THE EMBO JOURNAL

# Molecular architecture of 40S translation initiation complexes on the Hepatitis C virus IRES

Zuben Brown, irina Abaeva, Swastik De, Christopher Hellen, Tatyana Pestova, and Joachim Frank
**DOI: 10.15252/embj.2022110581**

*Corresponding author(s): Joachim Frank (jf2192@cumc.columbia.edu) , Tatyana Pestova (tatyana.pestova@downstate.edu)*

## Review Timeline:


| | |
|---|---|
| Submission Date: | 5th Jan 22 |
| Editorial Decision: | 3rd Feb 22 |
| Revision Received: | 28th Apr 22 |
| Editorial Decision: | 25th May 22 |
| Revision Received: | 1st Jun 22 |
| Accepted: | 14th Jun 22 |

*Editor: Stefanie Boehm*


## 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. Depending on transfer agreements, referee reports obtained elsewhere may or may not be included in this compilation. Referee reports are anonymous unless the Referee chooses to sign their reports.)

**1st Editorial Decision** **3rd Feb 2022**

Thank you again for submitting your manuscript reporting IRES initiation complex cryo-EM structures to The EMBO Journal. We have now received reports from three experts in the field, which are included below for your information. In light of the referees' comments, we would now invite you to prepare and submit a revised version of the manuscript.

As you will see, the reviewers appreciate the structures and proposed model for complex, but also raise several major concerns that must be resolved before the study can be considered further for publication. In particular, the structures and modeling for the initial steps of IRES-binding should be carefully reviewed and the referees' concerns addressed. In line with this, please also revise the text to ensure that all interpretations and conclusions are fully supported by the experimental data (for example referee #1-point 1, 2, referee #2- point 6) and clarify the concerns of referee #2 and #3 regarding structure 14. In addition, referee #3 is concerned about the reported non-canonical eIF5B-bound intermediate, please provide additional support for this in a functional context. Please also carefully consider all other referee comments and revise the manuscript and figures as appropriate, as well as providing a detailed response to each comment. In case there are any changes to the raw data, please remember to again provide referee access to these files. We have copied the already provided source data to an independent server, such that the previous public link can be disabled if wanted.

Please note that it is our policy to allow only a single round of major revision. Acceptance depends on a positive outcome of a second round of review and therefore on the completeness of your responses included in the next, final version of the manuscript. It is thus important to clarify any questions and concerns at this stage and I encourage you to review the referees' comments and to contact me to discuss specific points or a preliminary revision plan in case there are any uncertainties regarding the revision.

-

---

Referee #1:

In this work, Brown et al. set out to structurally characterize in detail the mechanism of HCV IRES-mediated translation initiation, which is known to occur through two possible pathways (eIF2 or eIF5B dependent). To do so, they determined 18 different cryo-EM structures of ribosomal complexes assembled in vitro from purified components and representing various steps of the initiation process, with resolutions up to 3.5 Å. To capture the early steps of HCV IRES recognition and binding, they relied on the use of a deletion mutant of HCV IRES lacking domain II (ΔdII), which can bind to the 40S subunit, but has difficulty loading its initiation codon-containing region into the mRNA binding channel. In addition, they could obtain various complexes of pre-initiation and initiation complexes featuring wild-type or ΔdII HCV IRES, Met-tRNAiMet and initiation factors, including eIF2 or eIF5B. Thus, this study provides detailed snapshots of all steps of IRES binding up to immediately before subunit junction. In particular, it reveals the conformational changes that accompany the dissociation of eIF2-GDP from the 48S complex and subsequent binding of eIF5B, which is also relevant to understanding canonical eukaryotic translation initiation.

Overall, this study is a valuable addition to the field of IRES-mediated and canonical translation initiation. For the most part, the data presented support the structural claims regarding the various conformational transitions that occur during the initiation process, and I would like to thank the authors for providing the necessary structural data (maps and coordinates) to facilitate the evaluation of their work. While I have no major comments regarding the latter parts of the study (when translation factors are involved), I nevertheless feel that the structures related to the early steps of the process (i.e., IRES binding prior to mRNA loading into the channel) are sometimes over-interpreted given the resolution and map quality for the regions described. In particular, I have some doubts about the modeling of certain features involved in the initial steps of IRES binding, especially in the 1ΔdII structure, as detailed in my comments below:

- Line 194-196: While it is clear that interactions between ES7 residues U1114-U1115 and IRES residues A296 and A136, respectively, are not yet formed in the 1ΔdII structure, the map provided by the authors seems to indicate that base-pairing between U1116-U1118 and G266-268 is already present, as clear density indicating the formation of a duplex between ES7 and IIId can be seen (especially when filtering the map at lower resolution to reveal the groove of the helix). The authors should therefore fix the model for this region and revise the text accordingly. A possibility would be to start from the conformation seen in 6ΔdII or earlier structures and refine the coordinates in Isolde with base pairs restrained, at least initially.

- Lines 199-201: "Structure 1ΔdII also shows another transient, previously undescribed hydrogen bond between U265 in domain IIId and Lys199 in eS1, a residue that instead interacts with IIIe in the fully bound IRES (Quade et al, 2015; Yamamoto et al, 2015)." The resolution of the map in this region does not allow such a statement to be made, even though it is clear that the two molecules come into contact.

- Line 211: there is no density to suggest the existence of a hydrogen bond between Glu75 and residues 266-267 in structure 2ΔdII even though some sort of contact occurs in this area.

- Lines 227-233: As mentioned above, I would argue that docking of domain IIId onto ES7 is present in all of the structures, including structure 1ΔdII. The flexibility of ES7 nevertheless enables the various conformations observed between 1ΔdII and 6ΔdII to be populated.

Additional minor points:

- Line 192: to improve readability, "The fully bound IRES forms Watson-Crick base pairs between ES7 nt U1114-1118 and IRES domains IIId (GGG266-268), IIIe (A296), and hIII1 (A136)..." should be replaced with "The fully bound IRES forms Watson-Crick base pairs between ES7 nt U1114, U1115 and C1116-1118, and IRES domains IIIe (A296), hIII1 (A136) and IIId (GGG266-268), respectively..."

- Line 198: G296 should be G295

- Lines 201-203: "These two contacts along with the domain IIIa/IIIc interactions with eS27 are the only bonds between the IRES and the 40S subunit in this complex (Table S4)." This statement is incomplete since U1116-U1118 and G266-268 are already formed.

- Lines 206: A296 contacts U1115 but, despite the limited map quality in this area, does not appear to do so via a stacking interaction. The text should be revised accordingly.

- Lines 206 and 217: the G268-C1116 base pair is already formed in the 2ΔdII structure and should be mentioned.

- Line 219: Contrary to what the text says, the A296/U115 interaction is modeled in the structure 4ΔdII provided, but the map may need to be filtered to lower resolution to better understand what is going on.

Referee #2:

The authors report cryo-EM structures of HCV IRES bound with the 40S subunit with initiation factors eIF1A, eIF2 and eIF5B and initiator tRNA to represent 43S and 48S initiation complexes. Comparison of cryo-EM states with mutant IRES (with domain 2 deleted) and wild-type IRES emphasizes a key role of domain 2 in stabilizing the IRES mRNA and 40S subunit conformation. These structures allow a more complete view of the HCV-IRES-driven translation initiation. The manuscript is well written and will be suitable for publication after addressing several issues, including interpretation of density in structure 14wt and the proposed linear mechanism, along with minor comments:

