Cell Reports
Report

Architecture of the Lsm1-7-Pat1 Complex: A Conserved Assembly in Eukaryotic mRNA Turnover

Humayun Sharif[1] and Elena Conti[1,*]
[1]Department of Structural Cell Biology, Max Planck Institute of Biochemistry, Am Klopferspitz 18, 82152 Martinsried, Germany
*Correspondence: conti@biochem.mpg.de
http://dx.doi.org/10.1016/j.celrep.2013.10.004

SUMMARY

The decay of mRNAs is a key step in eukaryotic gene expression. The cytoplasmic Lsm1-7-Pat1 complex is a conserved component of the 5′-to-3′ mRNA decay pathway, linking deadenylation to decapping. Lsm1-7 is similar to the nuclear Sm complexes that bind oligo-uridine tracts in snRNAs. The 2.3 Å resolution structure of *S. cerevisiae* Lsm1-7 shows the presence of a heptameric ring with Lsm1-2-3-6-5-7-4 topology. A distinct structural feature of the cytoplasmic Lsm ring is the C-terminal extension of Lsm1, which plugs the exit site of the central channel and approaches the RNA binding pockets. The 3.7 Å resolution structure of Lsm1-7 bound to the C-terminal domain of Pat1 reveals that Pat1 recognition is not mediated by the distinguishing cytoplasmic subunit, Lsm1, but by Lsm2 and Lsm3. These results show how the auxiliary domains and the canonical Sm folds of the Lsm1-7 complex are organized in order to mediate and modulate macromolecular interactions.

INTRODUCTION

RNA degradation modulates the steady-state levels of cellular transcripts and has emerged as a powerful mechanism for altering the abundance of proteins in response to changes in physiological conditions (reviewed in Schoenberg and Maquat, 2012). In eukaryotes, cytoplasmic mRNA turnover generally starts with the shortening of the poly(A) tail at the 3′ end of the message (reviewed in Chen and Shyu, 2011). The short stretch of adenosines that is left by the action of the deadenylases (Ccr4-Not and Pan2-Pan3) is the foothold for initiating two alternative decay pathways: the degradation of the RNA body in the 3′-to-5′ direction (via the exosome-Ski complex) or the removal of the 5′ cap structure and degradation in the 5′-to-3′ direction (via the decapping factors and Xrn1) (reviewed in Garneau et al., 2007). Genetic, biochemical, and structural data have shown that the core enzymes and regulators in mRNA turnover are evolutionarily conserved and have revealed the presence of intricate interaction networks (see the reviews above).

The conserved Lsm1-7-Pat1 complex plays an important role in coupling deadenylation and decapping in the 5′-to-3′ decay pathway (Bouveret et al., 2000; Tharun et al., 2000, Tharun, 2009; Haas et al., 2010; Ozgur et al., 2010; Totaro et al., 2011). Lsm1-7-Pat1 preferentially associates with the 3′ end of oligoadenylated mRNAs in vivo (Tharun et al., 2000; Tharun and Parker, 2001), protecting the last 20–30 nucleotides of the message (He and Parker, 2001). Lsm1-7-Pat1 subunits are required for normal rates of decapping in vivo (Bouveret et al., 2000; Tharun et al., 2000) and colocalize to discrete cytoplasmic foci known as P bodies along with all other 5′-to-3′ decay factors (Tharun et al., 2000; Pilkington and Parker, 2008; Haas et al., 2010; Ozgur et al., 2010). Lsm1-7 is composed of seven Sm-like proteins (numbered 1–7) and is related to the nuclear Sm complexes involved in binding small nuclear RNAs (snRNAs) (reviewed in Wilusz and Wilusz, 2005). Pat1 is a multifunctional protein. It binds the decapping complex Dcp1-Dcp2 (Pilkington and Parker, 2008; Braun et al., 2010; Nissan et al., 2010; Ozgur et al., 2010) as well as another decapping activator, Dhh1 (DDX6) (Braun et al., 2010; Haas et al., 2010; Nissan et al., 2010; Ozgur et al., 2010; Sharif et al., 2013). Pat1 has also been shown to interact with the Xrn1 exoribonuclease in yeast (Bouveret et al., 2000; Nissan et al., 2010) and with the Ccr4-Not deadenylase in *Drosophila* (Haas et al., 2010). Although many of these interactions are likely to be transient, the association of Pat1 with Lsm1-7 is sufficiently stable to allow the purification of the endogenous octameric complex from yeast (Bouveret et al., 2000; Chowdhury et al., 2007). In vitro, Lsm1-7-Pat1 binds directly polyuridine oligonucleotides with enhanced affinity when flanked by a short oligo-adenosine tail (Chowdhury et al., 2007). In yeast, Lsm1-7-Pat1 preferentially binds short oligo-uridine stretches located close to the 3′ end of endogenous mRNAs (Chowdhury et al., 2007; Mitchell et al., 2013).

Sm folds are uridine-specific RNA binding domains (Achsel et al., 2001). Structural studies have revealed how nuclear Sm complexes assemble into heteroheptameric rings around specific U-rich sequences of the U1 and U4 snRNAs (Pomeranz Krummel et al., 2009; Weber et al., 2010; Leung et al., 2011). In contrast to the spliceosomal Sm proteins, the two canonical Lsm complexes (Lsm1-7 and Lsm2-8) form rings spontaneously in the absence of RNA (Achsel et al., 1999; Salgado-Garrido et al., 1999). The cytoplasmic Lsm1-7 and the nuclear Lsm2-8 complexes share six of their seven subunits (reviewed in Wilusz and Wilusz, 2005). The distinguishing cytoplasmic subunit Lsm1 harbors critical determinants for the RNA binding properties of the complex not only in the canonical Sm domain but also in the distinctive C-terminal domain (Tharun et al.,

A

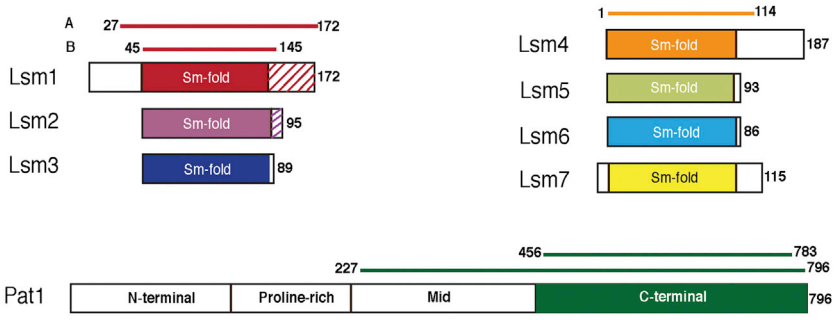


Figure 1. Structural Analysis of Lsm1-7 and Lsm1-7-Pat1 Complexes

(A) A schematic representation of the domain arrangements of the yeast Lsm1-7 and Pat1 proteins. Color-filled rectangles highlight the Sm folds of the Lsm proteins and the folded domain of Pat1. Dashed rectangles highlight the C-terminal extensions of Lsm1 and Lsm2. The residue numbers indicate the constructs used in this study.

