

Manuscript EMBO-2011-77992

The structural basis of Edc3- and Scd6-mediated activation of the Dcp1:Dcp2 mRNA decapping complex

Simon A. Fromm, Vincent Truffault, Julia Kamenz, Joerg E. Braun, Niklas A. Hoffmann, Elisa Izaurralde and Remco Sprangers

Corresponding author: Remco Sprangers, Max Planck Institute for Developmental Biology

Review timeline:

Submission date:	26 April 2011
Editorial Decision:	25 May 2011
Revision received:	11 August 2011
Editorial Decision:	29 August 2011
Revision received:	29 September 2011
Editorial Decision:	12 October 2011
Accepted:	12 October 2011

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

1st Editorial Decision

25 May 2011

Thank you for submitting your manuscript for consideration by The EMBO Journal. It has been now evaluated by three referees and I enclose their reports below. As you will see the referees find the solution structure of Edc3 LSm domain bound to DCP2 LIM and its implications for decapping and P-body formation to be interesting. However, the referees raise two sets of major concerns, these are the critical technical concerns raised by referee #2 and the in vivo evidence that the interactions involving the Dcp2 C-terminal domain are important for mRNA decay and P-body formation as required by referee #1 and #3. Both these concerns are important and central to the main conclusions of the study and must be satisfactorily addressed before the manuscript can be further considered for The EMBO Journal. Nevertheless, given the interest in the study, should you be able to address the concerns we would be happy to consider a revised version of the manuscript for publication in The EMBO Journal.

I should remind you that it is EMBO Journal policy to allow a single round of revision only and that, therefore, acceptance or rejection of the manuscript will depend on the completeness of your responses included in the next, final version of the manuscript. When you submit a revised version to the EMBO Journal, please make sure you upload a letter of response to the referees' comments. Please note that when preparing your letter of response to the referees' comments that this will form part of the Review Process File, and will therefore be available online to the community. For more details on our Transparent Editorial Process initiative, please visit our website: <http://www.nature.com/emboj/about/process.html>

Thank you for the opportunity to consider your work for publication. I look forward to your revision.

Yours sincerely,

Editor
The EMBO Journal

REFEREE COMMENTS

Referee #1

In this paper, Fromm et al identified an LSM interaction motif (LIM-1) in the *S. pombe* Dcp2 protein located downstream of the Nudix domain, which mediates the interaction of Dcp2 protein with the LSM domain of Edc3. The solution structure of the two domains revealed that the Edc3 LSM domain interacts with residues 242-291 of Dcp2 in an unexpected manner distinct from previous LSM modes of binding. The LSM domain of another activator of decapping, Scd6 also interacts in a similar way. They further show that Dcp2 contains multiple LIM domains and may accommodate multiple interactions of the Edc3 and Scd6 LIM domains in the C-terminal extension of Dcp2 and may contribute to the formation of P-body and enhancement of decapping. Interestingly, the LIM domain seems to reside in the Dcp1 protein in metazoans. These findings are potentially interesting and could shed light on the formation of P-bodies and the regulation of RNA decay, but several points listed below should be addressed including the biological significance of the Dcp2 C-terminal extension in cells.

1. Page 11, end of first paragraph. Modeling the Edc3 LSM domain onto the Dcp1:Dcp2 LIM-1 domain reveals steric clashes and the authors suggest "This observation clearly indicates that Dcp1:Dcp2 cannot adopt the conformation found in the closed crystal form while bound to Edc3 or Scd6". Since the Dcp1:Dcp2 structures were derived in the absence of other proteins, is it possible the regulator region could adopt a conformation to accommodate the Edc3 LSM domain? Similarly, the authors suggest: "Edc3 binding is incompatible with this closed conformation". Is there evidence that it is compatible with the open conformation?

2. The data in figure 6A is presented relative to Dcp1:Dcp2 complex containing the truncated Dcp2 1-289 protein. How active is this truncated Dcp2 complex? (what percent of substrate is decapped in the assay conditions used?) In addition, it is surprising that 10 fold molar excess of Edc3 LSM domain is required to see a stimulation of Dcp1:Dcp2 decapping. Do the authors have an explanation for the requirement of the excess protein. And was a similar excess required in the structural analysis? These points raise questions about the biological significance of the observed interactions. Is the Dcp2-C-terminal domain functional in yeast cells as the biochemical data may indicate in the formation of P-bodies and/or mRNA decay?

3. In Fig.1D and Fig.3B, inputs should be included to show all the proteins were expressed at equivalent levels in *E. coli*.

4. Could the authors expand within the discussion whether LSM motifs in proteins other than Edc3 and Scd6 could also interact with Dcp2 (or Dcp1 in metazoans), or whether the LSM in these two proteins may be unique?

5. The authors also postulate that the LIM-LSM interaction may contribute to P body formation. But, since Dcp2 is not required for P body assembly, it's not obvious why this would be the case. Are there other Pbody proteins that contain LIM domains?

6. Is the following statement at the bottom of page 6 correct?
"the loop preceding β -strand 5 and β -strand 5"

7. Middle of page 14: In the statement, "metazoan Dcp1:Edc4:Dcp2 decapping", do the authors mean Edc3?

8. Abstract: "mNRA" should be "mRNA"

Referee #2

This manuscript reports an NMR-based study of the interactions between various components of the mRNA de-capping complex, including particularly the Edc3 LSm domain and a fragment from Dcp2 that is disordered in isolation, part of which apparently forms a helix upon interacting with the Ecd3 LSm domain. I am not an expert in this area of biology, but it seemed to me that the structures and interactions described would represent a step forward in understanding the mechanisms that regulate this system and may therefore be of sufficiently widespread interest to justify publication in EMBO Journal. However, as matters stand there are very serious deficiencies in the description of the NMR structure determination, and possibly in the structure determination itself, that would need to be corrected before the manuscript could be considered acceptable.

1) One major problem that I have with the manuscript is that there is essentially no description given as to how the complex between the Edc3 LSm domain and the Dcp2 peptide was actually solved, despite the fact that this structure is absolutely the central result of the whole study. Generic descriptions of the NMR experiments and structure calculations are given, but these are extremely brief and they relate only to solving a single species by NMR, with no additional details whatsoever given to describe how the significantly more difficult task of solving the structure of the complex was achieved. It is simply not the case that solving the structure of a complex is like solving the structure of a slightly larger protein, as the authors appear to believe. Key questions that are at present totally ignored in the manuscript include:

a) How was the complex prepared? The experimental section refers to co-expression in deuterated media for preparing the Dcp1 : Dcp2 complex, but says nothing at all about how the Edc3 LSm domain : Dcp2 complex, which is the main object of the study, was prepared. Were the components made separately and then mixed in one step, or was one component titrated into the other (as in Figure 1)? How was 1:1 stoichiometry achieved or checked? Were both components isotope labelled with ^{13}C and ^{15}N , or was one component labelled and the other not, so as to allow filtered experiments to detect selectively different types of signals? Did different samples have different labelling patterns?

b) How were the components of the complex differentially assigned? The description of the assignment method is referred back recursively to a much earlier publication on a quite different system (*Pyrococcus horikoshii* PhS018, described in Coles et al., *Structure*, 14, 1489, 2006), but the earlier system is a single protein, not a complex. In the present case, were experiments done that allowed the resonances from the two components to be distinguished, or to be observed separately, e.g. using filtered experiments? This is of course linked to the question of how samples were isotope labeled.

c) Also related to this is the still more important question of how intermolecular NOE enhancements were distinguished from intramolecular NOEs. Were filtered NOE experiments carried out? If all NOEs were measured in samples where both components were uniformly labeled, how unambiguous were the assignments of the intermolecular NOEs? On this note, it is quite incredible that the very existence of the 22 intermolecular NOEs, which are absolutely the key data that define the nature of the interface that is the main result of the entire paper, is disclosed only in a footnote to a Supplementary Table. This is very poor. There should be a Supplementary Figure that illustrates some of the intermolecular NOEs in the spectra so that the reader can judge the quality of these crucial data.

d) How were the NOEs calibrated to yield specific distance constraints? We are told absolutely nothing about this, not even whether they were assigned to intensity bins (strong, medium, weak etc.) or set to individually calibrated ranges. If the intermolecular NOEs were measured in different experiments from the intramolecular NOEs (e.g. using filtered experiments), how were the intensities from one experiment calibrated relative to those from others?

e) Separate structural statistics should be given for the individual components of the complex, in addition to the combined statistics that are all that is given at present. Also, ensemble views should be presented (at least in the Supplementary Material) to allow readers to assess the precision of the structures visually.

f) How many ensemble members were calculated in total? Footnote 1 of Tables S2 shows that the final ensembles contained 19 structures for the free protein and 21 for the complex, but it is important to know from how large a set were these final structures chosen, and on what criteria were they selected? It is often a hallmark of poorly defined structures that the calculations only converge poorly so that only a small proportion of the total number of ensemble members calculated can be used.

2) Another problem that I have concerns the very low number of NOEs used to determine the structures. According to Supplementary Table S1, for the free Ecd3 LS_m domain there are just 356 NOEs assigned, of which only 168 are medium or long-range (the other 157 are intra-residue or sequential and thus have little influence on the outcome of the structure calculations). This is in the region of 3.5 constraints per residue, of which probably only 1.5 per residue or thereabouts are structurally significant. This is an extraordinarily low number, and it leaves me puzzled as to how a reasonably high-resolution structure was actually obtained. The situation for the Ecd3 LS_m domain : Dcp2 complex is little better: there are just 436 NOEs across roughly 110 ordered residues, of which 199 are medium- or long-range), which again leads to a figure of roughly 2 active NOE constraints per residue. I do not understand why so few NOEs were assigned, or how this number was enough to solve the structure. Given that 3D ¹³C-NOESY spectra were acquired, I would have expected that very many more NOEs would have been assigned.

One possible reason for the low number of NOEs apparently measured could potentially be that the protein was only available at low concentration, however as far as I could see we are not told what the actual protein concentration was in the NMR samples. This information must be added to the manuscript. Another possible reason could be that the assignments may have been highly incomplete. Again, we are told nothing about the extent of assignment. Assignment data are deposited at BMRB, but there should still be a summary statement of their completeness in the manuscript - and of course at this stage, prior to release, I have no access to the deposition and cannot tell anything whatsoever about the extent of assignment.

3) A related question is how the hydrogen bond constraints were defined. These are very powerful constraints, and as the number of active NOEs is so low it is likely that the H-bond constraints may be playing a very significant role in defining the structure. At present nothing at all is said about how these were assigned, indeed the fact that they were even used is only revealed by the existence of entries for H-bond constraints in Supplementary Tables S1 and S2. Are they assigned based on evidence that particular NH signals exchange slowly with solvent? If so, how were these slowly exchanging NH's identified? How were acceptors defined for each H-bond? Or are the constraints all based on structural expectations, i.e. are they non-experimental? The Footnote to Table S3 states that details of the way in which H-bond constraints were handled during the structure calculations is described in Materials and Methods, yet in fact nothing whatsoever is said about H-bond constraints in Materials and Methods.

