

ARTICLE

Received 13 Feb 2015 | Accepted 27 May 2015 | Published 8 Jul 2015

DOI: 10.1038/ncomms8646

OPEN

# Cryo-EM structure of Hepatitis C virus IRES bound to the human ribosome at 3.9-Å resolution

Nick Quade<sup>1</sup>, Daniel Boehringer<sup>1</sup>, Marc Leibundgut<sup>1</sup>, Joop van den Heuvel<sup>2</sup> & Nenad Ban<sup>1</sup>

Hepatitis C virus (HCV), a widespread human pathogen, is dependent on a highly structured 5'-untranslated region of its mRNA, referred to as internal ribosome entry site (IRES), for the translation of all of its proteins. The HCV IRES initiates translation by directly binding to the small ribosomal subunit (40S), circumventing the need for many eukaryotic translation initiation factors required for mRNA scanning. Here we present the cryo-EM structure of the human 40S ribosomal subunit in complex with the HCV IRES at 3.9 Å resolution, determined by focused refinement of an 80S ribosome-HCV IRES complex. The structure reveals the molecular details of the interactions between the IRES and the 40S, showing that expansion segment 7 (ES7) of the 18S rRNA acts as a central anchor point for the HCV IRES. The structural data rationalizes previous biochemical and genetic evidence regarding the initiation mechanism of the HCV and other related IRESs.

<sup>1</sup>Department of Biology, Institute of Molecular Biology and Biophysics, Otto-Stern-Weg 5, ETH Zürich, Zürich 8093, Switzerland. <sup>2</sup>Research Group Recombinant Protein Expression, Helmholtz Centre for Infection Research, Inhoffenstraße 7, Braunschweig 38124, Germany. Correspondence and requests for materials should be addressed to N.B. (email: ban@mol.biol.ethz.ch).

**H**epatitis C virus (HCV) is a widespread human pathogen infecting an estimated 150 million people worldwide<sup>1</sup>. HCV virions contain a single-stranded RNA genome and rely on a structured region of its 5'-untranslated region, referred to as the internal ribosome entry site (IRES), for translation of all the encoded proteins. Canonical translation initiation and many IRESs depend on a large number of initiation factors (eIFs) such as eIF1, eIF1A, eIF2, eIF3 and eIF4A and G<sup>2</sup>. In contrast, the HCV IRES requires only eIF2 and eIF3 for start codon recognition, binding directly to the 40S subunit and eIF3 (refs 3,4). Translation initiation by the HCV IRES begins with binding of the IRES to the 40S ribosomal subunit. The IRES then recruits eIF3 and the ternary complex of eIF2-tRNA<sub>i</sub><sup>Met</sup>-GTP is bound at the start codon. On GTP hydrolysis eIF2 as well as eIF3 dissociate and the 60S subunit joins to form a translation competent 80S complex with the initiator tRNA positioned at the P-site<sup>5</sup> (Supplementary Fig. 1).

Due to the central role of the HCV IRES in recruiting host ribosomes for translation of viral mRNA, its architecture and the mode of interactions with the ribosome have been extensively investigated. The HCV IRES as well as several HCV-like IRESs share a common conserved fold encompassing around 300 bases, which form two major domains (II and III)<sup>6</sup> (Fig. 1a). Electron microscopic (EM) studies have revealed that domain III binds in an elongated conformation at the back of the 40S<sup>7-9</sup>. It is responsible for high-affinity binding and proper positioning of the mRNA in the mRNA channel by establishing direct interactions through subdomains IIIa, c, d, e and f<sup>3</sup> (Fig. 1b). eIF3 binding has been shown to be mediated by domain IIIb and the junction IIIabc<sup>3,9</sup>. Domain II, on the other hand, forms an L-shaped structure that reaches across the 40S subunit and into the intersubunit space<sup>7,8,10</sup>. It has been demonstrated that domain II induces a conformational change in the 40S<sup>7</sup> and is involved in release of eIF2 (ref. 11) and transition from translation initiation to elongation<sup>12</sup>.

