

Molecular mechanism of translational stalling by inhibitory codon combinations and poly(A) tracts

Petr Tesina, Laura N. Lessen, Robert Buschauer, Jingdong Cheng, Colin Chih-Chien Wu, Otto Berninghausen, Allen R. Buskirk, Thomas Becker, Roland Beckmann, Rachel Green.

Editor: Hartmut Vodermaier

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 2019 2nd October 2019

Thank you again for submitting your manuscript on translational stalling mechanisms to The EMBO Journal. We have now received the reports of three expert referees, copied below for your information. As you will see, all referees appreciate the importance of the subject and consider the presented findings interesting in principle. At the same time, they raise a number of issues that would need to be satisfactorily addressed prior to publication. Among the most salient concerns is the major point of referee 1 (regarding effects of CGA codons in the P site), as well as the combined concerns with the Asc1-related analyses.

Should you be able to adequately clarify these key issues, as well as the other main criticisms/suggestions well-explained in all three reports, then we should be happy to consider a revised manuscript further for publication in The EMBO Journal.

--

REFEREE REPORTS

Referee #1:

The manuscript by Tesina & Lessen et al. presents a mechanistic dissection of how certain mRNA codon pairs are prone to induce stalling of the eukaryotic ribosome. These experiments build on the genetic analysis of codon-pair induced decreases in translation efficiency in yeast identified by the Grayhack group. Here the authors use a combination of biochemistry in a reconstituted yeast translation system, ribosome profiling, and cryo-EM to determine at which step these stalling events occur, and the likely molecular basis for the stalling. The authors find that certain codon pairs allow translation to proceed through the step of peptide bond formation and mRNA translocation, but the subsequent mRNA decoding event on the 2nd codon is impaired. Structurally, this seems to be due to nonfunctional conformations of the mRNA in the A site of the small ribosomal subunit. In three separate cases, the A-site codon is mispositioned relative to what would be needed for proper mRNA decoding.

Overall, this is an interesting paper, and the authors can fix a few minor issues before publication. However, there is one major problem that the authors must address before publication. This has to do with the effects of the CGA codon in the P site. In both the CGA-CCG and CGA-CGA codon pair cryo-EM structures, the authors find that the ICG tRNA(Arg) anticodon adopts a "wide" A-I purine-purine base pair conformation. This causes the mRNA extending into the A site to be displaced from its normal conformation seen in canonical mRNA decoding complexes. In the case of CGA-CCG, the normal "kink" between the $+3$ (wobble) position of the P site and the $+4$ nucleotide is misplaced to the +4/+5 position, and the backbone is shifted by 2.6 Å. In the case of the CGA-CGA codon pair, the +4 nucleotide is also displaced, although the following density is not resolved. This raises a fundamental question that the authors do not address. Why, then, does CGA in the P site not affect all codon pairs? In other words, perhaps the "wide" A-I purine-purine pairing is a dead-end conformation, once it traverses from the A site to the P site. Although the A site can accommodate the "wide" conformation, as seen in the Murphy and Ramakrishnan structure (which is a 30S complex, not a functional 70S complex), the P site seems to prevent proper accommodation, as the contacts between the small ribosomal subunit (40S in this case) and the P-site anticodon stemloop are too constrained to allow the tRNA to adapt to the conformational constraints of forming a "wide" A-I pair. Thus the mRNA has to move out of the way, leading to a mispositioned A-site codon. The authors should be more explicit about this and point out that it is still possible (their data suggests likely) that the proper A-I pair would be in the Hoogsteen conformation, rather than the "wide" conformation, to allow for translation to continue.

I would say that the authors' biochemical data supports the Hoogsteen model, as well. In Figures 1B-C and Figure 2, the authors see that elongation through the CGA codon in the P site is inhibited, but never completely. (The authors should double-check their labeling in Figure 1C, by the way.) This might make sense if the "wide" A-I is more of a dead end, but a Hoogsteen A-I can be elongated without much problem. Although it would require future experiments to unravel, it may also be that a "wide" A-I can rearrange on the ribosome in the P site to a Hoogsteen A-I to allow for further elongation, but the non-optimal codons in the A site that Grayhack found and that they study here (CGA and CCG) do not allow for this to happen. In summary, the authors should address the general problem with the "wide" A-I codon-anticodon pair in the P site more directly in the Discussion, along the lines laid out above.

There are a number of minor issues the authors should address:

1. The bibliography is corrupted in a number of places. Perhaps the bibliography library was not fully up-to-date, among other problems. Ikeuchi et al. is duplicated, for example.

2. Figure 1C. Please be sure the labels are right.

3. Red-green traces are hard for color-blind readers. Please consider other choices.

4. The authors cite unpublished data on lines 290-291. If it's not in press by the time this paper is accepted, the authors should post on bioRxiv, which does not preclude publication in Science or elsewhere.

5. The authors use "data not shown" on line 367. Should add to supplement.

6. The title on line 268-269 is confusing. Please clarify.

7. Please use the names for ribosomal proteins as defined in Ban et al., 2014 (see: https://bangroup.ethz.ch/research/nomenclature-of-ribosomal-proteins.html).