(B) A table with data collection and refinement statistics. Values for the highest-resolution shell are given in parenthesis. Structure validation was carried out with Molprobity (Chen et al., 2010).

B

Dataset	$Lsm1\text{-}7_A$	$Lsm1_B\text{-}7\text{-}Pat1_C$
Space group	$P2_1$	$P6_3$
Cell dimensions		
a , *b*, *c* (Å)	61.80, 90.57, 68.46	258.69, 258.69, 47.10
α , β, γ (°)	90.0, 100.8, 90.0	90.0, 90.0, 120.0
Data Collection		
Wavelength (Å)	0.9980	0.8729
Resolution (Å)	90.53-2.30 (2.42-2.30)	84.68-3.70 (3.90-3.70)
R_{merge}	4.9 (41.2)	7.0 (68.9)
$I / \sigma I$	13.5 (3.1)	11.3 (1.4)
Completeness	99.3 (96.5)	99.4 (97.6)
Multiplicity	2.1 (2.0)	1.8 (1.6)
Refinement		
Resolution (Å)	53.99 (2.3)	43.40 (3.70)
No. of unique reflections	62641	36400
R_{work} / R_{free} (%)	21.1 / 25.8	24.9 / 29.5
B-factors	45.5	129.3
No. of atoms		
Proteins	4889	6277
Water	108	
Stereochemistry		
R.m.s.d. bond lengths (Å)	0.0080	0.0060
R.m.s.d. bond angles (°)	1.223	1.533
Ramachandran favored (%)	98.38	94.51
Ramachandran outliers (%)	0.00	0.24

2005; Chowdhury and Tharun, 2008; Chowdhury et al., 2012). The expectation is that Lsm1 might also specify other cytoplasmic-specific interactions. In this work, we address how the Sm folds and the distinct auxiliary domains of Lsm1-7 are structured and how they mediate the interaction with Pat1.

RESULTS AND DISCUSSION

Reconstitution of a Recombinant Lsm1-7-Pat1 Core Complex

The Lsm1–Lsm7 proteins contain a central Sm-like domain flanked by N-terminal and C-terminal extensions (Figure 1A). The Lsm1-7 complex is expected to be formed by hetero-oligomeric building blocks similar to those of the nuclear Sm complex, and Lsm2-3 corresponds to SmD1-D2, Lsm6-5-7 to SmF-E-G, and Lsm4-1 to SmD3/B (Kambach et al., 1999; Raker et al., 1999; Salgado-Garrido et al., 1999; Bouveret et al., 2000; Zaric et al., 2005; Mund et al., 2011). We obtained recombinant *S. cerevisiae* Lsm2-3 and Lsm5-6-7 from coexpression constructs (a gift of K. Nagai and Y. Kondo). *S. cerevisiae* Lsm1 and Lsm4 were expressed individually. In the latter case, the polypeptide (residues 1–114, hereby referred to as Lsm4) lacked the long Q- and N-rich C-terminal extension (Figure 1A). This region of Lsm4 is proteolytically sensitive in vitro (K. Nagai and Y. Kondo, personal communication) and is not required for normal mRNA decay rates in vivo (Decker et al., 2007). Reconstitution of the Lsm1-7 complex included a mild denaturing step (Zaric et al., 2005) that most likely overcomes the tendency of these proteins to oligomerize unspecifically when at high concentrations in vitro.

S. cerevisiae Pat1 (796 residues) is a multidomain protein with N-terminal, proline-rich, Mid, and C-terminal domains (Figure 1A) (Haas et al., 2010; Nissan et al., 2010). The C-terminal domain is required and sufficient to bind Lsm1-7 in both yeast (Nissan et al., 2010) and humans (Braun et al., 2010). The Mid domain has also been shown to contribute to Lsm1-7 binding (Pilkington and Parker, 2008; Braun et al., 2010) and, in the case of the *Drosophila* ortholog, provides a major interaction site (Haas et al., 2010). We expressed and purified a portion of *S. cerevisiae* Pat1 including both the Mid and C-terminal domains (residues 220–796), formed a complex with the reconstituted Lsm1-7, and subjected the octameric assembly to limited proteolysis (Figure S1A). Treatment with the protease elastase resulted in the accumulation of truncated Pat1 and

Lsm1 proteins, whereas all other subunits of the complex remained stable (Figure S1A). N-terminal sequencing and mass spectrometry analysis mapped the proteolytic fragment of Pat1 to the C-terminal domain (residues 450–796, hereby referred to as $Pat1_C$). Full-length Lsm1 (172 residues) was proteolyzed into different fragments that started at residues 27 or 45 and ended at residues 145 or 159. Previous studies have shown that the N-terminal extension of Lsm1 is functionally dispensable in yeast (Tharun et al., 2005). In contrast, the conserved C-terminal extension (also known as C-terminal domain or CTD) is required for RNA binding in vitro and for Lsm1 function in vivo (Tharun et al., 2005; Chowdhury et al., 2012). Therefore, we engineered a construct of Lsm1 containing residues 27–172 (referred to as $Lsm1_A$) and purified the corresponding $Lsm1_A$-7 and $Lsm1_A$-7-$Pat1_C$ complexes for structural analysis.

Lsm1-7 Is an Sm-like Heptameric Ring

We obtained crystals of yeast $Lsm1_A$-7 diffracting at 2.3 Å resolution and solved the structure by molecular replacement with the coordinates of known Sm-like rings (Leung et al., 2011; Mund et al., 2011). The final model is refined to an R_{free} of 25.8% and an R_{work} of 21.1% with good stereochemistry (Figures 1B and S1B). The model includes most of the polypeptide chains (see the Supplemental Information) and, in addition, includes ten residues of the tag engineered in Lsm2 for purification purposes (Figure 2A). Each of the seven Lsm proteins contains the characteristic Sm fold, a β barrel of five antiparallel and highly bent β strands with an N-terminal α helix (helix α1) on top. The seven Sm domains pack side by side in a ring-like architecture with a flat surface on top (the so-called proximal face, where the α helix is positioned) and the so-called tapered or distal surface at the bottom (Figure 2A). The overall oligomeric structure of the Lsm1-7 ring is generally similar to that of the nuclear U4 snRNP core (Leung et al., 2011). The oligomerization is based on the same repeating principle, namely the β4 strand of one subunit packing against the β5 strand of the neighboring subunit (Figure 2A, left). This results in an intermolecular β sheet that scaffolds the ring via extensive hydrophobic contacts.

