THE EMBO JOURNAL

Structure of a eukaryotic cytoplasmic pre-40S ribosomal subunit

Alain Scaiola, Cohue Peña, Melanie Weisser, Daniel Böhringer, Marc Leibundgut, Purnima Klingauf-Nerurkar, Stefan Gerhardy, Vikram Govind Panse, Nenad Ban.

Review timeline:

Submission date:	25th October 2017
Editorial Decision:	23rd November 2017
Additional correspondence	1st of December
Revision received:	18th December 2017
Accepted:	10th January 2018

Editor: Anne Nielsen.

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 23rd November 2017

Thank you for submitting your manuscript for consideration by The EMBO Journal. Your study was sent out to three reviewers and we have now heard back from two of them. Since they are both supportive of your work, pending adequate revision, I have made an editorial decision at this stage already to avoid unnecessary delays. Should we still receive a report from the third referee, I will forward it to you so you can incorporate any additional points in the revised study

As you will see from the reports below, both referees express interest in the findings reported in your manuscript although ref #2 raises a number of technical points that will need to be fully clarified. Based on these positive recommendations, I would like to invite you to submit a revised version of the manuscript, addressing the comments of all reviewers. I should add that it is EMBO Journal policy to allow only a single round of revision, and acceptance of your manuscript will therefore depend on the completeness of your responses in this revised version.

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://emboj.embopress.org/about#Transparent_Process

We generally allow three months as standard revision time. As a matter of policy, competing manuscripts published during this period will not negatively impact on our assessment of the conceptual advance presented by your study. However, we request that you contact the editor as soon as possible upon publication of any related work, to discuss how to proceed. Should you foresee a problem in meeting this three-month deadline, please let us know in advance and we may be able to grant an extension.

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

REFEREE REPORTS

Referee #1:

This manuscript present the structure of the yeast pre-40S subunit at a late stage of cytoplasmic maturation. This is a welcome and timely addition to the collection of recently published structures of pre-ribosomes. In particular, high resolution structures of the pre-40S have been elusive, presumably because of its inherent flexibility and heterogeneity. Here, the authors have used a nice genetic trick, a catalytic null of the endonuclease Nob1, to capture particles enriched for this late intermediate. The work provides sufficient molecular detail to begin to understand the interactions and specific functions of the various associated factors and, as the authors suggest, should lay the foundation for subsequent functional studies. The manuscript is clearly written and the data are well presented. I have only a few minor comments that should be addressed before publication.

1) The authors point out that Dim1 is too far away from its target methylation site to be able to methylate and suggest that within this particle, dimethylation may have already taken place. Because this manuscript is devoted to characterizing this particle rather than presenting functional analysis of maturation, the authors should determine the status of Dim1-dependent dimethylation within the particle they are presenting. This should be a straight forward primer extension experiment.

2) 1st paragraph of Discussion. The authors describe "an extreme rotation of the cytoplasmic pre-40S head that exceeds any 40S head rotation observed during translation." I could not find this described in the Results. If it is a major point to be discussed, it should be documented in the results.

3) Figure 6. This figure is very difficult to interpret and currently does not add to the manuscript. As arranged, one expects it to represent a series of events but, after studying it, there is only a single species in the box that has been rotated four different ways. In addition, the pre-40S to the left of the box is shown in a different perspective from the left-most image in the box. It's really hard to understand their relationship. And it's not clear what the gray color of the head indicates - that it is mature? The authors should think carefully about what they are trying to show in this figure. There are not enough intermediates to show a pathway. Perhaps they should focus only on what's boxed, one species, and explain the conformational changes that must take place to produce a mature subunit.

4) The legend for Figure 1 refers to the lower panel of D. This panel is labeled "E" in the figure and should be referred to as such.

5) Figure 4 needs additional labels and changes. Tsr1 should be labeled in Panel A. Without zooming in significantly, the color of uS12 looks like that of Tsr1 and should be changed. Rio2 should be labeled. The lower case letters are confusing. Change 'c" to "Panel C" to make clear that these are calling out regions described in the other panels and be consistent in use of capital or lower case letters.

6) In the Discussion, the section headed "Interdependent AF departure events lead to 40S maturation" needs additional citations. The work provided here is a nice structure, but it's the work of other labs that as built up our understanding of the function of AFs.

Referee #3:

Strunk et al. previously used cryo-EM to visualize the molecular architecture of late pre-40S ribosomes carrying Rio2, Dim1, Nob1, Pno1, Tsr1, Ltv1 and Enp1. The resolution of their structre (18 Å) was recently extended to around 9 Å with individual assembly factors resolved to around 8-10 Å after focused classification (Johnson et al Structure 2017). Conformational heterogeneity was noted in Tsr1, Rio2, Dim1. Tsr1 and Dim1 were shown to physically block binding of eIF5B and 60S subunits. Pno1 positions Nob1 at the 3' end of the small subunit rRNA, while uS3 bridges the head and beak and is held in an immature position by Ltv1/Enp1.