Although high-resolution X-ray and NMR structures of many fragments of the HCV IRES are available<sup>13-17</sup>, interactions of the HCV and CSFV IRESs with the ribosome have only been investigated by cryo-EM at around 8 Å resolution in heterologous complexes with porcine or rabbit ribosomes<sup>9,10</sup>. Here we present a 3.9 Å cryo-EM structure of the homologous complex of the HCV IRES bound to the human 40S ribosome, which was solved by focused refinement of the HCV IRES bound to 80S ribosomes. This represents the first near-atomic structure of the HCV IRES that allows visualization of the network of molecular interactions formed between the 40S subunit and the HCV IRES. We identify expansion segment 7, which is tightly bound between two IRES domains, as a central anchor point for the HCV IRES. Furthermore, we show that the HCV IRES and the Cricket paralysis virus (CrPV) IRES, although structurally unrelated, share interactions with ribosomal proteins.

## Results

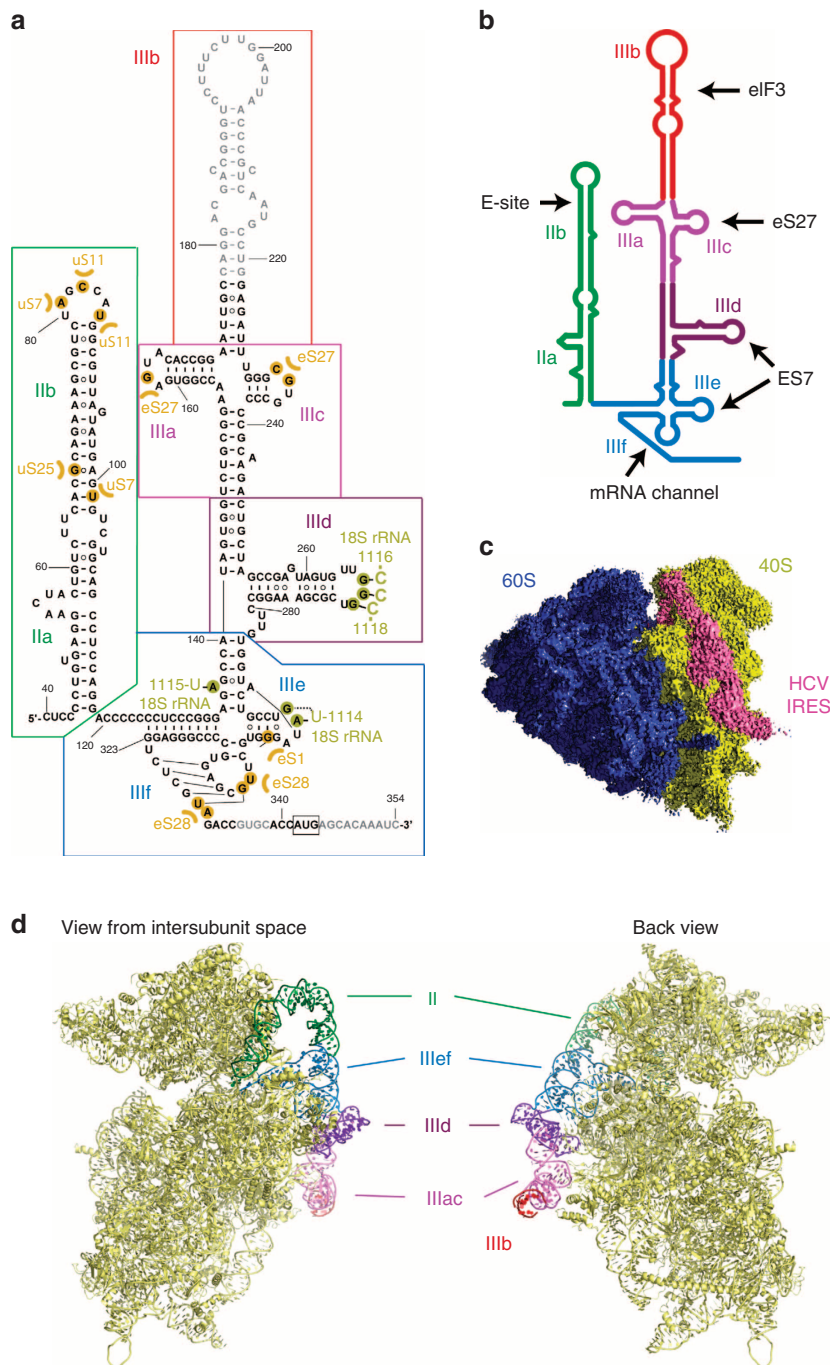
### Cryo-EM reconstruction of the ribosome-bound HCV IRES.