The order of the subunits in the ring is Lsm1-2-3-6-5-7-4, as predicted previously (Kambach et al., 1999; Raker et al., 1999; Salgado-Garrido et al., 1999; Bouveret et al., 2000; Zaric et al., 2005; Mund et al., 2011) (Figure 2A, left). Specificity in the hetero-oligomerization is dictated by subunit-specific interactions, typically electrostatic contacts along the outer and inner circumferences of the ring. Several conserved electrostatic pairs are observed within the Lsm2-3 and Lsm6-5-7 building blocks (for example, $Glu18^{Lsm2}$-$Arg61^{Lsm3}$, $Asp22^{Lsm2}$-$Arg24^{Lsm3}$, and $Glu29^{Lsm5}$-$Arg87^{Lsm7}$) (Figures 2B and S2). In addition, conserved electrostatic pairs also occur between building blocks (for example, $Lys19^{Lsm3}$-$Glu54^{Lsm6}$, $Lys41^{Lsm7}$-$Glu23^{Lsm4}$, and $Arg62^{Lsm1}$-$Glu48^{Lsm4}$) (Figures 2B and S2) and most likely contribute to the ability of the Lsm subunits to assemble in a preformed ring without the need of an RNA molecule to nucleate hetero-oligomerization (Achsel et al., 1999; Salgado-Garrido et al., 1999). We note that the corresponding Sm building blocks lack some of the charged pairs at the equivalent positions, suggesting why mixed assemblies of Sm and Lsm proteins might not be favored.

The Lsm1-7 Ring Is Complemented by the Distinct C-Terminal Extensions of Lsm1 and Lsm2

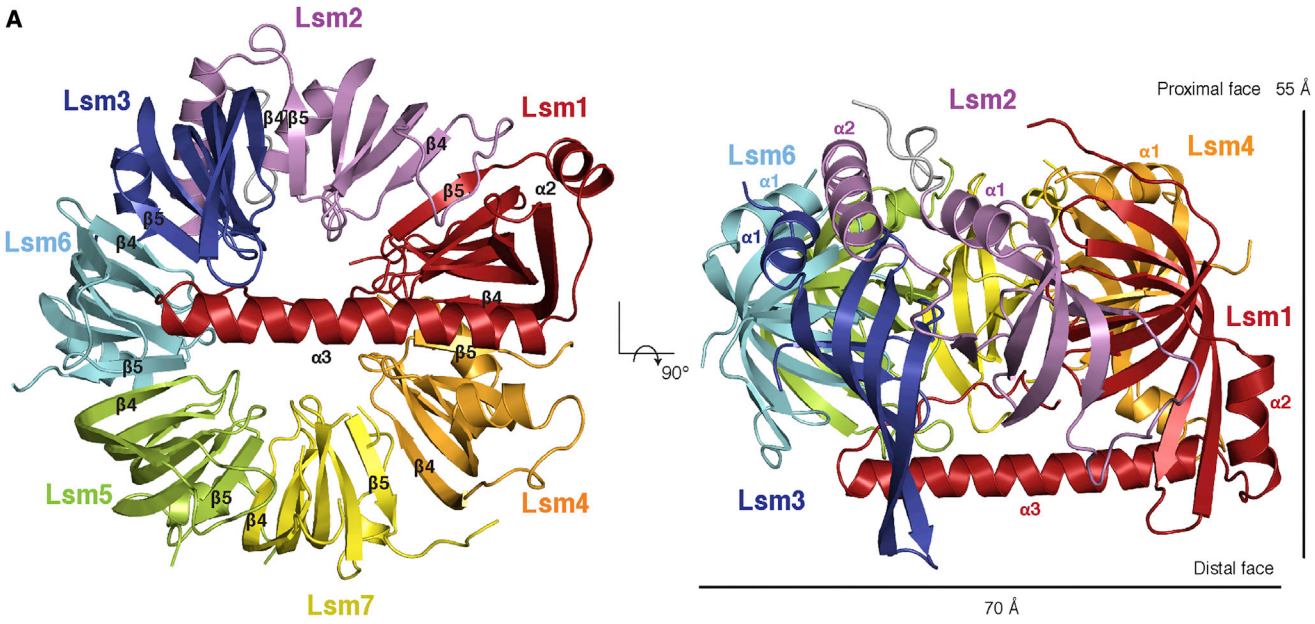
The Lsm1-7 ring has distinct structural features in comparison to known Sm rings. First, the C-terminal extension of Lsm2 (residues 72–94) forms a short α helix (α2) that lies on the proximal face of the ring between the α1 helices of Lsm2 and Lsm3 (Figure 2A, right). The second and most striking feature is the C-terminal extension of Lsm1 (residues 115–172). This domain starts at the proximal face of the ring, wraps around the outer surface of the Lsm1 β barrel with a short α helix (α2), and reaches the distal face (Figure 2A, right). Here, it forms a long α helix (α3) that traverses the diameter of the ring, interacting with Lsm1 and Lsm4 on one side and with Lsm3 and Lsm6 on the other (Figure 2A, left). Then, the polypeptide makes a sharp bend (at the conserved Gly162) and stretches in an antiparallel fashion on top of helix α3 (Figure 3, left). Several evolutionarily conserved interactions hook the very C terminus of Lsm1 inside the ring; $Tyr172^{Lsm1}$ fits in a pocket created between Lsm1 and Lsm4, whereas $Asp170^{Lsm1}$ interacts electrostatically with $Arg59^{Lsm1}$. The extensive interactions we observe in the structure rationalize why the C-terminal extension of Lsm1 is able to function even when in trans (Chowdhury et al., 2012).

