Open

ORIGINAL ARTICLE

npg

Lsm2 and Lsm3 bridge the interaction of the Lsm1-7 complex with Pat1 for decapping activation

Donghui Wu[1], Denise Muhlrad[2], Matthew W Bowler[3, 4], Shimin Jiang[1], Zhou Liu[5], Roy Parker[2], Haiwei Song[1, 5, 6]

[1]Institute of Molecular and Cell Biology, 61 Biopolis Drive, Singapore 138673; [2]Department of Biochemistry and Howard Hughes Medical Institute, University of Colorado, Boulder, Boulder CO 80303, USA; [3]European Molecular Biology Laboratory, 6 rue Jules Horowitz, BP 181, 38042, Grenoble, France; [4]Unit of Virus Host-Cell Interactions, UJF-EMBL-CNRS, UMI 3265, 6 rue Jules Horowitz, 38042 Grenoble Cedex 9, France; [5]Life Sciences Institute, Zhejiang University, 388 Yuhangtang Road, Hangzhou, Zhejiang 310000, China; [6]Department of Biochemistry, National University of Singapore, 14 Science Drive, Singapore 117543

The evolutionarily conserved Lsm1-7-Pat1 complex is the most critical activator of mRNA decapping in eukaryotic cells and plays many roles in normal decay, AU-rich element-mediated decay, and miRNA silencing, yet how Pat1 interacts with the Lsm1-7 complex is unknown. Here, we show that Lsm2 and Lsm3 bridge the interaction between the C-terminus of Pat1 (Pat1C) and the Lsm1-7 complex. The Lsm2-3-Pat1C complex and the Lsm1-7-Pat1C complex stimulate decapping *in vitro* to a similar extent and exhibit similar RNA-binding preference. The crystal structure of the Lsm2-3-Pat1C complex shows that Pat1C binds to Lsm2-3 to form an asymmetric complex with three Pat1C molecules surrounding a heptameric ring formed by Lsm2-3. Structure-based mutagenesis revealed the importance of Lsm2-3-Pat1C interactions in decapping activation *in vivo*. Based on the structure of Lsm2-3-Pat1C, a model of Lsm1-7-Pat1 complex is constructed and how RNA binds to this complex is discussed.

Introduction

mRNA degradation is paramount in the regulation of gene expression and the eradication of aberrant mRNAs. In eukaryotes, two major conserved mRNA decay pathways, namely, 5′-3′ pathway and 3′-5′ pathway, degrade mRNAs [1, 2]. Both pathways are initiated with poly(A) tail shortening (deadenylation) by deadenylases, which converts polyadenylated mRNAs to oligoadenylated mRNAs [3-5]. In the 5′-3′ pathway, oligoadenylated or deadenylated mRNAs are subjected to decapping by the Dcp1/Dcp2 holoenzyme and then digestion by the 5′-3′ exonuclease Xrn1 [6-9].

Decapping is a critical event in mRNA decay as the bulk of mRNA turnover, at least in yeast, is catalyzed by this step [1]. Decapping is also involved in several

specialized decay pathways such as nonsense-mediated mRNA decay, AU-rich element-mediated mRNA decay, histone mRNA decay, 3′ uridylation-mediated decay [10, 11] and microRNA-mediated mRNA decay [1, 12-15]. More recently, decapping has been shown to modulate expression of long noncoding RNAs to regulate inducible genes [16]. Given the importance of decapping in mRNA decay, it thus becomes a key node of many control inputs including both decapping inhibitors and activators [1]. The identified decapping inhibitors include poly(A) binding protein (Pab1) and the components of translation initiation machinery, while the decapping activators consist of Lsm1-7 complex, Pat1, Dhh1, Edc1-3, and in metazoans, Edc4 [1] and PNRC2 [17].

Degradation of mRNAs and their translation often act in a competitive manner, at least in yeast [18]. Non-translating mRNAs together with the decapping holoenzyme (Dcp1/Dcp2), decapping activators (Pat1, Lsm1-7 complex, Dhh1 and Edc3) and 5′-3′ exonuclease Xrn1 can aggregate into P bodies [18]. P bodies have been observed in many eukaryotic species such as yeast, nematodes, insect cells and mammalian cells [19, 20]. In ad-

dition to their role in mRNA decapping in normal mRNA decay, P bodies are also involved in nonsense-mediated mRNA decay, AU-rich element-mediated mRNA decay, microRNA-mediated translational repression, general translation repression and mRNA storage [19, 20].

Pat1 is a key protein that functions in P body assembly, decapping activation and translational repression. It contains three domains: the N-terminal domain, the middle domain and the C-terminal domain in yeast [21] (Figure 1A). Pat1 has conserved interactions with Dcp2, Lsm1-7 complex and Dhh1 (human ortholog DDX6/RCK) in *Saccharomyces cerevisiae*, *Drosophila melanogaster* and humans [22-24]. Among these conserved interactions, the C-terminal domain of Pat1 (denoted as Pat1C hereafter) contributes to the interactions with both Dcp2 and Lsm1-7 complex, and also affects the stimulation of decapping by Pat1 [22, 23]. The crystal structure of the human Pat1C has been solved, which showed that Pat1C folds into an α-α superhelix [23]. Besides its involvement in decapping, human Pat1 was also found to be tightly associated with Ccr4-Caf1-Not deadenylation complex and thus may serve as the scaffold to bridge decapping and deadenylation [25, 26].

The physical interaction of Pat1 with the Lsm1-7 com-plex *in vivo* has been demonstrated by co-purification of the Lsm1-7 complex with Pat1 from yeast [27]. The purified Lsm1-7-Pat1 complex has intrinsic affinity for the 3′ end oligoadenylated mRNAs over polyadenylated mRNAs, thus protecting this end from decay by the exosome while activating decapping [27]. Moreover, Lsm1-7 complex binds preferentially to deadenylated mRNAs carrying a U-tract at their 3′ terminal end over those that do not [27]. Lsm1-7 complex also binds certain viral mRNAs with a 5′ poly(A) tract, thereby stabilizing these mRNAs by inhibiting both 3′-5′ and 5′-3′ decay [28]. Besides its role in general mRNA decay, Lsm1-7 complex is also involved in histone mRNA decay [15, 29], uridylation-mediated mRNA decapping [11, 30] and microRNA (miRNA) biogenesis [31-35] by recognizing and binding to the 3′ poly(U) tract of the target RNAs in these processes.

An unresolved issue is how Pat1 interacts with the Lsm1-7 complex and the functional consequences of this interaction. The interaction of Pat1C with the Lsm1-7 complex has been reported to require the Lsm1 subunit [21, 23]. However, all these observations are based on yeast-two hybrid or co-immunoprecipitation, neither of which can rule out the possibility of indirect interactions.
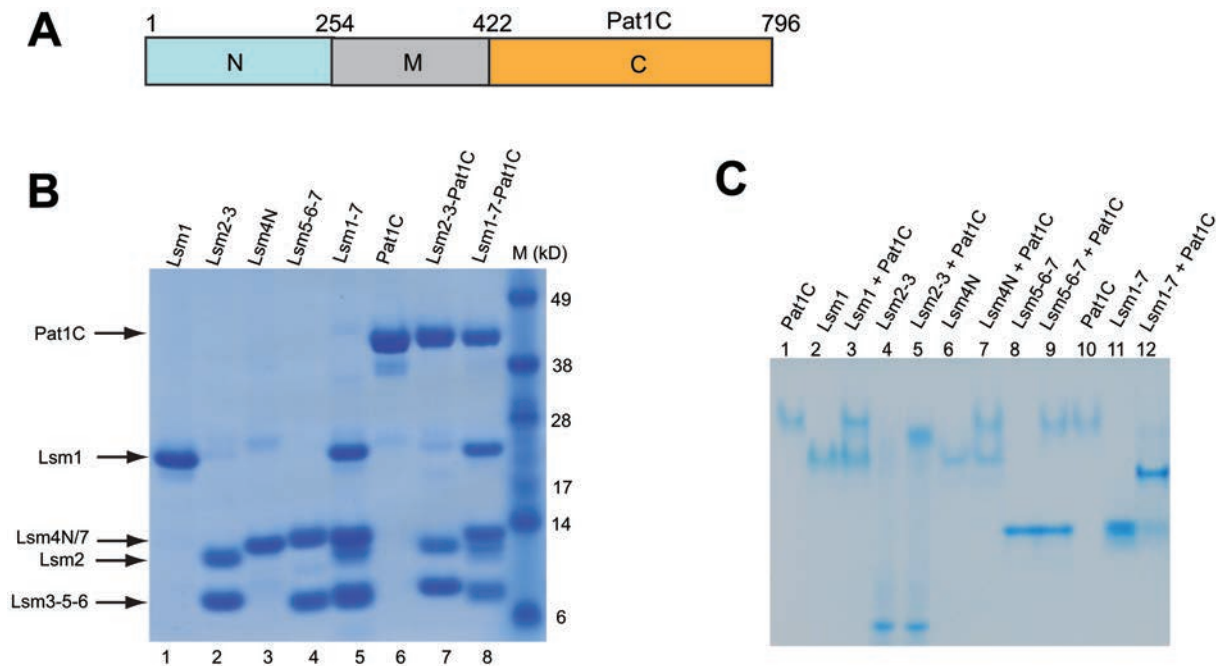

Figure 1 Pat1C interacts with Lsm1-7 complex through Lsm2 and Lsm3 bridging. **(A)** Domain organization of yeast Pat1. **(B)** Purified *S. cerevisiae* proteins, including Lsm1, Lsm2-3 subcomplex, Lsm4N, Lsm5-6-7 subcomplex, Pat1C, and the reconstituted Lsm1-7, Lsm2-3-Pat1C and Lsm1-7-pat1C complexes, are visualized by NuPAGE gradient gel. The positions of each protein are shown by arrows. **(C)** Native PAGE was used to detect the interaction of individual Lsm proteins or subcomplexes with Pat1C.