1. Show density for the base-paired residues described: ES7 nt U1114-1118 and IRES domains IIId (GGG266-268), IIIe (A296), and hIII1 (A136) as well as a stacking interaction between domain IIIe (G295) and U1115 of ES7.
2. Other stabilizing interactions that are described in the text (e.g. U265) are not shown. Include density for these interactions in Supplementary Figures.
3. Codon-anticodon (unpaired) in structure 14wt appear incorrectly modeled (with base 3 nearly perpendicular to the bases it is expected to stack on) and density does not seem to be sufficient to distinguish between purine and pyrimidine bases (Fig. S3G right panel). The conservative interpretation of this density is the paired codon-anticodon. To distinguish between different fits (shifted vs. paired codon), authors could fit and refine structural models for the two alternative frames, then calculate the local fits (CC) of the codon-anticodon helix for each scenario.
4. Remove the first period symbol in: "the tRNA acceptor stem binding domain IV. (Figure 5C)."
5. Some sentences are too long and/or confusing. Please proofread the paper and simplify those occurrences. Here is a couple of examples:
a) "In contrast, when interacting with eIF5B, tRNA establishes these contacts only in the closed position of the head, whereas in the open pre-48S ICs tRNA is stabilized by contacts between eIF5B domain IV and the acceptor stem, tRNA anticodon U35 and mRNA A(+1), and a single 18S rRNA nucleotide (C1701) that partially stacks with C33 (instead of the stacking of both U1248/C1701 as in the closed case) (Figure 6B)."
b) Confusing sentence: "These observations suggest that the 60S subunit might stimulate the interaction between A415 and the β13-β14 loop to reposition this loop away from switch 1 so that a transition from GTP- to GDP-bound conformation."
6. In discussion and Fig. 7, the authors present the specific order of events in a linear pathway (domain III binds first). But it's not clear how this order can be deduced based on wt and mutant HCV IRES. Although some interesting observations can be made using the mutant (domain-II deletion) IRES structures, these structures should be interpreted with caution: they cannot report on the binding sequence of wt-IRES if domain II of the latter normally binds first.
The alternative mechanism, in which domain II binds first, is more likely. Note that the head of free 40S head is dynamic and the opened state is sampled without the IRES. The IRES binding via domain II to the open conformation would enrich the 40S*IRES open state, as observed in the authors' data. Furthermore, no closed 40S (intermediates) were captured with the wt IRES (which could support the domain-III-first sequence), and no "inducing the head of the ribosome to open" is required with wt IRES. It is also possible that the two pathways (domain II or domain III first) are sampled in parallel. Discuss this alternative pathway and include it in the illustration.

Referee #3:

Brown et al. investigated the discrete stages of HCV IRES-mediated translation initiation on the mammalian 40S small ribosomal subunit by cryo-EM. For this, they used in vitro reconstituted initiation complexes (ICs) that contain initiator-tRNA, initiation factor eIF1A, IF3, and either eIF2 or eIF5B in combination with wild-type (wt) HCV IRES or a domain II deletion mutant (delta-dII) HCV IRES. From the four samples, the authors are able to classify for and reconstruct an entire series of cryo-EM maps, 18 of which are interpreted as atomic structures. Based on the presence of bound factors, observed conformational changes, published biochemical data and existing structural evidence, the reconstructions are attributed to discrete initiation stages that include initial binding of HCV IRES, followed by 40S head opening and sequential positioning of the IRES-mRNA into the P- E- and A site of the mRNA channel. Corresponding to later stages of initiation, the authors identify 40S-HCV IRES structures that are similar to canonical IF2-mediated 48S ICs and suggest, based on exclusively structural evidence, a pathway that includes non-canonical eIF5B-mediated loading of the initiator tRNA onto the 40S via a pre-48S IC intermediate.

Although the manuscript represents a comprehensive structural overview of the HCV IRES-mediated initiation stages, the work is not suitable for publication in EMBO without extensive revision. This is due to the overall form of the presentation, the lack of supporting functional data for the purely structural interpretations, and because the quality of the EM data does not sufficiently support many details of the structural interpretation.

While the introduction of the manuscript is very well written and the results concisely summarized in the discussion, the results section is overwhelming due to confusing argumentation and poorly prepared and overloaded main and supplementary figures. Rather than including all data, the authors should have more carefully chosen which structural interpretations are functionally important and are actually supported by the experimental maps. Another concern is the balance between the length of sections addressing novel findings and those that recapitulate established concepts. To be publishable in EMBO, the results section of the manuscript would have to be rewritten and considerably shortened.

Specific comments

Overview of cryo-EM analysis...

Major points:
* The authors reconstitute in vitro the four complexes used for cryo-EM analysis in the presence of eIF3, however in none of the structures eIF3 is visible, although known to bind at the periphery of the HCV IRES. The authors argue that this is due to grid preparation and does not affect the data interpretation because all complexes can be assembled without eIF3. However, how can they be sure that this does not influence the different conformations of the HCV IRES RNA described below and the order in which they occur?
* In the refinement Table S3, the clashscore values reported for most structures are unusually high (>10), which is indicative of poor local model building, while the remaining geometric values are reasonable due to the imposed restraints during refinement. With modern building/refinement software and high-resolution models at hand, the model quality should be locally checked and improved, rather than just providing poorly refined rigid body docking models. Due to this, the contact details between the 40S and the HCV IRES listed in Table S4 are not very meaningful. These should be updated accordingly after improvement of the models.

Minor points:
* In Figure 1B, all 18 structures are shown. This panel is highly redundant, especially for structures 1-9, which are closely related. Showing the major structures in larger size instead would be more informative.
* In Figure 1B eIF2 and eIF5 are shown with identical color, which is confusing at a first glance.

Stepwise binding of the IRES...

Major points:
* In this section, the authors describe numerous alternative conformations observed for the 40S-bound delta-dII HCV IRES mutant and suggest an order of events for HCV docking to the 40S that may also occur in the wt IRES (structures #1 - #9). This interpretation, including the proposed series of events, is not valid because such conformations are never observed for the wt IRES where domain II is present. This data does not add to the story and should not be included. Furthermore, at the obtained >4A resolutions an atomic interpretation seems difficult, as exemplified by structure #5, where HCV IRES residues 265-267 that contact the 40S are obviously not reasonably base-pairing with the 18S rRNA residues 1117-1119.

Accommodation of the IRES in the mRNA-binding channel
Major points:
* Here the authors show how the IRES mRNA becomes loaded in the mRNA binding cleft upon HCV domain 2 and eIF1A ternary complex binding (structures #9 - #11). Whereas the general trend is indeed revealed by the EM density at lower local resolution, the details depicted in Fig. 3I are an over-interpretation: Due to poor resolution, eIF1A can only be rigid-body docked, and the stacking of W70 on the mRNA residue 345 and the rRNA decoding base 1825 are not clearly visible in the EM density.

Furthermore, the connection between the IRES and the mRNAs is poorly built (structure #11, IRES residues 335-343).

The structure of eIF2-containing 48S complexes assembled on the HCV IRES

Major points:
* As in the ternary complex described in the previous paragraph (structure #11), eIF1A and the surrounding areas are not resolved at atomic resolution in structure #12, and therefore the details shown in panels 4E & 4F and the corresponding text (P12, lines 361-364) are not representing the experimental observations (although in structure #12 the codon-anticodon interactions are well resolved). For all atomic details shown, clear corresponding density should be visible. For panels 4G and 4H (reorganization of contacts between eS28 and the mRNA), unfortunately the depicted area is only resolved in the ternary complex (structure #11), while partially disordered in the 48S IC (structure #12), which precludes a direct comparison.
* In the last paragraph (P13, lines 378-386), the authors interpret a 40S complex with bound HCV IRES and initiator tRNA but lacking all initiation factors (structure #13) as a mimic of "an intermediate state after eIF2 dissociation and prior to the binding of eIF5B." However, in the same paragraph potential technical problems or spontaneous occurrence of GTP hydrolysis are discussed. As the physiological relevance of this intermediate state is even questioned by the authors themselves why is it prominently presented?

Minor points:
* In panel 4B, the residue numbering of eIF2alpha is shifted relative to the numbering in the PDB. Please correct the coordinates or the labels of the residues.

The structure of eIF5B-containing 48S complexes...

Major points:
* This section includes the most unexpected result of the manuscript, describing the structure of non-canonical eIF5B-bound 48S initiation complexes in pre-48S and closed states in the context of the HCV IRES, prior to 60S subunit joining. In these complexes, the initiator tRNA is held in place by domain IV of eIF5B, which is bound to the 40S in GTP conformation. However, as before the level of detail in the structural interpretation needs to be reconsidered in the context of the EM map quality, which locally varies a lot.
* The authors should provide some biochemical evidence that this non-canonical intermediate is not an artefact of in vitro complex formation.
* The major concern resides in the interpretation of the pre-48S, where the authors claim that the initiator tRNA bound to the P-site is not base-pairing with the start codon but shifted by one residue (structure #14, depicted in Figures 5H, 6B and S3G). The corresponding EM density does not support this interpretation, it is rather consistent with a conformation in which the start codon interacts with the anticodon of the tRNA in a cognate manner and not via the modelled mismatch, although the tRNA may still not be in fully accommodated conformation compared to the closed eIF5B 48S IC. The claim that the start codon is shifted cannot be made considering that the residues flanking the start codon upstream and downstream are poorly resolved it is not possible to assign register of the mRNA. Further, in the current structure, all three residues of the codon are built in highly unfavorable orientations and do not explain the observed density in a satisfying manner.
* The authors discuss the details of interactions between eIF1A and domain IV of eIF5B (depicted in Figure 5D). However, as these are poorly resolved areas in both structures (#14 & #15), analysis at such detail is not possible.