The question of H-bonds is particularly relevant to the conformation of Dcp2 in the bound state, which is helical in the calculated structures. Comparison of Tables S1 and S2 shows that there were 16 more H-bond constraints (44-28) used for the complex calculations than for the free Ecd3 LS_m domain. Although it is not stated how these were distributed, it seems very likely that some of these may have been within the helix in the bound Dcp2 chain. Given that it is this helical conformation of the bound Dcp2 that the authors state is "unprecedented" and one of the key novel features disclosed in the paper, it is absolutely essential that the experimental data on which this helical conformation is based are carefully described, especially that which allowed the authors to define any helical H-bonds that they may have constrained. At present there is no description of this at all.

4) The discussion of the titration of Dcp2 into Edc3 (near bottom of page 5, describing Figure 1) is completely illogical. The authors say that because the system was in slow exchange, assignments could not be transferred from the free to the bound state of the Edc3 LS_m domain, and that therefore the interface could not be mapped. It is true that slow exchange makes it more difficult to obtain assignments for the bound state, as in such cases bound-state assignments need to be determined

completely independently, without being able to use data from the free state to help the process. However, this certainly does not mean that assignment is impossible. Indeed, it very clearly must have been possible in the present case, since the authors went on to solve the solution structure of the complex, the first and absolutely necessary stage of which is to assign the signals. There is thus a fundamental contradiction between their adjacent statements that a) the interface could not be mapped due to the impossibility of making assignments, and b) they solved this problem by determining the solution structure. I simply cannot understand how they were even able to write this down. It also means that, despite their statement to the contrary, it must be entirely trivial to produce a figure showing the chemical shift changes for backbone amide groups of Edc3 (as a function of sequence) upon addition of Dcp2.

5) It seems that the Dcp2 fragment used in the NMR experiments comprised residues 242-291, but the part of this fragment that becomes structured on binding to Edc3 is just residues 254-266. This implies that some 37 residues of the Dcp2 fragment remain unfolded in the complex. How did the signals from these residues affect the NMR experiments? Were they assigned (presumably not)? Did they show NOEs? Were their signals essentially identical in the presence or absence of Edc3? How did the presence of these signals from disordered regions affect the assignment process for signals from the ordered regions?

6) The authors state that the structure of the Edc3 LSm domain is unchanged when Dcp2 binds, but no data is given. I would expect to see a backbone rmsd calculated for Edc3 between the average structures for each of the two ensembles (free and bound) to back up this statement.

Referee #3

In this manuscript, the group of E. Izzauralde explores the basis for interaction occurring between decapping factor and Edc3/Scd6. The results are interesting but preliminary. The main weakness is that the authors do not provide support for the biological roles of the proposed interactions. Currently, the manuscript focuses on structural data with a single biochemical experiment suggesting biological relevance. Given that the authors analyze yeast (*S. pombe*) factors and that the RNA degradation process can easily be analyzed in simple model species (e.g., *S. pombe* or *S. cerevisiae*) one would expect that they provide *in vivo* evidence for the role of the Dcp2 C-terminal region in P bodies formation (in relation with the clustering hypothesis presented in Figure 4C) and support for the role of the Dcp2 C-terminal tail in decapping activation. This additional functional information and a more careful writing of the manuscript (see below) are required to make this manuscript of interest to the broad readership of the EMBO Journal.

- Throughout the manuscript, the authors have an incomplete and biased presentation of the published literature. Surprisingly, a significant fraction of the unquoted references have been published in the EMBO Journal whose long-term impact is thus undermined by this situation.

* The authors use "LSm interaction motif (LIM)" to identify the repeated sequence present in Dcp2 that mediate interaction with Edc3/Scd6. However, these sequences have been noticed earlier and defined as "Helical Leucine-rich Motifs (HLMs)" (Gaudon et al., EMBO J. (1999) 18:2229-40). To avoid generating confusion and giving credit to this earlier publication, the original name should be kept.

* Credits for the description of Dcp2 as the decapping enzyme is given to a review emanating from a lab which erroneously presented Dcp1 as the decapping enzyme for years, while not a single reference present some of the groups active in this area that were primarily involved in establishing this fact (Wang et al. Proc Natl Acad Sci U S A. (2002) 99:12663-8; van Dijk et al. EMBO J. (2002) 21:6915-24, see also Lykke-Andersen J. Mol Cell Biol. (2002) 22:8114-21).

* Similarly, the description of decapping activators (3 references to the work of a single lab) involved additional groups (e.g., Boeck et al. Mol Cell Biol. (1998) 18:5062-72; Bouveret et al. EMBO J. (2000) 19:1661-71; Bonnerot et al. Mol Cell Biol. (2000) 20:5939-46 for Pat and LSm proteins) that are not mentioned.

* Page 3, line 10: The presence of an EVH1 domain in Dcp1 was originally identified by Callebaut

(FEBS Lett. (2002) 519:178-80).

* Page 4, second paragraph: The authors should mention the functional redundancy between Scd6 and Edc3 identified by the Jacquier group (Decourty et al. Proc Natl Acad Sci U S A. (2008) 105:5821-6).

* Page 6, first line: The authors may want not to forget the pioneering work of the Nagai group on Sm domain-Sm domain interaction (Kambach et al. Cell (1999) 96:375-87) and the early work of the Suck group (Toro et al. EMBO J. (2001) 20:2293-303, Thore et al. J. Biol Chem. (2003) 278: 239-47) for different modes of RNA binding by Sm proteins.

* Page 9, line 18: Beside Tarassov et al. and Nissan et al., evidence for interaction of Dcp2 with Dhh1, Pat1 and Scd6 have been reported by Decker et al. (J Cell Biol. (2007) 179:437-49); Fromont-Racine et al. (Nat Genet. (1997) 16:277-82) and Decourty et al. (Proc Natl Acad Sci U S A. (2008) 105:5821-6).

- Page 5, first lines: "...and found that a Dcp2 region located between residues 254 and 266 is sufficient for the interaction with the Edc3 LSM domain (Figure 1D, lane 3 vs. 1 and 2)." is not correct. The data show that the region is necessary but not that it is sufficient. The data leading to this conclusion should be shown especially as the sequence requirements for binding remain unclear.

- Page 6, line 7: What does "we solved the high-resolution solution structure" mean? The authors should be more explicit.

- Page 9, line 7 and Figure 4 A and B: In contrast to their statement, the authors have not shown that a "hexameric" complex is formed. Gel filtration or ultracentrifugation data should be provided to support this conclusion.

- Page 10, top paragraph and Figure 5 C: As a specificity control for the competition between Scd6 and Edc3, the authors should demonstrate that addition of an unrelated protein does not induce the dissociation of Scd6 LSM and LIM-I.

- Page 11: The authors conclude, "the closed Dcp2 structure is incompatible with Edc3 binding". Given that Edc3 can bind different sequences present in the C-terminal extension of Dcp2, the evidence presented does not support this statement. Moreover, given the redundancy of interaction, the biological consequences of this conclusion are unclear. Altogether, this part is weak and does not deserve a full paragraph of the Results section. It would be more appropriate to present it in a couple of sentences in the Discussion. (See also contradiction with second paragraph of Discussion.)

- Last section of Results and Figure 7 are not entirely convincing, especially the similarity between the sequence present in metazoan Dcp1 and HLMS/LIMs.

- Page 13: The authors discuss the sequence requirements for interaction between Dcp2 and LSM domains. As they identify sequences that do and do not interact and have obtained some structural information, they should probably be more explicit on the features required for binding or preventing interaction.

- Page 14: "However, it should be noted that the activation effect is likely more prominent in a cellular context in which additional factors are present that influence the Dcp2 inter-domain orientation." Without data or references to support this statement, it remains only wishful thinking. This is critical as this relates to the main conclusion of the manuscript.

Referee #1

1. Page 11, end of first paragraph. Modeling the Edc3 LSm domain onto the Dcp1:Dcp2 LIM-1 domain reveals steric clashes and the authors suggest "This observation clearly indicates that Dcp1:Dcp2 cannot adopt the conformation found in the closed crystal form while bound to Edc3 or Scd6". Since the Dcp1:Dcp2 structures were derived in the absence of other proteins, is it possible the regulator region could adopt a conformation to accommodate the Edc3 LSm domain?

The interaction between the Dcp2 helix (residues 257-266) and the Dcp2 regulatory domain provides the contacts that lock the decapping complex in the closed conformation. The residues between Dcp2 residues 257 and 266 that are involved in this intramolecular interaction are to a large extent the same residues that interact with the Edc3 LSm domain. As an example: the side chain of Leu 260 in Dcp2 contacts Val 63 in Dcp2 in the closed conformation of the decapping complex. In the complex with Edc3, Leu 260 makes close contacts with the core of the Edc3 Lsm domain (Figure 2B). This Dcp2 helix thus needs to dissociate from the regulatory domain in order to allow for the interaction with the Edc3 LSm domain.

The size of the Edc3 LSm domain (64 residues) is comparable to the size of the Dcp2 regulatory domain (95 residues). As such, it is hard to imagine how structural changes in the Dcp2 regulatory domain could accommodate the Edc3 or Scd6 LSm domains.

We have now changed the wording of the main text in order to clarify this point (page 13).

Similarly, the authors suggest: "Edc3 binding is incompatible with this closed conformation". Is there evidence that it is compatible with the open conformation?

There are two crystal structures of the open conformation of the Dcp1:Dcp2 complex. In both these structures the Dcp2 region that interacts with the LSm domains is invisible. This suggests that these residues are disordered in the open conformation and that they are thus available for the interaction with Edc3. That the interaction of Edc3 with Dcp2 is compatible with the open conformation of the decapping complex is supported by our NMR data. In Figure S3B we show the spectrum of the Dcp2 catalytic domain plus the region that interacts with Edc3. Structurally this construct resembles the open conformation of the decapping complex, as the regulatory region (that is required for the closure of the complex) is not present. Addition of the Edc3 LSm domain leads to the formation of an Dcp2:Edc3 complex, whereby the disordered region adopts an helical conformation. As such, the two available open conformations of the decapping complex are both compatible with Edc3 binding. We added a sentence to the manuscript to indicate this (page 13).

2. The data in figure 6A is presented relative to Dcp1:Dcp2 complex containing the truncated Dcp2 1-289 protein. How active is this truncated Dcp2 complex? (what percent of substrate is decapped in the assay conditions used?)

All of the decapping complexes are catalytically active. During the assay, around 15% of the capped mRNA substrate was converted into product.

In addition, it is surprising that 10 fold molar excess of Edc3 LSm domain is required to see a stimulation of Dcp1:Dcp2 decapping. Do the authors have an explanation for the requirement of the excess protein.