In the current manuscript, the authors succeed in trapping a pre-40S particle poised to undergo final 20S pre-rRNA cleavage by exploiting a dominant negative, catalytically inactive Nob1 mutant allele in yeast. This tool has allowed them to determine the structure of a pre-40S ribosomal subunit intermediate to an overall resolution of 3.4 Å using single-particle cryo-EM, a significant improvement on previous work. With the enhanced overall resolution, the authors have been able to obtain new mechanistic insights into the process of 40S subunit maturation. In addition, some genetic analysis is included that confirms prior work in human cells (Zemp et al. 2014) showing that the Hrr25 kinase likely starts the cascade of assembly factor release events by acting upstream of Enp1 and Ltv1. The EM data appear to be generally of high quality and reveal some interesting new insights into 40S subunit biogenesis. However, I have some concerns that the model is not adequately refined to justify the conclusions the authors wish to draw based on the statistics presented in Table S4 and the presentations in the Figure panels. Additional FSC curves need to be presented to demonstrate absence of overfitting of the model. The presentation and annotation of the figures should be improved to justify key conclusion that the authors wish to draw. In particular, EM density should be included in the figures in several places. In general the manuscript could have been more carefully assembled-there are typos, the Figures are not always discussed in the appropriate order or are incorrectly or incompletely annotated. For me, the section on model fitting and homology modeling is a little confusing and would benefit from the inclusion of an additional Table.

Major points: Refinement and model building

1. In Appendix Table S4, please add the specific detector and microscope used.

2. Given the overall resolution, the clash score of 14.12 seems very high after refinement. No percentile (ranking) is provided from Molprobity. The high clashscore indicates that further work is required with model refinement. Are there specific areas of the map where the RNA is particularly poorly defined? Please state this specifically in the text.

3. Page 20: "The refinement... rRNA base pairs" Please clarify what restraints were used for the RNA? e.g. base pair restraints, stacking etc

4. Please provide an overall Fourier Shell Correlation in Table S4.

5. Why is the 80S-like particle included in Table S4 and in the classification scheme in Figure S1? On p6, the authors suggest that there is additional density at the 3' end of the 18S rRNA representing partial density for the unlceaved 20S rRNA. This is interesting if it is true, but no views of this feature are provided in the manuscript making the claim difficult to properly assess. Either include a view of this feature or remove the 80S images for clarity. What is the local resolution around this proposed feature?

6. Figure S2: The FSC curves diverge significantly. The authors should include FSC curves between the model refined in the reconstruction from only half of the particles and the reconstruction from that same half; and between that same model and the reconstruction from the other half of the particles to show absence of overfitting.

7. Figure S2B-D: Please indicate the relative orientations of the images with respect to each other.

8. Figure S3: what is the local resolution for the Dim1 density? Please include this in the text on p7. Is it not possible that Dim1 might still come close to its methylation target in a better-resolved map where the protein is properly fitted?

9. Figure S5: The model does not fit well into the density, particularly in Panel S5B. This raises concerns, along with the refinement statistics in Table S4, that the model has not been well refined. In panel S5A, is there density for the side-chain of R73 or is just sticking into a nearby density? N634 is not within the map in panel B and E60 is barely visible. Does the helix in Figure S5A interact only with the phosphate backbone of the RNA? Are there any flipped bases compared with the mature 40S?

Additional points

10. The figure panels are not discussed in the correct order in the text which leads to a lot of confusion. Specific panels should be referenced as Figure S5C not Figure S5. Also some of the figures are incorrectly annotated eg p10, line 1, likely refers to Figure 4E. This all needs to be carefully revised.

11. Page 6: "The Enp1 and Ltv1 factors ... mature 40S subunit." It would be good to reference Fig2A while describing the localization of these factors.

12. Figure 2D: the text on p8 says that Chaetomium Rio2 crystal structures were docked into the density, but the text on p19 says that A. fulgidus Rio2 homology models were used for model building. This is confusing. Please provide a Table to clarify which models from which species were used for the fitting of the assembly factors, and whether yeast homology models were then built and refined. What exactly is included in the final refined model?

13. Figure 4A: the way in which the figure is annotated is quite confusing as lower case is used for the figure while the panels are labeled in upper case. Please try to make this clearer for the reader.

14. Figure 4A: Given that the distortion of H44 is a major point of the paper, supporting EM density should be shown in the figure.

15. What are the black arrows meant to indicate in Figure 4D?

16. In Figure 4E, it is important to show EM density to support the conclusions regarding displacement of the decoding center bases that the authors wish to draw.

17. In Figure 4C, show EM density to support the conclusions.

18. Figure 5. Please show density to support the statements in the figure legend and on p10, particularly with respect to the proposed flipping out of base A1749.

19. Figure S4: Please label Ltv1 in the figure for clarity

20. Please help the reader by referring to specific figures in the discussion.

21. Page 2: additional examples of publications on cytoplasmic quality control and proofreading should be cited at the end of the first paragraph and not just the authors' own contributions and their review in this area.

22. I feel that it would be more appropriate to modify the title- "structure of a (not the) eukaryotic cytoplasmic pre-40S ribosomal subunit.

23. Several references are not properly formatted

24. Consider replacing the article by Scheres cited on page 10, second line from the bottom with "Processing of Structurally Heterogeneous Cryo-EM Data in RELION." PMID: 27572726.