To investigate the molecular interactions between the HCV IRES and the human ribosome, we reconstituted the 80S-HCV IRES complex from human 80S ribosomes purified from HEK293-6E cells and prefolded HCV IRES. Although our complex contained both ribosomal subunits, initial single particle cryo-EM reconstructions indicated that the IRES in the 80S complex contacts exclusively the 40S subunit and adopts a conformation previously observed in 40S-HCV IRES complexes<sup>7</sup> (Fig. 1c). The general features of the structure are very similar to a previous structure of the 80S ribosome stalled in initiation by cycloheximide bound to HCV IRES<sup>8</sup> except that the stalled

80S-HCV IRES complex showed additional density connecting domain IIa of the HCV IRES to the L1 stalk of the 60S subunit. This interaction is not observed in our structure even at low thresholds. The additional density contacting the L1 stalk might be attributable to HCV IRES domain I or the RNA aptamer used for purification located at the 5' end of the IRES in domain IIa, which were not present in our construct. Formed in the absence of initiation factors or tRNAs, our complex corresponds to the initial binding complex of IRES to the 40S subunit (Supplementary Fig. 1), in agreement with kinetic data that show that binding occurs in a single step and is followed by 80S assembly<sup>18</sup>. Consequently, we focused the alignment on the 40S subunit with bound IRES molecule to obtain a 3.9 Å resolution reconstruction (Supplementary Figs 2,3). Our reconstruction of the human 40S-HCV IRES complex revealed clear density corresponding to the HCV IRES on the back of the 40S. On the basis of structures of individual IRES domains an atomic model could be built for most of the IRES (Fig. 1d, Supplementary Fig. 4), except for a part of domain IIIb (bases 178-221), for which density is weak, apparently due to partial disorder in the absence of eIF3<sup>9</sup>. The single-stranded mRNA region of the IRES was also poorly resolved in our density and was only partially interpreted.

**ES7 acts as a central anchor point for the HCV IRES.** The most prominent binding site of the HCV IRES on the 40S subunit is ES7 of the 18S rRNA (bases 1,114-1,118). Specifically, domains IIIId and the pseudoknot domain IIIef of the IRES contact two sides of the ES7 stem loop (Fig. 2a). These IRES domains have been identified biochemically to provide most of the binding affinity of the HCV IRES for the 40S subunit<sup>5</sup> and are critical for the initial binding of the HCV IRES<sup>3,5</sup>. A detailed view of this contact area explains the molecular basis of the nanomolar affinity of the HCV IRES for the 40S subunit by revealing an intricate network of interactions between ES7 and domains IIIId and IIIe (Fig. 2b,c). The apical loop of domain IIIId forms a kissing loop interaction with ES7, which becomes ordered on HCV IRES binding in comparison with the structure of the free porcine 40S subunit, where ES7 exhibits considerable flexibility<sup>19</sup> (Fig. 2d). The highly conserved<sup>20</sup> bases <sup>266</sup>GGG<sup>268</sup> in the loop of HCV IRES domain IIIId form Watson-Crick base pairs with residues <sup>1116</sup>CCC<sup>1118</sup> within ES7 (Fig. 2b, Supplementary Fig. 4b). This interaction explains the previous observations that mutations of these guanosine residues lead to reduced IRES activity, and that activity can be rescued by compensatory mutations in ES7 (ref. 21). Intriguingly, a stem-loop structure involved in translation termination/reinitiation of caliciviruses containing three guanosine residues has been shown to directly interact with the same bases of ES7 (ref. 22), suggesting that the observed binding mechanism may also be utilized by other RNAs.