Minor points:
* P14, line 401: "14w" should read "14wt"
* P15, lines 429-431: the sentence "These observations ... to GTP-bound conformation." is incomplete.
* P16, lines 472-475: Figure 5G should be mentioned somewhere.
* P16, line 488: "Figure 5F" should probably read "Figure 5G".
* Figure 5 contains too many panels. Especially panel 5F is too small and difficult to understand. Panels 5A & 5B do not convey the main message in spite of their large size. Additional labels and a more plastic 3D representation of the cartoon may help the reader.
* In both eIF5B-bound 48S IC structures, the GTP in the active site of the GTPase domain of eIF5B is in a distorted conformation. More care should be taken to model and refine the nucleotide and to keep it correctly coordinated to the active site residues and the magnesium ion. For this, high resolution templates are available in the PDB.

Comparison of eIF5B-containing pre-48S initiation complexes with eIF2-containing 43S scanning complexes
* Although a comparison with existing 43S structures is obvious, this paragraph is excessively long and considering the structural issues mentioned above, the text may be shortened and the corresponding Figure 6 should be revised.

**We are very grateful to the reviewers for their careful reading of our manuscript and for their insightful comments and suggestions.**

## Reviewer #1

### Major points:

**1.** Line 194-196: While it is clear that interactions between ES7 residues U1114-U1115 and IRES residues A296 and A136, respectively, are not yet formed in the 1ΔdII structure, the map provided by the authors seems to indicate that base-pairing between U1116-U1118 and G266-268 is already present, as clear density indicating the formation of a duplex between ES7 and IIId can be seen (especially when filtering the map at lower resolution to reveal the groove of the helix). The authors should therefore fix the model for this region and revise the text accordingly. A possibility would be to start from the conformation seen in 6ΔdII or earlier structures and refine the coordinates in Isolde with base pairs restrained, at least initially.

We thank the reviewer for the suggestion to remodel this region. After closer examination, we agree that there is an interaction between IRES domain IIId and ES7 in structure $1_{\Delta dII}$. The fully bound IRES, described in Quade *et al,* (2015) and Yamamoto *et al,* (2015), has three Watson-Crick base-pairs between IRES domain IIId $G_{266}$-$G_{268}$ and ES7 $C_{1116}$-$C_{1118}$. However, after remodeling domain IIId in structure $1_{\Delta dII}$, we observed that only two of these three interactions had formed ($C_{1116}$-$C_{1117}$ and $G_{267}$-$G_{268}$). This was confirmed by running multiple refinements using PHENIX both with and without restraints and in all cases the third interaction ($C_{1118}$/$G_{266}$) was not formed. All three contacts between domain IIId and ES7 are established in structure $2_{\Delta dII}$ and all following complexes.

We have updated the text (Lines 203-204 of the revised manuscript): " Initially, structure $1_{\Delta dII}$ showed only two of the six canonical interactions with ES7, namely the base-pairs between domain IIId ($G_{267\text{-}268}$) and ES7 nt. $C_{1116\text{-}1117}$."

**2.** Lines 199-201: "Structure 1ΔdII also shows another transient, previously undescribed hydrogen bond between U265 in domain IIId and Lys199 in eS1, a residue that instead interacts with IIIe in the fully bound IRES (Quade et al, 2015; Yamamoto et al, 2015)." The resolution of the map in this region does not allow such a statement to be made, even though it is clear that the two molecules come into contact.

We appreciate the improvement offered here. Although the canonical interaction (see Quade 2015 & Yamamoto 2015) of eS1 α helix 6 Lys199 and domain IIIe is not present in structure $1_{\Delta dII}$, there is clear density for positioning α helix 6 next to IRES domain IIId. We have rewritten this section to make it clear what the new interaction is, but at the same time presented this interaction in a more conservative manner.

We have updated the text (Lines 204-207 of the revised manuscript): " The position of domain IIId in structure $1_{\Delta dII}$ also allows some interaction between this domain and α-helix 6 in eS1, a region that contains Lys199, a residue that interacts with domain IIIe in the fully-bound IRES (Quade *et al,* 2015; Yamamoto *et al,* 2015).", and table S4 to include these interactions. Density corresponding to domain IIId and its interacting ribosomal elements is now presented in Figure S3B.

**3.** Line 211: there is no density to suggest the existence of a hydrogen bond between Glu75 and residues 266-267 in structure 2ΔdII even though some sort of contact occurs in this area.

This statement was eliminated from the revised manuscript.

**4.** Lines 227-233: As mentioned above, I would argue that docking of domain IIId onto ES7 is present in all of the structures, including structure 1ΔdII. The flexibility of ES7 nevertheless enables the various conformations observed between 1ΔdII and 6ΔdII to be populated.

As we stated in the response to the reviewer's first comment, the reviewer is correct that docking of domain IIId onto ES7 is present in all structures, including structure 1ΔdII. However, in contrast to the following structures, structure 1ΔdII showed only two interactions between IIId and ES7 (between G267-G268 and C1116-C1117). Although we feel strongly that structures 1ΔdII-6ΔdII represent a likely sequence of binding events during association of the HCV IRES with the 40S subunit, we nevertheless understand that there is no definite proof for this. Thus, we consider it wholly appropriate to present our view as a suggestion.

**Minor points:**

- Line 192: to improve readability, "The fully bound IRES forms Watson-Crick base pairs between ES7 nt U1114-1118 and IRES domains IIId (GGG266-268), IIIe (A296), and hIII1 (A136)"…" should be replaced with "The fully bound IRES forms Watson-Crick base pairs between ES7 nt U1114, U1115 and C1116-1118, and IRES domains IIIe (A296), hIII1 (A136) and IIId (GGG266-268), respectively".."

This sentence has been rewritten (Lines 193-197 of the revised manuscript): "The IRES fully bound to 80S ribosomes forms six interactions with ES7: four Watson-Crick base pairs between ES7 nt. U1115 and C1116-1118 and IRES domains hIII1 (A136) and IIId (GGG266-268), respectively; a reverse Hoogsteen base-pair between ES7 nt. U1114 and domain IIIe (A296); and a stacking interaction between ES7 nt. U1114 and domain IIIe (G295) (Quade et al, 2015; Yamamoto et al, 2015)."

- Line 198: G296 should be G295.

This change has been made.

- Lines 201-203: "These two contacts along with the domain IIIa/IIIc interactions with eS27 are the only bonds between the IRES and the 40S subunit in this complex (Table S4)." This statement is incomplete since U1116-U1118 and G266-268 are already formed.

This sentence has been modified (Lines 207-209 of the revised manuscript): " These contacts between ES7 and domain IIId, along with the domain IIIa/IIIc interactions with eS27 are the only bonds between the IRES and the 40S subunit in this complex."

This sentence is now preceded by a description of base-pairing between domain IIId ($G_{267-268}$) and ES7 nt. $C_{1116-1117}$ and this interaction is therefore mentioned in a summarized form.

- Lines 206: A296 contacts U1115 but, despite the limited map quality in this area, does not appear to do so via a stacking interaction. The text should be revised accordingly.

This entire section describing structures $1_{\Delta dII}$-$6_{\Delta dII}$ was rewritten and this statement was eliminated from the text.

- Lines 206 and 217: the G268-C1116 base pair is already formed in the 2ΔdII structure and should be mentioned.

The establishment of the G268-C1116 base pair in the 2ΔdII structure is now clearly stated in the revised text (Lines 209-211): " Structure $2_{\Delta dII}$ maintains all of these contacts but repositioning of domain IIId by 4.4 Å relative to structure $1_{\Delta dII}$ allows formation of the third Watson-Crick base-pair, between ES7 ($C_{1118}$) and domain IIId ($G_{266}$)."

- Line 219: Contrary to what the text says, the A296/U115 interaction is modeled in the structure 4ΔdII provided, but the map may need to be filtered to lower resolution to better understand what is going on.

The reviewer is correct - the model has been revised, and the text was modified accordingly (Lines 215-216 of the revised manuscript): " The subsequent complex, structure $4_{\Delta dII}$, maintains all previous contacts and forms the fifth interaction with ES7, between domain IIIe ($A_{296}$) and ES7 ($U_{1114}$)."