Referee #2:

Tesina, Lessen et al studied the mechanistic basis for ribosome stalling at two distinct sequence elements that result in frameshifting and recruitment of quality control sequences. They demonstrated slower elongation rates at the Arg-Pro CGA-CGG and Arg-Arg CGA-CGA inhibitory codon pairs in the yeast S. cerevisiae and defined two molecular defects in translation of these pairs. An initial in vitro analysis pointed to two defects in translation of these codon pairs: a kinetic defect

in decoding that is partially suppressed by increased charged tRNA and a strong endpoint defect that is not suppressed by tRNA. Using high resolution ribosome profiling, the authors then provided additional evidence that the kinetic defect is likely due to defects in tRNA binding or accommodation. They demonstrate the in vivo enrichment of ribosomes with an empty A site at these pairs (and using different combinations of inhibitors argue against the idea that peptide bond formation is rate limiting). Using cryo-EM analysis of ribosomes stalled on these codon pairs, they provided evidence of the fundamental structural defects that may explain the second observation that some ribosomes cannot elongate. The CryoEM points to a central observation of this manuscript, a distorted mRNA structure, in which the normal kink in mRNA between the A and P site codons is displaced. (albeit this is not the expertise of this reviewer) The authors also examine the structures of ribosomes stalled on polyA sequences, making the observation that the mRNA in the A site involves $a \pi$ stacking array, incompatible with decoding and similar to the conformation of single stranded poly(A) stretches.

These results provide important insights into the mechanisms that cause translation stalling and/or frameshifting. The combination of approaches and citations of the literature result in corroboration of the conclusions and strengthens the paper. Overall, this paper contributes substantial new information and new ideas, perhaps most importantly the role of the mRNA in stalling.

Major Concerns.

1. The primary concern of this reviewer is that the distorted mRNA structures could arise as a consequence of collisions in which the leading ribosome is lacking an A site tRNA. Thus, they may not be a direct consequence of the stall itself. On the other hand, the altered mRNA structures do account for the end point defect in the in vitro reaction (on which there is no room for ribosome collisions). It would be useful if the authors directly addressed this possibility.

2. The results showing that in vivo the asc1 mutant exhibits both an overall increase in 21 nt footprints and an even greater defect in decoding the CGA-CCG and CGA-CGA codon pairs are very interesting. However, results showing that ribosomes prepared from asc1-delta mutants exhibit reduced rates and endpoints on both inhibitory and optimal pairs are not particularly informative. While many of the effects of Asc1 can be attributed to its regulation of NoGo decay, the high rates of frameshifting at these codon pairs in an asc1 mutant (Wang J, Zhou J et al, 2018) cannot be explained. Thus, the exacerbated defect in decoding CGA-CCG and CGA-CGA codon pairs in the asc1 mutant may lead to frameshifting (consistent with a lot of literature).

3. The rationale for looking at Asc1 is not explained entirely well. The role of Asc1 as a mediator of inhibitory effects was initially reported by Kuroha, K... Inada (2010). EMBO reports 11(12): 956- 961. and then reported in Brandman, O... Weissman (2012). Cell 151(5): 1042-1054. Moreover, the prevailing idea is that that Asc1 is required to activate the NoGo Decay (Sitron, C. S., J. H. Park and O. Brandman (2017). RNA 23(5): 798-810), in part because because effects on increased downstream expression are accounted for by increased mRNA (Sitron et al, 2017 and Wang, 2018), and consistent with the observation that asc1 mutants only suppress inhibitory codon pairs after amino acid 44 (Wolf and Grayhack, 2015).

Minor points 1. Figure 1C-the graph is labelled CGA-CGA and CGC-CGC, but should be CGA-CCG and CGC-CCA.

2. p 11, line 255 "di-codon pair" redundant

3. Figure 5. The mRNA is not visible in B. Can you include information on the colors used to depict each item in the figure legend? For instance, the 18S rRNA is a different color in each of panels B, E and H. Similarly, the rRNA nucleotides A1756, A 1755 and C 1637 are different colors in C, F and I. This is confusing.

4. As indicated above, I am not expert in CryoEM or structural analysis, but I found it difficult to obtain information from Fig. 7D.

5. There are issues with several references- either repeats of the same reference (marked as a and b) or missing information. Also, one citation is a manuscript under revision (Buschauer)- this may be

done before publication or is a matter for the editor.

Referee #3:

The paper by Tesina et al. addresses one of the most exciting and outstanding questions in translation, namely what determines ribosome stalling at inhibitory codon combinations and on the poly(A) tracts. They combine biochemical experiments in a fully reconstitutes yeast translation system, analysis of translational stalling in vivo by ribosome profiling experiments, and structural analysis of the isolated ribosome complexes by cryo-EM to show why the ribosomes stall. It is a very important contribution from two outstanding labs. The paper provides convincing evidence as to the mechanisms of stalling and should certainly be published. Having said that, I do have some questions/suggestion which - I hope - should help the authors to improve the manuscript. 1. The biochemical experiments lack an important control, namely the positive control showing the endpoint and the rate of reaction with a non-stalling codon pair, i.e. one of those that are read efficiently based on the Grayhack paper. This control is essential to judge the activity of the in vitro translation system and should be provided.

2. The statistical significance and the number of replicates should be provided in all Figs related to biochemical and ribosome profiling studies (in fairness, they are provided in some, but not in all cases). It is not clear whether the differences in rates in Fig. 1, Fig 2C, 4A and the EV Figs are significant, in fact some of the differences appear quite small to be consistent with the in vivo effects (Grayhack). The potential discrepancy between the in vivo and in vitro data should be addressed in the Discussion.

3. The "rescue" effect of UCG tRNA on reading the CGA codon (Fig. 2C) is very subtle if at all (no statistics). The respective sentence on p. 9 (...this non-natural tRNA did partially rescue the endpoint defects...) should be removed.

4. The data concerning the effect of stalling codons on peptide bond formation is not very convincing. The lack of ANS effect is not informative (as any lack of an effect). The question should best be analyzed in the biochemical assay using a non-stalling codon pair as a control. Fig. EV1F is confusing: how can authors obtain saturating MFR levels given that in Fig. EV1C the MFR peptide clearly goes down after 30 sec?

5. Fig. 4 that shows the potential effect of Asc1 is probably the weakest point of the (otherwise very strong) paper. There are some technical issues, i.e. the statistical evaluation of Fig. 4B. Also the conclusions based on Fig. 4C are somewhat stretched due to the lack controls, e.g. a comparison to non-stalling codons or to another ribosome mutant (or other wt strain) that would control of unspecific effects, such as the reproducibility of libraries, etc. In my opinion, this section does not contribute much to the paper and could be deleted.