In this study, we reconstituted and biochemically characterized the Lsm1-7-Pat1 complex. Our results showed that subunits Lsm2 and Lsm3 bridge the interaction of the Lsm1-7 complex with Pat1C, and that *in vitro* reconstituted Lsm2-3-Pat1C and Lsm1-7-Pat1C were able to stimulate decapping to a similar extent. Importantly, both complexes exhibited stronger decapping stimulation activities than Lsm2-3 complex, Lsm1-7 complex or Pat1C alone, suggesting that the interaction of the Lsm complex with Pat1 *in vivo* directly enhances its ability to promote decapping. To shed light on the structural basis of Pat1 interacting with the Lsm1-7 complex, we determined the crystal structure of the Lsm2-3-Pat1C complex and found that three Pat1C molecules bind to an Lsm2-3 heptameric ring in an asymmetric manner. Structure-guided mutagenesis revealed the importance of Lsm2-3-Pat1C interactions in decapping activation *in vivo*. Based on the structure of Lsm2-3-Pat1C, we constructed a model of the Lsm1-7-Pat1 complex, which has implications for its function.

Results and Discussion

Pat1C interacts with the Lsm1-7 complex via Lsm2 and Lsm3

To study the interaction between Pat1C and the Lsm proteins, Lsm1, Lsm4N (C-terminal region truncated), Lsm2-3 subcomplex, and Lsm5-6-7 subcomplex were expressed and purified. The Lsm1-7 complex was reconstituted following a previous strategy [36]. The purity of these proteins was evaluated by a NuPAGE gel (Figure 1B, lanes 1-6).

To identify which subunits of the Lsm1-7 complex interact directly with Pat1C, native-PAGE was used. As shown in Figure 1C, incubation of Lsm2-3 with Pat1C (lane 5) induced a clear band shift of Pat1C as compared with Pat1C alone (lane 1 or 10), which indicates that Lsm2-3 can interact with Pat1C to form a Lsm2-3-Pat1C complex. Similarly, incubation of the Lsm1-7 complex with Pat1C (lane 12) generated a totally new band. The bands corresponding to the Lsm1-7 complex (lane 11) and Pat1C alone (lane 1 or 10) almost disappeared completely, demonstrating that the Lsm1-7 complex interacts strongly with Pat1C to form the Lsm1-7-Pat1C complex. In contrast, incubation of Pat1C with either Lsm1 (lane 3) or Lsm4N (lane 7) or Lsm5-6-7 (lane 9) showed no band shifts, indicating that Pat1C has very weak or no interactions with these proteins. Together, these results show that Lsm2-3 is sufficient to form an interaction with Pat1C and suggest that these two subunits bridge the interaction of the Lsm1-7 complex with Pat1C.

Lsm1-7-Pat1C and Lsm2-3-Pat1C complexes directly stimulate decapping

To study whether Lsm2-3-Pat1C and Lsm1-7-Pat1C could directly enhance decapping, *in vitro* reconstituted and purified Lsm2-3-Pat1C and Lsm1-7-Pat1C (Figure 1B, lanes 7 and 8) were examined for their effects on the Dcp1/Dcp2 decapping enzyme activity [37]. In this assay, the ability of Dcp1/Dcp2 to release labeled m^{7}GDP from the 5′ end of an mRNA is monitored by TLC.

As shown in Figure 2A, Lsm2-3 (lane 2), Lsm1-7 (lane 3) and Pat1C (lane 5) alone exhibited weak stimulation of Dcp1/Dcp2 decapping activity. In contrast, Lsm2-3-Pat1C (lane 4) and Lsm1-7-Pat1C (lane 6) complexes strongly stimulated the decapping activity of Dcp1/Dcp2. This suggests that the formation of an efficient decapping activation complex requires the participation of both Pat1 and Lsm proteins. It has been observed that m^{7}GDP can comigrate with inorganic phosphate (Pi) on the TLC plates used to monitor decapping [38]. Therefore, nucleotide diphosphate kinase, which can only convert m^{7}GDP, but not Pi, to m^{7}GTP, was used to confirm that Lsm2-3-Pat1C and Lsm1-7-Pat1C indeed stimulate decapping to releases m^{7}GDP as shown in Figure 2A (lanes 9 and 10).

We also observed that Lsm2-3-Pat1C and Lsm1-7-Pat1C stimulate the decapping activity of Dcp1/Dcp2 in a dose-dependent manner (Figure 2B). Surprisingly, the Lsm2-3-Pat1C complex exhibited stronger decapping stimulation activity than the Lsm1-7-Pat1C complex, which might be due to the enhanced stoichiometry of Pat1C in this complex as compared to the Lsm1-7-Pat1C complex (see below). The observation that the Lsm-Pat complexes stimulate decapping to a greater extent than individual Lsm1-7, Lsm2-3, or Pat1, argues that the interaction of the Lsm complex with Pat1 directly enhances its stimulation of decapping *in vivo*. This is consistent with the requirement of the Lsm1-7 subunits as being required for optimal decapping *in vivo* [39, 40] and the accumulation of P bodies in *lsm1Δ* mutants [41], the latter of which is consistent with a role for the Lsm1-7 complex late in the decapping pathway.

Lsm1-7-Pat1C and Lsm2-3-Pat1C complexes show similar RNA binding preference in vitro

The Lsm1-7-Pat1 complex has the binding preference for a U-rich sequence, and an *in vitro* analysis shows that the complex only binds to U$_{20}$ but not other 20-mer homo-oligoribonucleotides [27]. To compare the homo-oligomeric RNA-binding ability of Lsm2-3-Pat1C and Lsm1-7-Pat1C, fluorescence anisotropy assay was performed. As shown in Figure 2C and 2D, the Lsm1-7-Pat1C complex has higher binding affinity for U$_{15}$ (Dissociation constant (Kd) = 21.5 nM) and lower bind-
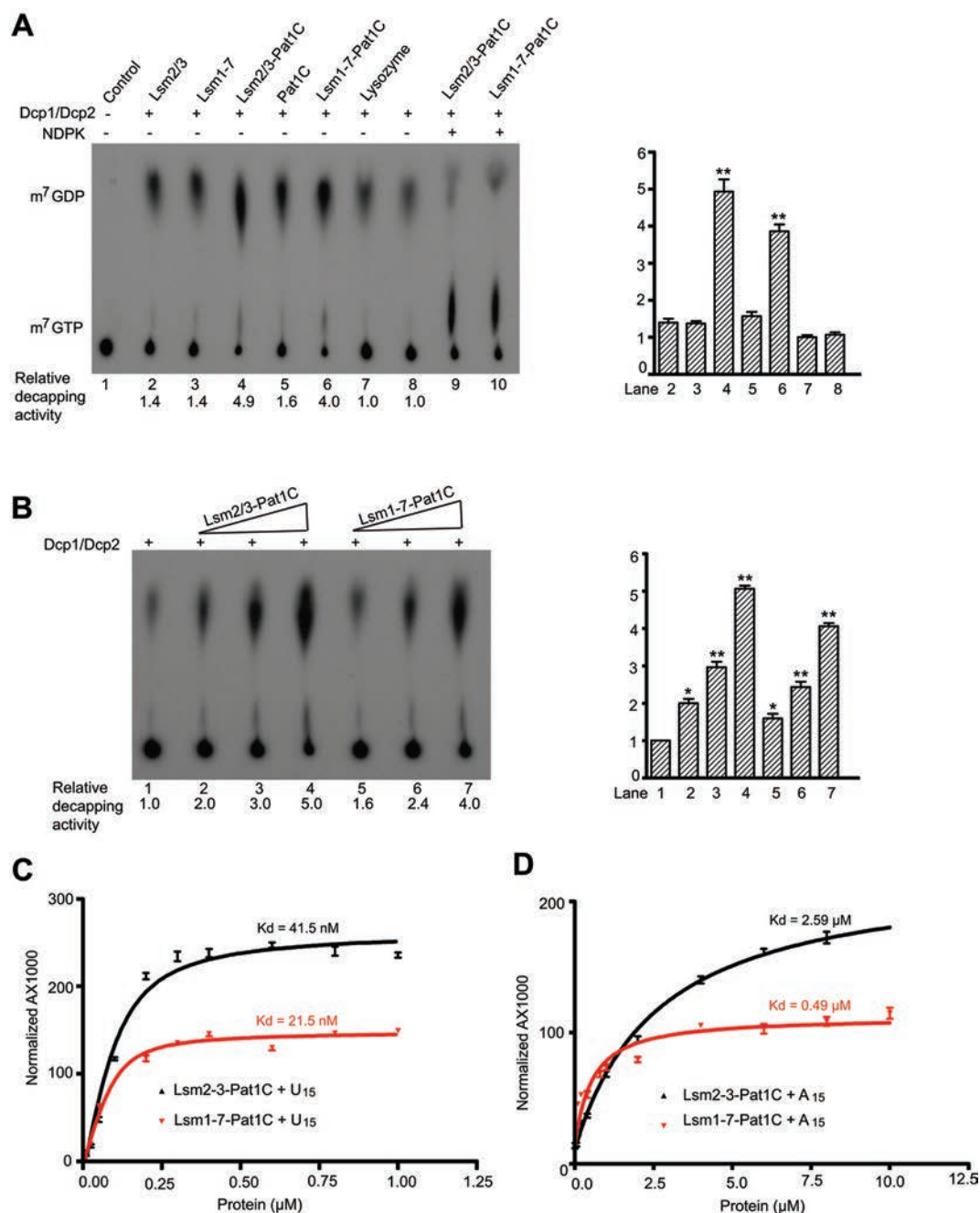
Figure 2 Lsm2-3-Pat1C and Lsm1-7-Pat1C have comparable decapping stimulation activity and similar RNA-binding properties. **(A)** Left panel, effects of Lsm2/3, Lsm1-7, Pat1C, Lsm1-7-Pat1C and Lsm2-3-Pat1C on the decapping activity of Dcp1/Dcp2. The enhanced decapping activities were quantified using the amount of cap-labeled RNA relative to that of Dcp1/Dcp2 alone (lane 8) from three independent measurements. Right panel, quantification of the stimulation effects shown in the left panel. Student's t-test was used to assess the statistically significant difference of the decapping activity of Dcp1/Dcp2 before (lane 8) and after addition of the indicated proteins (lanes 2-7). **$P < 0.01$. **(B)** Left panel, Lsm1-7-Pat1C and Lsm2-3-Pat1C stimulate decapping in a dose-dependent manner. Right panel, quantification of the stimulation effects shown in the left panel. Student's t-test was used to assess the statistically significant difference of the decapping activity of Dcp1/Dcp2 before (lane 1) and after addition of the indicated proteins (lanes 2-7). *$P < 0.05$; **$P < 0.01$. The columns and bars represent the means and standard deviations, respectively, calculated from three independent experiments. **(C, D)** Quantitative measurements of RNA-binding affinities of Lsm1-7-Pat1C and Lsm2-3-Pat1C in solution by fluorescence anisotropy. Kd values were determined by fitting the experimental data to a binding equation describing a single-site binding model. The Kd values and their corresponding errors are the mean and standard deviation of three independent experiments.