## Reviewer #2

**1.** Show density for the base-paired residues described: ES7 nt U1114-1118 and IRES domains IIId (GGG266-268), IIIe (A296), and hIII1 (A136) as well as a stacking interaction between domain IIIe (G295) and U1115 of ES7.

These densities are now shown in Figure S3B-D of the revised manuscript.

**2.** Other stabilizing interactions that are described in the text (e.g. U265) are not shown. Include density for these interactions in Supplementary Figures.

Following the suggestions from reviewer #1, comment 1, we remodeled the interaction between IRES domain IIId and 18S rRNA expansion segment (ES) 7 which led us to reinterpret them and to revise the text relating to interactions between the IRES and ES7. Now we report no novel interactions and instead relate the 'building up' of the set of canonical interactions reported by Quade et al, (2015) and Yamamoto et al, (2015).

**3.** Codon-anticodon (unpaired) in structure 14wt appear incorrectly modeled (with base 3 nearly perpendicular to the bases it is expected to stack on) and density does not seem to be sufficient to distinguish between purine and pyrimidine bases (Fig. S3G right panel). The conservative interpretation of this density is the paired codon-anticodon. To distinguish between different fits (shifted vs. paired codon), authors could fit and refine structural models for the two alternative frames, then calculate the local fits (CC) of the codon-anticodon helix for each scenario.

Both reviewer #2 (this comment) and reviewer #3 raised similar questions about structure 14wt and the assignment of nucleotides in the P site. We are very grateful to the reviewers for raising these concerns as they led us to reevaluate the data, thereby improving the modelling in the final structure 14wt complex completely in line with the suggestions made by these reviewers, and correcting a significant interpretational error. The P site triplet of structure 14wt is now the A(+1), U(+2), G(+3) codon as they suggested.

To clarify the identity of the P-site codon we examined all available maps that showed the eIF5B containing pre-48S complex. These maps were: structure 14wt (3.9 Å, 60,578 particles), and the related structure 44wt (4.2 Å, 29,072 particles) showing eIF5B in a slightly different position on the 40S subunit intersubunit interface, as well as the consensus map structure 38wt (3.5 Å, 199,047 particles) that structure 14wt and structure 44wt were derived from (See Figure S1F). Even though in the higher resolution consensus map the identity of the P site codon is ambiguous, we noticed that the upstream G(-7), U(-6), G(-5) region was modelled incorrectly (see Figure R1 below). By correcting this region the mRNA register is changed by 1 and so the P site is now occupied by A(+1), U(+2), G(+3). Examination of the unsharpened maps of structure 14wt, 38wt, and 44wt also confirm the placement of these nucleotides in the new position.

The text corresponding to discussion of structure 14wt in the original sections " The structure of eIF5B-containing 48S complexes assembled on the IRES" and " Comparison of eIF5B-containing pre-48S initiation complexes with eIF2-containing scanning 43S complexes", as well as the new section "Dynamics of ribosomal contacts of mRNA along the mRNA-binding channel throughout the initiation process" was revised accordingly.

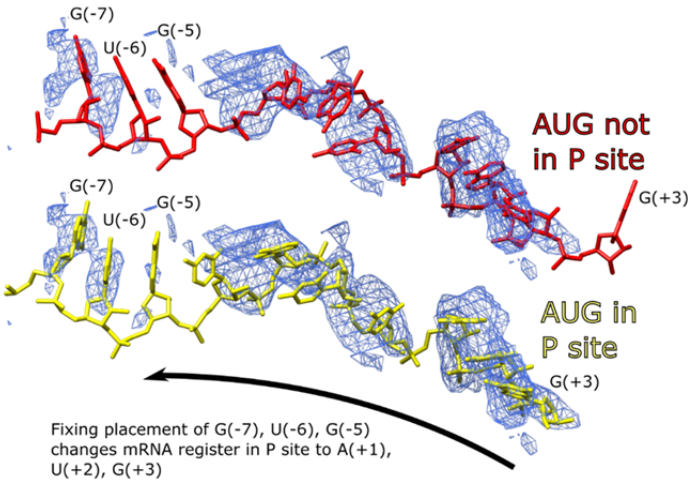


**Figure R1.** Difference in register of mRNA depending on the modelling of upstream nucleotides.

**4.** Remove the first period symbol in: "the tRNA acceptor stem binding domain IV. (Figure 5C)."

This has been done - note that the original Figure 5C became (Figure EV4C) in the revised manuscript.

**5.** Some sentences are too long and/or confusing. Please proofread the paper and simplify those occurrences. Here is a couple of examples:

a) "In contrast, when interacting with eIF5B, tRNA establishes these contacts only in the closed position of the head, whereas in the open pre-48S ICs tRNA is stabilized by contacts between eIF5B domain IV and the acceptor stem, tRNA anticodon U35 and mRNA A(+1), and a single 18S rRNA nucleotide (C1701) that partially stacks with C33 (instead of the stacking of both U1248/C1701 as in the closed case) (Figure 6B)."

This statement has been clarified and rewritten (Lines 511-519 of the revised manuscript): "In contrast, in the open pre-48S IC, tRNA is stabilized by contacts between eIF5B domain IV and the acceptor stem, codon-anticodon base-pairing, and by stacking of the aromatic rings of ASL $C_{33}$ with $C_{1701}$ of 18S rRNA (Figure 7B, left panel). Interestingly, in the scanning 43S complex, the $C_{33}$ ribose instead forms a lone pair-π interaction with $U_{1248}$ of 18S rRNA (Figure 7B, middle panel), whereas in canonical 48S ICs (e.g. Simonetti et al., 2020), as well as in eIF5B-containing 48S ICs (structure $15_{wt}$), $C_{33}$ is involved in both interactions (Figure 7B, right panel). Thus, whereas some aspects of eIF5B-containing pre-48S complexes are analogous to those of eIF2-containing scanning 43S complexes, the overall orientation and the specific interactions of tRNA in them differ."

b) Confusing sentence: "These observations suggest that the 60S subunit might stimulate the interaction between A415 and the β13-β14 loop to reposition this loop away from switch 1 so that a transition from GTP- to GDP-bound conformation."

This sentence has been completed (Lines 401-404 of the revised manuscript): " These observations suggest that the 60S subunit might stimulate the interaction between $A_{415}$ and the β13-β14 loop to reposition this loop away from switch 1 so that a transition from the GTP- to the GDP-bound conformation is possible."

6. In discussion and Fig. 7, the authors present the specific order of events in a linear pathway (domain III binds first). But it's not clear how this order can be deduced based on wt and mutant HCV IRES. Although some interesting observations can be made using the mutant (domain-II deletion) IRES

structures, these structures should be interpreted with caution: they cannot report on the binding sequence of wt-IRES if domain II of the latter normally binds first.

The alternative mechanism, in which domain II binds first, is more likely. Note that the head of free 40S head is dynamic and the opened state is sampled without the IRES. The IRES binding via domain II to the open conformation would enrich the 40S*IRES open state, as observed in the authors' data. Furthermore, no closed 40S (intermediates) were captured with the wt IRES (which could support the domain-III-first sequence), and no "inducing the head of the ribosome to open" is required with wt IRES. It is also possible that the two pathways (domain II or domain III first) are sampled in parallel. Discuss this alternative pathway and include it in the illustration.

We thank the reviewer for raising these questions and agree that cryo-EM data alone are not sufficient to build such a pathway with an absolute certainty. Thus, we ordered structures 1ΔdII-6ΔdII based on the number of IRES-40S subunit contacts and displacement from the position of the canonically-bound IRES and appropriately presented our view merely as a suggestion.

Although we appreciate the possibility of the alternative mechanism of binding, in which domain II binds first, there is no experimental evidence supporting it. Importantly, the presence or absence of domain II does not affect the binding affinity of the IRES to the 40S subunit: Kieft et al, (2001) (PMID: 11233977) and Ji, Doudna et al (PMID: 15563596) both reported that the 40S subunit binding affinities of the wt IRES and the ΔdII IRES are nearly identical (Kd = 1.9 nM and 3.1 nM, respectively), indicating that domain II does not play a role in binding of the IRES to the 40S subunit. In contrast, removing of linked domains IIIa/IIIc, which in our structures contact eS27 via nt. 163 and 233 and undergo minimal structural change regardless of the conformation of the rest of the IRES, completely abrogates binding. For these reasons we respectfully decline the suggestion to discuss the domain II originated mechanism as a viable alternative.