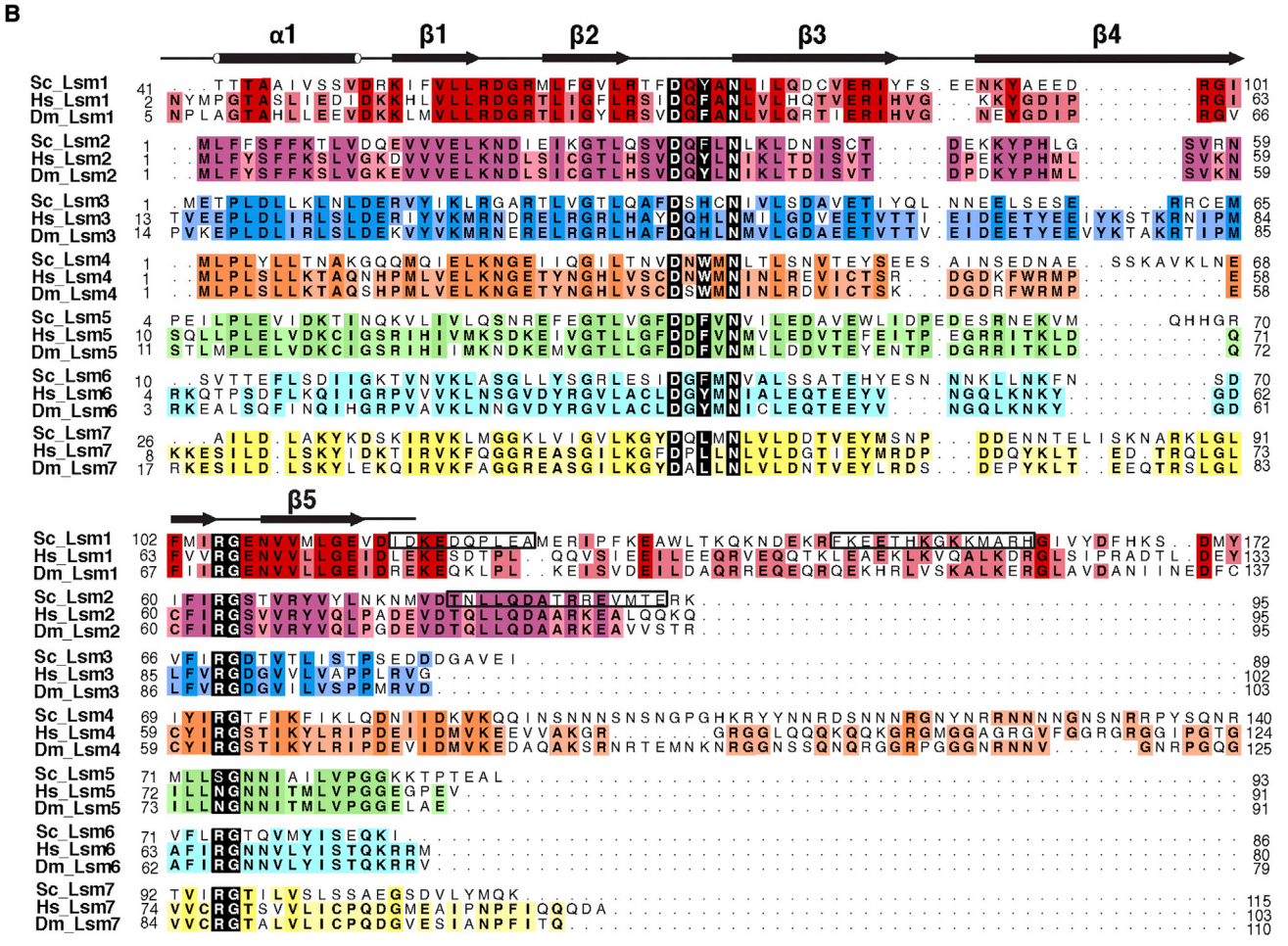
Deletion of the C-terminal extension of Lsm1 has been shown to decrease the RNA binding affinity of the complex but not to prevent the specific recognition of U tracts (Chowdhury et al., 2012). Superposition of the $Lsm1_A$-7 structure with that of the U4 snRNP core (Leung et al., 2011) allowed us to examine the putative RNA binding path (Figure 3). In the nuclear Sm complex, RNA binds at the consecutive uridine binding pockets that line the inner circumference of the ring and then threads through the entire central channel to exit with the 3′ end at the distal surface (Leung et al., 2011) (Figure 3, right). The uridine-binding pockets are created by the so-called Sm1 and Sm2 sequence motifs and are also present in Lsm1-7 (Figure 2B). Therefore, the Lsm proteins are expected to engage U bases with stacking and hydrogen-bonding interactions similar to those observed in the nuclear Sm complex (Leung et al., 2011). In contrast, RNA cannot exit the Lsm1-7 ring with the same path observed in the U4 snRNP structure because it would clash against the C-terminal extension of Lsm1 (Figures 3 and S3). It is possible that, although it poses considerable steric hindrance at the exit site of the channel, the Lsm1 extension might still allow the RNA 3′ end to thread through a narrow hole leading to the distal face of the ring. However, it is also possible that the C-terminal extension prevents the RNA from threading through the entire channel. In this case, we note that the unhindered part of the channel would be able to fit two to three additional nucleotides after the last uridine expected from the U4 snRNP structure (U836) (Figure 3).

The Pat1 C-Terminal Domain Protrudes on the Side of the Lsm1-7 Ring

Next, we addressed how Lsm1-7 binds Pat1. The $Lsm1_A$-7-$Pat1_C$ complex yielded diffracting crystals, but we found that only $Lsm1_A$-7 was present in the asymmetric unit. Inspection of the lattice suggested the presence of a possibly unfavorable crystal contact involving the C-terminal extension of Lsm1. Given that the Lsm1 C-terminal extension is not required for

Cell PRESS
Open ACCESS





(legend on next page)