On the other face of ES7, bases <sup>1114</sup>UU<sup>1115</sup> are flipped out and contact several residues of IRES domain IIIe and a helical region adjacent to it (Fig. 2c,d, Supplementary Fig. 4c-e). In this helical region, HCV IRES base A136 is flipped out and forms a Watson-Crick base pair with residue U1115 of ES7, while A296 and G295 from the loop of IRES domain IIIe interact with U1114 of ES7. A296 forms a reverse Hoogsteen base pair with U1114, and G295 engages in stacking interactions. Bases A136, G295 and A296 are highly conserved among HCV-like IRESs and mutations in G295 and A296 lead to severe defects in translation initiation<sup>23</sup>. The unusual conformation of the IRES at this interaction site is further stabilized by a Watson-Crick base pair between U297 on domain IIIe and A288 in the helical region (Supplementary Fig. 5), which leaves A136 flipped out in the correct position to interact with ES7. The importance of the tertiary interaction

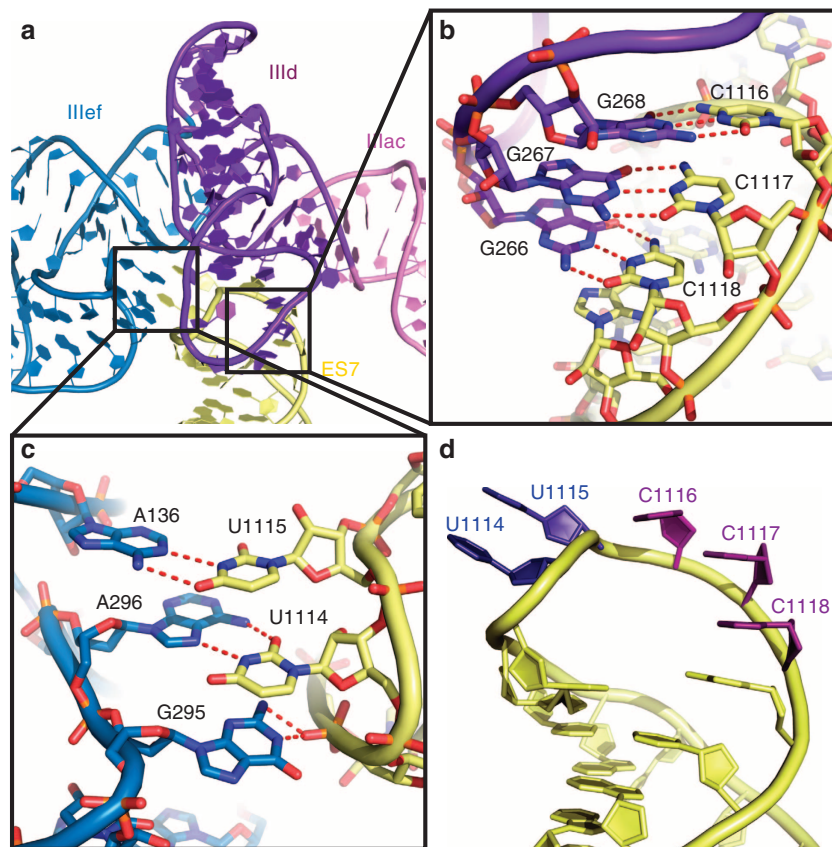


**Figure 1 | HCV IRES bound to human ribosome.** (a) Secondary structure diagram of the HCV IRES. The various domains are indicated by different colours and roman numbers. Canonical base pairs are marked by lines and non-standard base pairs by circles. Interactions of the IRES with ribosomal proteins are highlighted in orange, interactions with the 18S rRNA in dark yellow. Bases not built in our structure are indicated in grey, the start codon is indicated by a black box. (b) Cartoon representation of the HCV IRES secondary structure. The interaction sites of the different domains are indicated. (c) EM density of 80S-HCV IRES complex showing that the HCV IRES does not contact the 60S subunit. The 40S is coloured in yellow, the 60S in blue and the HCV IRES in pink. (d) Structure of the HCV IRES bound to the human 40S. Views from the intersubunit side (left) and the solvent side (right).