ing affinity for A_{15} (Kd = 0.49 μM). Similarly, Lsm2-3-Pat1C displays higher binding affinity for U_{15} (Kd = 41.5 nM) and lower binding affinity for A_{15} (Kd = 2.6 μM). Both complexes have no measurable binding to C_{15} (data not shown). These data indicate that Lsm1-7-Pat1C and Lsm2-3-Pat1C have similar RNA-binding preference *in vitro*.

Structural determination of Lsm2-3-Pat1C complex

The structure of Lsm1-7-Pat1C complex is important for understanding the mechanism by which it activates decapping. However, attempts to obtain crystals for this complex were not successful after an extensive crystallization screening. As Lsm2-3-Pat1C complex showed an even stronger activity than Lsm1-7-Pat1C complex in decapping stimulation and exhibited a RNA-binding preference similar to that of Lsm1-7-Pat1C complex, we set out to solve the crystal structure of the Lsm2-3-Pat1C complex at a resolution of 3.15 Å. The final refined model includes four Lsm3 molecules (chains A, B, E and F), three Lsm2 molecules (chains C, D and G) and three Pat1C molecules (chains H, I and J). Some residues are not visible in the electron density map and assumed to be disordered. These include residues 81-89 in chains A and B, residues 49-55 and 94-95 in chain C, residues 50-58 and 94-95 in chain D, residues 48-59 and 81-89 in chain E, residues 53-56 and 80-89 in chain F, residues 49-55 and 94-95 in chain G, residues 422-469, 751-757 and 795-796 in chains H, I and J, respectively, and residue 794 in chain I. The statistics of data collection and refinement are summarized in Supplementary information, Table S1.

Overall architecture of Lsm2-3-Pat1C complex

As shown in Figure 3A, three Pat1C molecules surround an Lsm2-3 heptameric ring in an asymmetric manner. This is in agreement with our initial analysis that non-crystallographic symmetry could not be identified from native Patterson map or self rotation functions. The N-terminal region of Pat1C associates with the Lsm2-3 heptameric ring through helix-helix interactions.

Pat1C is composed of 15 α helices and forms a right-handed α-α superhelix domain (Figure 3B). Comparison of yeast Pat1C with human Pat1C using the Dali server [42] shows that the overall orientation of the α-α superhelix between the two proteins is similar with *Z* score of 15.1; however, structural diversity is demonstrated by the high $C_α$ root mean square deviation value of 3.5 Å (Figure 3B). One notable difference is that yeast Pat1C contains two long helixes (α14 and α15) located in its C-terminal region that are absent in human Pat1C (Figure 3B). The C-terminal region may be important in maintaining the overall architecture of the yeast Pat1C superhelix as de-

letion of residues 697-763 in yeast Pat1, which consists of α13 (residues 702-711), α14 (residues 720-747) and part of α15 (residues 761-783), disrupted the recruitment of Lsm1 and presumably the whole Lsm1-7 complex to P bodies [21].

Lsm2 and Lsm3 proteins form a hetero-heptamer with a central cavity, which consists of four Lsm3 proteins and three Lsm2 proteins, giving rise to an asymmetrical interaction network with six Lsm2/Lsm3 interfaces and one Lsm3/Lsm3 interface. Each Lsm protein contains a conserved Sm fold, which is featured by an N-terminal α helix followed by a highly bent five-stranded β sheet. Moreover, the C-terminal extension of Lsm2 forms a second α helix. The doughnut-shaped heptamer has two faces, namely the helix and loop faces, and a central cavity connecting the two faces. The heptamer has an outer diameter of ~60 Å at the helix face, an outer diameter of ~26 Å at the loop face, an inner diameter of ~19 Å at its narrowest part of the cavity and a thickness of ~42 Å.

The oligomeric state of Lsm2-3 in solution

To examine whether the heptameric Lsm2-3 complex observed in the structure of Lsm2-3-Pat1C complex exists in solution as a heptamer or whether its formation is induced by the binding of Pat1C, sedimentation velocity of analytical ultracentrifuge was employed to study the oligomeric state of Lsm2-3 under three different concentrations. The data were analyzed using the continuous c(S) and c(M) distributions, giving an average molecular weight of 80 700 Dalton, which is close to the theoretical molecular weight of 78 448 Dalton for the heptameric Lsm2-3 observed in the structure (Figure 3C and Supplementary information, Table S2). Intriguingly, the crystal structure of yeast Lsm3 has an octameric ring organization [43], which has also been confirmed by our solution study [44]. Thus these data suggest that, at least *in vitro*, incorporation of Lsm2 shifts the homo-octameric Lsm3 ring to the hetero-heptameric Lsm2-3 ring and the binding of Pat1C to Lsm2-3 has no effect on the heptameric ring formation. An interesting unresolved issue is whether the Lsm2-Lsm3-Pat1 complex can form *in vivo* and has any specific biological role.

