

[s8a.m6.o1] Crystallographic dissection of the spliceosomal snRNP structures. K. Nagai, C. Kambach, S. Walke, R. Young, C. Oubridge, S. Price, P. Evans and J. Li. *MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.*

Keywords: proteins, nucleic acids.

The protein coding sequence of the majority of eukaryotic genes is interrupted by non-coding intervening sequences (introns). Following transcription into mRNA precursors the introns are excised from pre-mRNA to form continuous coding sequences by two-step trans-esterification reactions within a macromolecular assembly called the spliceosome. The major components of the spliceosome are four large RNA-protein complexes called the U1, U2, U4/U6 and U5 small nuclear ribonucleoprotein particles (snRNPs). These snRNPs contain seven Sm proteins (B/B', D1, D2, D3, E, F and G) in common, which assemble around the Sm site present in four of the major spliceosomal small nuclear RNAs (snRNAs) and form a globular core domain. In addition to the Sm proteins each snRNP contains some specific protein components. For example the U1 snRNP contains three specific proteins, namely U1 70k, U1A and U1C proteins. The 12S U2 snRNP contains two specific proteins, namely the U2B'' and U2A' proteins. We are trying to understand the molecular mechanism of pre-mRNA splicing through crystallographic, biochemical and genetic studies of the snRNPs (1).

The Sm proteins share a common sequence motif in two segments, Sm1 and Sm2, separated by a short variable linker. We have determined the crystal structures of two Sm protein complexes, D3B and D1D2 (2). All four Sm proteins, D3, B, D1 and D2, have a common fold containing an N-terminal helix followed by a strongly bent five stranded anti-parallel beta sheet, and the D1D2 and D3B dimers superpose closely in their core regions, including the dimer interfaces. The crystal structures suggest that the seven Sm proteins could form a closed ring and the snRNAs may be bound in the positively charged central hole (2).

The U1A protein contains two copies of the RNP motif (or RRM or RBD), one of the most commonly found RNA binding motifs. We have determined the crystal structure of the RNA-binding domain of the U1A protein bound to a 21-nucleotide RNA hairpin (3). The ten-nucleotide RNA loop binds to the surface of the beta sheet as an open structure, and the AUUGCAC sequence of the loop interacts extensively with the protein. These interactions include stacking of RNA bases with aromatic side chains of the protein and many direct and water-mediated hydrogen bonds. The structure reveals the stereochemical basis for sequence-specific RNA recognition by the RNP domain. We have also determined the crystal structure of a ternary complex between the U2B''/U2A' protein complex and hairpin-loop IV of U2 snRNA (4). Unlike its close homologue, the U1A protein, U2B'' binds to its cognate RNA only in the presence of U2A', which contains leucine-rich repeats. The crystal structure reveals how protein-protein interaction regulates RNA binding specificity and how replacing only a few key residues allows the U2B'' and U1A proteins to discriminate between their cognate and non-cognate RNA hairpins by forming alternative networks of interactions.

Our next goal is to assemble all the protein subunits and the in vitro transcribed snRNAs to reconstitute the U1 and 12S U2 snRNPs and crystallise them for structural determination by X-ray crystallography. This will provide considerable insight into the architecture of the snRNPs and will help us understand the molecular mechanism of pre-mRNA splicing (1).

[1] Kambach C, Walke S, Nagai K: Structure and assembly of the spliceosomal small nuclear ribonucleoprotein particles. *Curr. Opin. in Structural Biology* 1999, 9:222-230

[2] Kambach C, Walke S, Young RA, Avis JM, De la Fortelle E, Raker VA, Lührmann R, Li J, Nagai K: Crystal structures of two Sm protein complexes and their implications for the assembly of the spliceosomal snRNPs. *Cell* 1999, 96:375-387.

[3] Oubridge C, Ito N, Evans PR, Teo HC, Nagai K: Crystal structure of the RNA-binding domain of the U1 small nuclear ribonucleoprotein A. *Nature* 1994, 348:515-520

[4] Price SR, Evans PR, Nagai K: Crystal structure of the spliceosomal U2B''-U2A' protein complex bound to a fragment of U2 small nuclear RNA. *Nature* 1998, 394:645-650.

[s8a.m6.o2] Structural studies of Sm-related proteins from archaea. I. Törö, C. Mayer, S. Thore, J. Basquin, H. Dreher, M. Dreher, B. Séraphin and D. Suck, *European Molecular Biology Laboratory, Meyerhofstrasse 1, Postfach 102209, 69012 Heidelberg, Germany*

Keywords: proteins, nucleic acids.

Sm and Sm-like (LSm) proteins are components of the small nuclear ribonucleoprotein particles (snRNPs) of eukaryotic cells which are involved in a variety of processes including pre-mRNA splicing, telomere replication, rRNA processing and tRNA maturation. In eukaryotes the core domains of the snRNPs contain seven distinct members of the Sm protein family associated with an RNA. Recently, Sm-related proteins of unknown function were discovered in several archaeal organisms using sequence-database searches. In contrast to eukaryotes, archaea appear to encode only one or two Sm-related proteins. We have undertaken a structural and functional characterization of these proteins.

We have cloned and expressed several Sm-related proteins from different archaeal sources including *Archaeoglobus fulgidus*, *Pyrococcus abyssi*, and *Sulfolobus solfataricus*. At present three of these proteins have been crystallized and the structure of the *A. fulgidus* protein AF-Sm2 has been determined by the SIRAD technique and refined to an R-factor of 20.5% using 1.95Å data. In addition, the structure of the *P. abyssi* protein PA-Sm1 has been solved by molecular replacement.

In the crystal AF-Sm2 forms a doughnut-shaped hexamer approx. 60Å in diameter and 30Å high. The overall fold of the monomer - a strongly bent 5-stranded beta-sheet with an N-terminal helix - and the intersubunit beta-sheet formation involving beta-strands 4 and 5 of neighbouring subunits is similar to that seen in the crystal structures of two dimeric complexes of human Sm proteins recently solved by Nagai and coworkers¹. This suggests a close evolutionary relationship, however, based on the dimeric complexes these authors have proposed a seven-membered, hetero-oligomeric ring structure for the human snRNP core domain. Interestingly, preliminary results from the X-ray analysis of the PA-Sm1 protein also indicate a heptameric arrangement. Solution studies and immunoprecipitation experiments are under way to further characterize the archaeal Sm-related proteins and to identify their RNA targets.

[1] Kambach, C., Walke, S., Young, R., Avis, J.M., de la Fortelle, E., Raker, V., Lührmann, R., Li, J. & Nagai, K. (1999) Crystal structures of two Sm protein complexes and their implications for the assembly of the spliceosomal snRNPs. *Cell* 96, 375-387.