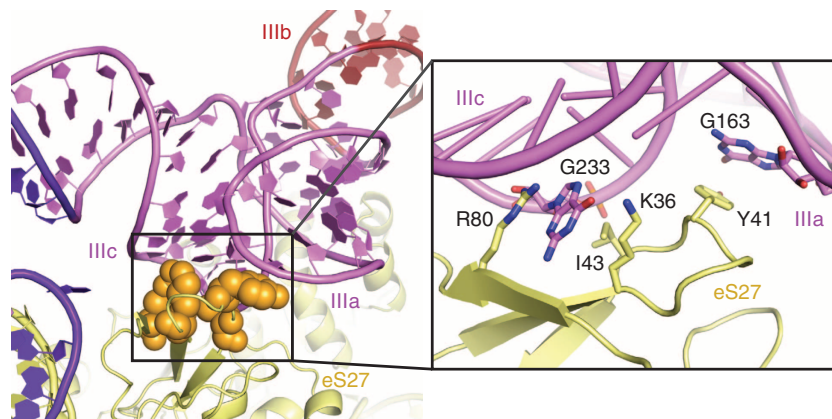
between A288 and U297 is supported by the strong conservation and covariation among HCV-like IRESs<sup>20</sup> and mutational studies showing that single base substitutions can have a significant effect on translation efficiency<sup>13,24</sup>. While this domain was visualized previously in a crystal structure<sup>13</sup>, our ribosome-bound structure now reveals how this complex fold enables high-affinity ribosome binding by stabilizing the conformation of A136 and the loop of domain IIIe to provide an interaction site for the flipped out U1114/U1115 of ES7.

In summary, the observed network of specific contacts between the IRES domain IIIe and IIIc with ES7 of the 18S rRNA now rationalizes how ES7 can serve as the central anchor point for the HCV IRES and other HCV-like IRESs on the 40S subunit.

**Binding of HCV IRES and eIF3 to eS27 is mutually exclusive.** The second most prominent interface between the HCV IRES and the ribosome involves domains IIIa and IIIc, which contact



**Figure 2 | ES7 acts as the main anchoring point for the HCV IRES.** (a) Overview over the architecture of the binding site. (b,c) Close-ups of the interactions between ES7 and domain IIIc (b) and IIIe (c). The structures are shown in cartoon and sticks, red dashes indicate hydrogen bonds. (d) Conformation of ES7 bound to HCV IRES. Bases interacting with domain IIIc are coloured in purple, and bases interacting with domain IIIef are coloured in dark blue.

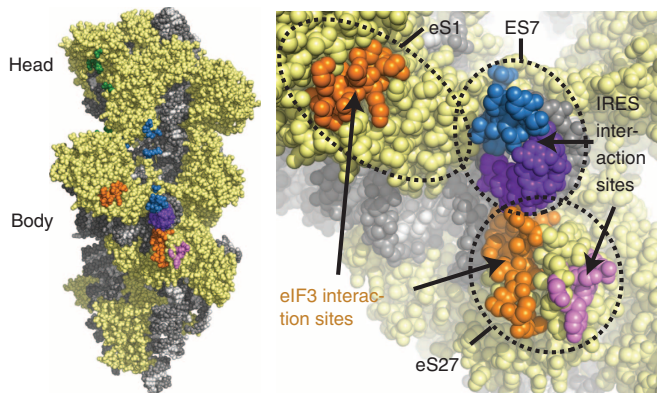


**Figure 3 | Domain IIIc binding site on eS27.** Orange spheres indicate amino acids interacting with the IRES. Inset panel: detailed interactions between bases G163 and G233 of domain IIIc with amino acids from eS27. The structure is shown as cartoon, interacting residues as sticks.