Interaction of Pat1C with Lsm2-3

Pat1C interacts with both Lsm2 and Lsm3, mainly via helix-helix interactions with α1 and α4 of Pat1C contacting both the N-terminal α helix of Lsm3 and the C-terminal α helix of Lsm2. The binding of Pat1C to Lsm2-3 is mediated through a network of hydrogen bonding and hydrophobic interactions, which buries a solvent-accessible surface of ~1 270 Å^2 (Figure 4A). Briefly, the OE1 and OE2 of Glu483 in Pat1C (α1) is hydrogen bonded
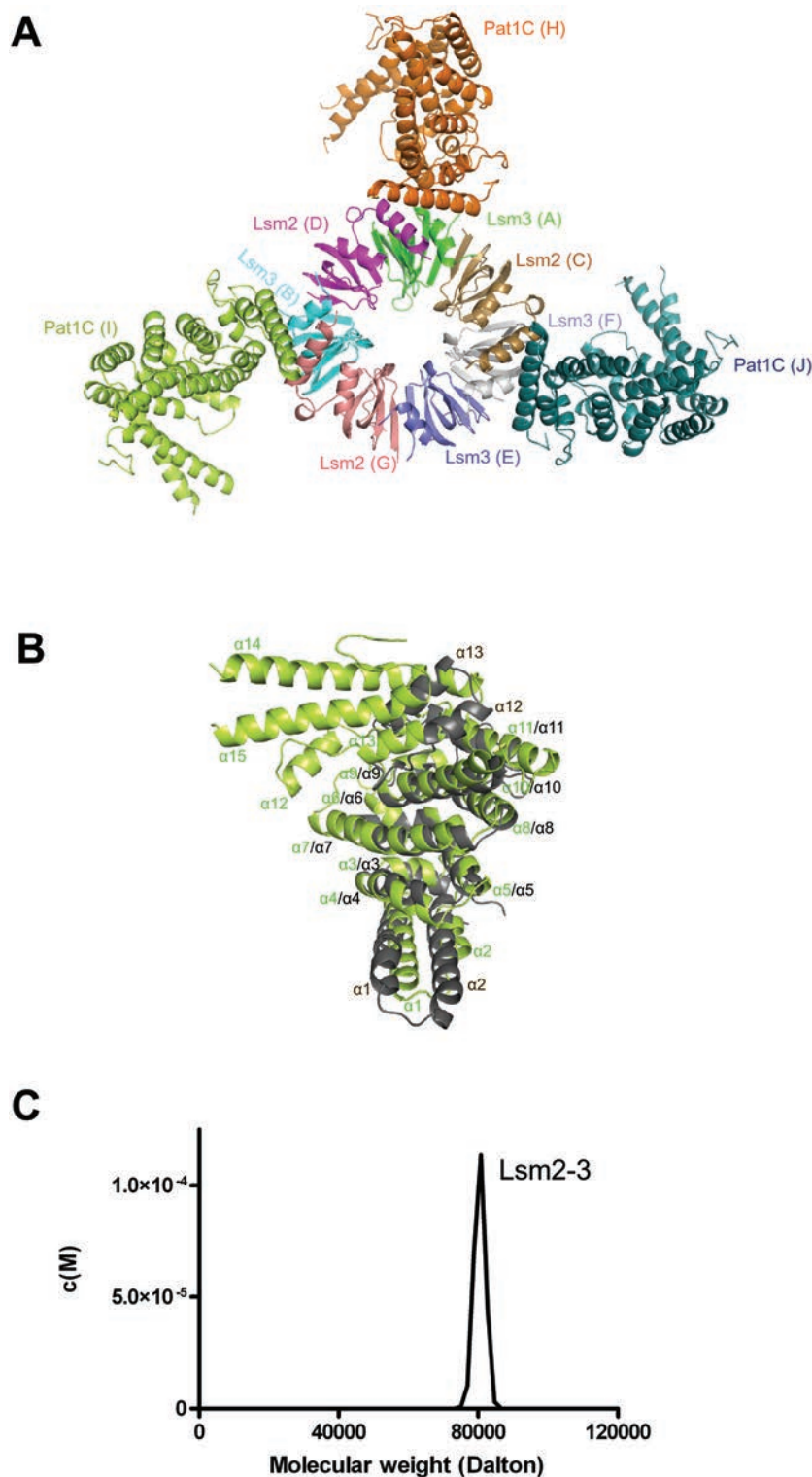
Figure 3 The structure of the Lsm2-3-Pat1C complex. **(A)** The overall structure of the Lsm2-3-Pat1C complex. Each chain is shown as a different color and labeled in parenthesis. **(B)** Stereo view of the superposition of yeast Pat1C with human Pat1C (PDB code 2XEQ chain A). The Dali server was used to superpose yeast Pat1C chain I (green) with human Pat1C (grey). **(C)** The Lsm2-3 subcomplex exists in solution as a heptamer. Sedimentation velocity analysis of Lsm2-3 proteins at three different protein concentrations (0.75, 1.0 and 1.5 mg/ml) was conducted and fitted based on the c(S) and c(M) distribution functions. The c(S) and c(M) distribution profile at these concentrations is similar and a representative profile is shown at a concentration of 0.75 mg/ml.

to the carbonyl group of Lys9 from Lsm3 (N-terminal α helix) and the side chain of Arg87 from Lsm2 (C-terminal α helix), respectively. Moreover, Lys9 of Lsm3 forms an ionic interaction with Glu487 of Pat1C (α1). The side chain of Lys531 from Pat1C (α4) forms a hydrogen bond with the carbonyl group of Leu12 (loop between α helix and β1) from Lsm3 while the side chains of Lys534 and Arg538 from Pat1C (α4) form ionic interactions with Asp13 of Lsm3 (loop between α helix and β1) and Asp84 of Lsm2 (C-terminal α helix), respectively. Besides hydrogen bonding interactions, the interactions of Lsm2-3 with Pat1C are strengthened by the hydrophobic interactions, which involve Leu479 and Leu490 from Pat1C (α1), Leu10 of Lsm3 (N-terminal α helix) and Met91 of Lsm2 (C-terminal α helix), respectively. It is worth noting that the residues involved in the interaction between Lsm2-3 and Pat1C are highly conserved in eukaryotes (Supplementary information, Figure S1). Moreover, the residues of Pat1C involved in interaction with Lsm2-3 correspond to a group of residues located in a highly conserved surface area on the structure of human Pat1C [23].

To probe the role of the Pat1C-Lsm2-3 interface, we constructed three deletion mutants of Lsm2-3 and one multiple-Ala mutant of Pat1C. His-tag pull-down assays (Figure 4B) show that deletion of the C-terminal α helix of Lsm2 (denoted as Lsm2Δα2-3 hereafter) (lane 5), or the N-terminal α helix of Lsm3 (denoted as Lsm2-3Δα1 hereafter) (lane 7), or both (denoted as Lsm2Δα2-3Δα1 hereafter) (lane 9) completely abolished the interaction of Lsm2-3 with Pat1C, while Ala substitution of seven residues (Leu479, Glu483, Glu487, Leu490, Lys531, Lys534 and Arg538) of Pat1C (denoted as Pat1C-7m) almost completely disrupted its association with Lsm2-3 (lane 3). Consistent with these results, analytical size exclusion chromatography (Supplementary information, Figure S2) showed that wild-type Pat1C interacts with wild-type Lsm2-3 efficiently to form a new peak on the elution profile while the combination of any of the mutant proteins with wild-type proteins or another mutant failed to induce the formation of this new peak, indicating that mutations have disrupted the interaction of Pat1C and Lsm2-3. These data support the role of residues in the interface in the interaction of Pat1C with Lsm2-3.

Interaction of Pat1C with Lsm2-3 affects decapping in vivo

To test whether the Pat1C interaction with Lsm2 and Lsm3 was biologically relevant *in vivo*, we created mutations predicted to disrupt the Pat1C-Lsm2-3 interface and expressed these mutant proteins from their own promoter in single copy vectors in yeast strains lacking the corresponding wild-type genes. These mutations include *Lsm2Δα2* and *Lsm3Δα1*, as well as the *Pat1-7m* allele (Ala substitution of seven residues Leu479, Glu483, Glu487, Leu490, Lys531, Lys534 and Arg538), which biochemically affects the interaction with Lsm2 and Lsm3 (Figure 4B and Supplementary information, Figure S2). As *Lsm2* and *Lsm3* genes are essential for viability, we first introduced the mutant *Lsm* genes into a deleted *lsm2Δ* or *lsm3Δ* strains by a plasmid shuffle (see Supplementary information, Data S1). We observed that *lsm2Δα2* and *lsm3Δα1* alleles were both able to complement the inviability of *lsm2Δ* and *lsm3Δ* strains, respectively (data not shown). This demonstrates that these mutations do not severely compromise the overall folding or function of the Lsm2 or Lsm3 proteins in splicing, which appears to be their essential role [45]. In contrast, as *pat1Δ* strain is viable, the Pat1-7m allele, or a wild-type control, was introduced into a *pat1Δ* strain to assess its phenotypic consequences. We examined the effects of these mutations on decapping by measuring the decay rate of the MFA2pG reporter, which is primarily degraded by decapping [46].

We observed that all three mutations disrupting the Pat1C-Lsm2-3 interface led to a slower decay rate of the MFA2pG mRNA (Figure 4C). These results indicate that the interaction between Pat1C and the Lsm1-7 complex is required for optimal decapping *in vivo*. This conclusion is consistent with the earlier work in mammalian cells wherein mutations in the same region of human Pat1 were shown to affect mRNA decapping [23].