The interaction between the Edc3 LSm domain and the decapping complex is very tight. An equimolar mixture of the Edc3 LSm domain and the decapping complex results in the complete formation of the tertiary Dcp1:Dcp2:Edc3 complex. The interaction of the Scd6 Lsm domain with the decapping complex is much weaker. We thus need an excess of Scd6 over Dcp2 in order to saturate the decapping complex with the Scd6 LSm domain.

To be able to directly compare the stimulatory effect Edc3 and Scd6 (and to prevent secondary effects resulting e.g. from viscosity or total amount of protein present in the assay), we used the same excess of activator in all experiments. The 10- fold excess of Edc3 was thus not required for activation, but used in order to be consistent with the Scd6 activated assays.

We have added this fact to the text in order to prevent confusion (page 13 and page S27).

And was a similar excess required in the structural analysis? These points raise questions about the biological significance of the observed interactions.

The interaction between the Edc3 LSM domain and Dcp2 is very tight. We now mention this more directly in the manuscript to emphasize the biological relevance of the interaction.

For the structural analysis, we co-expressed the Edc3 LSM domain and the Dcp2 linear sequence motif. Subsequently, the complex was purified using a histidine affinity tag located on the Dcp2 peptide. During the (long) purification, the protein and the peptide stayed in complex, and we obtained a complex that contains Edc3 and Dcp2 in an 1:1 ratio.

We have now added a detailed method section in the supplementary material that discusses all expression and purification protocols.

Is the Dcp2- C- terminal domain functional in yeast cells as the biochemical data may indicate in the formation of P- bodies and/or mRNA decay?

In the revised version of the manuscript we included new experiments that address the functional role of the Dcp2 C-terminal residues in fission yeast. These experiments clearly show that the C-terminal residues of Dcp2 are necessary for the recruitment of the Dcp1:Dcp2 decapping complex to P- bodies (page 13-14, Figures 6 and S8).

The Edc3 protein is able to cluster multiple decapping complexes. These bridging interactions are however not required for P-body formation, as we still observe P-bodies (though Edc3 or Lsm7) in yeast strains that lack all motifs that can interact with Edc3. This finding is not too surprising, as P-body formation is a redundant process, where multiple intermolecular interactions contribute to the clustering of decapping complexes. Removal of any of these single interactions is not sufficient to disrupt P-body formation. A role of Dcp2 in the formation of P-bodies has been reported (Teixeira & Parker, 2007) and our *in vitro* experiments provide a plausible mechanism for how Dcp2 and Edc3 can contribute to the formation of larger assemblies.

3. In Fig.1D and Fig.3B, inputs should be included to show all the proteins were expressed at equivalent levels in E. coli.

The experiment shown in Figure 1D was obtained by overexpression of the Edc3 LSM domain and different versions of the decapping complex. The decapping complex was selectively purified from the cell lysate and was analyzed by SDS- PAGE. As the Edc3 LSM domain is small and contains no tags, it is hard to visualize quantitatively in the cell lysate by SDS-PAGE analysis.

We thus decided to repeat the experiment shown in Figure 1D. To make sure that an equal amount of the Edc3 LSM domain was present in all experiments we now expressed the different versions of the decapping complex and the Edc3 LSM domain separately. We then supplemented each of the decapping complexes with the same amount of Edc3 LSM domain and performed the pull down experiments. The results we obtained are identical with the ones obtained in the previous experiment: Dcp2 residues 255- 266 are required for the interaction with Edc3.

The inputs of Figure 3B are shown in Supplementary Figure 6. This clearly indicates that the MBP-Dcp2 sequences were present in the input of all pull down experiments.

4. Could the authors expand within the discussion whether LSM motifs in proteins other than Edc3 and Scd6 could also interact with Dcp2 (or Dcp1 in metazoans), or whether the LSM in these two proteins may be unique?

The question of the referee addresses two interesting points. First, whether additional proteins exist that can interact with the linear motifs we identified here in Dcp2 (Dcp1), and secondly, whether additional proteins exist that possess similar linear sequence motifs that can interact with the Edc3 and or Scd6 LSM domains.

Most LSM domain containing proteins form ring like structures in solution. This multimerization occludes the LSM residues that are important for the interaction with the helical motifs. Of the few proteins that contain a monomeric N-terminal LSM domain only Edc3 (LSM16) and Scd6 (LSM13) have, to our knowledge, been found in P-bodies. Thus, it is not possible to predict whether other LSM domain containing proteins interact with the Dcp2 (Dcp1) linear motifs.

We do now expand in the discussion on additional proteins that might use linear motifs to interact with the Edc3 LSM domain. We explicitly mention the Rps28B protein that depends on Edc3 for the degradation of its own mRNA. Rps28B contain a C- terminal extension that harbors a sequence motif that resembles an HLM and that can thus potentially interact with the Edc3 LSM domain. It is possible that the Edc3 dimer interacts with one of its two LSM domains with Rps28B-mRNA complex, whereas the other LSM domain interact with Dcp2. In that way, a direct link between an mRNA and the decapping complex can be established through the Edc3 protein and the linear motifs in Dcp2 and the Rps28B protein (page 17 in the discussion).

5. The authors also postulate that the LIM- LSM interaction may contribute to P body formation. But, since Dcp2 is not required for P body assembly, it's not obvious why this would be the case. Are there other Pbody proteins that contain LIM domains?

We have shown *in vitro* that one Edc3 dimer is able to cluster multiple decapping complexes. Clearly a complex network of additional Edc3:Dcp2 interactions is plausible, which could result in the formation of large decapping assemblies.

Deletion of the LSM interaction motifs in Dcp2 prevents the accumulation of the decapping complex in P- bodies. However, as the referee indicates, Dcp2 is not required for P-body formation. In agreement with that, we observe that P-bodies are present in yeast cells in the absence of the Dcp2 linear sequence motifs.

It should be noted that P-bodies are formed by multiple, redundant protein:protein and protein:RNA interactions. As such there is no single protein that when deleted results in the complete loss of observable P-bodies, as the referee mentions. Previous experiments have, however, shown that Dcp2 can contribute to P-body formation (Teixeira & Parker, 2007). As mentioned above, it is thus plausible that the Edc3 mediated clustering of decapping complexes that we observe *in vitro* contributes to P-body formation *in vivo*.

We apologize for the confusing statement in the previous version of the manuscript and have modified the text accordingly.

6. Is the following statement at the bottom of page 6 correct? "the loop preceding β -strand 5 and β -strand 5"

The statement is correct. We nevertheless rewrote the sentence. It now reads: "the surface of the protein formed by β -strands 2 and 5, and the loop preceding β -strand 5"

7. Middle of page 14: In the statement, "metazoan Dcp1:Ecd4:Dcp2 decapping", do the authors mean Edc3?

This should read Edc4 as the Edc4 protein is required to bridge the Dcp1-Dcp2 interaction in metazoan. To prevent confusion, we used the name Ge-1 instead of Edc4.

8. Abstract: "mNRA" should be "mRNA" Corrected.

Referee #2

This manuscript reports an NMR-based study of the interactions between various components of the mRNA de-capping complex, including particularly the Edc3 LSM domain and a fragment from Dcp2 that is disordered in isolation, part of which apparently forms a helix upon interacting with the Ecd3 LSM domain. I am not an expert in this area of biology, but it seemed to me that the structures and interactions described would represent a step forward in understanding the mechanisms that regulate this system and may therefore be of sufficiently widespread interest to justify publication in EMBO Journal. However, as matters stand there are very serious deficiencies in the description of the NMR structure determination, and possibly in the structure determination itself, that would need to be corrected before the manuscript could be considered acceptable.

We apologize to the reviewer for not writing a more detailed method section concerning the sample preparation and structure determination. We do realize that this was a weak point of our manuscript and have now added an extensive materials and methods section to the supplementary information. This includes a list of all expression plasmids and samples that were used, including the preparation and labeling. Also, a clear explanation of the structure calculation protocols is now added.

1) One major problem that I have with the manuscript is that there is essentially no description given as to how the complex between the Edc3 LSM domain and the Dcp2 peptide was actually solved, despite the fact that this structure is absolutely the central result of the whole study. Generic descriptions of the NMR experiments and structure calculations are given, but these are extremely brief and they relate only to solving a single species by NMR, with no additional details whatsoever given to describe how the significantly more difficult task of solving the structure of the complex was achieved. It is simply not the case that solving the structure of a complex is like solving the structure of a slightly larger protein, as the authors appear to believe. Key questions that are at present totally ignored in the manuscript include:

a) How was the complex prepared? The experimental section refers to co-expression in deuterated media for preparing the Dcp1 : Dcp2 complex, but says nothing at all about how the Edc3 LSM domain : Dcp2 complex, which is the main object of the study, was prepared. Were the components made separately and then mixed in one step, or was one component titrated into the other (as in Figure 1)? How was 1:1 stoichiometry achieved or checked? Were both components isotope labelled with ^{13}C and ^{15}N , or was one component labelled and the other not, so as to allow filtered experiments to detect selectively different types of signals? Did different samples have different labelling patterns?

In summary, the Edc3:Dcp2 complex was prepared (as also mentioned in response to referee #1) by co-expression of both the Edc3 LSM domain and Dcp2 residues 242-291 (see table S2,S3). Due to the co-expression of the two proteins, both Edc3 and Dcp2 were labeled with NMR active nuclei. The protein complex that formed in the *E. coli* cells was purified using a histidine tag on Dcp2. Due to the strong interaction of the Edc3 LSM domain with Dcp2 and due to the removal of a potential excess of free LSM domain or a potential excess of free Dcp2 during the purification, the NMR samples contained the LSM domain and Dcp2 in a 1:1 ratio. This stoichiometry of the complex was confirmed with NMR spectra that showed no signals of the free Edc3 LSM domain or of free Dcp2. For the structure determination of the complex, samples that were 100% ^{15}N labeled or 100% $^{15}\text{N}^{13}\text{C}$ labeled were prepared (see Table S3).

In line with the referees suggestion, we have now included a detailed description of the sample preparation in the supplementary information.

*b) How were the components of the complex differentially assigned? The description of the assignment method is referred back recursively to a much earlier publication on a quite different system (*Pyrococcus horikoshii* PhS018, described in Coles et al., *Structure*, 14, 1489, 2006), but the earlier system is a single protein, not a complex. In the present case, were experiments done that allowed the resonances from the two components to be distinguished, or to be observed separately, e.g. using filtered experiments? This is of course linked to the question of how samples were isotope labeled.*

As the molecular weight of the structured part of the complex of the Edc3 LSm domain and the Dcp2 peptide is small (7.6 kDa) and because the interaction between the two components is very strong, we could treat the complex as a single protein in our resonance assignment. This is possible as the Edc3:Dcp2 complex is an asymmetric heterodimer, and no formal ambiguity exists between inter- and intra- molecular contacts (in contrast to symmetrical homodimers, where such ambiguity always exists). Isotope filtered experiments may be useful in solving heterodimeric complexes, but in this case, where the amount of resonance overlap is very small, they were not required.