## Reviewer #3

**Overview of cryo-EM analysis**

**Major points:**

* The authors reconstitute in vitro the four complexes used for cryo-EM analysis in the presence of eIF3, however in none of the structures eIF3 is visible, although known to bind at the periphery of the HCV IRES. The authors argue that this is due to grid preparation and does not affect the data interpretation because all complexes can be assembled without eIF3. However, how can they be sure that this does not influence the different conformations of the HCV IRES RNA described below and the order in which they occur?

Chemical/enzymatic probing showed that the structure of HCV and related IRESs is not altered by binding of eIF3 to the apex of domain III (Sizova et al., 1998; PMID 9573242). Moreover, the structure of ribosomal complexes assembled on the HCV-like CSFV IRES containing bound eIF3 did not reveal any specific conformational changes in the complex that could be attributed to the presence of eIF3 (Hashem et al., 2013). eIF3 is not essential for initiation on the HCV IRES and could have been omitted from the reconstitution mixtures but was included to make reaction mixtures more closely resembling the native process. We therefore consider that the trade-off between high resolution and the presence of eIF3 in the complexes is appropriate and justifiable. We have clarified the explanation of these points in lines 142-151 of the revised text.

Line 144-154: None of the structures obtained contained eIF3. During initiation on HCV-like IRESs, eIF3 interacts with the apical region of IRES domain III rather than with the 40S subunit (Hashem et al, 2013). This interaction is sensitive to the process of grid preparation and is more stable when grids have thicker ice so that imaging complexes that contain eIF3 requires the intentional selection of regions with sufficiently thick ice (e.g., Hashem et al, 2013; Neupane et al, 2020); however, our study aimed to determine the details of ribosomal interactions with the IRES, initiation factors and Met-tRNAiMet at high resolution, which relies on imaging in regions with thinner ice. Importantly, however, the absence of eIF3 does not affect data interpretation because all studied complexes can be assembled efficiently without eIF3 (Pestova et al, 1998b; 2008) and because the structure of the IRES is not affected by binding of eIF3 (Sizova et al., 1998)."

* In the refinement Table S3, the clashscore values reported for most structures are unusually high (>10), which is indicative of poor local model building, while the remaining geometric values are reasonable due to the imposed restraints during refinement. With modern building/refinement software and high-resolution models at hand, the model quality should be locally checked and improved, rather than just providing poorly refined rigid body docking models. Due to this, the contact details between the 40S and the HCV IRES listed in Table S4 are not very meaningful. These should be updated accordingly after improvement of the models.

This was a critically important observation and we are very grateful for the time and effort of the reviewer in bringing this to our attention. We reprocessed all models using PHENIX and were able to improve the clashscore for all models from an average of 14.4 to 5.0 (see following table R2). The previous high clashscores were caused by a setting in PHENIX that favored improving geometry over minimizing clashes. The previous models and current models only show very minor differences in geometry and have a minimal RMSD difference of 0.5 Å (see Table R2). To reflect the new statistics, we have updated Table S3 in its entirety.

For interest, the cause of the high clashscores was an update to PHENIX that reduced the default value for the *nonbonded_weight* setting from 1000 to 100. Interestingly, the patch notes do not include any information about why this change was made and other researchers have noted that reverting back to the previous default value improves the clashscore during refinement. See link: http://www.phenix-online.org/pipermail/phenixbb/2019-October/024432.html

Table R2. Previous and current clashscore for each complex; RMSD between previous and current models.

| | Clashscore | | |
|---|---|---|---|
| **Structure** | **Previous** | **Current** | **RMSD** |
| 1ΔdII | 13.73 | 5.12 | 0.653 |
| 2ΔdII | 15.52 | 5.13 | 0.504 |
| 3ΔdII | 16.03 | 5.7 | 0.486 |
| 4ΔdII | 17.31 | 5.37 | 0.462 |
| 5ΔdII | 15.7 | 5.27 | 0.525 |
| 6ΔdII | 17.19 | 6.16 | 0.51 |
| 7ΔdII | 14.81 | 5.28 | 0.491 |
| 8ΔdII | 13.84 | 9.66 | 0.444 |
| 9ΔdII | 20.63 | 5.43 | 0.575 |
| 10wt | 13.03 | 4.63 | 0.439 |
| 11wt | 11.99 | 4.35 | 0.393 |
| 12wt | 8.81 | 3.73 | 0.691 |
| 12ΔdII | 9.32 | 3.53 | 0.506 |
| 13wt | 14.31 | 4.83 | 0.939 |
| 13ΔdII | 25.59 | 4.35 | 0.521 |
| 14wt | 14.01 | 4.43 | 0.644 |
| 15wt | 8.88 | 3.89 | 0.569 |
| 15ΔdII | 9.41 | 4.26 | 0.361 |
| Average | 14.45 | 5.06 | 0.54 |

**Minor points:**

* In Figure 1B, all 18 structures are shown. This panel is highly redundant, especially for structures 1-9, which are closely related. Showing the major structures in larger size instead would be more informative.

Figure 1 has been redesigned and now includes only structures 1ΔdII, 10wt, 11wt, 12wt, 14wt and 15wt. All other structures are now shown in Figure EV1.

* In Figure 1B eIF2 and eIF5 are shown with identical color, which is confusing at a first glance.

We thank the reviewer for this suggestion. We have changed the color of eIF5B (to cyan) in Figure 1.

**Stepwise binding of the IRES...**

**Major points:**

* In this section, the authors describe numerous alternative conformations observed for the 40S-bound delta-dII HCV IRES mutant and suggest an order of events for HCV docking to the 40S that may also occur in the wt IRES (structures #1 - #9). This interpretation, including the proposed series of events, is not valid because such conformations are never observed for the wt IRES where domain II is present. This data does not add to the story and should not be included. Furthermore, at the obtained >4A resolutions an atomic interpretation seems difficult, as exemplified by structure #5, where HCV IRES residues 265-267 that contact the 40S are obviously not reasonably base-pairing with the 18S rRNA residues 1117-1119.

This comment is related to the analogous question from Reviewer #2, and as in our response to that question, we agree that cryo-EM data alone are not sufficient to confidently build such a pathway. The rationale for the suggested pathway is clearly explained in the text: we ordered structures 1ΔdII-6ΔdII based on the number of IRES-40S subunit contacts and displacement from the position of the canonically-bound IRES, and ordered structures 6ΔdII-7ΔdII, which have a full complement of IRES/40S contacts, on the basis of conformational changes in the head, from closed to fully open. Our model is hypothetical, and we consider that we have appropriately presented it as such.

We note that none of the individual components in our models differ radically from every available prior structures with respect to IRES residues 265-267 and 18S rRNA 1117-1119: the local ribosomal environment and the conformation of domain IIId match the published literature. The unique aspect of our report is their disposition and separation relative to each other. With respect to these specific nucleotides: they are located within hydrogen bonding distance and the positions of domain IIId and ES7 are very clear in the unsharpened map. To aid interpretation, we have uploaded the unsharpened maps for all structures to the EMDB as well as made them available to the reviewers.

**Accommodation of the IRES in the mRNA-binding channel**

To clarify the changing interactions of the viral mRNA in the mRNA-binding cleft during initiation and to highlight the roles of eIF1A and IRES domain II in this process, we have now presented the relevant data as a free-standing section “Dynamics of ribosomal contacts of mRNA along the mRNA-binding channel throughout the initiation process” (lines 439-484) and have expanded the former Figure 3C as the new Figure 6.

**Major points:**

* Here the authors show how the IRES mRNA becomes loaded in the mRNA binding cleft upon HCV domain 2 and eIF1A ternary complex binding (structures #9 - #11). Whereas the general trend is indeed revealed by the EM density at lower local resolution, the details depicted in Fig. 3I are an over-interpretation: Due to poor resolution, eIF1A can only be rigid-body docked, and the stacking of W70 on the mRNA residue 345 and the rRNA decoding base 1825 are not clearly visible in the EM density. Furthermore, the connection between the IRES and the mRNAs is poorly built (structure #11, IRES residues 335-343).

We thank the reviewer for this suggestion. We performed extensive additional modelling to address these comments regarding the clashscore for all models and to improve the register of nucleotides for the pre-48S IC. As a result, we prepared completely new models of mRNA covering the region 335-343.