We also examined the effects of these mutations on mRNA decay products stabilized by a polyG tract inserted into the 3′ UTR. One mRNA decay product corresponding to the 3′ end of the mRNA is produced when the polyG tract blocks Xrn1 progression from the 5′ end, and is generally decreased when decapping is inhibited [47, 48]. Consistent with an effect on decapping, using probes overlapping the 3′ side of the polyG tract, we observed that the *lsm2Δα2* and *lsm3Δα1* mutations led to a reduction in the amount of this 5′ to 3′ decay product from both the PGK1pG and MFA2pG reporter mRNAs, providing additional evidence that these mutations decreased mRNA decapping (Figure 5A). Consistent with this, we also observed an increase in the fraction of deadenylated full-length mRNAs in these mutants, which occurs when mRNA decapping is slowed by defects in the Lsm1-7-Pat1 complex [40]. The Pat1-7m mutation also led to an increase in deadenylated mRNA for the MFA2 and PGK1 mRNAs, although the 5′ to 3′ decay product was still produced to similar extent, suggesting that the effect of Pat1-7m mutation is weaker *in vivo* than those of the *lsm2Δα2* and *lsm3Δα1*
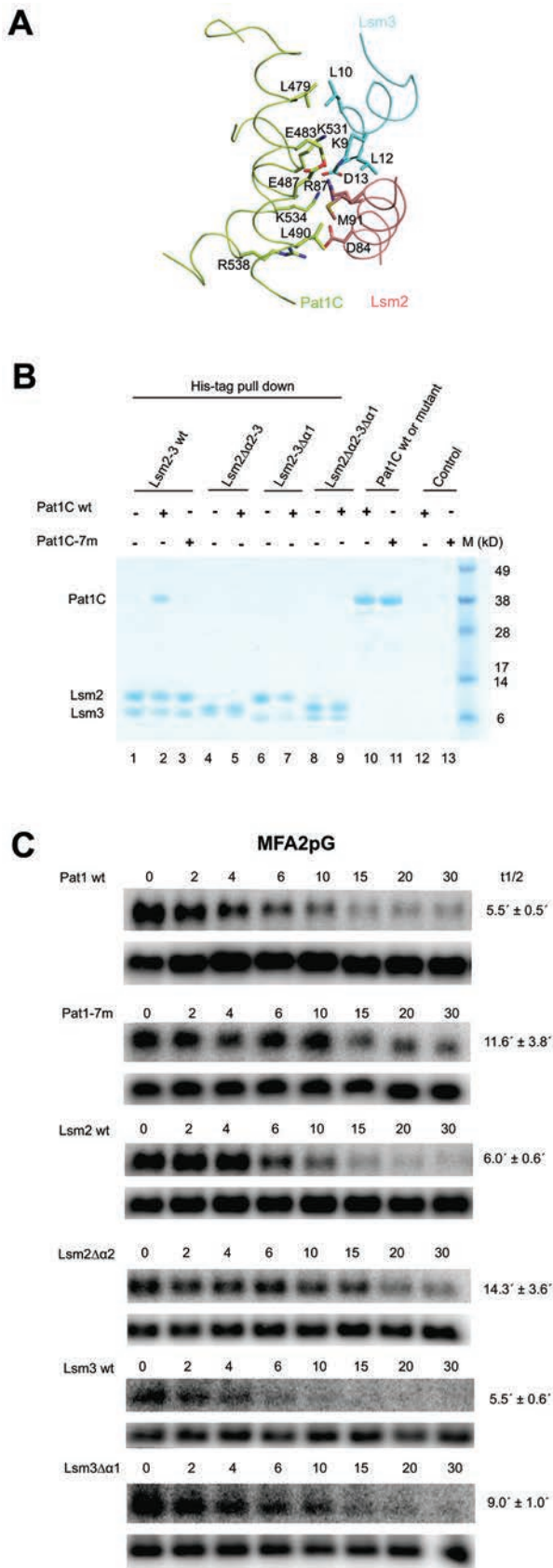
mutations.

We also examined how disruption of the Pat1C-Lsm2-3 interaction affected 3′ to 5′ degradation of the mRNAs by the exosome, which increases when decapping is inhibited and thereby provides another assay for the *in vivo* function of the Pat1-Lsm1-7 complex [49]. The degree of 3′ to 5′ degradation of these reporter mRNAs can be estimated by the accumulation of an mRNA decay product produced by 3′ to 5′ degradation of the mRNA to the 3′ side of the polyG tract, which we detect by an oligo probe. We observed that the *lsm2Δα2*, *lsm3Δα1*, and *pat1-7m* mutations all led to an increase in the mRNA decay product produced by degradation in the 3′ to 5′ direction for the PGK1pG mRNA (an mRNA where these 3′ to 5′ products are more easily observed due to a slower intrinsic rate of decapping [48]) (Figure 5B). This observation provides additional evidence that an effective interaction between Pat1 and Lsm2-3 is required for the Lsm1-7-Pat1 complex to promote decapping and bias mRNA decay away from 3′ degradation by the exosome. Taken together, the effects of mutations disrupting the Pat1-Lsm2-3 interface on mRNA decapping and 3′ to 5′ degradation *in vivo* demonstrate that this interaction is functionally significant in modulating mRNA degradation in cells.

A model of the Lsm1-7-Pat1C complex

A critical goal is to understand the structure of the entire Lsm1-7-Pat1 complex, which has currently remained resistant to study by crystallography. We have therefore developed a model of the Lsm1-7-Pat1C complex in several steps. First, sequence alignment shows that each Lsm protein of the Lsm1-7 complex is analogous to one of the seven Sm proteins [50-52] and thus, the order of

Figure 4 The role of the residues in the Lsm2-3-Pat1C interface. **(A)** Stereo view of the inter-subunit interfaces of the Lsm2-3-Pat1C complex. The representative interface of Pat1C-Lsm2-3 (Pat1C: chain I, Lsm2: chain G, Lsm3: chain B) is shown in the same color scheme as in **Figure 3A**. **(B)** His-tag pull-down assay was used to detect the interactions of Pat1C wild type (Pat1C wt) or Pat1C mutant (Pat1C-7m) with Lsm2-3 wild type (Lsm2-3 wt), Lsm2Δα2-3, Lsm2-3Δα1 or Lsm2Δα2-3Δα1. Lanes 1-9 show His-tag pull-down assay, lanes 10-11 show the purity of Pat1C wt and Pat1C-7m and lanes 12-13 show the control. **(C)** The panels show the half life of MFA2pG reporter mRNA probed with an oligo covering the polyG tract and 3′ sequences (oRP140). Half life measurements are given on the right of each panel in minutes from an average of at least three experiments. Standardization for each lane was done using the SCR1 RNA (orp100 probe) shown below each panel.
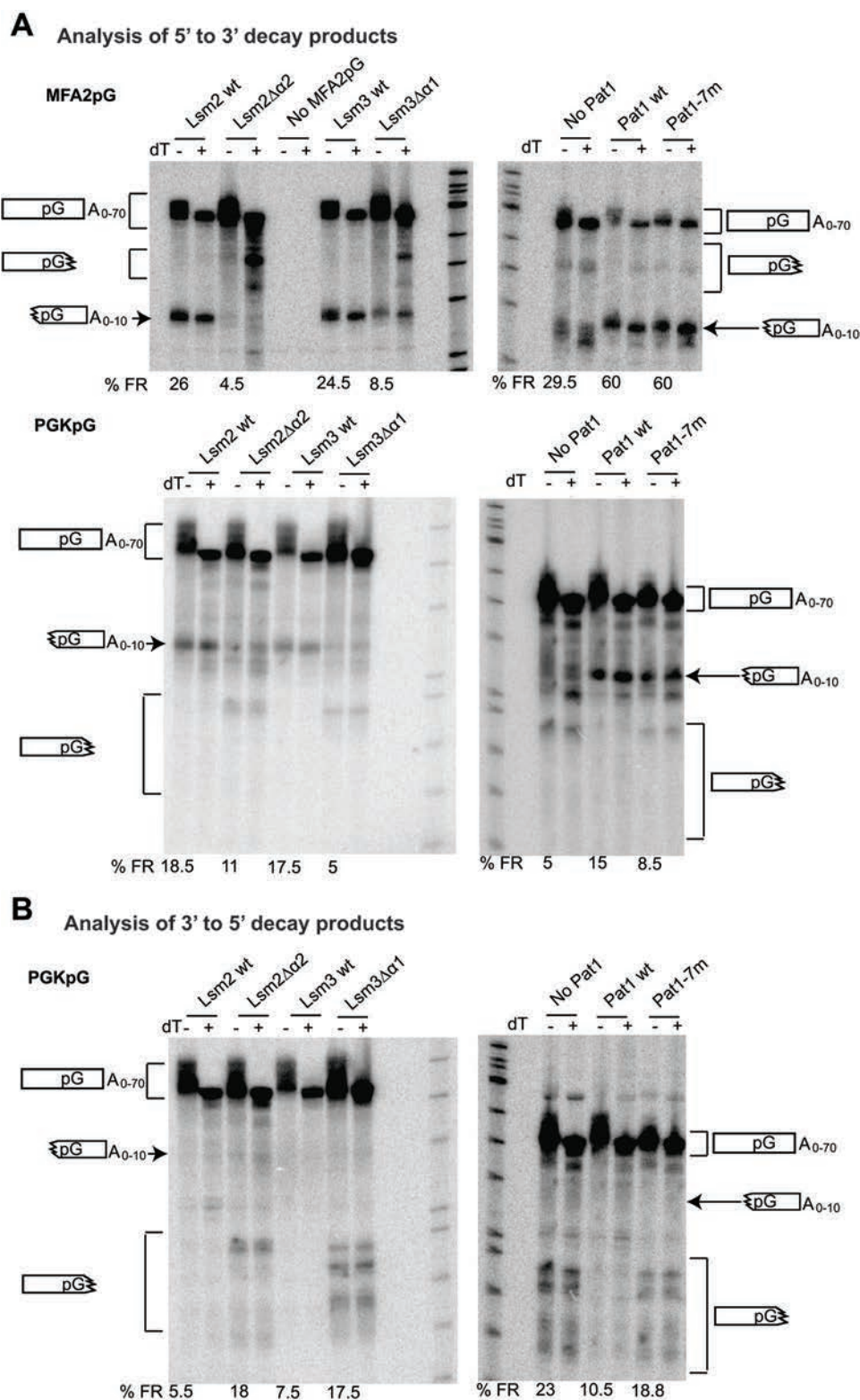
Figure 5 Alteration of the Pat1C-Lsm2-3 interface affects mRNA decapping and 3′ to 5′ degradation *in vivo*. **(A)** Acrylamide northern blots showing the 3′ fragments from either MFA2pG or PGK1pG mRNAs (oRP140 or oRP141 probes, respectively). **(B)** Acrylamide northern blots showing the 5′ fragments of the PGK1pG mRNA(oRP154). Cleavage with oligo(dT) is indicated above each lane. Cartoons on each end of the panels represent the mRNA species present in each band of the gels. The percent of mRNA fragment for each lane is given below the lane and is from at minimum of three experiments.