Additional correspondence 1st of December 2017

I wanted to let you know that we have now received the last referee report for your manuscript. As you'll see, the referee is positive about the study and only raises a few minor textual clarifications that should be included in the revised version.

Feel free to contact me with any questions about the revision.

.....

Ref #2:

In an interesting study, Scaiola et al. present the cryo-EM structure of an assembly intermediate of the eukaryotic ribosomal 40S subunit. In order to stall 40S maturation at a specific step, the authors generated a conditional Nob1 mutant yeast strain. The catalytically inactive Nob1 mutant lacks nuclease activity and is thus not able to promote the final 20S pre-rRNA cleavage step and subsequent events. Accordingly, affinity purified pre-40S particles contained 20S pre-rRNA and, as shown by mass spectrometry, the eight 40S assembly factors Hrr25, Ltv1, Enp1, Rio2, Tsr1, Dim1, Pno1 and Nob1. The affinity purified pre-40S particles were subjected to cryo-EM and the authors obtained a cryo-EM reconstruction with an overall resolution of 3.4 A. This allowed the authors to interpret and analyze the structure of a cytoplasmatic pre-40S particle at unprecedented resolution. As expected for a cryo-EM reconstruction of a not entirely rigid particle the local resolution varies from 2.8 to 4.9 A. However, careful model building and fitting of previously obtained X-ray structures of assembly factors allowed the authors to solve the structure, which provides important insights into the late maturation steps of the 40S subunit and the role of the various assembly factors.

Minor Points:

1. For the general reader it may be helpful to indicate the local resolution range also in the abstract. Essentially all previous papers of similar complexes report only the overall resolution, but this is to some extent misleading, because structural information of important parts of the particle is obtained at significantly lower resolution only.
2. The sentence (page 3, 2nd sentence) "Tsr1 is a GTPase-like protein, which adopts a four-domain fold similar to translational GTPases such as EF-Tu and eIF5B..." is misleading. EF-Tu consists of three domains.
3. In M&M Data processing and reconstruction the authors write that they removed images representing 80S particles. However, in the text they describe that they used these images to obtain an 80S reconstruction.

1st Revision - authors' response 18th December 2017

Referee #1:

1) The authors point out that Dim1 is too far away from its target methylation site to be able to methylate and suggest that within this particle, dimethylation may have already taken place. Because this manuscript is devoted to characterizing this particle rather than presenting functional analysis of maturation, the authors should determine the status of Dim1-dependent dimethylation within the particle they are presenting. This should be a straight forward primer extension experiment.
This primer extension experiment was actually performed previously on the same construct (Nob1-D15N) in Lebaron et al., 2012 (Fig. 1C). We should have referred to this paper and we apologize for this oversight. Additionally, Johnson et al., 2017, observed the unmethylated helix 45 in a different conformation compared to the conformation we observe. Therefore, these results support our conclusions that, in the pre-40S particle we observe, helix 45 has already been methylated. We modified the manuscript to place our findings in the context of the previous work:
“
The Dim1 rRNA dimethylase reported to be part of the pre-40S could only be identified in a small fraction of the final set of 40S-like particles after additional local classification (Expanded View Table EV1 and Expanded View Figs. EV1 and EV3A-B and Appendix Table S4). The reconstruction of this subset of particles reached an overall resolution of 4.3 Å and showed a pre-40S subunits containing Dim1 and helix 45 in its mature position (Expanded View Fig. EV3B), opposed to the displaced conformation described previously (Johnson et al., 2017). In our conformation, Dim1 (average local resolution of 6.1 Å) is at a position too distant from its target helix 45 for methylation (Boehringer et al., 2012;Johnson et al., 2017) (Expanded View Fig. EV3A-B). This suggests that Dim1 might have already methylated helix 45 in this particle, as shown by previous primer extension analysis on the Nob1-D15N particles (Lebaron et al., 2012), and remains only weakly attached to the rRNA.

“

2) 1st paragraph of Discussion. The authors describe "an extreme rotation of the cytoplasmic pre-40S head that exceeds any 40S head rotation observed during translation." I could not find this described in the Results. If it is a major point to be discussed, it should be documented in the results.
This conformational change was indeed not directly indicated in the Results. The manuscript was modified and a new figure panel 3A was added to highlight this feature of our structure.
“
On the solvent-exposed side of the pre-40S ribosome, densities corresponding to Enp1 and Ltv1 connect the beak of the highly rotated pre-40S head (Fig. 3A-B) with rRNA helix 16 on the pre-40S body (Fig. 2A and 2B).
“

3) Figure 6. This figure is very difficult to interpret and currently does not add to the manuscript. As arranged, one expects it to represent a series of events but, after studying it, there is only a single species in the box that has been rotated four different ways. In addition, the pre-40S to the left of the box is shown in a different perspective from the left-most image in the box. It's really hard to understand their relationship. And it's not clear what the gray color of the head indicates - that it is mature? The authors should think carefully about what they are trying to show in this figure. There are not enough intermediates to show a pathway. Perhaps they should focus only on what's boxed, one species, and explain the conformational changes that must take place to produce a mature subunit.
The Figure 6 summarizes in a single view the conformational changes that are expected to be interconnected. In other words departures of individual maturation factors would affect the conformation of rRNA in a way to promote dissociation of another maturation factor. However, the order of events regarding which groups of maturation factors depart first and which depart last cannot be established based on our results. We clarified this point with the following statement in the figure legend:
“
Figure 6: A model for cytoplasmic maturation of the pre-40S subunit through series of interdependent conformational changes.