Regarding the interpretation in Figure 3I (now renumbered as Figure 3C): for a more conservative interpretation, we removed the A(+4) nucleotide from the figure as its placement is not clearly defined. The position and placement of W70 and A1825 are however very clear despite the lower

occupancy of eIF1A in this map. As a consistency check we rigid body fitted the entire model for another 48S IC (PDB:6YAL from Simonetti et al, 2020) and based on alignment against the ribosome alone, W70 and A1825 were positioned within clear density present in structure 11wt. For this reason we have retained W70 and A1825 (with annotation) in Figure 3C, lower panel.

**The structure of eIF2-containing 48S complexes assembled on the HCV IRES**
**Major points:**

* As in the ternary complex described in the previous paragraph (structure #11), eIF1A and the surrounding areas are not resolved at atomic resolution in structure #12, and therefore the details shown in panels 4E & 4F and the corresponding text (p12, lines 361-364) are not representing the experimental observations (although in structure #12 the codon-anticodon interactions are well resolved). For all atomic details shown, clear corresponding density should be visible. For panels 4G and 4H (reorganization of contacts between eS28 and the mRNA), unfortunately the depicted area is only resolved in the ternary complex (structure #11), while partially disordered in the 48S IC (structure #12), which precludes a direct comparison.

We thank the reviewer for this comment. We endeavored to be conservative in our analysis and noted that the rigid body fitting of our refined eIF1A models matched the structure of eIF1A from canonical cap dependent 48S ICs (e.g., PDB:6YAL from Simonetti et al, 2020) as the eIF1A model is a good fit in the unsharpened maps. However, since the occupancy of eIF1A in structure 12wt is low, upon map sharpening the region becomes disordered and analysis (and assessment) more difficult. For this reason we have removed those figures and the text that mentions the formation of a stacking triple for structure 12wt.

* In the last paragraph (P13, lines 378-386), the authors interpret a 40S complex with bound HCV IRES and initiator tRNA but lacking all initiation factors (structure #13) as a mimic of "an intermediate state after eIF2 dissociation and prior to the binding of eIF5B." However, in the same paragraph potential technical problems or spontaneous occurrence of GTP hydrolysis are discussed. As the physiological relevance of this intermediate state is even questioned by the authors themselves why is it prominently presented?

Structure 13wt can only have been formed via eIF2-mediated delivery of initiation tRNA, because domain II is not present in the ribosomal E site and eIF2 is responsible for its displacement from this site. This structure and the related structure 13ΔdII are therefore of interest, even if the reason for dissociation of factors is not established. These two structures are described cautiously in a single short paragraph and are shown as two panels out of a total of 18 in figure EV1A.

**Minor points:**
* In panel 4B, the residue numbering of eIF2alpha is shifted relative to the numbering in the PDB. Please correct the coordinates or the labels of the residues.

We thank the reviewer for spotting this mistake. The numbering shown in panel 4B (now Figure 4A) is correct and based on uniprot P05198. This is consistent with yeast (e.g., Hussain et al, 2014) and human numbering (e.g., Brito-Querido et al, 2020). We have updated the PDB numbering to the correct value.

**The structure of eIF5B-containing 48S complexes...**

**Major points:**

* This section includes the most unexpected result of the manuscript, describing the structure of non-canonical eIF5B-bound 48S initiation complexes in pre-48S and closed states in the context of the HCV IRES, prior to 60S subunit joining. In these complexes, the initiator tRNA is held in place by domain IV of eIF5B, which is bound to the 40S in GTP conformation. However, as before the level of detail in the structural interpretation needs to be reconsidered in the context of the EM map quality, which locally varies a lot.

We thank the reviewer for this comment. We have endeavored to be reasonably conservative with our interpretations of the structure given the differences in map quality for different structures as well as within a single structure. In the case of our modelling of eIF5B domain IV, we would like to note that the position relative to other domains is novel (e.g., rotation of domain IV relative to eIF5B domains II/III/G) but the actual structure of domain IV matches very closely to all other available eIF5B structures. Table R3 (see below) reports the RMSD for each domain of eIF5B compared with structures from the literature as well as the predicted structure determined using AlphaFold. The RMSD of our reported domain IV (structure 14wt-15wt) is very low compared to that of eIF5B bound to the yeast 80S ribosome (average RMSD 2.9 Å) and even lower when compared with the predicted human eIF5B structure using AlphaFold (average RMSD 1.3 Å). The placement of eIF5B domain IV relative to the tRNA was informed using the unsharpened map in which identifying the position is much simpler. Originally we did not provide this map to the reviewers, but due to the importance of this question, we have uploaded all unsharpened maps to the EMDB and have made them available to the reviewers as these can help visualize the correct placement of various domains such as eIF5B domain IV.

Table R3. RMSD between published eIF5B or eIF1A structures compared with structure 14wt and structure 15wt. Yeast eIF5B (PDB: 6WOO; Wang et al, 2020), human eIF5B (https://alphafold.ebi.ac.uk/entry/O60841; Jumper et al, 2021), human eIF1A from 43S IC (PDB: 6ZWM; Brito-Querido et al, 2020) and rabbit eIF1A from 48S IC (PDB: 6YAL; Simonetti et al, 2020).

| | Structure 14 | Structure 15 |
|---|---|---|
| Global RMSD 6WOO | 5.58 | 5.38 |
| G domain | 3.07 | 2.74 |
| Domain II | 4.89 | 4.81 |
| Domain III | 1.78 | 1.44 |
| helix 12 | 1.92 | 1.91 |
| Domain IV | 2.91 | 2.86 |
| Domain average | 2.91 | 2.75 |
| Global AlphaFold | 4.67 | 4.07 |
| G domain | 2.82 | 2.46 |
| Domain II | 3.08 | 2.98 |
| Domain III | 1.78 | 1.41 |
| helix 12 | 2.15 | 1.96 |
| Domain IV | 1.43 | 1.13 |
| Domain average | 2.25 | 1.99 |
| 6ZMW eIF1A | 3.73 | 3.79 |

| | | |
|---|---|---|
| 6YAL eIF1A | 2.10 | 1.91 |

* The authors should provide some biochemical evidence that this non-canonical intermediate is not an artefact of in vitro complex formation.

Since the eIF5B-mediated (eIF2-independent) mechanism of initiation on the HCV IRES is very well established and widely accepted (see Introduction and references therein), we assume that this comment specifically concerns the eIF5B-containing open pre-48S complex (structure 14wt). The identification of this open-state intermediate would have prompted us to immediately compare the susceptibility of eIF5B-containing 48S complexes formed on the wt and ΔdII HCV IRESs to dissociation by eIF1. However, these experiments have already been done (Pestova et al., 2008), and gratifyingly, they reported high susceptibility of eIF5B-containing complexes assembled on the wt IRES, which is highly compatible with the formation of structure 14wt. We were very excited by the identification of this structure because it can provide a plausible explanation for the previously observed and reported effect. The consistency of our structural finding and the previously reported results is clearly explained in Lines 539-545 of the revised manuscript.

* The major concern resides in the interpretation of the pre-48S, where the authors claim that the initiator tRNA bound to the P-site is not base-pairing with the start codon but shifted by one residue (structure #14, depicted in Figures 5H, 6B and S3G). The corresponding EM density does not support this interpretation, it is rather consistent with a conformation in which the start codon interacts with the anticodon of the tRNA in a cognate manner and not via the modelled mismatch, although the tRNA may still not be in fully accommodated conformation compared to the closed eIF5B 48S IC. The claim that the start codon is shifted cannot be made considering that the residues flanking the start codon upstream and downstream are poorly resolved it is not possible to assign register of the mRNA. Further, in the current structure, all three residues of the codon are built in highly unfavorable orientations and do not explain the observed density in a satisfying manner.

Both this reviewer and reviewer #2 raised similar questions about structure 14wt and the assignment of nucleotides in the P site. For the convenience of this reviewer, we will reiterate the response that we gave above to reviewer #2.

We are very grateful to the reviewers for raising these concerns as they led us to reevaluate the data, thereby improving the modelling in the final structure 14wt complex completely in line with the suggestions made by these reviewers and correcting a significant interpretational error. The P-site triplet of structure 14wt is now the A(+1), U(+2), G(+3) codon as they suggested.