Lsm proteins is derived from that of the Sm ring structure [53, 54]. Second, we used the recent crystal structures of Lsm4N and the Lsm5-6-7 complex from *Schizosaccharomyces pombe* [44, 55] and human SmB (corresponding to Lsm1) in the spliceosomal U4 snRNP core domain [54] to generate homology models of Lsm4, Lsm5, Lsm6, Lsm7 and Lsm1 of *S. cerevisiae* in the SWISS-MODEL workspace [56]. Third, Pat1C (chain H), Lsm2 (chain D) and Lsm3 (chain A) in the Lsm2-3-Pat1C structure were kept while the other two Pat1C molecules (chains I and J) were removed. Finally, based on the order of Sm proteins (Supplementary information, Figure S3) and their sequence alignment counterparts in Lsm1-7 proteins, chains B, G, F, C and E in the Lsm2-3-Pat1C structure (Figure 3A) are replaced by the homology models of Lsm1, Lsm4, Lsm5, Lsm6 and Lsm7 through superposition to yield the final model (Figure 6A).

To examine how Lsm1-7-Pat1 complex might bind RNA, we mapped the electrostatic potential on the molecular surface of the modeled Lsm1-7-Pat1C complex and compared it to the electrostatic potential distribution of the human Sm ring in the U4 snRNP core domain (Figure 6B and 6C and Supplementary information, Figure S3). Both the helix and loop faces of the human Sm ring have large positively charged patches. The central cavity of this complex is dominated by positively charged residues, with its width sufficient to accommodate a single-stranded RNA. Consistent with this analysis, stem loops II and III of U4 snRNA bind to the helix and loop faces of the human Sm ring, respectively, with its Sm site passing through the central channel [54]. In contrast, the inner cavity in the Lsm1-7-Pat1C complex is narrowed by the protrusion of the Lsm3 loop region from the loop face (Figure 6B), which may disfavor the passing of RNA through the cavity. Moreover, the negatively charged patches in the loop face mainly contributed by subunits Lsm2, Lsm3, Lsm4, Lsm5 and Lsm7 may also occlude the RNA via charge-charge repulsions. Together, this analysis suggests that Lsm1-7-Pat1C complex may only bind RNA from the helix face.

In the helix face of the Lsm1-7-Pat1C complex, the surfaces of subunits Lsm3, Lsm5 and Lsm6 are dominated by negatively charged patches while the other subunits are neutral or positively charged (Figure 6C). This uneven charge distribution may give rise to the difference in RNA-binding affinity with respect to each subunit. Consistent with this possibility, Lsm1 and/or Lsm4 strongly cross-link to RNA as compared to other subunits [27]. Moreover, shorter RNA substrates bind more weakly than the longer RNAs to the Lsm1-7-Pat1 complex [27] and human Pat1C interacts with U_{30} but fails to interact with U_{20} and U_{15} [23]. The minimum RNA length requirement for detectable binding to both Pat1 alone and the Lsm1-7-Pat1 complex suggests that Pat1 and Lsm1-7 may act in a cooperative manner for binding to the RNA substrate through specific interactions with the helix face of the Lsm1-7 complex and then with a region of Pat1C.

Examination of the Pat1C surface reveals that a patch formed by the interface of helix 1 in Pat1C with Lsm2 and Lsm3 on the loop face side of the complex is a likely site for RNA binding (Figure 6B). In this RNA-binding model, RNA would interact with the helix face of the Lsm1-7 complex, then transverse the positively charged region on the helix face of Lsm2 and bend around to the loop face in the positively charged channel between Lsm2/Lsm3 and Pat1C. Consistent with this model, this patch contains several conserved positively charged amino acids, three (K475, K476 and R538) of which are analogous to residues required for RNA binding of human Pat1C [23]. The combination of this putative RNA-binding surface in Pat1C with the Lsm1-7 complex provides a possible explanation for the preference of Pat1 for larger RNAs.

Closing remarks

Herein, we have shown the structure of the C-terminal domain of Pat1 interacting with the C-terminal helix of Lsm2 and the N-terminal helix of Lsm3. In contrast, no direct interactions were detected between the C terminal domain of Pat1 and other components of the Lsm1-7 complex (Figure 1C). Moreover, this interaction is required for the formation of Lsm2-3-Pat1C complex *in vitro* and for the stimulation of decapping *in vivo*. As Lsm2 and Lsm3 proteins are also part of the Lsm2-8 complex, this interaction may explain why Pat1 can be detected in the U6-containing penta-snRNP [57], although whether Pat1 has any function in conjunction with the U6 sn-RNP remains to be determined. Pat1 appears to have additional interactions with the Lsm1-7 complex as the middle domain of Pat1 can interact with the Lsm1-7 complex [21-23]. Finally, modeling of the Lsm1-7 complex suggests a potential mechanism for RNA binding that involves the helix and loop faces of the Lsm1-7 ring and a channel on the Pat1 C-terminal domain. This putative extended RNA-binding surface provides a possible molecular explanation for the preferred binding of the Lsm1-7-Pat1 complex to longer RNAs.

Materials and Methods

Reconstitution of Lsm1-7 and Lsm1-7-Pat1C complexes

Reconstitution of Lsm1-7 complex followed a similar strategy as described by Zaric and coworkers [36] with some modifications. Briefly, each purified Lsm protein and protein subcomplex including Lsm1, Lsm2-3, Lsm4N and Lsm5-6-7 were mixed
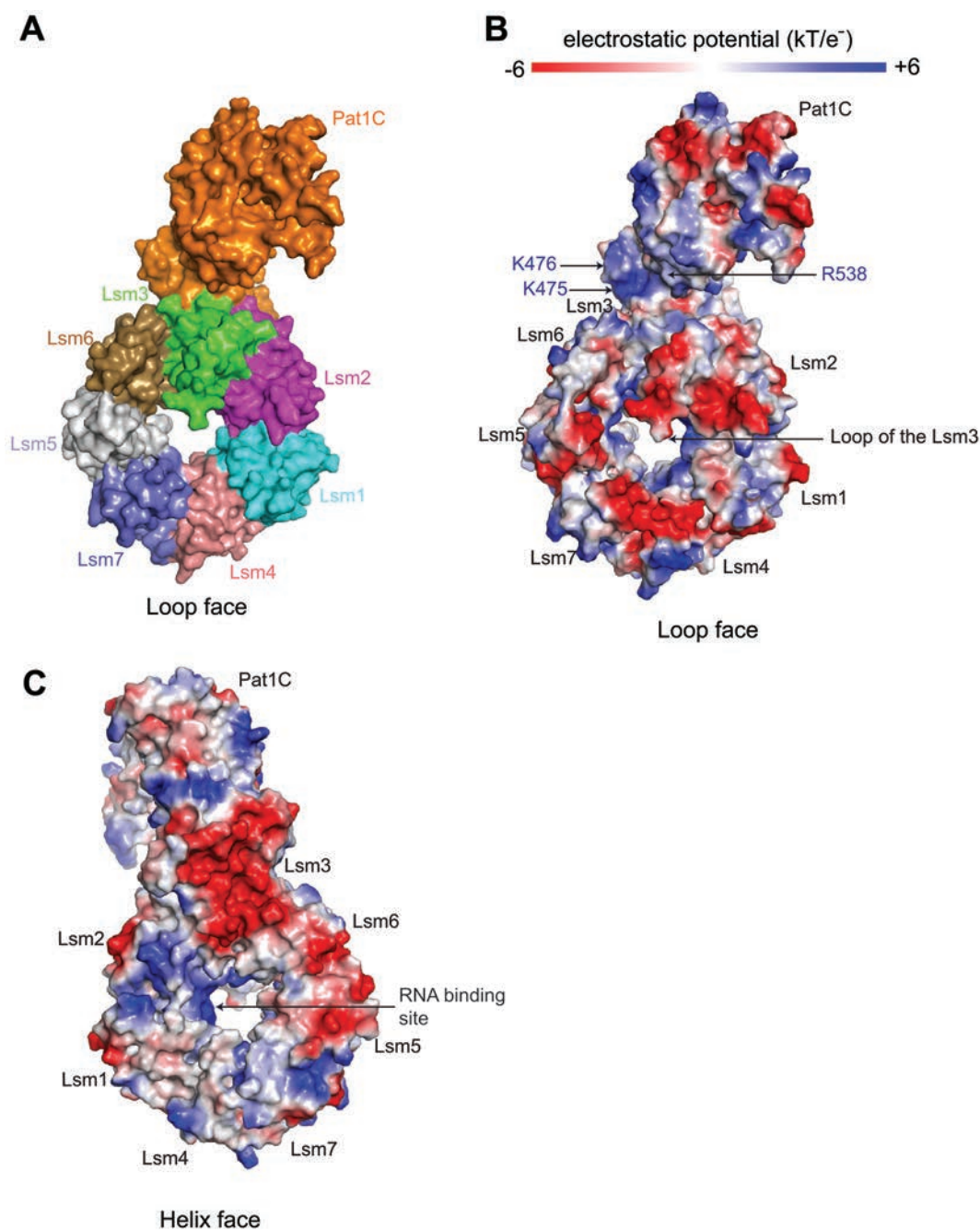
Figure 6 A model of the Lsm1-7-Pat1C complex. **(A)** Surface representation of the modeled Lsm1-7-Pat1C complex showing the loop face. **(B)** Electrostatic potential showing the loop face of the Lsm1-7-Pat1C model. **(C)** Electrostatic potential showing the helix face of the Lsm1-7-Pat1C model. The view in **B** is as in **A** and the complex in **C** is rotated 180 ° along a vertical axis relative to the view in **B**.

