure of the homologous Lsm1-7 complex with adenylated RNA via cryo-EM. An added benefit to this alternative approach would be training and experience with phase plate microscopy, as the method would likely be needed for imaging such a small complex.