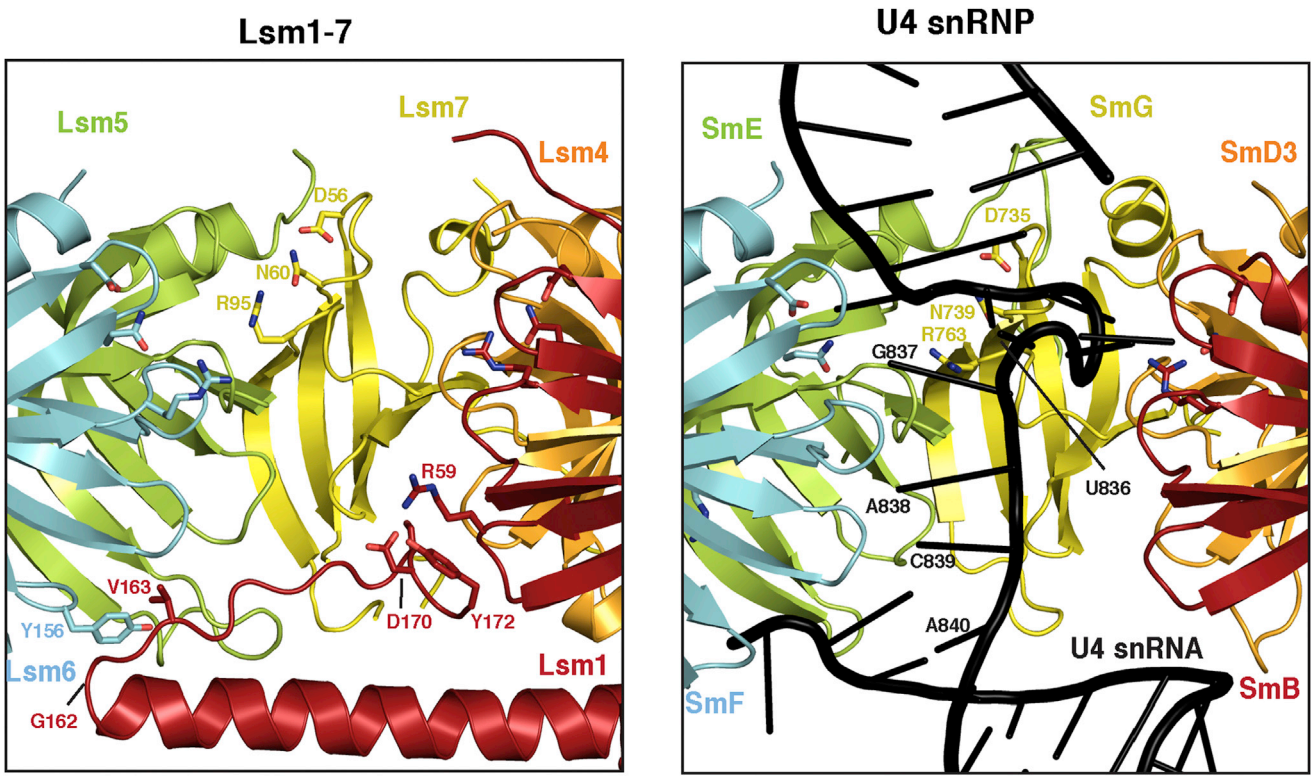


Figure 3. The C-Terminal Extension of Lsm1 Obstructs the RNA Exit Site

The C-terminal extension of Lsm1 binds in the central channel. On the left, a zoom-in of the $Lsm1_A$-7 structure is shown in the same orientation as in the left panel of Figure 2A. On the right is the corresponding zoom-in view of the U4 snRNP core structure (Leung et al., 2011) after optimal superposition of the Sm subunits, and RNA is shown in black. The Asp and Asn residues of the Sm1 motif and the Arg residue of the Sm2 motif indicate the position of the uridine-binding pocket of Lsm7 (left) and SmG (right). Lsm2-3 and SmD1-D2 have been removed for clarity. The Lsm1 C-terminal extension is 12–15 Å away from the corresponding position of U836 in the U4 snRNP structure and would clash against a nucleotide at the corresponding position of A840.

the assembly of the Lsm ring or Pat1 binding (Chowdhury et al., 2012), we generated an Lsm1 construct encompassing residues 45–145 (referred to as $Lsm1_B$), which corresponds to its smallest proteolytic fragment (Figure S1A). The $Lsm1_B$-7-$Pat1_C$ complex yielded crystals diffracting to 3.7 Å resolution. We solved the structure by molecular replacement using $Lsm1_A$-7 and a homology model of $Pat1_C$ based on the crystal structure of the human ortholog previously crystallized in isolation (Braun et al., 2010). The structure is refined to an R_{free} of 29.5% and an R_{work} of 24.9% with good stereochemistry (Figures 1B and S1B). The final model includes essentially all the residues of $Lsm1_B$-7 and residues 471–783 of Pat1.

S. cerevisiae $Pat1_C$ is an elongated domain formed by helical hairpins related to the ARM repeat and HEAT repeat family of proteins (Figures 4A and S4A). Each hairpin is composed of two antiparallel α helices (termed A and B) connected by loops or additional helical segments (Andrade et al., 2001). The hairpins pack side by side with a right-handed twist, forming a superhelix with a layer of six A helices on one side and a layer of five B helices on the other. However, the hairpins of Pat_C are rather irregular in both length and curvature. A comparison of the structure of yeast $Pat1_C$ with that of the human ortholog (Braun et al., 2010) shows that the first four hairpins are very similar, whereas the last two differ substantially in the position and orientation of the individual helices (Figure S4B). The $Pat1_C$ superhelix binds with the N-terminal repeats to the outer surface of Lsm1-7 and projects into solvent, the last hairpin being positioned more than 40 Å away from Lsm1-7 (Figure 4A).

Figure 2. Canonical and Idiosyncratic Features of the Lsm1-7 Ring

(A) The crystal structure of *S. cerevisiae* $Lsm1_A$-7 is shown in two orientations related by a 90° rotation around a horizontal axis. In the left panel, the complex is viewed on the distal face. The β4 and β5 strands of each subunit are labeled. The right panel is a side view of the complex. The N-terminal helices (α1 and α2) in the C-terminal extension of Lsm2 are labeled as well as the helices in the C-terminal extension of Lsm1 (α2 and α3). The additional residues of the Lsm2 tag are highlighted in gray in the right panel.

(B) Structure-based sequence alignment of Lsm1-7 orthologs from *S. cerevisiae* (Sc), *H. sapiens* (Hs), and *D. melanogaster* (Dm). Conserved residues are in colored boxes. In black are the central residues of the Sm1 motif (the D-x-Φ-x-N sequence, Φ being a hydrophobic residue) and the Sm2 motif (the R-G sequence). The canonical secondary structures of the Sm cores are above the sequences. The additional helices in the extensions of Lsm1 and Lsm2 are boxed around the corresponding sequences.