The boxed region of the scheme represents a single species and highlights the conformationally coupled changes upon the release of individual assembly factors in rRNA conformation between the pre-40S (white, changing part in pink) and the mature 40S (gray, changed part in dark gray) without implying a sequential order. The tilt and rotation of the head as well as conformational changes in helix 34, helix 44 and helix 28 are interconnected such that conformational changes in the beak, the subunit interface and the platform lead to the release of AFs from their respective binding sites.
”

4) The legend for Figure 1 refers to the lower panel of D. This panel is labeled "E" in the figure and should be referred to as such.
We apologize of this oversight and fixed the reference to the panels D and E:
“
D-E. *The kinase activity of Hrr25 is essential for cell viability and is required for cytoplasmic release and recycling of Enp1 and Tsr1.*
D. *The PGAL1-HRR25 strain was transformed with indicated plasmids and spotted in 10-fold dilutions on selective and repressive glucose-containing plates and grown at indicated temperatures for 3-7 days.*
E. *Yeast strains expressing endogenous Enp1-GFP or Tsr1-GFP were transformed with plasmids carrying either a galactose-inducible wild-type HRR25 gene or a dominant-negative hrr25-K38A kinase dead mutant gene. Strains were then grown on galactose-containing medium at 25°C and GFP constructs were visualized by fluorescence microscopy.*
“

5) Figure 4 needs additional labels and changes. Tsr1 should be labeled in Panel A. Without zooming in significantly, the color of uS12 looks like that of Tsr1 and should be changed. Rio2 should be labeled. The lower case letters are confusing. Change 'c" to "Panel C" to make clear that these are calling out regions described in the other panels and be consistent in use of capital or lower case letters.

We agree that the labels were misleading and changed the figure panel accordingly.

6) In the Discussion, the section headed "Interdependent AF departure events lead to 40S maturation" needs additional citations. The work provided here is a nice structure, but it's the work of other labs that as built up our understanding of the function of AFs.
Additional references were added to this part of the discussion as requested by the referee.

Referee #2

1. For the general reader it may be helpful to indicate the local resolution range also in the abstract. Essentially all previous papers of similar complexes report only the overall resolution, but this is to some extent misleading, because structural information of important parts of the particle is obtained at significantly lower resolution only.
We agree that only indicating the overall resolution of the reconstruction may be misleading considering that the individual factors have been observed at varying resolutions. Therefore, we decided not to state the overall resolution of the reconstruction in the abstract and provide detailed information in the results section of the manuscript.
”
Final maturation of eukaryotic ribosomes occurs in the cytoplasm and requires the sequential removal of associated assembly factors and processing of the immature 20S pre-RNA. Using cryo-electron microscopy (cryo-EM), we have determined the structure of a cytoplasmic pre-40S particle poised to initiate final maturation. The structure reveals the extent of conformational rearrangements of the 3’ major and 3’ minor domains of the ribosomal RNA that take place during maturation, as well as the roles of the assembly factors Enp1, Ltv1, Rio2, Tsr1, and Pno1 in the process. Altogether, we provide a structural framework for the coordination of the final maturation events that drive a pre-40S particle towards the mature form capable of engaging in translation.
“

2. The sentence (page 3, 2nd sentence) "Tsr1 is a GTPase-like protein, which adopts a four-domain fold similar to translational GTPases such as EF-Tu and eIF5B..." is misleading. EF-Tu consists of three domains.
We agree that the sentence can be confusing and adapted the manuscript to clarify the comparison between Tsr1 and EF-Tu :

“
Tsr1 is a GTPase-like protein composed of four domains, where the first three domains form a fold similar to translational GTPases such as EF-Tu and eIF5B, but lack the residues to bind and hydrolyze GTP (McCaughan et al., 2016).
“

3. In M&M Data processing and reconstruction the authors write that they removed images representing 80S particles. However, in the text they describe that they used these images to obtain an 80S reconstruction
The 80S particles were removed from the 40S subclass, but were used for refinement as a separate subclass. The manuscript was modified to explain the process more clearly:
“
After removing images representing ice crystals and carbon edges, we separately selected approximatively 145000 80S-like particles and 236000 40S-like particles. The 40S-like particles were refined using a mature 40S density as reference followed by a 3D classification without alignment of the particles. The 165000 good particles containing Tsr1 were selected for a refinement using the full pixel size (1.39 Å/pixel). After post-processing and sharpening, the reconstruction reached an overall resolution of 3.4 Å (Expanded View Fig. EV2A).
Further focused classifications were performed with different masks around densities corresponding to ENP1, Rio2, Dim1, and Pno1 generated manually using UCSF Chimera (Pettersen et al., 2004). The 80S-like particles that were separated after the 2D classification were refined using the full pixel size (1.39 Å/pixel). The reconstruction reached an overall resolution of 3.1 Å after post-processing and sharpening (Expanded View Fig. EV1 and Appendix Fig. S1A-B).
“

Referee #3:

Major points: Refinement and model building
1. In Appendix Table S4, please add the specific detector and microscope used.
We used a FEI Titan Krios cryo-transmission electron microscope equipped with a FEI Falcon 3EC Direct Electron Detector. This information was added to the Table S4 (now Table EV1) as suggested by the referee.