6. The authors initially carry out their biochemical experiments at limiting tRNA concentrations (which they call kcat/KM conditions). Then, they increase the tRNA concentration 10-fold and imply that these conditions are at tRNA saturation. However, this is not shown, i.e. the conditions could be kcat/KM even at higher tRNA concentrations. To show that the conditions are saturating kinetically, the authors would have to titrate the kobs until it is not changing with tRNA concentration. This is certainly not necessary for the conclusions of this paper, in particular after it

becomes clear that the tRNA does not seem to bind to the A site with the CGA-CGA or CGA-CCG codons. Furthermore, the notion of "saturating tRNA conditions" contradicts the cryo-EM data. The authors should carefully revise the wording to avoid misunderstanding here. A minor point: it is very difficult to actually find the concentrations of components throughout the text. Perhaps the authors should indicate them more clearly in Fig legends.

7. P. 8 "Energetics of tRNA binding (a second order event)". What the authors mean is the kinetics of tRNA binding (a second order event). The energetics may be irrelevant to the phenomena described in the paper and is not addressed anywhere in the paper.

8. The cryo-EM structures did not reveal any complexes with the A site occupied. How does this agree with the results of the biochemical and profiling experiments that indicate at least some binding to the A site, albeit at a reduced endpoint? For the preparation of the complexes for cryo-EM analysis, the authors used an mRNA reporter with two CGA-CCG or CGA-CGA codon pairs. Can they see whether the ribosome stopped at the 1st or the 2nd codon pair? The authors argue that the conformation of the PTC is not affected. It would be important to document the density at the PTC. The authors use an A-P tRNA complex for comparison, however, the correct comparison

would be with a conventional P-E tRNA complex. Can the authors provide such comparison as well?

9. Fig. 5. The multitude of colors in ribosome structures of A,D,G, bottom panels, is overwhelming. Why not use a unifying color code for 40S and 60S, which should differ from the color code used for tRNAs? The details are really difficult to see. Fig 5E, what is the gray element pointing into the A site?

10. For the disome, the 2nd ribosome is shown to stall in the Post state, which is different from the stalled disome complexes reported so far. Where does the 2nd ribosome stall on the mRNA? The authors should see the presumed frameshifting based on the mRNA position. Does the POST state of the 2nd ribosome agree with the profiling experiments? Why is the resolution so much poorer for both ribosomes in the disome and in particular for the 1st ribosome?

11. Concerning the poly(A)-induced stalling, how certain the authors are that the accumulation of Lys residues in the polypeptide exit tunnel does not attenuate the stalling effect? Did they test this experimentally?

12. P. 18, Discussion. The authors compare their P-site orientation of the I:A base pair with that in the A site (Ramakrishnan 2004). But aren't also the ribosomes coming from different organisms, yeast vs. bacterial? How informative is such a comparison? Is this part of the decoding center conserved enough to warrant the discussion?

Minor points:

p. 3, in the Introduction, the authors mention synonymous codon choice as a factor affecting translational efficiency and co-translational protein folding. I have to say that the choice of references appears strange to this reviewer. There are two groups worldwide who contributed most to the understanding of this question, Patricia Clark and Anton Komar. Neither of them is cited, although the paper by the Komar group (Mol Cell 2016) would be very relevant here (together with Pechmann and Frydmann citation), whereas the Thanaraj and Argos paper is just a computer prediction lacking any experimental validation (there are many of them). Similarly, the experimental work by Clark should be cited for the efficiency.

Fig. 1B,C and all other figures showing rate constants. Indicate the dimention on k, is it per sec or per min?

Fig. 3: Please provide statistical analysis in the Fig. legend.

Appendix Fig S5, the data on puromycin reactivity in B look strange, because free peptide does not seem to appear with time. The result should be quantified to check whether the sum of substrate and product match for each lane.

Fig EV4: The colors are weird; please change to systematic color code.

Fig EV5. Panel C is actually not very convincing, as the effect on the mRNA conformation seems much less than what the one in the main Figs. Perhaps some other way of presenting the differences would help. Panel D, please provide the mRNA sequence +4 to +14.

Fig EV6, add (min) to "time". The reason for diluting the sample is entirely unclear - please explain in the legend.

1st Revision - authors' response 8th November 2019

Referee #1:

The manuscript by Tesina & Lessen et al. presents a mechanistic dissection of how certain mRNA codon pairs are prone to induce stalling of the eukaryotic ribosome. These experiments build on the genetic analysis of codon-pair induced decreases in translation efficiency in yeast identified by the Grayhack group. Here the authors use a combination of biochemistry in a reconstituted yeast translation system, ribosome profiling, and cryo-EM to determine at which step these stalling events occur, and the likely molecular basis for the stalling. The authors find that certain codon pairs allow translation to proceed through the step of peptide bond formation and mRNA translocation, but the subsequent mRNA decoding event on the 2nd codon is impaired. Structurally, this seems to be due to nonfunctional conformations of the mRNA in the A site of the small ribosomal subunit. In three separate cases, the A-site codon is mispositioned relative to what would be needed for proper mRNA decoding.