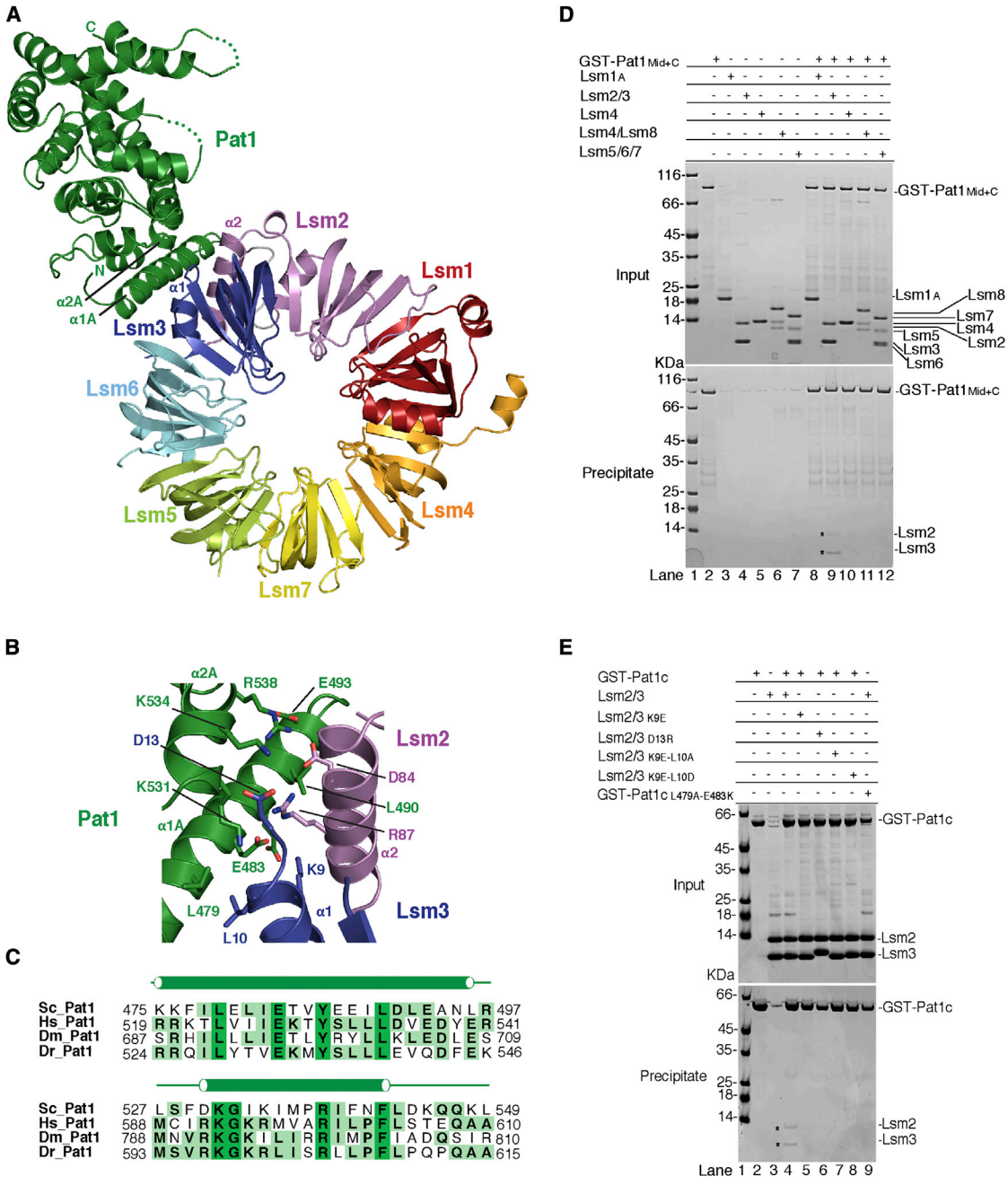


Figure 4. Pat1$_C$ Binds the Lsm1-7 Ring at Lsm2 and Lsm3

(A) Structure of Lsm1$_B$-7-Pat1$_C$ viewed with the ring in the same orientation and colors as in Figure 2A, left. Pat1 is shown in green, and the interacting helices of the HEAT-repeat-like superhelix is indicated.

(B) A zoom-in view of the interaction interface between Pat1, Lsm2, and Lsm3 with conserved interacting residues highlighted and indicated.

(C) Evolutionary conservation of the Lsm2-3 binding region of Pat1$_C$ (helices 1A and 2A).

(D) Pull-down experiments of GST-tagged Pat1 residues (Mid + C-terminal domains) with untagged Lsm2-3, Lsm5-6-7, Lsm4, and Lsm1. Input samples (top) and samples precipitated on glutathione-agarose beads (bottom) were analyzed on 4%–12% Bis-Tris NuPage gel with 2-(N-morpholino)ethanesulfonic acid running buffer. The proteins corresponding to the bands are indicated on the right side of both panels. Asterisks indicate the precipitated bands.

(E) Pull-down experiments of GST-tagged Lsm2-3 (WT and mutants) with WT Pat1$_C$ and of GST-tagged Pat1$_C$ WT and mutants with WT Lsm2-3. The experiments were carried out and are shown as described in (C).

Conserved Interactions of Pat1 with Lsm2 and Lsm3

$Pat1_C$ binds Lsm1-7 at the Lsm2 and Lsm3 subunits (Figure 4A). Helices 1A and 2A of $Pat1_C$ dock onto the C-terminal extension of Lsm2 (helix α2) and onto the canonical helix of the Lsm3 core (α1) (Figure 4B). A comparison of the yeast $Lsm1_B$-7-$Pat1_C$ structure with that of human $Pat1_C$ and of yeast $Lsm1_A$-7 in isolation shows that the interacting regions do not undergo significant conformational changes upon binding. Pat1, Lsm2, and Lsm3 are involved in an intricate set of electrostatic interactions. $Glu483^{Pat1}$ interacts with $Arg87^{Lsm2}$ and $Lys9^{Lsm3}$. Salt bridges also occur between $Arg538^{Pat1}$-$Asp84^{Lsm2}$ and $Lys534^{Pat1}$-$Asp13^{Lsm3}$. In addition, hydrophobic contacts engage $Leu479^{Pat1}$, $Tyr486^{Pat1}$, and $Leu490^{Pat1}$ with $Arg87^{Lsm2}$ and $Leu10^{Lsm3}$. All these residues are evolutionarily conserved (Figures 2B and 4C), suggesting that the metazoan orthologs share a similar recognition mechanism. Indeed, substitution of some of the equivalent residues in quadruple mutations of human Pat1 have been shown to impair Lsm1 binding in coimmunoprecipitation assays (Braun et al., 2010).

The structural analysis predicts that the Lsm2-Lsm3 subcomplex is sufficient for $Pat1_C$ binding. We tested this hypothesis in GST pull-down assays. A GST-Pat1 polypeptide encompassing both the Mid and C-terminal domains was indeed able to precipitate Lsm2-3 and not Lsm1, Lsm4, or Lsm5-6-7 (Figure 4D). Next, we engineered specific mutations. Consistent with the structure, GST-$Pat1_C$ was unable to precipitate Lsm2-3 upon mutation of Lsm2 Lys9Glu, Leu10Asp. Furthermore, mutation of Leu479Ala, Glu483Lys in GST-$Pat1_C$ impaired the interaction with wild-type (WT) Lsm2-3 (Figure 4E). The structural analysis also suggests a possible mechanism for the recognition of the Mid domain of Pat1. In both Lsm1-7 structures, part of the tag of Lsm2 is well ordered and wedges between the α1 and α2 helices of Lsm2 (Figures 2A, 4A, and S1B). The Asn-Leu-Tyr-Phe-Gln sequence of the tag interacts with evolutionarily conserved residues of Lsm2 (Leu2, Lys8, and Thr9 on α1 and Leu82 and Ala85 on α2). Interestingly, the Mid domain of Pat1 contains a similar stretch of amino acids (Asp-Phe-Tyr-Phe-Gln, residues 304–308 in *S. cerevisiae* Pat1) that are highly conserved and embedded in a predicted unstructured region, both of which are typical features of short linear motifs (Davey et al., 2012). Thus, it is possible that the tag serendipitously mimics a short linear motif in the Mid domain of Pat1.