2. Given the overall resolution, the clash score of 14.12 seems very high after refinement. No percentile (ranking) is provided from Molprobity. The high clashscore indicates that further work is required with model refinement. Are there specific areas of the map where the RNA is particularly poorly defined? Please state this specifically in the text.
The template for the 40S subunit we used (4v88 from Ben-Shem et al., 2011) was refined to higher resolution than our reconstruction and can be currently considered as the best structural information on the yeast 40S subunit. In our structure we took all parts of the 4v88 structure that fitted our density and only modified the regions where clear differences were observed. In the revised manuscript we further improved the model in other areas and re-refined it, resulting in improved clashscore of 11.65, which is comparable to the statistics observed for 4v88.

Refinement and model validation statistics (MolProbity implemented in PHENIX.REFINE)	
Resolution range (Å)	50-3.5
Spacegroup	P1
a=b=c (Å)	278
α=β=γ (°)	90
Number of reflections	1’048’658
Model resolution (Å) (FSC threshold = 0.5)	3.6
Average B-factor (Å2)	116.7
Clashscore (all atoms)	11.65
R-factor (%)	27.9
wxc weighting value	1.2
Protein	
Good rotamers (%)	93.06
Rmsd (bonds)	0.006
Rmsd (angles)	0.92
Ramachandran plot (%)	
favored	90.50
allowed	7.98
outliers	1.52
RNA	
Correct sugar puckers (%)	98.8
Good backbone conformation (%)	65.3

“

3. Page 20: "The refinement... rRNA base pairs" Please clarify what restraints were used for the RNA? e.g. base pair restraints, stacking etc
We only used base pair restraints in our refinement and the text was modified to clarify the possible misunderstanding.
“
To regularize the geometry and optimize the fit of the final atomic model to cryo-EM map, the structure was then refined in reciprocal space using PHENIX.REFINE (Adams et al., 2010) against structure factors back-calculated from the high-resolution EM density as described earlier (Greber et al., 2014) (Expanded View Table EV1). The refinement was stabilized in areas of weaker density

by applying Ramachandran restraints and secondary structure restraints for protein α-helices and β-strands as well as RNA base pair restraints for the rRNA.
"

4. Please provide an overall Fourier Shell Correlation in Table S4.
All the FSC curves are shown in Figure S2 (now EV2). The overall resolution for the EM map (3.4 Å) and the model (3.6 Å) were added in the Table with their respective FSC threshold (0.143 for FSC between cryo EM half-set, 0.5 for model versus map).

5. Why is the 80S-like particle included in Table S4 and in the classification scheme in Figure S1? On p6, the authors suggest that there is additional density at the 3' end of the 18S rRNA representing partial density for the unlceaved 20S rRNA. This is interesting if it is true, but no views of this feature are provided in the manuscript making the claim difficult to properly assess. Either include a view of this feature or remove the 80S images for clarity. What is the local resolution around this proposed feature?
The 80S-like reconstruction was included as it is part of the dataset and relevant to the construct and more specifically to the gels in Figure 1, as they explain the presence of 25S rRNA. An additional appendix figure S1 is provided to show the additional density at the 3' end of the 18S rRNA, as well as the local resolution of the reconstruction. The local resolution around this feature is roughly 6-7 Å which isn't enough to resolve any features of a single-stranded RNA, but sufficient to claim its presence.
6. Figure S2: The FSC curves diverge significantly. The authors should include FSC curves between the model refined in the reconstruction from only half of the particles and the reconstruction from that same half; and between that same model and the reconstruction from the other half of the particles to show absence of overfitting.
The additional requested FSC curves were added to the Fig. S2 (now Fig. EV2) to show that the model is not overfitted.

7. Figure S2B-D: Please indicate the relative orientations of the images with respect to each other.
The relative orientations were added, as well as labels to indicate which side of the particle is showed.

8. Figure S3: what is the local resolution for the Dim1 density? Please include this in the text on p7. Is it not possible that Dim1 might still come close to its methylation target in a better-resolved map where the protein is properly fitted?
The average local resolution for Dim1 is 6.1 Å according to "post_process" of Relion 2.1 and this information was added to the manuscript. Dim1 was rigid-body fitted into the density and shows that it is too far from the methylation site to be interacting with its target rRNA bases. Attempts to bring it closer to the methylation site would lead to clashes with the rRNA. The unmethylated helix 45 was shown in Johnson et al., 2017 to have a different conformation compared to the mature 40S or the conformation observed in our structure. The manuscript was modified to highlight this aspect of the methylation and to reference Lebaron et al., 2012, which performed primer extension experiments on Nob1-D15N particles and showed that the A1781 and A1782 are already methylated. cf comments for the point #1 of the referee #1.