Overall, this is an interesting paper, and the authors can fix a few minor issues before publication. However, there is one major problem that the authors must address before publication. This has to do with the effects of the CGA codon in the P site. In both the CGA-CCG and CGA-CGA codon pair cryo-EM structures, the authors find that the ICG tRNA(Arg) anticodon adopts a "wide" A-I purine-purine base pair conformation. This causes the mRNA extending into the A site to be displaced from its normal conformation seen in canonical mRNA decoding complexes. In the case of CGA-CCG, the normal "kink" between the $+3$ (wobble) position of the P site and the $+4$ nucleotide is misplaced to the $+4/+5$ position, and the backbone is shifted by 2.6 Å. In the case of the CGA-CGA codon pair, the $+4$ nucleotide is also displaced, although the following density is not resolved. This raises a fundamental question that the authors do not address. Why, then, does CGA in the P site not affect all codon pairs? In other words, perhaps the "wide" A-I purine-purine pairing is a dead-end conformation, once it traverses from the A site to the P site. Although the A site can accommodate the "wide" conformation, as seen in the Murphy and Ramakrishnan structure (which is a 30S complex, not a functional 70S complex), the P site seems to prevent proper accommodation, as the contacts between the small ribosomal subunit (40S in this case) and the P-site anticodon stem-loop are too constrained to allow the tRNA to adapt to the conformational constraints of forming a "wide" A-I pair. Thus the mRNA has to move out of the way, leading to a mispositioned A-site codon. The authors should be more explicit about this and point out that it is still possible (their data suggests likely) that the proper A-I pair would be in the Hoogsteen conformation, rather than the "wide" conformation, to allow for translation to continue.

I would say that the authors' biochemical data supports the Hoogsteen model, as well. In Figures 1B-C and Figure 2, the authors see that elongation through the CGA codon in the P site is inhibited, but never completely. (The authors should double-check their labeling in Figure 1C, by the way.) This might make sense if the "wide" A-I is more of a dead end, but a Hoogsteen A-I can be elongated without much problem. Although it would require future experiments to unravel, it may also be that a "wide" A-I can rearrange on the ribosome in the P site to a Hoogsteen A-I to allow for further elongation, but the non-optimal codons in the A site that Grayhack found and that they study here (CGA and CCG) do not allow for this to happen. In summary, the authors should address the general problem with the "wide" A-I codon-anticodon pair in the P site more directly in the Discussion, along the lines laid out above.

We thank the referee for this interesting point. It is indeed possible that rearrangement from the observed Ianti-Aanti into the Hoogsteen Isyn-Aanti can occur providing a valid alternative to support and explain some of our results. We now discuss this possibility in the Discussion as requested:

"It is in principle possible that under normal conditions, a Isyn-Aanti conformation (Hoogsteen base pair with the hypoxanthine base of inosine rotated by 180° respective to its ribose) is adopted to minimize geometric strain and allow for further elongation. In this particular case, the inhibitory codon pairs studied here could prevent this Hoogsteen base pair formation and thereby create a dead-end situation. Since the Isyn-Aanti *conformation has not been observed in the context of decoding by ICG tRNAArg, this alternative would require closer investigation in the future."*

There are a number of minor issues the authors should address:

1. The bibliography is corrupted in a number of places. Perhaps the bibliography library was not fully up-to-date, among other problems. Ikeuchi et al. is duplicated, for example.

Thank you for catching this – the bibliography has been corrected.

2. Figure 1C. Please be sure the labels are right. *This has been corrected.*

3. Red-green traces are hard for color-blind readers. Please consider other choices.

We did consider this as we prepared the figures, but because red and green are associated with stop and go, or inhibitory and optimal in our case, we thought it would make for easier interpretation of the data. Therefore, we chose variations of the two colors that are accessible to colorblind readers – vermillion for red and bluish green for green. (Color Universal Design (CUD) - How to make figures and presentations that are friendly to Colorblind people [https://jfly.uni-koeln.de/color/\)](https://jfly.uni-koeln.de/color/)

4. The authors cite unpublished data on lines 290-291. If it's not in press by the time this paper is accepted, the authors should post on bioRxiv, which does not preclude publication in Science or elsewhere.

We will do exactly as suggested by the reviewer and deposit these data on bioRxiv to be cited upon publication.

5. The authors use "data not shown" on line 367. Should add to supplement.

We have added this result as Appendix Fig S5C.

6. The title on line 268-269 is confusing. Please clarify.

This has been corrected.

7. Please use the names for ribosomal proteins as defined in Ban et al., 2014 (see: https://bangroup.ethz.ch/research/nomenclature-of-ribosomal-proteins.html).

We accidentally used the old nomenclature on one instance when discussing our Asc1 results. Since the whole chapter and corresponding discussion on Asc1 was deleted from the revised manuscript (see below) we should now consistently use only the new nomenclature of ribosomal proteins throughout the whole manuscript.

Referee #2:

Tesina, Lessen et al studied the mechanistic basis for ribosome stalling at two distinct sequence elements that result in frameshifting and recruitment of quality control sequences. They demonstrated slower elongation rates at the Arg-Pro CGA-CGG and Arg-Arg CGA-CGA inhibitory codon pairs in the yeast S. cerevisiae and defined two molecular defects in translation of these pairs. An initial in vitro analysis pointed to two defects in translation of these codon pairs: a kinetic defect in decoding that is partially suppressed by increased charged tRNA and a strong endpoint defect that is not suppressed by tRNA. Using high resolution ribosome profiling, the authors then provided additional evidence that the kinetic defect is likely due to defects in tRNA binding or accommodation. They demonstrate the in vivo enrichment of ribosomes with an empty A site at these pairs (and using different combinations of inhibitors argue against the idea that peptide bond formation is rate limiting). Using cryo-EM analysis of ribosomes stalled on these codon pairs, they provided evidence of the fundamental structural defects that may

explain the second observation that some ribosomes cannot elongate. The CryoEM points to a central observation of this manuscript, a distorted mRNA structure, in which the normal kink in mRNA between the A and P site codons is displaced. (albeit this is not the expertise of this reviewer) The authors also examine the structures of ribosomes stalled on polyA sequences, making the observation that the mRNA in the A site involves a \Box stacking array, incompatible with decoding and similar to the conformation of single stranded poly(A) stretches.