Concluding Remarks

The Lsm1-7 complex contains an Sm ring with auxiliary structural features. The C-terminal extension that is characteristic of Lsm1 partially occupies the internal channel of the Sm-like ring. This extension approaches the RNA binding pockets of Lsm1-7, providing a rationale for the observation that it enhances the RNA binding properties of the core (Chowdhury et al., 2012). However, the basis for the specific recognition of U tract RNAs presenting a short oligo-A tail is currently unclear and an important question for future studies. The C-terminal extension that is characteristic of Lsm2, and the canonical Sm domains of Lsm2-Lsm3 create preformed protein-protein interaction sites. The C-terminal domain of Pat1 binds a composite surface of Lsm2 and Lsm3 with a rather rigid recognition mechanism between folded domains. We speculate that the unstructured Mid domain of Pat1 (Braun et al., 2010) might flexibly dock to an adjacent pocket of Lsm2 and enhance binding affinity. Counterintuitively, the binding determinants for Pat1 are not provided by Lsm1, the subunit of the cytoplasmic Lsm1-7 complex that differs from the nuclear Lsm2-8 complex. This finding has several implications. First, the localization of these proteins to distinct subcellular compartments (Reijns et al., 2009) is likely to provide a key contribution to binding specificity. Second, the interaction surfaces of Lsm2-3 that we identified for the cytoplasmic Lsm1-7-Pat1 complex might also be involved in protein-protein recognition in the nucleus in the context of the nuclear Lsm2-8 complex. An interesting candidate for Lsm2-8 binding is the splicing factor Prp8, which appears to contain a sequence similar to the Lsm2-Lsm3 binding region of Pat1 (data not shown). Given that, in human cells, Pat1 is a shuttling protein with transient nuclear localization (Marnef et al., 2012), it is also possible that Pat1 itself might interact with Lsm2-8, rationalizing how Pat1 might exert its nuclear functions.

EXPERIMENTAL PROCEDURES

Protein Purification and Binding Assays

S. cerevisiae Lsm1-7 complexes were formed by mixing purified Lsm1, Lsm4, Lsm2-3, and Lsm5-6-7 in a 2:2:1:1 ratio and reconstituted essentially as described previously (Zaric et al., 2005). All Pat1 fragments were cloned as either TEV-cleavable His_6-ZZ-tagged or His_6-GST-tagged proteins and purified with standard procedures. The octameric complex was reconstituted by incubating the individually purified proteins in a 1:1.5 molar ratio of Lsm1-7 and Pat1 for 1 hr at 4°C. The complex was purified further by size-exclusion chromatography (Superdex 200) in a buffer containing 20 mM Tris (pH 7.4), 150 mM NaCl, and 1 mM dithiothreitol. For in vitro pull-down experiments, point mutations were introduced with QuikChange site-directed mutagenesis according to the manufacturer's instruction (Stratagene). Mutants were purified by similar protocol as for the WT Lsm1-7 and Lsm2-3. The pull-down assays were carried out as described previously (Sharif et al., 2013). Protocols are detailed in the Supplemental Information.

Crystal Structure Determination

$Lsm1_A$-7 and Lsm_B1-7-$Pat1_C$ yielded crystals (conditions are detailed in the Supplemental Information) that diffracted to 2.3 Å and 3.7 Å resolution, respectively, with Swiss Light Source and European Synchrotron Radiation Facility synchrotron radiation. Data were processed with XDS (Kabsch, 2010), and the structures were solved by molecular replacement with Phaser (McCoy et al., 2007). The atomic models were built with Coot (Emsley et al., 2010) and refined with PHENIX (Adams et al., 2010). The data collection and refinement statistics are summarized in Figure 1B.

ACCESSION NUMBERS

The coordinates and structure factors have been deposited in the Protein Data Bank under accession numbers 4C92 for $Lsm1_A$-7 and 4C8Q for $Lsm1_B$-7-$Pat1_C$.

SUPPLEMENTAL INFORMATION

Supplemental Information contains Supplemental Experimental Procedures and four figures and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2013.10.004.

ACKNOWLEDGMENTS

We are grateful to Kiyoshi Nagai and Yasushi Kondo at the Medical Research Council Laboratory of Molecular Biology (Cambridge, UK) for sharing their

yeast Lsm2-8 expression constructs and protocols. We thank the Max Planck Institute (MPI) Crystallization Facility for screenings and optimization, the MPI Core Facility for mass spectrometry and N-terminal sequencing, and the beamline scientists at the Swiss Light Source and European Synchrotron Radiation Facility for excellent assistance with data collection. We also thank Steffen Schüssler and Marc Baumgärtner for help with purification and the members of our lab for discussions and critical reading of the manuscript. This study was supported by the Max-Planck-Gesellschaft, the European Commission (ERC Advanced Investigator grant 294371 and Marie Curie ITN RNPnet), and the Deutsche Forschungsgemeinschaft (DFG SFB646, SFB1035, GRK1721, FOR1680, and CIPSM) to E.C.

Received: September 24, 2013
Revised: October 3, 2013
Accepted: October 3, 2013
Published: October 17, 2013

REFERENCES

Achsel, T., Brahms, H., Kastner, B., Bachi, A., Wilm, M., and Lührmann, R. (1999). A doughnut-shaped heteromer of human Sm-like proteins binds to the 3′-end of U6 snRNA, thereby facilitating U4/U6 duplex formation in vitro. EMBO J. *18*, 5789–5802.

Achsel, T., Stark, H., and Lührmann, R. (2001). The Sm domain is an ancient RNA-binding motif with oligo(U) specificity. Proc. Natl. Acad. Sci. USA *98*, 3685–3689.

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. *66*, 213–221.

Andrade, M.A., Petosa, C., O'Donoghue, S.I., Müller, C.W., and Bork, P. (2001). Comparison of ARM and HEAT protein repeats. J. Mol. Biol. *309*, 1–18.

Bouveret, E., Rigaut, G., Shevchenko, A., Wilm, M., and Séraphin, B. (2000). A Sm-like protein complex that participates in mRNA degradation. EMBO J. *19*, 1661–1671.

Braun, J.E., Tritschler, F., Haas, G., Igreja, C., Truffault, V., Weichenrieder, O., and Izaurralde, E. (2010). The C-terminal alpha-alpha superhelix of Pat is required for mRNA decapping in metazoa. EMBO J. *29*, 2368–2380.

Chen, C.-Y.A., and Shyu, A.-B. (2011). Mechanisms of deadenylation-dependent decay. Wiley Interdiscip Rev RNA *2*, 167–183.

Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. *66*, 12–21.

Chowdhury, A., and Tharun, S. (2008). lsm1 mutations impairing the ability of the Lsm1p-7p-Pat1p complex to preferentially bind to oligoadenylated RNA affect mRNA decay in vivo. RNA *14*, 2149–2158.