We did prepare one sample of a ^{15}N labeled Edc3 LSm domain with an unlabeled Dcp2 peptide, which confirmed the separation of resonances into resulting from either Edc3 or Dcp2. A plot of this spectrum is shown in Figure 1E of the manuscript.

We have now described these issues clearly in the supplementary information.

c) Also related to this is the still more important question of how intermolecular NOE enhancements were distinguished from intramolecular NOEs. Were filtered NOE experiments carried out? If all NOEs were measured in samples where both components were uniformly labeled, how unambiguous were the assignments of the intermolecular NOEs? On this note, it is quite incredible that the very existence of the 22 intermolecular NOEs, which are absolutely the key data that define the nature of the interface that is the main result of the entire paper, is disclosed only in a footnote to a Supplementary Table. This is very poor. There should be a Supplementary Figure that illustrates some of the intermolecular NOEs in the spectra so that the reader can judge the quality of these crucial data.

As mentioned in point b, due to the small size of the complex, we had little to no problems with resonance overlap and as such had no problem distinguishing intra- from intermolecular NOEs in the 3D ^{15}N or ^{13}C based NOESY spectra. The number of 22 intermolecular NOEs is sufficient to dock the 6 residues folded part of the Dcp2 helix unambiguously on the Edc3 LSm domain. We have now included a plot in the supplementary material that indicates the assigned NOE restraints on the structure (Figure S2E). Both ends of the helix are clearly positioned on the Edc3 LSm domain due to the unambiguous NOE restraints. This accuracy is reflected in the low rmsd of the ensemble of structures of the complex (Figure S2B).

d) How were the NOEs calibrated to yield specific distance constraints? We are told absolutely nothing about this, not even whether they were assigned to intensity bins (strong, medium, weak etc.) or set to individually calibrated ranges. If the intermolecular NOEs were measured in different experiments from the intramolecular NOEs (e.g. using filtered experiments), how were the intensities from one experiment calibrated relative to those from others?

All NOE intensities were divided by the intensity of the corresponding HSQC peak. Then the NOE peak intensities were binned into four different distance categories in a standard manner.

This is now clearly mentioned in the Supplementary methods.

e) Separate structural statistics should be given for the individual components of the complex, in addition to the combined statistics that are all that is given at present. Also, ensemble views should be presented (at least in the Supplementary Material) to allow readers to assess the precision of the structures visually.

We have added the requested information to the structural statistic table (Table S1C). We have also included a picture of the ensemble of the free Edc3 LSm domain and of the Edc3:Dcp2 complex in the supplementary information (Figure S2A,B).

f) How many ensemble members were calculated in total? Footnote 1 of Tables S2 shows that the final ensembles contained 19 structures for the free protein and 21 for the complex, but it is important to know from how large a set were these final structures chosen, and on what criteria were they selected? It is often a hallmark of poorly defined structures that the calculations only converge poorly so that only a small proportion of the total number of ensemble members calculated can be used.

We agree with the referee that this information is indeed crucial to be able to judge the structural quality. For both structures an ensemble of 50 structures was calculated. Based on lowest restraint violations, we selected 19 structures for the Edc3 LSm domain and 21 structures for the Edc3:Dcp2 complex. The convergence of the calculations is thus very good, which confirms the high quality of our structures. This information has now been included in the supplementary methods.

2) Another problem that I have concerns the very low number of NOEs used to determine the structures. According to Supplementary Table S1, for the free Ecd3 LSm domain there are just 356 NOEs assigned, of which only 168 are medium or long- range (the other 157 are intra-residue or sequential and thus have little influence on the outcome of the structure calculations). This is in the region of 3.5 constraints per residue, of which probably only 1.5 per residue or thereabouts are structurally significant. This is an extraordinarily low number, and it leaves me puzzled as to how a reasonably high-resolution structure was actually obtained. The situation for the Ecd3 LSm domain : Dcp2 complex is little better: there are just 436 NOEs across roughly 110 ordered residues, of which 199 are medium- or long- range), which again leads to a figure of roughly 2 active NOE constraints per residue. I do not understand why so few NOEs were assigned, or how this number was enough to solve the structure. Given that 3D ¹³C-NOESY spectra were acquired, I would have expected that very many more NOEs would have been assigned.

One possible reason for the low number of NOEs apparently measured could potentially be that the protein was only available at low concentration, however as far as I could see we are not told what the actual protein concentration was in the NMR samples. This information must be added to the manuscript. Another possible reason could be that the assignments may have been highly incomplete. Again, we are told nothing about the extent of assignment. Assignment data are deposited at BMRB, but there should still be a summary statement of their completeness in the manuscript - and of course at this stage, prior to release, I have no access to the deposition and cannot tell anything whatsoever about the extent of assignment.

As expected for a small protein concentrated at round 1mM (Table S3), the NMR data were of very good quality. Further it is important to notice that 5 different 3D-NOESY spectra (see supplementary methods) have been measured, providing the set of NOEs necessary for high-resolution solution structure determination.

The structured region of the Edc3-LSm domain covers only around 58 residues, so there are over 6 distance restraints per structured residue, nearly 3 of which are medium- or long-range. The figures are similar for the complex, which has around 70 structured residues.

The structure determination strategy has now been more extensively described in the supplementary methods. This protocol is based on establishing well-defined local conformation via back-calculation of expectation NOESY spectra. In this process, many local (intra-residue and sequential) NOEs have been used to define dihedral restraints, rather than applied as distance restraints. For this reason the number of intra-and sequential NOEs is small, but the number of dihedral restraints is high. Together with the medium- and long- range contacts, these data are sufficient to calculate a precise structure. A statement of the percentage of resonances assigned has been added to the manuscript: 93.49 % (backbone) and 87.53 % (side- chains) for the Edc3- LSm domain; 97.47 % and 91.74 % for the Edc3:Dcp2 complex.

3) A related question is how the hydrogen bond constraints were defined. These are very powerful constraints, and as the number of active NOEs is so low it is likely that the H-bond constraints may be playing a very significant role in defining the structure. At present nothing at all is said about how these were assigned, indeed the fact that they were even used is only revealed by the existence of entries for H-bond constraints in Supplementary Tables S1 and S2. Are they assigned based on evidence that particular NH signals exchange slowly with solvent? If so, how were these slowly exchanging NH's identified? How were acceptors defined for each H-bond? Or are the constraints all based on structural expectations, i.e. are they non-experimental? The Footnote to Table S3 states that details of the way in which H-bond constraints were handled during the structure calculations is described in Materials and Methods, yet in fact nothing whatsoever is said about H-bond constraints in Materials and Methods.

The question of H- bonds is particularly relevant to the conformation of Dcp2 in the bound state, which is helical in the calculated structures. Comparison of Tables S1 and S2 shows that there were 16 more H- bond constraints (44- 28) used for the complex calculations than for the free Edc3 LSm domain. Although it is not stated how these were distributed, it seems very likely that some of these may have been within the helix in the bound Dcp2 chain. Given that it is this helical conformation of the bound Dcp2 that the authors state is "unprecedented" and one of the key novel features disclosed in the paper, it is absolutely essential that the experimental data on which this helical conformation is based are carefully described, especially that which allowed the authors to define any helical H- bonds that they may have constrained. At present there is no description of this at all.

Under the conditions used for both the free protein and the complex, amide protons not involved in hydrogen bonding interactions gave only weak signals in ¹⁵N-HSQC-based experiments. For this reason it was not critical to explicitly measure amide proton exchange rates. H- bonds restraints were only applied in regions of canonical secondary structure, as identified by combining secondary shift information (TALOS) with expected NOE patterns (mainly HN-HN, HN-H α , and H α -H α NOEs). A complete description of the use and implementation of the H-bonds is provided in the supplementary methods. As mentioned in question 2, it is the high amount of dihedral restraints that explains the apparently low number of NOE restraints; one must see that 1 dihedral restraint might account for 2-5 NOEs.

4) The discussion of the titration of Dcp2 into Edc3 (near bottom of page 5, describing Figure 1) is completely illogical. The authors say that because the system was in slow exchange, assignments could not be transferred from the free to the bound state of the Edc3 LSm domain, and that therefore the interface could not be mapped. It is true that slow exchange makes it more difficult to obtain assignments for the bound state, as in such cases bound-state assignments need to be determined completely independently, without being able to use data from the free state to help the process. However, this certainly does not mean that assignment is impossible. Indeed, it very clearly must have been possible in the present case, since the authors went on to solve the solution structure of the complex, the first and absolutely necessary stage of which <i>is to assign the signals</i>. There is thus a fundamental contradiction between their adjacent statements that a) the interface could not be mapped due to the impossibility of making assignments, and b) they solved this problem by determining the solution structure. I simply cannot understand how they were even able to write this down. It also means that, despite their statement to the contrary, it must be entirely trivial to produce a figure showing the chemical shift changes for backbone amide groups of Edc3 (as a function of sequence) upon addition of Dcp2.

Clearly there is some confusion in terminology here, which we have tried to resolve. In the first part of the paper, we assigned the resonances of the free Edc3 LSm domain and solved the structure. We were then interested in obtaining the binding site of Dcp2 on the Edc3 LSm domain. To that end, we added unlabeled Dcp2 to ¹⁵N labeled Edc3 LSm domain. Both the spectra of the free Edc3 LSm domain and the Edc3 LSm domain in complex with Dcp2 are shown in Figure 1E. This data was unfortunately not sufficient to identify which residues in the Edc3 LSm domain interact with Dcp2. This was mainly due to the fact that around 50- 75% of the resonances experience chemical shift changes. Due to this large number of changes, it is was not possible to define a small subset of Edc3 LSm residues that directly interact with the Dcp2 sequence. This is what we mean when we say that we cannot usefully map the interface based on the data shown in Figure 1E.

Clearly, we were able to assign the resonances in the complex, however, this was done from scratch and not aided by the assignments of the free protein (we argue that this is a formally more robust approach). This resonance assignment was a normal part of the process of solving the complex structure.

We now have rewritten the part of the manuscript that caused the confusion and resolved the ambiguity of the statement that we intended to make.

5) It seems that the Dcp2 fragment used in the NMR experiments comprised residues 242-291, but the part of this fragment that becomes structured on binding to Edc3 is just residues 254-266. This implies that some 37 residues of the Dcp2 fragment remain unfolded in the complex. How did the signals from these residues affect the NMR experiments? Were they assigned (presumably not)? Did they show NOEs? Were their signals essentially identical in the presence or absence of Edc3? How

did the presence of these signals from disordered regions affect the assignment process for signals from the ordered regions?