These results provide important insights into the mechanisms that cause translation stalling and/or frameshifting. The combination of approaches and citations of the literature result in corroboration of the conclusions and strengthens the paper. Overall, this paper contributes substantial new information and new ideas, perhaps most importantly the role of the mRNA in stalling.

Major Concerns.

1. The primary concern of this reviewer is that the distorted mRNA structures could arise as a consequence of collisions in which the leading ribosome is lacking an A site tRNA. Thus, they may not be a direct consequence of the stall itself. On the other hand, the altered mRNA structures do account for the end point defect in the in vitro reaction (on which there is no room for ribosome collisions). It would be useful if the authors directly addressed this possibility.

The reviewer raises an important point. However, here we directly demonstrate kinetic and endpoint defects on ribosome complexes prepared in vitro that certainly only have one ribosome, though these data do not report on mRNA structure per se. Moreover, and more importantly, the distorted mRNA structures that we observed here for CGA-CCG, CGA-CGA and poly(A) mRNAs were solved in each case with an isolated monomeric 80S fraction (see Appendix Figs S1, S2 and S6). As such, the mRNA conformations that we describe cannot be a result of or depend on ribosome collisions. In the case of the poly(A) stalled ribosomes, we observed the same mRNA conformation in both the monomeric ribosome and the first stalling ribosome of the disome structure. In order to be clear about this point, we now also mention the use of the 80S RNCs in figure legends of Figs 4 and 5.

2. The results showing that in vivo the asc1 mutant exhibits both an overall increase in 21 nt footprints and an even greater defect in decoding the CGA-CCG and CGA-CGA codon pairs are very interesting. However, results showing that ribosomes prepared from asc1-delta mutants exhibit reduced rates and endpoints on both inhibitory and optimal pairs are not particularly informative.

While many of the effects of Asc1 can be attributed to its regulation of NoGo decay, the high rates of frameshifting at these codon pairs in an asc1 mutant (Wang J, Zhou J et al, 2018) cannot be explained. Thus, the exacerbated defect in decoding CGA-CCG and CGA-CGA codon pairs in the asc1 mutant may lead to frameshifting (consistent with a lot of literature).

While we felt that our data were consistent with our overall observations (Asc1-deficient ribosomes are inefficient at decoding, as demonstrated by kinetics and profiling, and such defects typically increase frameshifting), we appreciate that the Asc1 data are a bit of a distraction from the main points of the paper. In light of related comments from multiple reviewers, we have removed this section from the manuscript.

3. The rationale for looking at Asc1 is not explained entirely well. The role of Asc1 as a mediator of inhibitory effects was initially reported by Kuroha, K... Inada (2010) . EMBO reports $11(12)$: 956-961. and then reported in Brandman, O... Weissman (2012). Cell 151(5): 1042-1054. Moreover, the prevailing idea is that that Asc1 is required to activate the NoGo Decay (Sitron, C. S., J. H. Park and O. Brandman (2017). RNA 23(5): 798-810), in part because because effects on increased downstream expression are accounted for by increased mRNA (Sitron et al, 2017 and Wang, 2018), and consistent with the observation that asc1 mutants only suppress inhibitory codon pairs after amino acid 44 (Wolf and Grayhack, 2015).

Asc1 data have been removed as discussed above in (2).

Minor points

1. Figure 1C-the graph is labelled CGA-CGA and CGC-CGC, but should be CGA-CCG and CGC-CCA.

Agreed. This has been corrected.

2. p 11, line 255 "di-codon pair" redundant

Agreed – this has been changed to "codon pair".

3. Figure 5. The mRNA is not visible in B. Can you include information on the colors used to depict each item in the figure legend? For instance, the 18S rRNA is a different color in each of panels B, E and H. Similarly, the rRNA nucleotides A1756, A 1755 and C 1637 are different colors in C, F and I. This is confusing.

We thank the referee for this valid point. To depict our results in a clear and broadly accessible way, we have reworked Figures 5, 6 and 7 to use a uniform coloring for ribosomal subunits, mRNA and tRNAs throughout the manuscript.

4. As indicated above, I am not expert in CryoEM or structural analysis, but I found it difficult to obtain information from Fig. 7D.

In line with the point discussed above, this figure has been completely reworked to improve the presentation clarity of our findings.

5. There are issues with several references- either repeats of the same reference (marked as a and b) or missing information. Also, one citation is a manuscript under revision (Buschauer)- this may be done before publication or is a matter for the editor.

We thank the referee for pointing this out. The bibliography has been corrected.

Referee #3:

The paper by Tesina et al. addresses one of the most exciting and outstanding questions in translation, namely what determines ribosome stalling at inhibitory codon combinations and on the $poly(A)$ tracts. They combine biochemical experiments in a fully reconstitutes yeast translation system, analysis of translational stalling in vivo by ribosome profiling experiments, and structural analysis of the isolated ribosome complexes by cryo-EM to show why the ribosomes stall. It is a very important contribution from two outstanding labs. The paper provides convincing evidence as to the mechanisms of stalling and should certainly be published. Having said that, I do have some questions/suggestion which - I hope - should help the authors to improve the manuscript.

1. The biochemical experiments lack an important control, namely the positive control showing the endpoint and the rate of reaction with a non-stalling codon pair, i.e. one of those that are read efficiently based on the Grayhack paper. This control is essential to judge the activity of the in vitro translation system and should be provided.

We appreciate these concerns. Based on the Grayhack paper, both controls that we refer to as "optimal", CGC CGC and CGC CCA, are "read efficiently" as they both exhibit high or intermediate expression in the original screen (their Supplementary Table S1). Additionally, the rates and endpoints from this study are comparable to those published using the same in vitro reconstituted system (Schuller et. al, 2017).

We have added a sentence to the text to clarify this point (line 145).

