Supplementary information, Data S1

Cloning, expression and purification

Reverse transcription-polymerase chain reaction (RT-PCR) was employed to amplify the genes encoding full length Lsm1, Lsm2, Lsm3, Lsm5, Lsm6 and Lsm7, a C-terminal truncated Lsm4 (residues 1-117, designated as Lsm4N) and an N-terminal truncated Pat1 (residues 422-796 designated as Pat1C) from *S. cerevisiae*. For co-expression of Lsm2-3, the *Lsm2* gene was inserted into the multiple cloning sites 1 (MCS1) of the pETDuet-1 vector (Novagen) with an N-terminal His_6-tag fused to Lsm2 while the *Lsm3* gene was inserted into the MCS2. For co-expression of Lsm5-6, a modified pETDuet-1 vector with a PreScission protease cleavage site introduced between the N-terminal His_6-tag and the MCS1 was employed, in which the *Lsm5* gene was inserted into the MCS1 while the *Lsm6* gene was inserted into the MCS2. The *Lsm7* gene was inserted into the MCS2 of the pACYCDuet-1 vector (Novagen). The genes encoding Lsm1, Lsm4N and Pat1C were cloned into the pGEX-6p-1 vector (GE Healthcare) with an N-terminal GST-tag. All the constructs were verified by automated DNA sequencing.

E. coli strain BL21-CodonPlus (DE3)-RIL (Stratagene) was used to express Lsm1, Lsm4N, Pat1C and the Lsm2-3 complex at 37°C in Luria broth (LB) media containing ampicillin. For co-expression of Lsm5-6-7, the modified pETDuet-1 vector expressing Lsm5-6 and the pACYCDuet-1 vector expressing Lsm7 were co-transformed into BL21 (DE3) star strain and grown in LB media containing ampicillin and chloramphenicol at 37 °C. At an OD_{600} of 0.6, cells were induced with 0.1 mM isopropylthio-β-galactoside (IPTG) and grown at 18°C for an additional 12 hours prior to harvest. Cell pellets of Lsm2-3 and Lsm5-6-7 were resuspended and sonicated in buffer A containing 20 mM Hepes pH 7.5, 200 mM NaCl, 2 mM β-

mercaptoethanol and 5mM imidazole. Cell pellets of Lsm1, Lsm4N and Pat1C were resuspended and sonicated in buffer B containing 20 mM Tris pH 8.0, 500 mM NaCl, 2 mM DTT. Cell debris was removed by centrifugation at 18,000 rpm at 4 °C. The supernatant containing His_6-tagged proteins was incubated with TALON Co^{2+} column (Clontech Inc.) pre-equilibrated with buffer A and the supernatant containing GST-tagged proteins was applied to a glutathione Sepharose 4B column (GE Healthcare) pre-equilibrated in buffer B. The His_6-tagged Lsm2-3 and Lsm5-6-7 proteins were eluted in buffer A containing 200 mM imidazole while the GST-tagged Lsm1, Lsm4N and Pat1C proteins were eluted in buffer B containing 20 mM reduced glutathione. The His_6-tagged Lsm5-6-7 proteins and the GST-tagged proteins including Lsm1, Lsm4N and Pat1C were cleaved with PreScission protease at 4 °C overnight. The cleaved Lsm5-6-7 was desalted into buffer A without imidazole and loaded into a second TALON Co^{2+} column to remove the cleaved His_6-tag. The cleaved Lsm1, Lsm4N and Pat1C were desalted into buffer B and loaded into a second glutathione Sepharose 4B column to remove the cleaved GST-tag. The protein samples Lsm1, Lsm2-3, Lsm4N, Lsm5-6-7 and Pat1C were further purified by gel filtration on a Superdex-200 26/60 column (GE Healthcare). Lsm1, Lsm2-3 and Lsm5-6-7 were eluted using buffer C containing 20 mM Hepes pH 7.5, 100 mM NaCl, 2 mM Dithiothreitol (DTT) while Lsm4N and Pat1C were eluted using buffer D containing 20 mM Tris pH 8.0, 300 mM NaCl, 2 mM DTT. The identities of these purified proteins were confirmed by electro-spray mass spectrometry. All protein samples were concentrated to ~10 mg/ml with the exception of Lsm4N at a concentration of ~3 mg/ml. Lsm2-3 and Pat1C were mixed at equal molar ratio and incubated on ice for 1 hour before loading onto Superdex-200 26/60 in buffer C for purification of the Lsm2-3-Pat1C complex. Selenomethionine(SeMet)-substituted Lsm2-3 and Pat1C

were expressed in a minimal medium containing 20 mg/l SeMet, and purified as above. SeMet-substituted Lsm2-3-Pat1C complex was reconstituted as described above and concentrated to ~10 mg/ml in 20 mM Hepes pH 7.5, 100 mM NaCl, 10 mM DTT. NuPAGE Bis-Tris 4-12% gradient gel in NuPAGE MES SDS running buffer (Invitrogen) was used to check the purity of the target proteins by Coomassie blue staining.

Native polyacrylamide gel electrophoresis (PAGE)

Preparation of native PAGE was similar to that of Sodium Dodecyl Sulfate (SDS)-PAGE, except that SDS and DTT were omitted during all steps. Electrophoresis was conducted at a constant voltage of 130 V at room temperature for a separation time of approximately 150 min. 3~5 μg of each protein sample was used for analysis. Before loading into the gel, all protein samples were incubated in the binding buffer of 20 mM Hepes pH 7.5, 100 mM NaCl, at room temperature for 1 h. After electrophoretic separation, the gel was stained and destained as for a normal SDS-PAGE gel.