The presence of unfolded regions in proteins can make the NMR data analysis difficult due to the fact that they can result in very strong signals in proton-nitrogen maps. Here, we performed our NMR experiments at basic pH (7.3) and at a slightly elevated temperature (303K). This causes a significant increase in the exchange rates of the amide protons with the bulk solvent. As a result, the intensity of protein resonances from unfolded parts of the protein (complex) were significantly reduced. The remaining very weak resonances resulting from the disordered 37 residues of Dcp2 did thus not influence our analysis. Apart from a small number of intra- residue NOEs, these resonances also showed no NOE contacts in our spectra.

The elevated pH and temperature that weakens resonances of disordered residues have now been explicitly mentioned in the materials and methods section in the supplementary information.

6) The authors state that the structure of the Edc3 LSm domain is unchanged when Dcp2 binds, but no data is given. I would expect to see a backbone rmsd calculated for Edc3 between the average structures for each of the two ensembles (free and bound) to back up this statement.

This information has now been provided in the supplementary information in Figures 2C and 2D. We present both an overlay of the ensembles of the free and Dcp2 complexed Edc3 LSm domains and an overlay of the average structures of the free and the Dcp2 complexed Edc3 LSm domain. In that figure, we also included the corresponding rmsd values. Apart from a minimal shift in the N-terminal turn in Edc3 there are no structural changes in the Edc3 LSm domain.

Referee #3

*In this manuscript, the group of E. Izzaualde explores the basis for interaction occurring between decapping factor and Edc3/Scd6. The results are interesting but preliminary. The main weakness is that the authors do not provide support for the biological roles of the proposed interactions. Currently, the manuscript focuses on structural data with a single biochemical experiment suggesting biological relevance. Given that the authors analyze yeast (*S. pombe*) factors and that the RNA degradation process can easily be analyzed in simple model species (e.g., *S. pombe* or *S. cerevisiae*) one would expect that they provide in vivo evidence for the role of the Dcp2 C-terminal region in P bodies formation (in relation with the clustering hypothesis presented in Figure 4C) and support for the role of the Dcp2 C-terminal tail in decapping activation. This additional functional information and a more careful writing of the manuscript (see below) are required to make this manuscript of interest to the broad readership of the EMBO Journal.*

- Throughout the manuscript, the authors have an incomplete and biased presentation of the published literature. Surprisingly, a significant fraction of the unquoted references have been published in the EMBO Journal whose long-term impact is thus undermined by this situation.

We apologize for not citing a number of important papers. We have corrected this in the new version of the manuscript.

** The authors use "LSm interaction motif (LIM)" to identify the repeated sequence present in Dcp2 that mediate interaction with Edc3/Scd6. However, these sequences have been noticed earlier and defined as "Helical Leucine-rich Motifs (HLMs)" (Gaudon et al., EMBO J. (1999) 18:2229- 40). To avoid generating confusion and giving credit to this earlier publication, the original name should be kept.*

In the revised version of the manuscript we acknowledge the findings of Gaudon *et al* and now refer to the linear sequence motifs that interact with Edc3 as HLM-1 and HLM-C1 to HLM-C5. As the location and number of the HLMs in *S. cerevisiae* and *S. pombe* is not conserved, we cannot use the same numbering for the HLMs as was used by Gaudon *et. al*.

** Credits for the description of Dcp2 as the decapping enzyme is given to a review emanating from a lab which erroneously presented Dcp1 as the decapping enzyme for years, while not a single reference present some of the groups active in this area that were primarily involved in establishing*

this fact (Wang et al. Proc Natl Acad Sci U S A. (2002) 99:12663-8; van Dijk et al. EMBO J. (2002) 21:6915-24, see also Lykke-Andersen J. Mol Cell Biol. (2002) 22:8114-21).

We have included the references to the papers that have identified Dcp2 as the active component of the decapping complex.

** Similarly, the description of decapping activators (3 references to the work of a single lab) involved additional groups (e.g., Boeck et al. Mol Cell Biol. (1998) 18:5062-72; Bouveret et al. EMBO J. (2000) 19:1661-71; Bonnerot et al. Mol Cell Biol. (2000) 20:5939-46 for Pat and LSm proteins) that are not mentioned.*

We now included some of these important contributions.

** Page 3, line 10: The presence of an EVH1 domain in Dcp1 was originally identified by Callebaut (FEBS Lett. (2002) 519:178-80).*

We acknowledge this fact in the manuscript now.

** Page 4, second paragraph: The authors should mention the functional redundancy between Scd6 and Edc3 identified by the Jacquier group (Decourty et al. Proc Natl Acad Sci U S A. (2008) 105:5821-6).*

This interesting observation is now mentioned in the introduction of the manuscript.

** Page 6, first line: The authors may want not to forget the pioneering work of the Nagai group on Sm domain-Sm domain interaction (Kambach et al. Cell (1999) 96:375-87) and the early work of the Suck group (Toro et al. EMBO J. (2001) 20:2293-303, Thore et al. J. Biol Chem. (2003) 278: 239-47) for different modes of RNA binding by Sm proteins.*

These references and the recently published paper by Leung, Nagai and Li (Leung et al. Nature (2011) 7348:536-9) have been added.

** Page 9, line 18: Beside Tarassov et al. and Nissan et al., evidence for interaction of Dcp2 with Dhh1, Pat1 and Scd6 have been reported by Decker et al. (J Cell Biol. (2007) 179:437-49); Fromont-Racine et al. (Nat Genet. (1997) 16:277-82) and Decourty et al. (Proc Natl Acad Sci U S A. (2008) 105:5821-6).*

We apologize for not having cited more references and now realize that we failed to include some very important contributions to the field. We corrected this in the revised version.

-Page 5, first lines: "...and found that a Dcp2 region located between residues 254 and 266 is sufficient for the interaction with the Edc3 LSm domain (Figure 1D, lane 3 vs. 1 and 2)." is not correct. The data show that the region is necessary but not that it is sufficient. The data leading to this conclusion should be shown especially as the sequence requirements for binding remain unclear.

We agree that the plot that is shown in Figure 1D proves that the region between residues 254 and 266 is required for the Edc3:Dcp2 interaction. In combination with Figure 1E (where we show that HLM-1 binds to the Edc3 LSm domain) and the structure of the complex of Edc3 and Dcp2 we can conclude that residues 254-266 are sufficient for the interaction between Dcp2 and Edc3. We modified the text accordingly.

-Page 6, line 7: What does "we solved the high-resolution solution structure" mean? The authors should be more explicit.

With high-resolution, we mean "well defined". The term "high-resolution" might be confusing as it is usually used for crystal structures. We modified this sentence to prevent confusion.

-Page 9, line 7 and Figure 4 A and B: In contrast to their statement, the authors have not shown that a "hexameric" complex is formed. Gel filtration or ultracentrifugation data should be provided to support this conclusion.

The data shown in figure 3D (previously Figure 4) shows that the dimeric Edc3 protein is able to interact with (at least) two decapping complexes (one containing Dcp1:Dcp2 and one containing His- Dcp1:Dcp2). This indicates that two Dcp1:Dcp2 decapping complexes (4 proteins) are brought in close special proximity by one Edc3 dimer (2 proteins). This strongly suggests that a His-Dcp1:Dcp2:Edc3:Edc3:Dcp2:Dcp1 complex is formed, which we referred to as hexameric. Any other model of a complex that contains Edc3 and multiple Dcp1:Dcp2 complexes is highly unlikely. Nevertheless, we have no direct data that proves that the complex contains exactly six proteins, as the referee points out.

Despite extensive efforts, we have not been able to produce full length Edc3 at concentrations sufficiently high to be able to analysis the complex using gel- filtration or ultracentrifugation experiments. In the revised version of our manuscript we no longer refer to the complex as a hexameric assembly.

-Page 10, top paragraph and Figure 5 C: As a specificity control for the competition between Scd6 and Edc3, the authors should demonstrate that addition of an unrelated protein does not induce the dissociation of Scd6 LSm and LIM-I.

In the supplementary information we now show that the addition of the Edc3 Yjef-N domain does not cause the dissociation of the Scd6 HLM-1 complex (Figure S7).

-Page 11: The authors conclude, "the closed Dcp2 structure is incompatible with Edc3 binding". Given that Edc3 can bind different sequences present in the C- terminal extension of Dcp2, the evidence presented does not support this statement. Moreover, given the redundancy of interaction, the biological consequences of this conclusion are unclear. Altogether, this part is weak and does not deserve a full paragraph of the Results section. It would be more appropriate to present it in a couple of sentences in the Discussion. (See also contradiction with second paragraph of Discussion.)

The interaction of Edc3 with the HLM-1 (the first interaction motif) is incompatible with the closed conformation of the decapping complex as found in the crystal structure (see also point 1 by referee #1). We have no evidence for a mechanism where the interaction of Edc3/Scd6 with the other HLM sequences in Dcp2 would result in an increase in catalytic activity of the decapping complex.

We now mention more explicitly in the discussion, that the biological consequence of the HLM: LSm domain interaction depends on the location of the HLM in Dcp2. In case Edc3/ Scd6 interacts with the first HLM, the activity of the decapping complex is modulated. In case Edc3 interacts with the more C-terminal HLMs, a proper localization of the decapping complex is assured. In addition, the multiple HLMs in Dcp2 allow for the formation of decapping complexes with a different Edc3/ Scd6 content, which might be important for the recognition of specific substrates (e.g. Rps28B, as we now mention in the discussion).

In line with the referees comments, we have moved the section regarding the modulation of the catalytic activity to the discussion.

-Last section of Results and Figure 7 are not entirely convincing, especially the similarity between the sequence present in metazoan Dcp1 and HLMs/LIMs.

We show in Figure 7A that the *D. melanogaster* Dcp1 motif-1 is able to specifically interact with the *D. melanogaster* Edc3 LSm domain. Based on the NMR data, this interaction is structurally similar to the *S. pombe* HLM-1:Edc3 complex. We thus conclude that the *S. pombe* Dcp2 HLM-1 : Edc3 LSm domain interaction is equivalent to the *D. melanogaster* Dcp1 motif-1 : Edc3 LSm domain interaction.

We agree with the referee that the sequences of the HLMs in *S. pombe* Dcp2 (SxxLLxxL) are not identical to the Motif-1 sequence in Dcp1 (ASIFNMLT). However, the HLMs that interact with the Edc3 and Scd6 LSm domains and the Dcp1 motif-1 sequence all show a propensity to form an amphipathic helical structure.

The divergence in sequence between the Dcp2 HLMs and Dcp1 motif-1 supports our statement that it is hard to identify the small linear motifs that interact with the Edc3/ Scd6 LSm domain based on the amino acid sequence only. It should be noted, however, that the variation between the HLMs and motif-1 might also be tolerated due to a variation of the binding surfaces of the *S. pombe* and *D. melanogaster* LSm domains. In the course of this work we have not tested if the *D. melanogaster* Dcp1 motif-1 is able to interact with the *S. pombe* Edc3 or Scd6 LSm domains or *vice versa*, as these potential interactions are biologically irrelevant.