2. The statistical significance and the number of replicates should be provided in all Figs related to biochemical and ribosome profiling studies (in fairness, they are provided in some, but not in all cases). It is not clear whether the differences in rates in Fig. 1, Fig 2C, 4A and the EV Figs are significant, in fact some of the differences appear quite small to be consistent with the in vivo effects (Grayhack). The potential discrepancy between the in vivo and in vitro data should be addressed in the Discussion.

We apologize for our lack of clarity on this point. These statistics are critical and we have now indicated the significance and replicates more clearly. P-values from Student's t-tests have been added and the exact number of replicates are now indicated by the individual data points on all graphs where rates or endpoints have been averaged.

For figure 1, the quantified rates and endpoints are for the individual TLC shown. The errors associated with the rates and endpoints for the two pairs are included in Figures 2A and 2B where multiple experiments are averaged and we have added statistical analysis (p-values from Student's ttests) to these graphs.

The statistical analysis for the rates and endpoints for the non-natural tRNA experiment (Figure 2C) has been added to the supplement (Figure EV2B-F), as again Figure 2C shows representative kinetic data for a single time course (this is also discussed in more detail in response to the next comment, below).

As for the correlation with in vivo data, we recognize that there are potential limitations for an in vitro reconstituted system where different steps may become rate limiting. Importantly, however, in vivo it is likely that relatively small differences for a given peptidyl transfer step are amplified by the iteration of these motifs in the original reporter, making a direct comparison difficult. The pairs we studied using our in vitro system are significantly inhibitory compared to their optimal controls consistent with both the previously published in vivo data as well as the profiling and structural data we present. We added a sentence to the discussion directly addressing this point "These findings are broadly consistent with previously published in vivo data (Gamble et al., 2016) although direct comparisons of the magnitude of defects in these systems are likely not of particular value given their very different limitations."

3. The "rescue" effect of UCG tRNA on reading the CGA codon (Fig. 2C) is very subtle if at all (no statistics). The respective sentence on p. 9 (...this non-natural tRNA did partially rescue the endpoint defects...) should be removed.

The graph shown in Figure 2C is one experiment. We have added the average rates and endpoints of repeated experiments to the supplement *(Figures EV2A and EV2B) as well as the individual kinetic analysis for the other repeats not displayed in the main figure (Figures EV2D-F). The statistics associated with these experiments are now shown and from these data we draw the conclusion that the non-natural tRNA does indeed partially rescue the defects.*

4. The data concerning the effect of stalling codons on peptide bond formation is not very convincing. The lack of ANS effect is not informative (as any lack of an effect). The question should best be analyzed in the biochemical assay using a nonstalling codon pair as a control. Fig. EV1F is confusing: how can authors obtain saturating MFR levels given that in Fig. EV1C the MFR peptide clearly goes down after 30 sec?

I believe we failed to effectively describe the profiling experiments which are based on some earlier results published by (Wu et al 2019). What we report in that study is that depending on which antibiotics are used to trap elongating ribosomes in the lysate preparation, different states of elongating ribosomes can be visualized by ribosome profiling. What we observe in this study is that the slow step for the codon pairs is captured effectively with tigecycline (which traps ribosomes that are struggling to bind tRNAs as 21 RPFs) and that library preparation with anisomycin (which additionally traps ribosomes that are struggling to form peptide bonds as 21 RPFs) does not increase the signal. These data indicate that the slow step for these problematic codon pairs is indeed decoding and not peptidyl transfer. These data are presented in Fig. 3. The supplemental figure (Supplemental Fig 3) is meant to show that if peptidyl transfer were rate limiting (as we have shown for PDE containing motifs), that our data would report on this. We have tried to clarify these points in the text.

Figure EV1F shows the kinetics of elongation on a simple MFR message (with only one arginine) to confirm that the addition of the first CGA arginine is not a problem in our system. Figure EV1D (we believe you meant D and not C) quantifies the strong accumulation of the MFR product on the MFRR messages for the CGA-CGA case and not the CGC-CGC case; these data show that it is the addition of the second CGA arginine that is strongly inhibited. We have removed Figure EV1D as we see how it is confusing and instead refer the readers to the build-up of the MFR product on the TLC in Figure 1B.

5. Fig. 4 that shows the potential effect of Asc1 is probably the weakest point of the (otherwise very strong) paper. There are some technical issues, i.e. the statistical evaluation of Fig. 4B. Also the conclusions based on Fig. 4C are somewhat stretched due to the lack controls, e.g. a comparison to non-stalling codons or to another ribosome mutant (or other wt strain) that would control of unspecific effects, such as the reproducibility of libraries, etc. In my opinion, this section does not contribute much to the paper and could be deleted.

We appreciate these concerns. Asc1 data have been removed as discussed above in comments to Reviewer 2.

6. The authors initially carry out their biochemical experiments at limiting tRNA concentrations (which they call kcat/KM conditions). Then, they increase the tRNA concentration 10-fold and imply that these conditions are at tRNA saturation. However, this is not shown, i.e. the conditions could be kcat/KM even at higher tRNA concentrations. To show that the conditions are saturating kinetically, the authors would have to titrate the kobs until it is not changing with tRNA

concentration. This is certainly not necessary for the conclusions of this paper, in particular after it becomes clear that the tRNA does not seem to bind to the A site with the CGA-CGA or CGA-CCG codons. Furthermore, the notion of "saturating tRNA conditions" contradicts the cryo-EM data. The authors should carefully revise the wording to avoid misunderstanding here. A minor point: it is very difficult to actually find the concentrations of components throughout the text. Perhaps the authors should indicate them more clearly in Fig legends.

We agree with this reviewer that the best way to show that the tRNA concentration is truly saturating is a titration. However, because we are not easily able to go higher in tRNA concentration, we lowered the tRNA concentration by 2-fold and did not see significant changes in the kobs at this 2-fold lower concentrations. As such, the concentration of tRNA utilized in this experiment is saturating. We mention that we previously established this concentration to be saturating in the text and have added the corresponding experiment to the expanded view (Figure EV2A). We also added the concentrations to the figure legends for clarity as suggested.