in equal molar ratio in 20 mM Hepes pH 7.5, 6 M urea at room temperature. After overnight incubation, the mixtures were loaded directly onto Superdex-200 26/60 column equilibrated with a buffer containing 20 mM Hepes pH 7.5, 100 mM NaCl to perform on-column refolding, following the strategy described previously [58]. The reconstituted Lsm1-7 complex was concentrated to

~10 mg/ml. For reconstitution of Lsm1-7-Pat1C complex, equal molars of Lsm1-7 complex and Pat1C were mixed and incubated on ice for 1 h before loading onto Superdex-200 26/60 column equilibrated with the same running buffer as that used for Lsm1-7 complex reconstitution.

Crystallization and heavy atom soaking

The hanging drop vapor diffusion method was used to grow crystals at 15 °C. Crystals of native and SeMet-substituted Lsm2-3-Pat1C proteins were grown by mixing 1 μl of each protein sample with 1 μl of 0.1 M Tris pH 7.0, 12% ethanol. Plate-shaped crystals were grown to full size in around 2 weeks. Crystals were transferred from mother liquor to 0.1 M Tris pH 7.0, 5% ethylene glycol, 30% 2-methyl-2,4-pentanediol (MPD) and plunged into liquid nitrogen.

Native crystals to be soaked by tantalum bromide cluster Ta_6Br_{12} (Jena Biosciences, Germany) were transferred to 4 °C and incubated at this temperature for a few hours. After incubation, crystals were transferred from mother liquor to 0.1 M Tris pH 7.0, 5% ethylene glycol, 10% 2-methyl-2,4-pentanediol supplemented with 1 mM Ta_6Br_{12} and soaked for 3 h before being flash-cooled into liquid nitrogen.

Data collection, structure determination and refinement

A native data set of Lsm2-3-Pat1C crystal and single-wavelength anomalous diffraction data set of SeMet-Lsm2-3-Pat1C crystal were collected on the beamline ID23-1 (ESRF, Grenoble, France) using diffraction cartography [59], as implemented in the beamline GUI MxCuBE [60, 61], to define the best volumes of the crystals. Multiple-wavelength anomalous diffraction (MAD) data sets of Ta_6Br_{12}-soaked Lsm2-3-Pat1C crystal were collected at a resolution of 3.7 Å on the beamline X06SA (PXI, SLS, Switzerland). All data sets were integrated with XDS [62] and merged and scaled with Scala from the CCP4 suite [63]. Phases of the Ta_6Br_{12}-soaked MAD data sets were calculated using the phasing module Autosol from the PHENIX program package [64]. After density modification, the electron density map was good enough to identify the three Pat1C molecules composed of α-α superhelix surrounding a seven-membered Lsm ring structure arranged in an asymmetrical manner.

A poly Ala model derived from Lsm3 (PDB code 3BW1 chain A) excluding loop4 (between β3 and β4) was built into the initial electron density map calculated at a resolution of 3.7 Å using O [65] to generate a poly Ala seven-membered ring model. For Pat1C, a poly Ala model was built manually in Coot [66]. Through superposition and combination, a poly Ala model consisting of three Pat1C molecules and one seven-member Lsm ring was generated. Due to the non-isomorphism between tantalum derivative data set and native data set, this poly Ala model was first subject to rigid body refinement in REFMAC5 [67] against the native data set at a resolution of 5 Å. Subsequently, through iterative refinement using PHENIX [64] and model building using Coot [66], the resolution gradually extends to 3.15 Å. The refinement strategy used in PHENIX includes simulation annealing, group B-factors refinement, tight non-crystallographic symmetry restraints on different protomers. Water was automatically added in the later stages of the refinement and manually edited with electron densities.

Se anomalous difference Fourier map at a resolution of 3.7 Å was calculated using the initial tantalum multiple-wavelength anomalous diffraction phases combined with those calculated from the poly Ala model. Twenty-three Se sites with peak height more than five were clearly identified. These Se sites were used independently to cross-check the register of the protein sequences in the model. Pymol (www.pymol.org) was used for structural represen-tation and APBS was used to generate electrostatic potential map [68].

Decapping assay

Uncapped RNA was transcribed using the Ampliscribe T7 transcription kit (Epicentre Technologies) with the control linear lamda DNA provided in the kit as a template. $[α-^{32}P]$ GTP-labeled capping of the 1.5-kb RNA fragment was carried out as described previously [69]. Yeast Dcp1/Dcp2 decapping enzyme was purified as described by Deshmukh et al. [37]. Decapping reactions were conducted at 22 °C for 30 min in 50 mM Tris pH 7.9, 30 mM $(NH_4)_2SO_4$, 5 mM $MgCl_2$. The reaction mixtures consisted of 20 fmol capped RNA, 10 pmol Dcp1/Dcp2, 20 pmol of lysozyme or other proteins being tested. For dose-dependent assays, 5, 10 and 20 pmol of Lsm2-3-Pat1C or Lsm1-7-Pat1C were used. Reactions were terminated by addition of 25 mM EDTA. The products of the reaction were spotted onto PEI-cellulose TLC plates (MERCK) and developed in 0.75 M LiCl. ImageJ 1.45m (Wayne Rasband, National Institutes of Health) was used for quantification of decapping activity.

Protein Data Bank deposition

The coordinates and diffraction data have been deposited into the Protein Data Bank with accession code 4N0A.

References

1 Coller J, Parker R. Eukaryotic mRNA decapping. *Annu Rev Biochem* 2004; **73**:861-890.

2 Parker R, Song H. The enzymes and control of eukaryotic mRNA turnover. *Nat Struct Mol Biol* 2004; **11**:121-127.

3 Wilson T, Treisman R. Removal of poly(A) and consequent degradation of c-fos mRNA facilitated by 3′ AU-rich sequences. *Nature* 1988; **336**:396-399.

4 Shyu AB, Belasco JG, Greenberg ME. Two distinct destabilizing elements in the c-fos message trigger deadenylation as a first step in rapid mRNA decay. *Genes Dev* 1991; **5**:221-231.

5 Decker CJ, Parker R. A turnover pathway for both stable and unstable mRNAs in yeast: evidence for a requirement for deadenylation. *Genes Dev* 1993; **7**:1632-1643.

6 Arribas-Layton M, Wu D, Lykke-Andersen J, Song H. Structural and functional control of the eukaryotic mRNA decapping machinery. *Biochim Biophys Acta* 2013; **1829**:580-589.

7 Hsu CL, Stevens A. Yeast cells lacking 5′→3′ exoribonuclease 1 contain mRNA species that are poly(A) deficient and partially lack the 5′ cap structure. *Mol Cell Biol* 1993; **13**:4826-4835.

8 Beelman CA, Stevens A, Caponigro G, et al. An essential component of the decapping enzyme required for normal rates

of mRNA turnover. *Nature* 1996; **382**:642-646.

9   Dunckley T, Parker R. The DCP2 protein is required for mRNA decapping in *Saccharomyces cerevisiae* and contains a functional MutT motif. *EMBO J* 1999; **18**:5411-5422.

10 Song MG, Kiledjian M. 3′ Terminal oligo U-tract-mediated stimulation of decapping. *Rna* 2007; **13**:2356-2365.

11 Rissland OS, Norbury CJ. Decapping is preceded by 3′ uridylation in a novel pathway of bulk mRNA turnover. *Nat Struct Mol Biol* 2009; **16**:616-623.

12 Fenger-Gron M, Fillman C, Norrild B, Lykke-Andersen J. Multiple processing body factors and the ARE binding protein TTP activate mRNA decapping. *Mol Cell* 2005; **20**:905-915.

13 Rehwinkel J, Behm-Ansmant I, Gatfield D, Izaurralde E. A crucial role for GW182 and the DCP1:DCP2 decapping complex in miRNA-mediated gene silencing. *RNA* 2005; **11**:1640-1647.

14 Behm-Ansmant I, Rehwinkel J, Doerks T, *et al*. mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. *Genes Dev* 2006; **20**:1885-1898.

15 Mullen TE, Marzluff WF. Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5′ to 3′ and 3′ to 5′. *Genes Dev* 2008; **22**:50-65.

16 Geisler S, Lojek L, Khalil AM, Baker KE, Coller J. Decapping of long noncoding RNAs regulates inducible genes. *Mol Cell* 2012; **45**:279-291.