-Page 13: The authors discuss the sequence requirements for interaction between Dcp2 and LSm domains. As they identify sequences that do and do not interact and have obtained some structural information, they should probably be more explicit on the features required for binding or preventing interaction

Our study identified a total of 4 sequences that interact with *S. pombe* Edc3/ Scd6 LSm domains. Based on these sequences and, more importantly, on the structure, we can conclude that the motif requires to possess helical propensity and that one side of the helix must contain hydrophobic side chains that can interact with the LSm domain. For the interaction with the *S. pombe* LSm domains, the presence of leucine residues in the HLMs seem to be favorable, albeit not strictly required. This is also reflected in the fact that single point mutations of central leucine residues in the HLMs weaken, rather than abolish, the inter- molecular interaction (Figure S3A). We have now added this more explicitly in the text.

-Page 14: "However, it should be noted that the activation effect is likely more prominent in a cellular context in which additional factors are present that influence the Dcp2 inter- domain orientation." Without data or references to support this statement, it remains only wishful thinking. This is critical as this relates to the main conclusion of the manuscript.

We agree with the referee and have deleted the sentence.

References

Decker CJ, Teixeira D, 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

Teixeira D, Parker R (2007) Analysis of P-body assembly in *Saccharomyces cerevisiae*. *Mol Biol Cell* **18**: 2274-2287

2nd Editorial Decision

29 August 2011

Thank you for submitting your revised manuscript for consideration by the EMBO Journal. It has now been seen by two of the original referees whose comments are shown below. As you will see from their comments they feel that that the manuscript has been strengthened but a number of issues still remain that need to be clarified and addressed. The additional text including the material and methods can be incorporated into the main text and we are able to increase the character count of the manuscript to accommodate the changes requested by the referees. Should you be able to satisfactorily address these remaining issues we would be happy to look at the manuscript once more.

When preparing your letter of response to the referees' comments, please bear in mind that this will form part of the Review Process File, and will therefore be available online to the community. For more details on our Transparent Editorial Process, please visit our website: <http://www.nature.com/emboj/about/process.html>

Thank you for the opportunity to consider your work for publication. I look forward to your revision.

Yours sincerely,

Editor
The EMBO Journal

REFEREE COMMENTS

Referee #2

The authors have made a very substantial improvement to the paper by adding much of the material requested by the referees. The changes that I had requested have mostly been made, and I feel that the paper could in principle now be acceptable for publication in EMBO. However, because so much new material has been added, I do have a number of new questions. These are mainly relatively small issues, but I feel they should be addressed.

Main Paper:

Page 6: The description of the NMR titration of Dcp2 into Edc3-LSm domain is more carefully described in the new version, and the apparent contradiction concerning the possibility or otherwise of assignment of the complex that had concerned me in the original version is now removed. There is still a small issue here that I feel could be improved by adjusting the wording: the fact that so many chemical shifts change during the titration, beyond just those directly involved in the binding site (the identity of which is of course known from the structure presented later in the paper) suggests to me that there is probably some change of conformation in the Edc3-LSm domain that results from binding of the Dcp2 peptide - in other words that some of the shift perturbations are secondary effects. The authors' conclusion elsewhere in the paper, from superposing the structural ensembles of free and bound Edc3-LSm domain, is that the two structures are almost identical, and I am guessing that they do not want to say anything about the titration results that might appear to contradict this. However, I don't think there is necessarily any contradiction here, as chemical shifts (particularly those of amide NH groups in proteins) are extremely sensitive to conformational changes, so if there are structural changes upon binding they could be small enough to be within the range suggested by the similarity of the ensembles. Nor do I think the authors' present revised wording avoids this issue at all, so I think they would be better to answer it directly. As a suggestion, one could replace the relevant paragraph (Pages 6-7) as follows:

In the reverse experiment we performed an NMR titration to identify residues in the Edc3 LSm domain that interact with Dcp2. To this end, we added NMR inactive Dcp2 to ¹⁵N-labeled Edc3 LSm domain (Figure 1E). Edc3 residues that come in close spatial proximity with Dcp2 are expected to experience chemical shift perturbations upon binding. In the event, more than 50% of the residues in the Edc3 LSm domain are perturbed on Dcp2 addition (Figure 1E red vs blue), preventing accurate identification of the binding site from these data. This could reflect a small secondary conformational effect in the Edc3 LSm domain upon Dcp2 binding, but given the marked sensitivity of amide group chemical shifts to even small structural movements, any such change may be very small.

I do not insist on this for publication, but I think it would be an improvement.

Supplementary Material

Fig. S2C: How was the superposition of the two ensembles constructed? Were the free and bound average structures for the Edc3-LSm domain superposed carrying their respective ensemble members with them? That seems the most likely, but I feel the method should be mentioned in the figure caption.

Page S25: The motive for using filtered experiments is not always associated with resolving symmetry-related ambiguities; such experiments are very often used just to relieve overlap. Indeed, the authors themselves used a ¹⁵N-filtered experiment to separate aromatic signals from NHs (Page S26). It would be more accurate to say (last lines of page S25): "Due to the asymmetric nature of the

complex and the low extent of signal overlap in the NOE spectra, ..."

Page S26: The authors should list the backbone experiments that were used and should specify the NOESY mixing times.

Page S26: Deriving χ^2 torsion angles for Leu residues is a non-trivial exercise, and at the same time these are important constraints for the structure determination (particularly for the interface). The authors should give more detail of how these angles were obtained. I doubt they could be obtained unambiguously from NOESY data alone unless very short mixing time data were available and very carefully analysed, but the present wording appears to suggest that only NOESY was used (given that an HNHB experiment is the only other one mentioned in this context, and that only gives information about χ^1).

Page S27/Fig. S3E: It seems from the displayed intermolecular NOE connectivities that the leucine methyl signals were stereoassigned, in that the Figure shows the two methyl groups of several individual leucines have different NOEs from one another. Is this real? If so, were the leucine methyl stereoassignments achieved experimentally (if so how), or were they the result of random swapping during the calculation protocol followed by choosing the assignment leading to the lowest energy? The latter is quite a widely used approach and not unreasonable to use, but it would imply a somewhat lower level of confidence in interpreting the results. Given that quite a high proportion of the intermolecular NOEs involve Leu methyls, the existence or otherwise of stereoassignments may have made a real difference to how well the interface is defined and should be commented on. A table listing the intermolecular NOEs would be useful, but since the constraints must be deposited at the PDB, in principle this information will later be available to those who wish to see it.

Page S27: How were H-bonds restrained (i.e. what were the max/min distances specified and between which atoms)?

Page S27: Why were pseudoatoms added for methyl and non-stereoassigned methylene groups? As far as I know, use of pseudoatoms is only required for calculations employing centre-averaging of such groups, whereas r-6 averaging (which the authors state they used) explicitly does not require pseudoatoms - that is one of the main reasons for using it. Also, do the authors really mean r-6 averaging, or r-6 summation? They are different, and I think the former is now almost never used, having been replaced by the latter.

Response letter to reviewers:

Page 12: I am not clear what is meant by the discussion on deriving dihedral constraints from local NOEs. Some older calculation programs used to do this explicitly (e.g. the REDAC strategy in older versions of the DIANA program), but I am not aware that this is done by XPLOR-NIH, the software program used here. What appears to be meant is that NOEs were taken out of the constraint list and replaced by derived torsion constraints following an analysis based on comparing experimental and back-calculated NOESY spectra. If so, this process should be described in the methods section, as to my knowledge it is far from being a standard approach and could have quite a substantial effect on local precision. This is not to say that it is wrong, just that it needs to be carefully described. On the same point, are any of the torsional constraints that are counted in Tables S1A/B/C derived from NOEs, or are those all derived from TALOS, and are the totals given for NOE enhancements before or after removal of any NOEs that were replaced by torsion constraints? It would appear from the letter that both numbers refer to totals after a process of deriving torsion constraints from NOEs, but this is completely unclear from reading the Supplementary Material alone.

Referee #3

Fromm et al. have provided a new version of the manuscript and a rebuttal letter that answer appropriately many of the comments raised by the referees.

On the functional issue, the authors now provide localization experiments indicating that deletion of

the C-terminal tail of Dcp2 prevents its accumulation in P-bodies. The basis for this phenotype is not entirely clear nor its physiological consequence (the largest deletion display only a weak growth phenotype). The functional part of this manuscript remains thus rather weak.

The authors also added, following the comment of one of the reviewers, that Rps28B may contain a conserved HLM. The authors should clarify this new part: First, the Rps28 HLM is not obvious from looking at the protein sequence; the authors should be more explicit. Second, Rps28A and RPS28B are expected to bind the Rps28B mRNA, the HLM element should be found both in both Rps28A and Rps28B. Third, the stem-loop involved in Rps28 binding is not universally conserved in "Rps28B mRNAs", thus the authors should explain why the putative Rps28 HLM is conserved; this seems counterintuitive.

Overall, the manuscript has been improved but it remains weak on the functional issue. Also, a few points, including some introduced during the correction, remain unclear (see above and below). Altogether, the paper is a little bit thin on biology and too structurally oriented but it should be of interest to two large communities (RNA decay/processing and structural biology) and it is based on solid experiments, and thus it might be published in The EMBO Journal once the various points listed herein have been corrected.

Some points to be clarified:

- Introduction, page 3: « this process irreversibly... » might not be correct, see Otsuka et al. Mol Cell Biol. (2009) 29, 2155-67.
- Introduction, page 3: « Pat and the LS_m1 » should be changed to « Pat1 and LS_m1 ».
- Results, page 7, line 2: « residues 255-266 » should be removed as up to this point in the manuscript no evidence was yet presented that residues 255-266 are sufficient for binding.
- Page 17, line 9: I didn't find data in Ito et al., 2001, related to interaction of Rps28 with the Rps28B mRNA nor the presence of a conserved HLM-like element in RPS28B.
- Page 18, line 11: « confirms previous findings...these cytoplasmic foci » is incorrect as it is possible that an unknown protein is essential for the formation of P bodies.
- Material and Methods: This section is very short and the authors depart significantly from the EMBO Journal guidelines: this section does not contain all information of immediate importance to understand the manuscript (e.g., decapping assays...). This should be revised in accordance to the guidelines, without transferring nearly everything in the "Supplementary Information".
- Figure 3 should include numbering; this would help to understand deletions made.
- Figure 5a: the legend and material and methods are poorly explicit on what is compared (initial rates of decapping, end point...). The authors should give more detail on the decapping assay that is the only true functional result and explain which controls were performed.
- Figure 6: many "D" should be changed in "delta".
- Figure S2 c and d: indicate « grey: Dcp2 ».