To clarify the identity of the P-site codon we examined all available maps that showed the eIF5B containing pre-48S complex. These maps were: structure 14wt (3.9 Å, 60,578 particles), and the related structure 44wt (4.2 Å, 29,072 particles) showing eIF5B in a slightly different position on the 40S subunit intersubunit interface, as well as the consensus map structure 38wt (3.5 Å, 199,047 particles) (See Figure S1F). Even though in the higher-resolution consensus map the identity of the P site codon is ambiguous, we noticed that the upstream G(-7), U(-6), G(-5) region was modelled incorrectly (see Figure R1 below). By correcting this region, the mRNA register is changed by 1 and so the P site is now occupied by A(+1), U(+2), G(+3). Examination of the unsharpened maps of structure 14wt, 38wt, and 44wt also confirm the placement of these nucleotides in the new position.

The text corresponding to discussion of structure 14wt in the original sections "The structure of eIF5B-containing 48S complexes assembled on the IRES" and "Comparison of eIF5B-containing pre-48S initiation complexes with eIF2-containing scanning 43S complexes", as well as the new section

"Dynamics of ribosomal contacts of mRNA along the mRNA-binding channel throughout the initiation process" was revised accordingly.

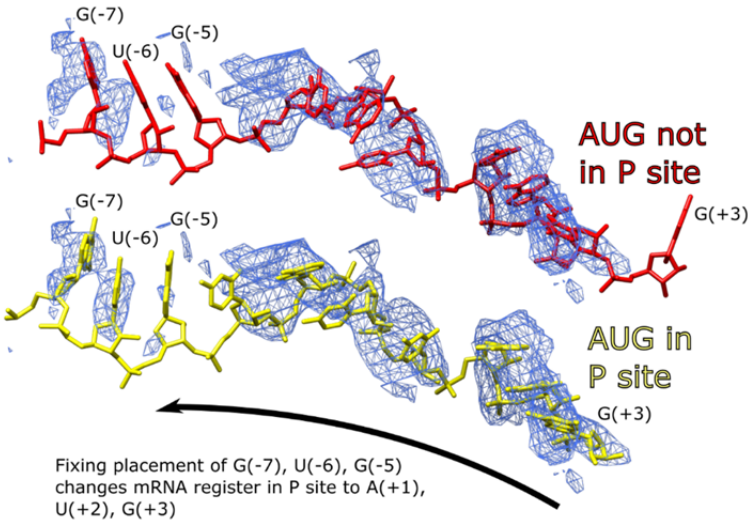


**Figure R1.** Difference in register of mRNA depending on the modelling of upstream nucleotides.

* The authors discuss the details of interactions between eIF1A and domain IV of eIF5B (depicted in Figure 5D). However, as these are poorly resolved areas in both structures (#14 & #15), analysis at such detail is not possible.

As outlined above, we relied upon the unsharpened maps to place the major domains of eIF5B, but during review unfortunately did not provide these maps to the reviewers, which likely hampered their ability to appropriately judge this material. We apologize for this oversight and now, in response to these concerns, provide the unsharpened maps, which we will also make available on the EMDB when this paper is published.

In the case of placement of eIF5B domain IV and eIF1A, we note that none of these domains are in a 'new' conformation, and that they match the other available published data (see Table R3). To aid in assessing the placement we have updated Figure 5D (Figure 5B in the revised manuscript) to show the unsharpened map and the position of the model within that density.

**Minor points:**

* P14, line 401: "14w" should read "14wt"

This has been corrected.

* P15, lines 429-431: the sentence "These observations ... to GTP-bound conformation." is incomplete.

This sentence has been completed.

* P16, lines 472-475: Figure 5G should be mentioned somewhere.

Figure 5 has been simplified, and Fig. 5G from the original manuscript is now designated as Fig. 5C. It is mentioned in Lines 471 of the revised text.

* P16, line 488: "Figure 5F" should probably read "Figure 5G".
Figure 5F from the original manuscript is now Fig. EV4E. It is mentioned in Line 432 of the revised text.

* Figure 5 contains too many panels. Especially panel 5F is too small and difficult to understand. Panels 5A & 5B do not convey the main message in spite of their large size. Additional labels and a more plastic 3D representation of the cartoon may help the reader.

Figure 5 has been simplified, and now contains only six images, shown as panels 5A - 5E. The former panels 5A, 5B, 5E and 5F are now shown as Figures EV4A, EV4B, EV4D and EV4E. Figures 5C and 5D have been revised, in the latter case to illustrate the revised analysis of the eIF1A-eIF5B interaction by including the unsharpened map and the position of the model within that density.

* In both eIF5B-bound 48S IC structures, the GTP in the active site of the GTPase domain of eIF5B is in a distorted conformation. More care should be taken to model and refine the nucleotide and to keep it correctly coordinated to the active site residues and the magnesium ion. For this, high resolution templates are available in the PDB.

We have remodelled the GTP- and GTP-binding region using a high-resolution crystal structure as a guide (PDB:4NCN), to better match both our density as well as the correct coordination of GTP.

**Comparison of eIF5B-containing pre-48S initiation complexes with eIF2-containing 43S scanning complexes**

* Although a comparison with existing 43S structures is obvious, this paragraph is excessively long and, considering the structural issues mentioned above, the text may be shortened and the corresponding Figure 6 should be revised.

We have shortened the text and simplified Figure 6. The structural issues raised by this reviewer concerning the structure of the eIF5B-containing pre-48S IC, specifically those concerning codon-anticodon base-pairing, have been resolved (see above) and we therefore consider that a comparison with existing 43S complexes is both valid and important.

Thank you again for submitting your revised manuscript. We have now received comments from the initial three referees (please see below). While referee #1 and #2 in principle now support publication, referee #3 currently does not. The main concerns of referee #3 relate to the readability of the manuscript, in particular for a non-specialist audience, and could well be resolved by further textual revision. It is crucial for a generalist journal such as The EMBO Journal that central findings of published manuscripts are accessible to a wide readership, which is why I would ask you to please review and revise the manuscript again in light of the referee's comments keeping the non-specialist reader in mind. In addition to her/his comments, I will attach a file with a number comments and questions from our side, which may help to identify some of the passages that require further explanation (EMBOJ-2022-110581R-data_edited_ms-SB). You may also want to consider a model overview figure similar to Fig. 8, yet for example also indicating closed vs. open and mRNA binding, to relate back to when discussing the context of specific structural findings. In addition to a careful textual revision to address this issue and referee #2's minor concern, I would also ask you to resolve a number of editorial issues that are listed in detail below.

---

Referee #1:

The authors have addressed all of the points raised and I think that the revised manuscript is clearer and more nuanced as a result of their efforts. I am therefore satisfied with the manner in which results are now presented and congratulate the authors on this solid and valuable addition to the field.

Referee #2:

In the revised manuscript, Brown and co-authors have addressed my criticisms. The manuscript has substantially improved. My only current criticism should be straightforward to address in the next revision. The way the authors describe the initiation mechanism often implies that the IRES or initiation factors "induce" a conformational change of the ribosome (e.g. "eIF1..., in cooperation with eRF1A, induces the open conformation of the 40S subunit"). This may be misleading, as the word "induce" means "to cause the formation of". The authors correctly point out that the ribosome can sample such states (e.g. head-opened states) even without IRES/factors - e.g. see the beginning of Discussion on page 18. Therefore, the factors "stabilize" a certain conformational state, rather than "cause" a conformational rearrangement. It would less misleading if the authors didn't use the ambiguous term "induce" and describe the functions of factors more accurately in all the instances they used "induce": for example, "eIF1..., in cooperation with eRF1A, stabilizes the open conformation of the 40S subunit".

Referee #3:

In the revised version of this manuscript, the authors addressed several concerns raised by the reviewers, in particular regarding the interpretation of the cryo-EM maps and the quality of the coordinates. Corresponding parts in the text were updated accordingly, and some of the claims were more carefully worded. Some of the main text figures were somewhat simplified as suggested.

Unfortunately the authors missed the opportunity to revise the manuscript such that it would become more focused and concise, and suitable for the broad EMBO readership. Consequently, the text is still difficult to follow, the interpretations are mechanistically vague, and the figures do not convey the findings well. The current version of the paper still overwhelms the reader with undigested structural information as the authors seem to have difficulty deciding which insights to emphasise. It is also disappointing that the authors have not supported the structural interpretations with biochemical experiments and instead argue that there is previously published biochemical evidence that explains all newly obtained structural insights.

Unfortunately, considering the extent of changes that would be necessary to improve this manuscript, it is not possible to provide sufficiently specific and extensive recommendations within the scope of a review.

Thank you for taking the time in continuing to shepherd this manuscript toward perfection. At your suggestion we have modified the text at a number of locations (see *track changes* in updated manuscript) to make the results and discussion more accessible to a non-specialist audience. We would like to thank **reviewers #1** and **#2** for their continued positive assessment of the manuscript.