17 Lai T, Cho H, Liu Z, *et al*. Structural basis of the PNRC2-mediated link between mRNA surveillance and decapping. *Structure* 2012; **20**:2025-2037.

18 Balagopal V, Parker R. Polysomes, P bodies and stress granules: states and fates of eukaryotic mRNAs. *Curr Opin Cell Biol* 2009; **21**:403-408.

19 Eulalio A, Behm-Ansmant I, Izaurralde E. P bodies: at the crossroads of post-transcriptional pathways. *Nat Rev Mol Cell Biol* 2007; **8**:9-22.

20 Parker R, Sheth U. P bodies and the control of mRNA translation and degradation. *Mol Cell* 2007; **25**:635-646.

21 Pilkington GR, Parker R. Pat1 contains distinct functional domains that promote P-body assembly and activation of decapping. *Mol Cell Biol* 2008; **28**:1298-1312.

22 Nissan T, Rajyaguru P, She M, Song H, Parker R. Decapping activators in *Saccharomyces cerevisiae* act by multiple mechanisms. *Molecular cell* 2010; **39**:773-783.

23 Braun JE, Tritschler F, Haas G, *et al*. The C-terminal alpha-alpha superhelix of Pat is required for mRNA decapping in metazoa. *EMBO J* 2010; **29**:2368-2380.

24 Fischer N, Weis K. The DEAD box protein Dhh1 stimulates the decapping enzyme Dcp1. *EMBO J* 2002; **21**:2788-2797.

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

26 Totaro A, Renzi F, La Fata G, *et al*. The human Pat1b protein: a novel mRNA deadenylation factor identified by a new immunoprecipitation technique. *Nucleic Acids Res* 2011; **39**:635-647.

27 Chowdhury A, Mukhopadhyay J, Tharun S. The decapping activator Lsm1p-7p-Pat1p complex has the intrinsic ability to distinguish between oligoadenylated and polyadenylated

RNAs. *RNA* 2007; **13**:998-1016.

28 Bergman N, Moraes KC, Anderson JR, *et al*. Lsm proteins bind and stabilize RNAs containing 5′ poly(A) tracts. *Nat Struct Mol Biol* 2007; **14**:824-831.

29 Herrero AB, Moreno S. Lsm1 promotes genomic stability by controlling histone mRNA decay. *EMBO J* 2011; **30**:2008-2018.

30 Norbury CJ. 3′ Uridylation and the regulation of RNA function in the cytoplasm. *Biochem Soc Trans* 2010; **38**:1150-1153.

31 Heo I, Joo C, Cho J, *et al*. Lin28 mediates the terminal uridylation of let-7 precursor microRNA. *Mol Cell* 2008; **32**:276-284.

32 Hagan JP, Piskounova E, Gregory RI. Lin28 recruits the TUTase Zcchc11 to inhibit let-7 maturation in mouse embryonic stem cells. *Nat Struct Mol Biol* 2009; **16**:1021-1025.

33 Heo I, Joo C, Kim YK, *et al*. TUT4 in concert with Lin28 suppresses microRNA biogenesis through pre-microRNA uridylation. *Cell* 2009; **138**:696-708.

34 Lehrbach NJ, Armisen J, Lightfoot HL, *et al*. LIN-28 and the poly(U) polymerase PUP-2 regulate let-7 microRNA processing in *Caenorhabditis elegans*. *Nat Struct Mol Biol* 2009; **16**:1016-1020.

35 van Wolfswinkel JC, Claycomb JM, Batista PJ, *et al*. CDE-1 affects chromosome segregation through uridylation of CSR-1-bound siRNAs. *Cell* 2009; **139**:135-148.

36 Zaric B, Chami M, Remigy H, *et al*. Reconstitution of two recombinant LSm protein complexes reveals aspects of their architecture, assembly, and function. *J Biol Chem* 2005; **280**:16066-16075.

37 Deshmukh MV, Jones BN, Quang-Dang DU, *et al*. mRNA decapping is promoted by an RNA-binding channel in Dcp2. *Mol Cell* 2008; **29**:324-336.

38 Wang Z, Kiledjian M. Functional link between the mammalian exosome and mRNA decapping. *Cell* 2001; **107**:751-762.

39 Bouveret E, Rigaut G, Shevchenko A, Wilm M, Seraphin B. A Sm-like protein complex that participates in mRNA degradation. *EMBO J* 2000; **19**:1661-1671.

40 Tharun S, He W, Mayes AE, *et al*. Yeast Sm-like proteins function in mRNA decapping and decay. *Nature* 2000; **404**:515-518.

41 Sheth U, Parker R. Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. *Science* 2003; **300**:805-808.

42 Holm L, Rosenstrom P. Dali server: conservation mapping in 3D. *Nucleic Acids Res* 2010; **38**:545-549.

43 Naidoo N, Harrop SJ, Sobti M, *et al*. Crystal structure of Lsm3 octamer from *Saccharomyces cerevisiae*: implications for Lsm ring organisation and recruitment. *J Mol Biol* 2008; **377**:1357-1371.

44 Wu D, Jiang S, Bowler MW, Song H. Crystal Structures of Lsm3, Lsm4 and Lsm5/6/7 from *Schizosaccharomyces pombe*. *PLoS ONE* 2012; **7**:e36768.

45 Pannone BK, Kim SD, Noe DA, Wolin SL. Multiple functional interactions between components of the Lsm2-Lsm8 complex, U6 snRNA, and the yeast La protein. *Genetics* 2001; **158**:187-196.

46 Muhlrad D, Decker CJ, Parker R. Deadenylation of the unstable mRNA encoded by the yeast MFA2 gene leads to decapping followed by 5′→3′ digestion of the transcript. *Genes Dev*

1994; **8**:855-866.

47 Hatfield L, Beelman CA, Stevens A, Parker R. Mutations in trans-acting factors affecting mRNA decapping in *Saccharomyces cerevisiae*. *Mol Cell Biol* 1996; **16**:5830-5838.

48 Cao D, Parker R. Computational modeling of eukaryotic mRNA turnover. *RNA* 2001; **7**:1192-1212.

49 He W, Parker R. The yeast cytoplasmic LsmI/Pat1p complex protects mRNA 3′ termini from partial degradation. *Genetics* 2001; **158**:1445-1455.

50 Achsel T, Brahms H, Kastner B, *et al.* 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* 1999; **18**:5789-5802.

51 Mayes AE, Verdone L, Legrain P, Beggs JD. Characterization of Sm-like proteins in yeast and their association with U6 snRNA. *EMBO J* 1999; **18**:4321-4331.

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

53 Pomeranz Krummel DA, Oubridge C, Leung AK, Li J, Nagai K. Crystal structure of human spliceosomal U1 snRNP at 5.5 A resolution. *Nature* 2009; **458**:475-480.

54 Leung AK, Nagai K, Li J. Structure of the spliceosomal U4 snRNP core domain and its implication for snRNP biogenesis. *Nature* 2011; **473**:536-539.

55 Mund M, Neu A, Ullmann J, Neu U, Sprangers R. Structure of the LSM657 complex: an assembly intermediate of the LSm1-7 and LSm2-8 rings. *J Mol Biol* 2011; **414**:165-176.

56 Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* 2006; **22**:195-201.

57 Stevens SW, Ryan DE, Ge HY, *et al.* Composition and functional characterization of the yeast spliceosomal penta-snRNP. *Mol Cell* 2002; **9**:31-44.

58 Wu D, Lim SC, Dong Y, *et al.* Structural basis of substrate binding specificity revealed by the crystal structures of polyamine receptors SpuD and SpuE from Pseudomonas aeruginosa. *J Mol Biol* 2012; **416**:697-712.

59 Bowler MW, Guijarro M, Petitdemange S, *et al.* Diffraction cartography: applying microbeams to macromolecular crystallography sample evaluation and data collection. *Acta Crystallogr D Biol Crystallogr* 2010; **66**:855-864.

60 Gabadinho J, Beteva A, Guijarro M, *et al.* MxCuBE: a synchrotron beamline control environment customized for macromolecular crystallography experiments. *J Synchrotron Radiat* 2010; **17**:700-707.

61 Brockhauser S, Svensson O, Bowler MW, *et al.* The use of workflows in the design and implementation of complex experiments in macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 2012; **68**:975-984.

62 Kabsch W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J Appl Crystallogr* 1993; **26**:795-800.

63 Potterton E, Briggs P, Turkenburg M, Dodson E. A graphical user interface to the CCP4 program suite. *Acta Crystallogr D Biol Crystallogr* 2003; **59**:1131-1137.

64 Zwart PH, Afonine PV, Grosse-Kunstleve RW, *et al.* Automated structure solution with the PHENIX suite. *Methods Mol Biol* 2008; **426**:419-435.

65 Jones TA, Zou JY, Cowan SW, Kjeldgaard M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* 1991; **47** (Part 2):110-119.

66 Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 2004; **60**:2126-2132.

67 Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* 1997; **53**:240-255.

68 Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci USA* 2001; **98**:10037-10041.

69 Zhang S, Williams CJ, Wormington M, Stevens A, Peltz SW. Monitoring mRNA decapping activity. *Methods* 1999; **17**:46-51.

(**Supplementary information** is linked to the online version of the paper on the *Cell Research* website.)