2nd Revision - authors' response

29 September 2011

Referee #2

The authors have made a very substantial improvement to the paper by adding much of the material requested by the referees. The changes that I had requested have mostly been made, and I feel that

the paper could in principle now be acceptable for publication in EMBO. However, because so much new material has been added, I do have a number of new questions. These are mainly relatively small issues, but I feel they should be addressed.

Main Paper:

Page 6: The description of the NMR titration of Dcp2 into Edc3-LSm domain is more carefully described in the new version, and the apparent contradiction concerning the possibility or otherwise of assignment of the complex that had concerned me in the original version is now removed. There is still a small issue here that I feel could be improved by adjusting the wording: the fact that so many chemical shifts change during the titration, beyond just those directly involved in the binding site (the identity of which is of course known from the structure presented later in the paper) suggests to me that there is probably some change of conformation in the Edc3-LSm domain that results from binding of the Dcp2 peptide - in other words that some of the shift perturbations are secondary effects. The authors' conclusion elsewhere in the paper, from superposing the structural ensembles of free and bound Edc3-LSm domain, is that the two structures are almost identical, and I am guessing that they do not want to say anything about the titration results that might appear to contradict this. However, I don't think there is necessarily any contradiction here, as chemical shifts (particularly those of amide NH groups in proteins) are extremely sensitive to conformational changes, so if there are structural changes upon binding they could be small enough to be within the range suggested by the similarity of the ensembles. Nor do I think the authors' present revised wording avoids this issue at all, so I think they would be better to answer it directly. As a suggestion, one could replace the relevant paragraph (Pages 6-7) as follows:

In the reverse experiment we performed an NMR titration to identify residues in the Edc3 LSm domain that interact with Dcp2. To this end, we added NMR inactive Dcp2 to ¹⁵N-labeled Edc3 LSm domain (Figure 1E). Edc3 residues that come in close spatial proximity with Dcp2 are expected to experience chemical shift perturbations upon binding. In the event, more than 50% of the residues in the Edc3 LSm domain are perturbed on Dcp2 addition (Figure 1E red vs blue), preventing accurate identification of the binding site from these data. This could reflect a small secondary conformational effect in the Edc3 LSm domain upon Dcp2 binding, but given the marked sensitivity of amide group chemical shifts to even small structural movements, any such change may be very small.

I do not insist on this for publication, but I think it would be an improvement.

We thank the referee for the suggested text and included this into our revised version of the manuscript.

Supplementary Material

Fig. S2C: How was the superposition of the two ensembles constructed? Were the free and bound average structures for the Edc3-LSm domain superposed carrying their respective ensemble members with them? That seems the most likely, but I feel the method should be mentioned in the figure caption.

The ensembles shown in Figure S2 were obtained by superposition of the secondary structure elements of the individual structures (residues V3-L58 for Figures S2A, S2C and S2D; residues V3-L58 and A254-S267 for Figure S2B). In each case, an average structure (not shown) was calculated from the superimposed ensemble by averaging the coordinates of the individual structures. The reported rmsd is the average (+/- standard deviation) rmsd of the individual members compared to the average structure. We now mention this explicitly in the caption of Figure S2.

Page S25: The motive for using filtered experiments is not always associated with resolving symmetry-related ambiguities; such experiments are very often used just to relieve overlap. Indeed, the authors themselves used a ¹⁵N-filtered experiment to separate aromatic signals from NHs (Page S26). It would be more accurate to say (last lines of page S25): "Due to the asymmetric nature of the complex and the low extent of signal overlap in the NOE spectra, ..."

The referee is correct. We have now added the low extend of signal overlap to the text.

Page S26: The authors should list the backbone experiments that were used and should specify the NOESY mixing times.

The sequential assignment was based on HNCA, HNCACB, CC(CO)NH-TOCSY, HNCOC and HN(CA)CO experiments. The CC(CO)NH experiment was also used to identify the amino acid type. An 80 ms mixing time was used in all NOESY experiments. This information is now added to the revised version of the manuscript (page 20 and S26).

Page S26: Deriving χ_2 torsion angles for Leu residues is a non-trivial exercise, and at the same time these are important constraints for the structure determination (particularly for the interface). The authors should give more detail of how these angles were obtained. I doubt they could be obtained unambiguously from NOESY data alone unless very short mixing time data were available and very carefully analysed, but the present wording appears to suggest that only NOESY was used (given that an HNHB experiment is the only other one mentioned in this context, and that only gives information about χ_1).

We followed the following procedure to determine the torsion angles (rotameric states) of side chains and to determine the stereospecific chemical shift assignment of the methylene and methyl groups.

1. The ϕ_1 angle and the stereospecific assignment of H β protons was determined based on an HNHB experiment and relative NOE intensities of the intra-residual HN-H β_1 , HN-H β_2 , Ha-H β_1 and Ha-H β_2 cross peaks (Wagner et al, 1987) in 3D HNH, HCH and CCH NOESY spectra.
2. For valine residues the ϕ_1 angle was determined based on the HNHB experiment and relative NOE intensities of intra-residual HN-H β , HN-H γ_1 , HN-H γ_2 and Ha-H β , Ha-H γ_1 , Ha-H γ_2 cross peaks in 3D HNH, HCH and CCH NOESY spectra. This can also provide the stereospecific assignment of the methyl groups.
3. For isoleucine and threonine residues the ϕ_1 angle was determined based on the HNHB experiment and relative NOE intensities of the intra-residual HN-H β , HN-H γ_2 and Ha-H β , Ha-H γ_2 cross peaks in 3D HNH, HCH and CCH NOESY spectra.
4. The determined ϕ_1 angles were used in the structure calculations. To that extend, the corresponding dihedral angle was restrained to +60 (+/- 30), -60 (+/- 30) or 180 (+/- 30) degrees depending on the rotameric state. The intra-residual NOE distances that defined the ϕ_1 angle were not used as distance restraints in the structures calculations to avoid the use of redundant information.
5. For leucine residues the ϕ_2 angle was determined based on relative intra-residual NOE intensities of the Ha-H δ_1 , Ha-H δ_2 , H β_1 -H δ_1 , H β_1 -H δ_2 , H β_2 -H δ_1 and H β_2 -H δ_2 cross peaks in 3D HCH and CCH NOESY spectra. This can also provide the stereospecific assignment for the methyl groups.
6. For isoleucine residues the ϕ_2 angle was determined based on relative intra-residual NOE intensities of the Ha-H δ_1 , H γ_2 -H δ_1 , H β -H δ_1 , Ha-H γ_{11} , H γ_2 -H γ_{11} , H β -H γ_{11} , Ha-H γ_{12} , H γ_2 -H γ_{12} and H β -H γ_{12} cross peaks in 3D HCH and CCH NOESY spectra. This can also provide the stereospecific assignment for the methylene γ protons.
7. The determined ϕ_2 angles were used in the structure calculations. To that extend, the corresponding dihedral angle was restrained to +60 (+/- 30), -60 (+/- 30) or 180 (+/- 30) degrees depending on the rotameric state. The inter-residual NOE distances that were used to define the ϕ_2 angle were not used as distance restraints in the structures calculations to avoid the use of redundant information.

During the structure refinement, we compare experimental NOE strips (derived from HNH, CCH, CNH and HCH NOESY spectra) with back-calculated NOE strips. The back-calculation was performed using in house written software that makes use of the full relaxation matrix (to include effects of spin diffusion) and the current structural model. This procedure allows us to identify potential inconsistencies due to wrongly assigned resonances, wrongly assigned NOE distances or wrongly assigned rotameric states. Potential errors were corrected in a novel round of structure refinement until no more inconsistencies were present.

It should be noted that the large number of NOESY spectra we recorded (that resolves potential spectral overlap) and the back-calculation of the NOESY spectra (taking spin diffusion into account) are fundamental for the procedure to function properly.

We now mention this complete procedure explicitly in the Supplementary material and mention the used strategy in the main text of the revised manuscript.

Page S27/Fig. S3E: It seems from the displayed intermolecular NOE connectivities that the leucine methyl signals were stereoassigned, in that the Figure shows the two methyl groups of several individual leucines have different NOEs from one another. Is this real? If so, were the leucine methyl stereoassignments achieved experimentally (if so how), or were they the result of random swapping during the calculation protocol followed by choosing the assignment leading to the lowest energy? The latter is quite a widely used approach and not unreasonable to use, but it would imply a somewhat lower level of confidence in interpreting the results. Given that quite a high proportion of the intermolecular NOEs involve Leu methyls, the existence or otherwise of stereoassignments may have made a real difference to how well the interface is defined and should be commented on. A table listing the intermolecular NOEs would be useful, but since the constraints must be deposited at the PDB, in principle this information will later be available to those who wish to see it.

As described above, we have stereo-specifically assigned the leucine methyl groups. We did not use a protocol that randomly swaps the methyl group assignments during the structure calculations and agree with the referee that such a protocol could compromise the structural quality.

Page S27: How were H-bonds restrained (i.e. what were the max/min distances specified and between which atoms)?

The H-bonds were restrained as described in (Truffault et al, 2001). In the structure calculations, H-bonds were treated as covalent bonds between the amide proton and the carbonyl oxygen. In X-PLOR, these additional bonds were added to the molecular structure through the PATCH statement. To ensure proper hydrogen bond geometry, the bond length was weakly (14 kcal/mol per Å²) restrained to 2.12 Å, whereas the bond angle was weakly (4 kcal/mol per rad²) restraint to 0 degrees. It should be noted that force constants used are designed to be very weak compared to other restraints. To prevent the hydrogen bond from being unrealistically short or long, an NOE distance restraint between the amide proton and the carbonyl oxygen was applied with a lower bound of 1.9 Å and an upper bound of 2.6 Å.

Although we referenced the paper that describes this method, we admit that the referencing was confusing. We now added the method by which we restrained the hydrogen bonds explicitly to the supplementary material.

Page S27: Why were pseudoatoms added for methyl and non-stereoassigned methylene groups? As far as I know, use of pseudoatoms is only required for calculations employing centre-averaging of

such groups, whereas r-6 averaging (which the authors state they used) explicitly does not require pseudoatoms - that is one of the main reasons for using it. Also, do the authors really mean r-6 averaging, or r-6 summation? They are different, and I think the former is now almost never used, having been replaced by the latter.

We apologize to the referee for the typing error in the methods section and the confusion that this caused. "Pseudoatoms (using r^{-6} averaging) were added for methyl groups and non stereo-specifically assigned methylene groups" should have read "Pseudoatoms *allowances* (using r^{-6} averaging) were added for methyl groups and non stereo-specifically assigned methylene groups". We do not explicitly add pseudo-atoms to the structure calculations.

Here, we use r^{-6} averaging and added a 0.8 Å pseudo atom allowance (pseudo-atom correction) to NOE distances that involve a) methyl groups, b) unresolved methylene groups and c) resolved, but non-stereospecifically assigned methylene groups (where we only used the stronger of the 2 cross peaks in the calculations). Formally, the pseudo-atom correction is different for the 3 cases, nevertheless, we took the simple, somewhat more conservative, approach of using a long (0.8 Å) allowance for all three cases.