9. Figure S5: The model does not fit well into the density, particularly in Panel S5B. This raises concerns, along with the refinement statistics in Table S4, that the model has not been well refined. In panel S5A, is there density for the side-chain of R73 or is just sticking into a nearby density? N634 is not within the map in panel B and E60 is barely visible. Does the helix in Figure S5A interact only with the phosphate backbone of the RNA? Are there any flipped bases compared with the mature 40S?
With the improved model as indicated by the improved refinement statistics, the residues in panel S5B now fit better into the density. Additionally a mistake in the labeling was corrected on panel S5B (now EV4B) referencing the correct base which is flipped out (A1754 instead of A1756). There is density for the side-chain of R73 of Tsr1, which is now shown better with slightly different view of the density and density level. The Tsr1 helix shown in Figure S5A only interacts with the phosphate backbone. The helix 44 is only displaced and twisted compared to the mature 40S and no flipped bases were observed.

Additional points
10. The figure panels are not discussed in the correct order in the text which leads to a lot of confusion. Specific panels should be referenced as Figure S5C not Figure S5. Also some of the figures are incorrectly annotated eg p10, line 1, likely refers to Figure 4E. This all needs to be carefully revised.
The figures and panels are now presented in order of reference in the text.

11. Page 6: "The Enp1 and Ltv1 factors ... mature 40S subunit." It would be good to reference Fig2A while describing the localization of these factors.
We agree with the referee's comment and added the reference to the Fig. 2A in the text.

12. Figure 2D: the text on p8 says that Chaetomium Rio2 crystal structures were docked into the density, but the text on p19 says that A. fulgidus Rio2 homology models were used for model building. This is confusing. Please provide a Table to clarify which models from which species were used for the fitting of the assembly factors, and whether yeast homology models were then built and refined. What exactly is included in the final refined model?
The reference to the *Chaetomium* Rio2 is a mistake on our part and we apologize for it. Comparison, figures and homology model were made using the *A. fulgidus* structure, which is bound to ATP. The *Chaetomium* structure is not bound to any nucleotide and shows an even more pronounce bending compare to our structure and *A. fulgidus*. The text was corrected to mention only *A. fulgidus* and avoid any confusion. Furthermore, an additional appendix table was created to highlight the PDB used as template model and their organism:
"
Interestingly, the structure of cytoplasmic pre-40S-bound Rio2 shows that the two halves of the Rio2 kinase domain, which form the nucleotide binding pocket, are rotated relative to each other when compared to the crystal structures of the A. fulgidus (LaRonde-LeBlanc et al., 2005, LaRonde-LeBlanc & Wlodawer, 2004) (Fig. 4E). As a result, the residues responsible for ATP binding are separated by greater distances, which most likely renders Rio2 unable to cleave ATP.

Appendix Table S5: Initial PDB templates used for docking and modeling

Structure	**Template**	**Organism of the template**
40S ribosome	4V88	*Saccharomyces cerevisiae*
Pno1	3AEV	*Pyrococcus horikoshii*
Enp1	5WWO	*Saccharomyces cerevisiae*
Ltv1	5WWO	*Saccharomyces cerevisiae*
Tsr1	5IW7	*Saccharomyces cerevisiae*
Rio2	1ZAO	*Archaeoglobus fulgidus*
Dim1	1ZQ9	*Homo sapiens*

"

13. Figure 4A: the way in which the figure is annotated is quite confusing as lower case is used for the figure while the panels are labeled in upper case. Please try to make this clearer for the reader.
The figure was modified according the referees comments and is now found in the Panel 4B

14. Figure 4A: Given that the distortion of H44 is a major point of the paper, supporting EM density should be shown in the figure.
An additional figure panel to Figure 4 was added to show the mature helix 44 overlaid with the immature helix in our low-pass filtered density.
15. What are the black arrows meant to indicate in Figure 4D?
The black arrow indicates the movements of the wHTH domain and the first lobe of the kinase domain needed to reach the close conformation of the crystal structure needed to cleave the ATP. An additional sentence was added to the figure legend to explain this:
"
Comparison of the yeast pre-40S Rio2 conformation (blue) and the crystal structure of ATP-bound Rio2 from A. fulgidus (brown, PDB : 1ZAO, (LaRonde-LeBlanc et al., 2005)). The structures are superimposed relative to their kinase domains. The ATP visualized in the crystal structure is shown in pink to highlight the binding pocket. The arrows represents the movements of the Rio2 domains needed to reach the close conformation of the crystal structure, which is able to cleave ATP

“

16. In Figure 4E, it is important to show EM density to support the conclusions regarding displacement of the decoding center bases that the authors wish to draw.
17. In Figure 4C, show EM density to support the conclusions.
18. Figure 5. Please show density to support the statements in the figure legend and on p10, particularly with respect to the proposed flipping out of base A1749.
EM Density for these three claims were shown in Figure S5B, S5C and S5D respectively (now EV4B, EV4C and EV4D respectively). Additional reference to this panel were added in the figure legend of Figure 4 to bring attention to this fact. The new figure panel 4A is also showing the displacement of the decoding center that indirectly shows the large movement of the decoding center bases.

19. Figure S4: Please label Ltv1 in the figure for clarity
The labels were added as suggested by the referee.

20. Please help the reader by referring to specific figures in the discussion.
Additional reference for the figures and panels were added in the discussion.