We do not believe that the use of the saturating tRNA conditions contradicts our cryo-EM data. We simply find that even at high concentrations of tRNA, we see an endpoint defect that is not rescued. These data are consistent with a model where the A site is unavailable for occupancy by the tRNA because of the unusual mRNA structures found there. These data are also very consistent with our ribosome profiling data (Figure 3).

7. P. 8 "Energetics of tRNA binding (a second order event)". What the authors mean is the kinetics of tRNA binding (a second order event). The energetics may be irrelevant to the phenomena described in the paper and is not addressed anywhere in the paper.

This has been corrected.

8. The cryo-EM structures did not reveal any complexes with the A site occupied. How does this agree with the results of the biochemical and profiling experiments that indicate at least some binding to the A site, albeit at a reduced endpoint? For the preparation of the complexes for cryo-EM analysis, the authors used an mRNA reporter with two CGA-CCG or CGA-CGA codon pairs. Can they see whether the ribosome stopped at the 1st or the 2nd codon pair? The authors argue that the conformation of the PTC is not affected. It would be important to document the density at the PTC. The authors use an A-P tRNA complex for comparison, however, the correct comparison would be with a conventional P-E tRNA complex. Can the authors provide such comparison as well?

As discussed above in response to point #6, these data are in general consistent with a model where the A site is unavailable for the tRNA due to obstruction by the mRNA structures found there. These data are also consistent with our ribosome profiling data (Figure 3). While some degree of tRNA binding likely occurs over time, it would not be stable enough during purification and cryo-EM sample preparation to represent a significant population of particles as evidenced by 3D classification (see Appendix).

We can see that the stalling stably occurs on the first copy of the inhibitory codon pair as documented by clear density of the nascent chain residues for which we could build a model. We have added this information to the main text and present the data as a new EV5A panel.

The state of the PTC is now documented in Fig EV5B showing both density and model.

We agree with the referee that the comparison with a "canonical" state of mRNA with an empty A site, yet on a ribosome with P/P E/E tRNA state would be ideal. Unfortunately, such comparison is difficult to perform because the "undecoded" canonical mRNA is usually intrinsically flexible. Unless stabilized in a particular conformation (such as in the cases we present), it yields only poorly resolved density which is not clearly interpretable.

9. Fig. 5. The multitude of colors in ribosome structures of A,D,G, bottom panels, is overwhelming. Why not use a unifying color code for 40S and 60S, which should differ from the color code used for tRNAs? The details are really difficult to see. Fig 5E, what is the gray element pointing into the A site?

We thank the reviewer for pointing this out. As discussed above in response to Referee #2, we have reworked Figures 5, 6 and 7 to depict our results in a clear and broadly accessible way using a uniform coloring for ribosomal subunits, mRNA and tRNAs throughout the manuscript.

10. For the disome, the 2nd ribosome is shown to stall in the Post state, which is different from the stalled disome complexes reported so far. Where does the 2nd ribosome stall on the mRNA? The authors should see the presumed frameshifting based on the mRNA position. Does the POST state of the 2nd ribosome agree with the profiling experiments? Why is the resolution so much poorer for both ribosomes in the disome and in particular for the 1st ribosome?

In principle the second (colliding) ribosome stalls approximately 30 nucleotides upstream of the first (stalling) one, as shown by us and others previously (Ikeuchi et al., 2019 and Juszkiewicz et al., 2018). As the stalling ribosome in this case stalls on the poly(A) tract, it is not possible to directly observe the frameshift even with much higher resolution since all the stabilized (well resolved) bases in this region are adenines.

We are not able to resolve "frame" using ribosome profiling, except in extreme cases, and so we are not able to determine by this method whether there is frameshifting on the poly(A) sequences in the transcriptome.

The resolution remained somewhat lower than for the other presented structures since the quantity and quality of the initial dataset were limited. However, since the obtained resolution (4.0 and 3.6 Å for the first and second ribosome, respectively) was more than sufficient to conclusively distinguish between hybrid and post states, we refrained from collecting more and better data.

11. Concerning the poly(A)-induced stalling, how certain the authors are that the accumulation of Lys residues in the polypeptide exit tunnel does not attenuate the stalling effect? Did they test this experimentally?

We appreciate these comments. Indeed, translation of not only poly Lys, but polybasic sequences in general have been shown to cause stalling during elongation as we cite in the manuscript (Lu & Deutsch, 2008) or as can found in related literature (Ito-Harashima, Kuroha et al., 2007; Dimitrova, Kuroha et al., 2009). In light of this reviewer comment, and some unassigned density in the PTC, we re-analyzed some of our profiling and find evidence of peptide-mediated stalling through defects in peptidyl transfer at poly-Lys sequences. When we compare ribosome profiling data from different library preparations, we see enhanced stalling on iterated *AAG lysines in libraries prepared with ANS/CHX, but not with those prepared with TIG/CHX. These data indicate that there are no stalling effects per se associated with "decoding" on AAG lysines but that there are defects associated with peptide bond formation. These data are now included as Fig. EV6B and C.*

12. P. 18, Discussion. The authors compare their P-site orientation of the I:A base pair with that in the A site (Ramakrishnan 2004). But aren't also the ribosomes coming from different organisms, yeast vs. bacterial? How informative is such a comparison? Is this part of the decoding center conserved enough to warrant the discussion?