The minor suggestion from **reviewer #2**, that we use a more precise descriptor when referring to the outcome of various factors binding to the ribosome, is a change we agree with and have made where appropriate.

The comments from **reviewer #3**, however, were neither specific nor constructive, so we have not been able to respond to them. The suggestion to perform additional 'biochemical experiments' is so broad as to be unactionable. It is unclear from the comments which stages are deemed currently unsupported by the literature. In the previous round of review both reviewers #1 and #2 were satisfied with the text and the coverage of supporting biochemical literature.

Again, we would like to thank you and the three reviewers for the positive effort that has resulted in a more accurate understanding of the HCV IRES initiation mechanism.

Thank you again for submitting the final revised version of your manuscript. As mentioned, you will likely hear again from me regarding the final textual edits of the transfer files in the next few days, but for now I happy to inform you that we have formally accepted the study for publication in The EMBO Journal.

## EMBO Press Author Checklist

| |
|---|
| Corresponding Author Name: Joachim Frank and Tatyana Pestova |
| Journal Submitted to: The EMBO Journal |
| Manuscript Number: EMBOJ-2022-110581 |

**USEFUL LINKS FOR COMPLETING THIS FORM**
The EMBO Journal - Author Guidelines
EMBO Reports - Author Guidelines
Molecular Systems Biology - Author Guidelines
EMBO Molecular Medicine - Author Guidelines

**Reporting Checklist for Life Science Articles (updated January 2022)**
This checklist is adapted from Materials Design Analysis Reporting (MDAR) Checklist for Authors. MDAR establishes a minimum set of requirements in transparent reporting in the life sciences (see Statement of Task: 10.31222/osf.io/9sm4x). Please follow the journal's guidelines in preparing your manuscript.
**Please note that a copy of this checklist will be published alongside your article.**

**Abridged guidelines for 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.
- ideally, figure panels should include only measurements that are directly comparable to each other and obtained with the same assay.
- plots 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. Any statistical test employed should be justified.
- Source Data should be included to report the data underlying figures according to the guidelines set out in the authorship 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 χ2 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.

**Please complete ALL of the questions below.**
**Select "Not Applicable" only when the requested information is not relevant for your study.**

### Materials

| Newly Created Materials | Information included in the manuscript? | In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|
| New materials and reagents need to be available; do any restrictions apply? | Not Applicable | |

| Antibodies | Information included in the manuscript? | In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|
| For **antibodies** provide the following information:<br>- Commercial antibodies: RRID (if possible) or supplier name, catalogue number and or/clone number<br>- Non-commercial: RRID or citation | Not Applicable | |

| DNA and RNA sequences | Information included in the manuscript? | In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|
| **Short novel DNA or RNA including primers, probes**: provide the sequences. | Not Applicable | |

| Cell materials | Information included in the manuscript? | In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|
| **Cell lines:** Provide species information, strain. Provide accession number in repository **OR** supplier name, catalog number, clone number, and/**OR** RRID. | Not Applicable | |
| **Primary cultures:** Provide species, strain, sex of origin, genetic modification status. | Not Applicable | |
| Report if the cell lines were recently **authenticated** (e.g., by STR profiling) and tested for mycoplasma contamination. | Not Applicable | |

| Experimental animals | Information included in the manuscript? | In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|
| **Laboratory animals or Model organisms:** Provide species, strain, sex, age, genetic modification status. Provide accession number in repository **OR** supplier name, catalog number, clone number, **OR** RRID. | Not Applicable | |
| **Animal observed in or captured from the field:** Provide species, sex, and age where possible. | Not Applicable | |
| Please detail **housing and husbandry conditions**. | Not Applicable | |

| Plants and microbes | Information included in the manuscript? | In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|
| **Plants:** provide species and strain, ecotype and cultivar where relevant, unique accession number if available, and source (including location for collected wild specimens). | Not Applicable | |
| **Microbes:** provide species and strain, unique accession number if available, and source. | Not Applicable | |

| Human research participants | Information included in the manuscript? | In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|
| If collected and within the bounds of privacy constraints report on age, sex and gender or ethnicity for all study participants. | Not Applicable | |

| Core facilities | Information included in the manuscript? | In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|
| If your work benefited from core facilities, was their service mentioned in the acknowledgments section? | Yes | ACKNOWLEDGEMENTS |

### Design

| Study protocol | Information included in the manuscript? | In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|
| If study protocol has been **pre-registered, provide DOI in the manuscript**. For clinical trials, provide the trial registration number **OR** cite DOI. | Not Applicable | |
| Report the **clinical trial registration number** (at ClinicalTrials.gov or equivalent), where applicable. | Not Applicable | |

| Laboratory protocol | Information included in the manuscript? | In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|
| Provide DOI OR other citation details if **external detailed step-by-step protocols** are available. | Not Applicable | |

| Experimental study design and statistics | Information included in the manuscript? | In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|
| Include a statement about **sample size** estimate even if no statistical methods were used. | Not Applicable | |
| Were any steps taken to minimize the effects of subjective bias when allocating animals/samples to treatment (e.g. **randomization procedure**)? If yes, have they been described? | Not Applicable | |
| Include a statement about **blinding** even if no blinding was done. | Not Applicable | |
| Describe **inclusion/exclusion criteria** if samples or animals were excluded from the analysis. Were the criteria pre-established?<br><br>If sample or data points were omitted from analysis, report if this was due to attrition or intentional exclusion and provide justification. | Not Applicable | |
| 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. Is there an estimate of variation within each group of data? Is the variance similar between the groups that are being statistically compared? | Not Applicable | |

| Sample definition and in-laboratory replication | Information included in the manuscript? | In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|
| In the figure legends: state number of times the experiment was **replicated** in laboratory. | Not Applicable | |
| In the figure legends: define whether data describe **technical or biological replicates**. | Not Applicable | |

## Ethics

| Ethics | Information included in the manuscript? | In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|
| Studies involving **human participants**: State details of **authority granting ethics approval** (IRB or equivalent committee(s), provide reference number for approval. | Not Applicable | |
| Studies involving **human participants**: 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. | Not Applicable | |
| Studies involving **human participants:** For publication of **patient photos**, include a statement confirming that consent to publish was obtained. | Not Applicable | |
| Studies involving experimental **animals**: State details of **authority granting ethics approval** (IRB or equivalent committee(s), provide reference number for approval. Include a statement of compliance with ethical regulations. | Not Applicable | |
| Studies involving **specimen and field samples:** State if relevant **permits** obtained, provide details of authority approving study; if none were required, explain why. | Not Applicable | |

| Dual Use Research of Concern (DURC) | Information included in the manuscript? | In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|
| Could your study fall under dual use research restrictions? Please check biosecurity documents and list of **select agents and toxins** (CDC): https://www.selectagents.gov/sat/list.htm | Not Applicable | |
| If you used a select agent, is the security level of the lab appropriate and reported in the manuscript? | Not Applicable | |
| If a study is subject to dual use research of concern regulations, is the name of the **authority granting approval and reference number** for the regulatory approval provided in the manuscript? | Not Applicable | |

## Reporting

The MDAR framework recommends adoption of discipline-specific guidelines, established and endorsed through community initiatives. Journals have their own policy about requiring specific guidelines and recommendations to complement MDAR.

| Adherence to community standards | Information included in the manuscript? | In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|
| State if relevant guidelines or checklists (e.g., **ICMJE, MIBBI, ARRIVE, PRISMA**) have been followed or provided. | Not Applicable | |
| 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. | Not Applicable | |
| 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. | Not Applicable | |

## Data Availability

| Data availability | Information included in the manuscript? | In which section is the information available? (Reagents and Tools Table, Materials and Methods, Figures, Data Availability Section) |
|---|---|---|
| Have **primary datasets** been deposited according to the journal's guidelines (see 'Data Deposition' section) and the respective accession numbers provided in the Data Availability Section? | Yes | All structural data uploaded to PDB and EMDB. Details in section: DATA AVAILABILITY |
| Were **human clinical and genomic datasets** deposited in a public access-controlled repository in accordance to ethical obligations to the patients and to the applicable consent agreement? | Not Applicable | |
| Are **computational models** that are central and integral to a study available without restrictions in a machine-readable form? Were the relevant accession numbers or links provided? | Not Applicable | |
| If publicly available data were reused, provide the respective **data citations in the reference list**. | Not Applicable | |