21. Page 2: additional examples of publications on cytoplasmic quality control and proofreading should be cited at the end of the first paragraph and not just the authors' own contributions and their review in this area.
Additional references were introduced to this part of the paper.

22. I feel that it would be more appropriate to modify the title- "structure of a (not the) eukaryotic cytoplasmic pre-40S ribosomal subunit.
We agree with the referee that “a eukaryotic cytoplasmic pre-40S ribosomal subunit” is more fitting and changed the title accordingly.

23. Several references are not properly formatted
All the references are now formatted according to the EMBO author guideline.

24. Consider replacing the article by Scheres cited on page 10, second line from the bottom with "Processing of Structurally Heterogeneous Cryo-EM Data in RELION." PMID: 27572726.
The reference was changed from the paper where Sjors Scheres describes the Relion program to the paper where the local classification is described in details, as suggested by the referee.

2nd Editorial Decision 10th January 2018

Thank you for submitting a revised version of your manuscript to The EMBO Journal. It has now been seen by two of the original referees whose comments are shown below. As you will see, they both find that all criticisms have been sufficiently addressed and I am therefore happy to inform you that your study is now accepted for publication here.

If you have any questions, please feel free to contact me. Thank you for your contribution to The EMBO Journal and congratulations on this nicely executed work!

REFEREE REPORTS

Referee #1:

The authors had adequately addressed the issues raised in my initial comments. I now recommend publication.

Referee #3:

The authors have satisfactorily addressed all the concerns raised by this reviewer and have significantly improved the clarity of the work presented in this manuscript.

EMBO PRESS

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND ↓

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

Corresponding Author Name: Nenad Ban
Journal Submitted to: EMBO Journal
Manuscript Number: EMBOJ-2017-98499

Reporting Checklist For Life Sciences Articles (Rev. June 2017)

This checklist is used to ensure good reporting standards and to improve the reproducibility of published results. These guidelines are consistent with the Principles and Guidelines for Reporting Preclinical Research issued by the NIH in 2014. Please follow the journal's authorship guidelines in preparing your manuscript.

A- Figures

1. Data

The data shown in figures should satisfy the following conditions:

- the data were obtained and processed according to the field's best practice and are presented to reflect the results of the experiments in an accurate and unbiased manner.
- figure panels include only data points, measurements or observations that can be compared to each other in a scientifically meaningful way.
- graphs include clearly labeled error bars for independent experiments and sample sizes. Unless justified, error bars should not be shown for technical replicates.
- if n< 5, the individual data points from each experiment should be plotted and any statistical test employed should be justified
- Source Data should be included to report the data underlying graphs. Please follow the guidelines set out in the author ship guidelines on Data Presentation.

2. Captions

Each figure caption should contain the following information, for each panel where they are relevant:

- a specification of the experimental system investigated (eg cell line, species name).
- the assay(s) and method(s) used to carry out the reported observations and measurements
- an explicit mention of the biological and chemical entity(ies) that are being measured.
- an explicit mention of the biological and chemical entity(ies) that are altered/varied/perturbed in a controlled manner.

- the exact sample size (n) for each experimental group/condition, given as a number, not a range;
- a description of the sample collection allowing the reader to understand whether the samples represent technical or biological replicates (including how many animals, litters, cultures, etc.).
- a statement of how many times the experiment shown was independently replicated in the laboratory.
- definitions of statistical methods and measures:
  - common tests, such as t-test (please specify whether paired vs. unpaired), simple $\chi 2$ tests, Wilcoxon and Mann-Whitney tests, can be unambiguously identified by name only, but more complex techniques should be described in the methods section;
  - are tests one-sided or two-sided?
  - are there adjustments for multiple comparisons?
  - exact statistical test results, e.g., P values = x but not P values < x;
  - definition of 'center values' as median or average;
  - definition of error bars as s.d. or s.e.m.

Any descriptions too long for the figure legend should be included in the methods section and/or with the source data.

In the pink boxes below, please ensure that the answers to the following questions are reported in the manuscript itself. Every question should be answered. If the question is not relevant to your research, please write NA (non applicable). We encourage you to include a specific subsection in the methods section for statistics, reagents, animal models and human subjects.

B- Statistics and general methods

Please fill out these boxes ↓ (Do not worry if you cannot see all your text once you press return)

1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size?	NA
1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used.	NA
2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria pre-established?	NA
3. Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. randomization procedure)? If yes, please describe.	NA
For animal studies, include a statement about randomization even if no randomization was used.	NA
4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing results (e.g. blinding of the investigator)? If yes please describe.	NA
4.b. For animal studies, include a statement about blinding even if no blinding was done	NA
5. For every figure, are statistical tests justified as appropriate?	NA
Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it.	NA
Is there an estimate of variation within each group of data?	NA
Is the variance similar between the groups that are being statistically compared?	NA