Fluorescence anisotropy assay

Fluorescence anisotropy measurements were performed with single-stranded RNA oligos U_{15}, A_{15} and C_{15} labeled at the 5' end with 6-carboxy-fluorescein (6-FAM) (Metabion International AG). 5' end labeled 6-FAM G_{15} was not used for measurement due to the difficulty of synthesis. 100 μl reactions were carried out in a 96-well plate at 25°C and measured using Safire II fluorescent plate reader (Tecan). 100 nM of the labeled oligonucleotide was incubated with increasing protein concentration (from 10 nM to 10 μM) at ambient temperature for 30 minutes prior to

measurements, in a buffer containing 20mM Hepes pH 7.5, 100mM NaCl. Measurement and calculation of dissociation constants (Kd) were carried out as described previously.

His-tag pull-down assay

The His-tagged Lsm2 wild type or C-terminal α helix truncation (residues 1-80; designated as Lsm2Δα2) was co-expressed with Lsm3 wild type or N-terminal α helix truncation (residues 14-89; designated as Lsm3Δα1) in the pETDuet-1 vector. The Pat1C mutant containing Ala substitutions of the seven residues including Leu479, Glu483, Glu487, Leu490, Lys531, Lys534 and Arg538 was generated using several rounds of two-step PCR and cloned into the pGEX-6p-1 vector. All these constructs were confirmed by automatic sequencing. These mutant proteins were expressed and purified using the same procedure as that described for the wild type proteins. For binding assays, his-tagged wild type and mutant Lsm2-3 proteins were incubated with 3μl of TALON metal affinity resin, respectively, at 4°C for 1 h. Subsequently, wild type or mutant Pat1C proteins were incubated with the beads immobilized with the His-tagged proteins at 4°C for 1 h. After extensive washing with the same binding buffer, the bound protein complexes were eluted by boiling in Laemmli sample buffer containing 5 mM EDTA and resolved on a NuPAGE Bis-Tris 4-12% gradient gel.

Analytical size exclusion chromatography

His-tagged Lsm2-3 wild type or mutants were mixed with wild type or mutant Pat1C at 4 °C for 1 h in the binding buffer as described for the His-tag pull-down assay with the molar ratio of Pat1C:Lsm2-3 at ~1.5:1 in each mixture. All the protein

mixtures were loaded onto Superdex^{TM} 200 HR 10/30 columns (GE Healthcare) equilibrated with the binding buffer and eluted at a flow rate of 0.8 ml/min.

Sedimentation velocity of analytical ultracentrifugation

Sedimentation velocity experiments were performed at 42000 r.p.m and at a temperature of 20 °C using a ProteomeLab XL-A analytical ultracentrifuge (Beckman Coulter) in quartz cells fitted with double-sector centerpieces. Absorption measurements were made at 150 s intervals at 280 nm until the boundaries reached the cell bottom. Prior to centrifugation, Lsm2-3 was dialyzed extensively into 20mM Hepes pH 7.5, 100mM NaCl. Three protein concentrations, 0.75, 1.0 and 1.5 mg/ml were used, which were measured using a NanoDrop Spectrophotometer 1000 with the molecular weight and extinction coefficient option. The SEDFIT program (SEDFIT version 12.52, http://www.analyticalultracentrifugation.com) was employed to calculate the protein partial specific volumes. The calculated protein partial specific volume of Lsm2-3 at 20 °C was 0.7325. SEDNTERP program (Sednterp version 1.09, http://www.rasmb.bbri.org) was used to calculate the solvent density and viscosity. The solvent density and viscosity were 1.00391 and 0.01026, respectively. The continuous c(S) distribution and continuous c(M) distribution methods from SEDFIT program (Schuck, 2000) were employed to analyze the data.

Analysis of Decapping *in vivo*

Wild-type constructs of Lsm2 (pRP2435) and Lsm3 (pRP2438) were constructed by PCR to include the endogenous promoters. Mutations in these genes were constructed using QuickChange (Stratagene) to give rise to *Lsm2*Δα2 (residues 1-80; pRP2436), and *Lsm3*Δα1 (residues 13-89; pRP2439). The original strains are

deleted for endogenous *Lsm2* or *Lsm3* but contained a plasmid. Thus the strains were transformed and selected on 5-FOA to shuffle in the WT or mutated versions of Lsm2-3. Confirmation of the new strains containing the wild-type or mutant versions of Lsm2 or Lsm3 was done and yielded new strains containing yRP2488 (Lsm2wt); yRP2489(Lsm2Δα2); yRP2891(Lsm3wt); and yRP2892(Lsm3Δα1). Strains were subsequently transformed with the reporter gene MFA2pG (pRP485) and PGK1pG (pRP469) for mRNA analysis as described above. Analysis of Pat1 was done in strain yRP1372 containing a Pat1Δ and integrated reporter mRNAs. Mutagenesis of WT Pat1 on plasmid pRP1478 was done by oligomutagenesis to create plasmid pRP2444 containing the 7 alanine mutations. Plasmids were transformed along with a vector control and analyzed by RNaseH cleavage with oRP1113 for the PGK1pG reporter and oligodT. Blots were probed with either oRP140(MFA2pG), or oRP141 and oRP154(PGK1pG)