The reviewer is correct, the previous study by Murphy and Ramakrishnan indeed characterizes a bacterial ribosome. We are convinced that the discussion is warranted for three reasons: First, in general, the basic ribosomal functions (such as decoding and peptidyl-transfer reaction) are extremely conserved in all kingdoms of life, both sequence-wise and structurally. Second, the previously observed I:A base pair structure in bacteria is to our knowledge the first and only available and therefore cannot be omitted. Third, as requested by Referee #1, we have now expanded our discussion on the issue of possible I:A base pair arrangements and can't ignore this seminal work by Murphy and Ramakrishnan. Therefore, we strongly believe that this comparison is informative.

Minor points:

p. 3, in the Introduction, the authors mention synonymous codon choice as a factor affecting translational efficiency and co-translational protein folding. I have to say that the choice of references appears strange to this reviewer. There are two groups worldwide who contributed most to the understanding of this question, Patricia Clark and Anton Komar. Neither of them is cited, although the paper by the Komar group (Mol Cell 2016) would be very relevant here (together with Pechmann and Frydmann citation), whereas the Thanaraj and Argos paper is just a computer prediction lacking any experimental validation (there are many of them). Similarly, the experimental work by Clark should be cited for the efficiency.

We agree with the reviewer and have added these citations to the introduction accordingly.

Fig. 1B,C and all other figures showing rate constants. Indicate the dimention on k, is it per sec or per min?

These are per second and the figure has been updated to label them accordingly.

Fig. 3: Please provide statistical analysis in the Fig. legend.

Student's t-test between each inhibitory and the corresponding optimal pair were performed for Figure 3B and the values have been added to the figure legend. Student's t-test between CHX-TIG and CHX-ANS were also performed for Figures 3C and EV3 and the values have been added to the figure legends.

Appendix Fig S5, the data on puromycin reactivity in B look strange, because free peptide does not seem to appear with time. The result should be quantified to check whether the sum of substrate and product match for each lane.

The reviewer is correct. To clarify this, we have repeated the experiment and added colorimetric quantification by band integrated density *measurement for Appendix Fig S5B. This analysis shows approximately constant overall sum value of both signals for each lane.*

Fig EV4: The colors are weird; please change to systematic color code.

As discussed above, we have now changed the colors accordingly.

Fig EV5. Panel C is actually not very convincing, as the effect on the mRNA conformation seems much less than what the one in the main Figs. Perhaps some other way of presenting the differences would help. Panel D, please provide the mRNA sequence $+4$ to $+14$.

As requested by the reviewer, we have added an additional view on the differences in revised Fig EV5E. We have also added the sequence information in the legend as requested.

Fig EV6, add (min) to "time". The reason for diluting the sample is entirely unclear - please explain in the legend.

The times are actually in seconds; the label has been corrected accordingly. Diluting the samples allows them to be slightly better resolved by TLC. This was indicated in the methods section and has been added to the figure legend as requested.

Accepted 6th December 2019

Thank you for submitting your revised manuscript for our consideration. Referee 3 has now assessed it once more, and found the new version and your responses to the original reports generally satisfactory. We shall therefore be happy to publish the study in The EMBO Journal, after a few minor editorial modifications.

REFEREE REPORTS

Referee #3 (Report for Author)

The authors did a very good job in revising their manuscript and answered all my concerns. I have no further questions.

EMBO PRESS

YOU MUST COMPLETE ALL CELLS WITH A PINK BACKGROUND $\bm{\Downarrow}$

PLEASE NOTE THAT THIS CHECKLIST WILL BE PUBLISHED ALONGSIDE YOUR PAPER

Journal Submitted to: EMBO Journal Corresponding Author Name: Rachel Green, Roland Beckmann , Thomas Becker Manuscript Number: EMBOJ-2019-103365R

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:

- è
- è 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, measuremen
- 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
- iustified
→ 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 chem
- → the exact sample size (n) for each experimental group/condition, given as a number, not a range; è 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:
- **3** a statement of how many times the experiment shown was independently replicated in the laboratory.
 3 definitions of statistical methods and measures:

* common tests, such as t-test (please specify whether paired v 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.
In the pink boxes bould be answered. If the question is not relevant to your research, please writ **Every question should be answered. If the question is not relevant to your research, please write NA (non applicable). Durage you to include a specific subsection in the methods section for statistics, reagents, animal models and h subjects.**

B- Statistics and general methods

1.a. How was the sample size chosen to ensure adequate power to detect a pre-specified effect size? 1.b. For animal studies, include a statement about sample size estimate even if no statistical methods were used. 2. Describe inclusion/exclusion criteria if samples or animals were excluded from the analysis. Were the criteria preestablished? 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. .
For animal studies, include a statement about randomization even if no randomization was used. 4.a. Were any steps taken to minimize the effects of subjective bias during group allocation or/and when assessing result (e.g. blinding of the investigator)? If yes please describe. 4.b. For animal studies, include a statement about blinding even if no blinding was done 5. For every figure, are statistical tests justified as appropriate? Do the data meet the assumptions of the tests (e.g., normal distribution)? Describe any methods used to assess it. Please fill out these boxes \bigvee (Do not worry if you cannot see all your text once you press return) NA NA NA NA or Figures 2, 3, EV2, EV3 unpaired, two-tailed Student's t tests were used as indicated in t legend and methods section. In all cases, two sets of measurements are being compared and none
of the samples are matched. When only two replicates are available the power of the statistical
test is lower due to the inabil es the data is normally distributed. Data with more than two replicates was tested for normality sing Shapiro-Wilk test. NA NA NA

USEFUL LINKS FOR COMPLETING THIS FORM

http://www.antibodypedia.com http://1degreebio.org http://www.equator-ne 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-recommentationshttp://figshare.com http://www.nchi.nlm.nih.gov/gap http://www.ebi.ac.uk/ega http://biomodels.net/

→ 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 a

http://datadryad.org

http://biomodels.net/miriam/
http://iii.biochem.sun.ac.za

C- Reagents

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

D- Animal Models

E- Human Subjects

F- Data Accessibility

G- Dual use research of concern