C- Reagents

USEFUL LINKS FOR COMPLETING THIS FORM

http://www.antibodypedia.com
http://1degreebio.org
http://www.equator-network.org/reporting-guidelines/improving-bioscience-research-repo

http://grants.nih.gov/grants/olaw/olaw.htm
http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/Useofanimals/index.htm
http://ClinicalTrials.gov
http://www.consort-statement.org
http://www.consort-statement.org/checklists/view/32-consort/66-title

http://www.equator-network.org/reporting-guidelines/reporting-recommendations-for-tun

http://datadryad.org

http://figshare.com

http://www.ncbi.nlm.nih.gov/gap

http://www.ebi.ac.uk/ega

http://biomodels.net/

http://biomodels.net/miriam/
http://jjj.biochem.sun.ac.za
http://oba.od.nih.gov/biosecurity/biosecurity_documents.html
http://www.selectagents.gov/

6. To show that antibodies were profiled for use in the system under study (assay and species), provide a citation, catalog number and/or clone number, supplementary information or reference to an antibody validation profile. e.g., Antibodypedia (see link list at top right), 1DegreeBio (see link list at top right).	NA
7. Identify the source of cell lines and report if they were recently authenticated (e.g., by STR profiling) and tested for mycoplasma contamination.	NA

* for all hyperlinks, please see the table at the top right of the document

D- Animal Models

8. Report species, strain, gender, age of animals and genetic modification status where applicable. Please detail housing and husbandry conditions and the source of animals.	NA
9. For experiments involving live vertebrates, include a statement of compliance with ethical regulations and identify the committee(s) approving the experiments.	NA
10. We recommend consulting the ARRIVE guidelines (see link list at top right) (PLoS Biol. 8(6), e1000412, 2010) to ensure that other relevant aspects of animal studies are adequately reported. See author guidelines, under 'Reporting Guidelines'. See also: NIH (see link list at top right) and MRC (see link list at top right) recommendations. Please confirm compliance.	NA

E- Human Subjects

11. Identify the committee(s) approving the study protocol.	NA
12. Include a statement confirming that informed consent was obtained from all subjects and that the experiments conformed to the principles set out in the WMA Declaration of Helsinki and the Department of Health and Human Services Belmont Report.	NA
13. For publication of patient photos, include a statement confirming that consent to publish was obtained.	NA
14. Report any restrictions on the availability (and/or on the use) of human data or samples.	NA
15. Report the clinical trial registration number (at ClinicalTrials.gov or equivalent), where applicable.	NA
16. For phase II and III randomized controlled trials, please refer to the CONSORT flow diagram (see link list at top right) and submit the CONSORT checklist (see link list at top right) with your submission. See author guidelines, under 'Reporting Guidelines'. Please confirm you have submitted this list.	NA
17. For tumor marker prognostic studies, we recommend that you follow the REMARK reporting guidelines (see link list at top right). See author guidelines, under 'Reporting Guidelines'. Please confirm you have followed these guidelines.	NA

F- Data Accessibility

18: Provide a "Data Availability" section at the end of the Materials & Methods, listing the accession codes for data generated in this study and deposited in a public database (e.g. RNA-Seq data: Gene Expression Omnibus GSE39462, Proteomics data: PRIDE PXD000208 etc.) Please refer to our author guidelines for 'Data Deposition'. Data deposition in a public repository is mandatory for: a. Protein, DNA and RNA sequences b. Macromolecular structures c. Crystallographic data for small molecules d. Functional genomics data e. Proteomics and molecular interactions	The final model containing the cytoplasmic Pre-40S with Enp1, Ltv1, Tsr1, Rio2, and Pno1 is available from the Protein Data Bank (PDB) under accession number 6FAI. The cryo-EM map of the unclassified pre-40S is deposited in the Electron Microscopy Data Bank (EMDB) under accession number EMD-4214, as well as the reconstructions from the focus classifications, which can be found under the accession numbers EMD-4216 (Rio1), EMD-4217 (Enp1), and EMD-4218 (Dim1), respectively. The map of the empty 80S is deposited in the EMDB under accession number EMD-4215.
19. Deposition is strongly recommended for any datasets that are central and integral to the study; please consider the journal's data policy. If no structured public repository exists for a given data type, we encourage the provision of datasets in the manuscript as a Supplementary Document (see author guidelines under 'Expanded View' or in unstructured repositories such as Dryad (see link list at top right) or Figshare (see link list at top right).	NA
20. Access to human clinical and genomic datasets should be provided with as few restrictions as possible while respecting ethical obligations to the patients and relevant medical and legal issues. If practically possible and compatible with the individual consent agreement used in the study, such data should be deposited in one of the major public access-controlled repositories such as dbGAP (see link list at top right) or EGA (see link list at top right).	NA
21. Computational models that are central and integral to a study should be shared without restrictions and provided in a machine-readable form. The relevant accession numbers or links should be provided. When possible, standardized format (SBML, CellML) should be used instead of scripts (e.g. MATLAB). Authors are strongly encouraged to follow the MIRIAM guidelines (see link list at top right) and deposit their model in a public database such as Biomodels (see link list at top right) or JWS Online (see link list at top right). If computer source code is provided with the paper, it should be deposited in a public repository or included in supplementary information.	NA

G- Dual use research of concern

22. Could your study fall under dual use research restrictions? Please check biosecurity documents (see link list at top right) and list of select agents and toxins (APHIS/CDC) (see link list at top right). According to our biosecurity guidelines, provide a statement only if it could.	NA