In addition, it should be noted that r^{-6} summation should be used in combination with ambiguous restraints. As our distance restraints are mainly unambiguous (apart from resolved, but non-stereospecifically assigned methylene groups, see above) we have not used r^{-6} summation for our calculations.

Response letter to reviewers:

Page 12: I am not clear what is meant by the discussion on deriving dihedral constraints from local NOEs. Some older calculation programs used to do this explicitly (e.g. the REDAC strategy in older versions of the DIANA program), but I am not aware that this is done by XPLOR-NIH, the software program used here. What appears to be meant is that NOEs were taken out of the constraint list and replaced by derived torsion constraints following an analysis based on comparing experimental and back-calculated NOESY spectra. If so, this process should be described in the methods section, as to my knowledge it is far from being a standard approach and could have quite a substantial effect on local precision. This is not to say that it is wrong, just that it needs to be carefully described. On the same point, are any of the torsional constraints that are counted in Tables S1A/B/C derived from NOEs, or are those all derived from TALOS, and are the totals given for NOE enhancements before or after removal of any NOEs that were replaced by torsion constraints? It would appear from the letter that both numbers refer to totals after a process of deriving torsion constraints from NOEs, but this is completely unclear from reading the Supplementary Material alone.

We apologize for the confusion and now (as also described above) mention the used procedure more accurately. As the referee understood correctly, we have replaced the intermolecular NOEs that were used to define the rotameric state with torsion angle restraints. This has been done

using an external procedure that is not implemented in the X-PLOR-NIH software. In our hands, this procedure significantly increases local precision.

As the referee indicated correctly, table S1A/B/C reports only on restraints that are used in the structure calculations. As such, the rmsd from dihedral restraints reports on backbone dihedral restraints derived from TALOS and on the side-chain dihedral restraints derived from the intra-residual NOEs/ HNHB data. The rmsd for distance restraints are without the intra-residual contacts that have been used to define dihedral angles. We have now added a footnote to Table S1 to clarify these points.

Referee #3

Fromm et al. have provided a new version of the manuscript and a rebuttal letter that answer appropriately many of the comments raised by the referees.

On the functional issue, the authors now provide localization experiments indicating that deletion of the C-terminal tail of Dcp2 prevents its accumulation in P-bodies. The basis for this phenotype is not entirely clear nor its physiological consequence (the largest deletion display only a weak growth phenotype). The functional part of this manuscript remains thus rather weak.

The authors also added, following the comment of one of the reviewers, that Rps28B may contain a conserved HLM. The authors should clarify this new part: First, the Rps28 HLM is not obvious from looking at the protein sequence; the authors should be more explicit.

The C-terminal residues of the *S. pombe* Rps28B protein contain a leucine rich sequence between residues 51 and 58 (EDDILVLL) that is reminiscent of the HLM-1 sequence in *S. pombe* Dcp2 (SSQLLSLL). We have added the exact residue number in the revised version of the manuscript. Interestingly, this sequence element is present in the Rps28 protein of many eukaryotes.

To make sure that the Rps28B indeed is able to interact with the Edc3 LSm domain (as we propose in the manuscript) we performed NMR titration experiments. To that extend, we added unlabeled Rps28B to ¹⁵N labeled Edc3 LSm domain. Interestingly, we observed significant chemical shift changes for residues in the Edc3 LSm domain that coordinate the HLM sequences from Dcp2. This unambiguously shows that a part of the Rps28B protein is able to interact specifically with the Edc3 LSm domain in a manner that is equivalent to the HLM-1:Edc3 interaction. We have not included this preliminary data in the current manuscript; it does, however, provide strong experimental evidence for the presence of a functional HLM in Rps28B.

Second, Rps28A and RPS28B are expected to bind the Rps28B mRNA, the HLM element should be found both in both Rps28A and Rps28B.

The sequences of the Rps28A and Rps28B proteins are very similar. The *S. pombe* Rps28A and Rps28B are identical, whereas the *S. cerevisiae* proteins differ in a single aminoacid at the N-terminus. Both the Rps28A and Rps28B proteins thus contain the putative HLM element.

Third, the stem-loop involved in Rps28 binding is not universally conserved in "Rps28B mRNAs", thus the authors should explain why the putative Rps28 HLM is conserved; this seems counterintuitive.

The current experimental data available for the interaction between the Rps28B protein and a stem-loop in the Rps28 3' UTR comes from experiments in *S. cerevisiae* (Badis et al, 2004).

We agree with the referee that the stem-loop structure in the 3' UTR of the Rps28B mRNA is not notably conserved. Nevertheless it is plausible that the *S. pombe* Rps28B mRNA contains a regulatory element in the 3' UTR. First, the *S. pombe* Rps28B 3' UTR is significantly longer than the Rps28A 3' UTR (as is the case in *S. cerevisiae*). Secondly, the *S. pombe* Rps28B 3' UTR is longer than the median 3' UTR found in the *S. pombe* genome and particularly long for a ribosomal protein (Wilhelm et al, 2008). Interestingly, the length of the 3' UTRs correlates with the stability of the transcript and it can very well be that one of the stem-loop structures in the *S. pombe* Rps28B 3' UTR serves as a binding platform for the Rps28 protein.

Overall, the manuscript has been improved but it remains weak on the functional issue. Also, a few points, including some introduced during the correction, remain unclear (see above and below). Altogether, the paper is a little bit thin on biology and too structurally oriented but it should be of interest to two large communities (RNA decay/processing and structural biology) and it is based on solid experiments, and thus it might be published in The EMBO Journal once the various points listed herein have been corrected.

Some points to be clarified:

- Introduction, page 3: « this process irreversibly... » might not be correct, see Otsuka et al. Mol Cell Biol. (2009) 29, 2155-67.

We agree with the referee and have changed the wording of the sentence to "The cleavage of the mRNA 5' cap structure (decapping) is an important step in gene expression as it removes a transcript from the translational pool."

- Introduction, page 3: « Pat and the LSm1 » should be changed to « Pat1 and LSm1 ».

We corrected the name of the protein throughout the manuscript.

-Results, page 7, line 2 « residues 255-266 » should be removed as up to this point in the manuscript no evidence was yet presented that residues 255-266 are sufficient for binding.

We changed "residues 255-266" to "residues 242-291" as a construct that contains these aminoacids was used to perform the NMR titration that is shown in Figure 1E.

- Page 17, line 9: I didn't find data in Ito et al., 2001, related to interaction of Rps28 with the Rps28B mRNA nor the presence of a conserved HLM-like element in RPS28B.

We apologize to the confusing placement of the reference. We intended to indicate that Ito et al., 2001, have described an interaction between Edc3 and Rps28B. We corrected this in the revised version of the manuscript.

- Page 18, line 11: « confirms previous findings...these cytoplasmic foci » is incorrect as it is possible that an unknown protein is essential for the formation of P bodies.

We agree with the referee and have changed the wording of the sentence to: “is in agreement with the fact that no single protein has been identified that is absolutely essential for the formation of these cytoplasmic foci in yeast”

- Material and Methods: This section is very short and the authors depart significantly from the EMBO Journal guidelines: this section does not contain all information of immediate importance to understand the manuscript (e.g., decapping assays....). This should be revised in accordance to the guidelines, without transferring nearly everything in the "Supplementary Information".

We have now moved the section that describes the decapping experiments back to the main manuscript and also extended the methods that describe the structural studies.

- Figure 3 should include numbering; this would help to understand deletions made.

We have added the residue numbers to Figure 3A.

- Figure 5a: the legend and material and methods are poorly explicit on what is compared (initial rates of decapping, end point...). The authors should give more detail on the decapping assay that is the only true functional result and explain which controls were performed.

We determined the end point of the reaction and determined the activity as the fraction of m7GDP that was released. We have added this information to the Figure legend of Figure 5A.

- Figure 6: many "D" should be changed in "delta".

Corrected.

- Figure S2 c and d: « grey: Dcp2 ».

We added this information to Figure S2c and d.

References

Badis G, Saveanu C, Fromont-Racine M, Jacquier A (2004) Targeted mRNA degradation by deadenylation-independent decapping. *Mol Cell* **15**: 5-15

Truffault V, Coles M, Diercks T, Abelmann K, Eberhardt S, Luttgen H, Bacher A, Kessler H (2001) The solution structure of the N-terminal domain of riboflavin synthase. *J Mol Biol* **309**: 949-960

Wagner G, Braun W, Havel TF, Schaumann T, Go N, Wuthrich K (1987) Protein structures in solution by nuclear magnetic resonance and distance geometry. The polypeptide fold of the basic

pancreatic trypsin inhibitor determined using two different algorithms, DISGEO and DISMAN. *J Mol Biol* **196**: 611-639

Wilhelm BT, Marguerat S, Watt S, Schubert F, Wood V, Goodhead I, Penkett CJ, Rogers J, Bahler J (2008) Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature* **453**: 1239-1243

3rd Editorial Decision

12 October 2011

Thank you for submitting your revised manuscript to The EMBO Journal. It has been re-evaluated by one of the referees who finds that you have satisfactorily addressed all their concerns, his/her comments are listed below. I am happy to accept the manuscript for publication in The EMBO Journal. You will receive the official acceptance letter in the next day or so.

Yours Sincerely,

Editor
The EMBO Journal

REFEREE COMMENTS

Referee #2

The authors have answered the points in my report on the revised version of this paper and I am now prepared to recommend publication in EMBO Journal. However, I do still have two points that I feel the authors should consider for the future.

1) It seems to me that the procedure for obtaining rotameric states of the various sidechains, particularly the leucines, should be more fully described and tested. I do not insist on this for the present paper, but I think it would be quite wrong if the authors were simply to use this procedure in future without ever describing it further, just referring back to this paper. While it seems to be essentially an extension of existing approaches, nonetheless I think it may represent an incremental step beyond what others have done in this area, particularly in the use of back-calculated NOEs to guide the process. Without such back-calculation I would think that some of the derived rotamer states might easily turn out to be wrong (indeed the authors say as much themselves), and I would like to see a clear verification that such errors are indeed avoided when back-calculation is used. In other words, I think the method needs to be validated. It would be good if it were also made available to others.

2) There is still confusion over the use of pseudoatom corrections. In fact no pseudoatom corrections are needed in conjunction with r-6 averaging - only multiplicity corrections are needed (see *J.Biomol.NMR*, 8, 292). Also, r-6 summation is not restricted to use with ambiguous constraints, it can be used under all circumstances where r-6 averaging is applicable, the only difference being that using r-6 summation obviates the need for making multiplicity corrections. However, I don't think any changes should now be made to the manuscript or Supplementary Information since the authors do at least now describe what they did in the calculations in this respect. Changing the corrections or the averaging method would mean recalculating all structures from scratch, and I doubt that this would lead to any significant changes or improvements.