Chowdhury, A., Mukhopadhyay, J., and Tharun, S. (2007). The decapping activator Lsm1p-7p-Pat1p complex has the intrinsic ability to distinguish between oligoadenylated and polyadenylated RNAs. RNA *13*, 998–1016.

Chowdhury, A., Raju, K.K., Kalurupalle, S., and Tharun, S. (2012). Both Sm-domain and C-terminal extension of Lsm1 are important for the RNA-binding activity of the Lsm1-7-Pat1 complex. RNA *18*, 936–944.

Davey, N.E., Van Roey, K., Weatheritt, R.J., Toedt, G., Uyar, B., Altenberg, B., Budd, A., Diella, F., Dinkel, H., and Gibson, T.J. (2012). Attributes of short linear motifs. Mol. Biosyst. *8*, 268–281.

Decker, C.J., Teixeira, D., and Parker, R. (2007). Edc3p and a glutamine/asparagine-rich domain of Lsm4p function in processing body assembly in Saccharomyces cerevisiae. J. Cell Biol. *179*, 437–449.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. *66*, 486–501.

Garneau, N.L., Wilusz, J., and Wilusz, C.J. (2007). The highways and byways of mRNA decay. Nat. Rev. Mol. Cell Biol. *8*, 113–126.

Haas, G., Braun, J.E., Igreja, C., Tritschler, F., Nishihara, T., and Izaurralde, E. (2010). HPat provides a link between deadenylation and decapping in metazoa. J. Cell Biol. *189*, 289–302.

He, W., and Parker, R. (2001). The yeast cytoplasmic Lsm1/Pat1p complex protects mRNA 3′ termini from partial degradation. Genetics *158*, 1445–1455.

Kabsch, W. (2010). Integration, scaling, space-group assignment and post-refinement. Acta Crystallogr. D Biol. Crystallogr. *66*, 133–144.

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

Leung, A.K.W., Nagai, K., and Li, J. (2011). Structure of the spliceosomal U4 snRNP core domain and its implication for snRNP biogenesis. Nature *473*, 536–539.

Marnef, A., Weil, D., and Standart, N. (2012). RNA-related nuclear functions of human Pat1b, the P-body mRNA decay factor. Mol. Biol. Cell *23*, 213–224.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J. Appl. Cryst. *40*, 658–674.

Mitchell, S.F., Jain, S., She, M., and Parker, R. (2013). Global analysis of yeast mRNPs. Nat. Struct. Mol. Biol. *20*, 127–133.

Mund, M., Neu, A., Ullmann, J., Neu, U., and Sprangers, R. (2011). Structure of the LSm657 complex: an assembly intermediate of the LSm1-7 and LSm2-8 rings. J. Mol. Biol. *414*, 165–176.

Nissan, T., Rajyaguru, P., She, M., Song, H., and Parker, R. (2010). Decapping activators in Saccharomyces cerevisiae act by multiple mechanisms. Mol. Cell *39*, 773–783.

Ozgur, S., Chekulaeva, M., and Stoecklin, G. (2010). Human Pat1b connects deadenylation with mRNA decapping and controls the assembly of processing bodies. Mol. Cell. Biol. *30*, 4308–4323.

Pilkington, G.R., and Parker, R. (2008). Pat1 contains distinct functional domains that promote P-body assembly and activation of decapping. Mol. Cell. Biol. *28*, 1298–1312.

Pomeranz Krummel, D.A., Oubridge, C., Leung, A.K.W., Li, J., and Nagai, K. (2009). Crystal structure of human spliceosomal U1 snRNP at 5.5 A resolution. Nature *458*, 475–480.

Raker, V.A., Hartmuth, K., Kastner, B., and Lührmann, R. (1999). Spliceosomal U snRNP core assembly: Sm proteins assemble onto an Sm site RNA nonanucleotide in a specific and thermodynamically stable manner. Mol. Cell. Biol. *19*, 6554–6565.

Reijns, M.A.M., Auchynnikava, T., and Beggs, J.D. (2009). Analysis of Lsm1p and Lsm8p domains in the cellular localization of Lsm complexes in budding yeast. FEBS J. *276*, 3602–3617.

Salgado-Garrido, J., Bragado-Nilsson, E., Kandels-Lewis, S., and Séraphin, B. (1999). Sm and Sm-like proteins assemble in two related complexes of deep evolutionary origin. EMBO J. *18*, 3451–3462.

Schoenberg, D.R.D., and Maquat, L.E.L. (2012). Regulation of cytoplasmic mRNA decay. Nat. Rev. Genet. *13*, 246–259.

Sharif, H., Ozgur, S., Sharma, K., Basquin, C., Urlaub, H., and Conti, E. (2013). Structural analysis of the yeast Dhh1-Pat1 complex reveals how Dhh1 engages Pat1, Edc3 and RNA in mutually exclusive interactions. Nucleic Acids Res. *41*, 8377–8390.

Tharun, S. (2009). Lsm1-7-Pat1 complex: a link between 3′ and 5′-ends in mRNA decay? RNA Biol. *6*, 228–232.

Tharun, S., and Parker, R. (2001). Targeting an mRNA for decapping: displacement of translation factors and association of the Lsm1p-7p complex on deadenylated yeast mRNAs. Mol. Cell *8*, 1075–1083.

Tharun, S., He, W., Mayes, A.E., Lennertz, P., Beggs, J.D., and Parker, R. (2000). Yeast Sm-like proteins function in mRNA decapping and decay. Nature *404*, 515–518.

Tharun, S., Muhlrad, D., Chowdhury, A., and Parker, R. (2005). Mutations in the Saccharomyces cerevisiae LSM1 gene that affect mRNA decapping and 3′ end protection. Genetics *170*, 33–46.

Totaro, A., Renzi, F., La Fata, G., Mattioli, C., Raabe, M., Urlaub, H., and Achsel, T. (2011). The human Pat1b protein: a novel mRNA deadenylation factor identified by a new immunoprecipitation technique. Nucleic Acids Res. *39*, 635–647.

Weber, G., Trowitzsch, S., Kastner, B., Lührmann, R., and Wahl, M.C. (2010). Functional organization of the Sm core in the crystal structure of human U1 snRNP. EMBO J. *29*, 4172–4184.

Wilusz, C.J., and Wilusz, J. (2005). Eukaryotic Lsm proteins: lessons from bacteria. Nat. Struct. Mol. Biol. *12*, 1031–1036.

Zaric, B., Chami, M., Rémigy, H., Engel, A., Ballmer-Hofer, K., Winkler, F.K., and Kambach, C. (2005). Reconstitution of two recombinant LSm protein complexes reveals aspects of their architecture, assembly, and function. J. Biol. Chem. *280*, 16066–16075.