ribosomal protein eS27 (Fig. 1a, Fig. 3, Supplementary Fig. 4f). The four-way junction between these domains assumes an anti-parallel conformation, placing domains IIIa and IIIc close together forming the binding platform for eS27 similar to previous structures<sup>9,10</sup>. This is in contrast to the crystal structure of this domain<sup>16</sup>, which assumes a parallel conformation positioning domain IIIb close to the stem connecting the junction to domain IIIc and separating domains IIIa and IIIc. It is possible that both conformations exist in solution<sup>25</sup> and the antiparallel

conformation is stabilized on ribosome binding. The binding of domains IIIa and IIIc to eS27 is mainly mediated by the apical loops via two flipped out guanosine residues: G163 in domain IIIa is stacked against the aromatic moiety of Y41, and G233 in domain IIIc forms specific interactions with residues R80, K36 and I43 of eS27. The apical loop of domain IIIa, including base G163 that interacts with the ribosome, is conserved among several HCV-like IRESs, such as CSFV<sup>20</sup>. Interestingly, eS27 is used during canonical initiation as a major contact point for the

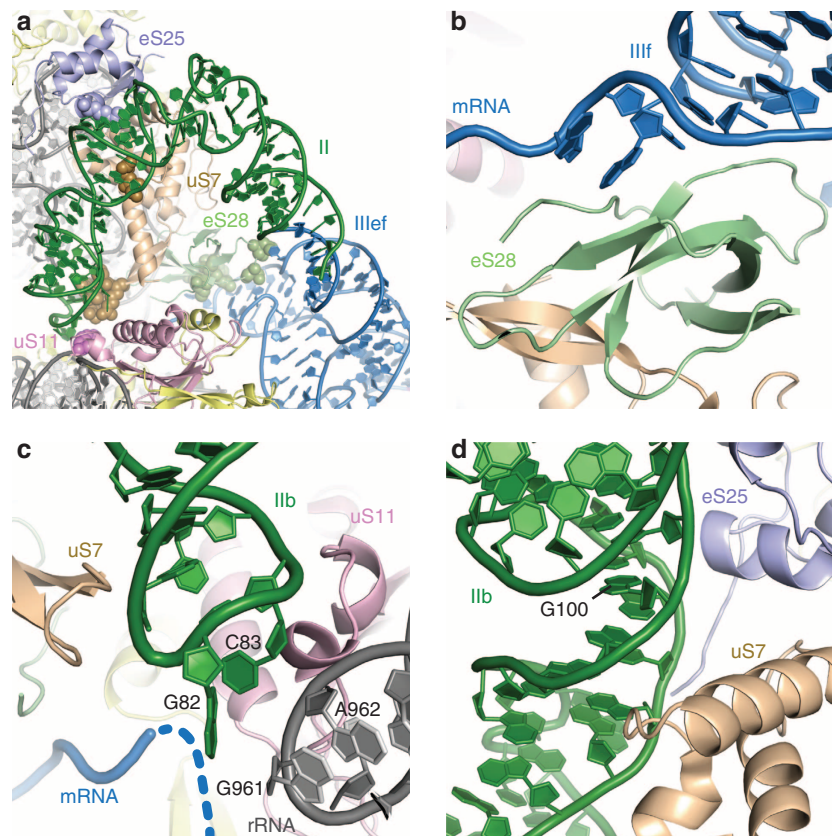
multisubunit eIF3 complex<sup>26</sup>. In the 11 Å resolution cryo-EM reconstruction of the 40S subunit in complex with eIF3, the initiation factor binds to the opposite face of eS27 compared to



**Figure 4 | The HCV IRES and eIF3 occupy distinct binding sites on the 40S.** Ribosomal proteins are coloured in yellow, the bases of the 18S rRNA are coloured in light grey, the backbone in dark grey. eIF3-binding sites are shown in orange, HCV IRES binding sites are coloured according to their interacting domain (green: domain II, blue: domain IIIef, purple: domain IIIId and pink: domain IIIac). The figure was generated by fitting the structure of the 40S bound to the HCV IRES into the density of 40S bound to eIF3 (ref. 26) (EMDB ID 5658).

the HCV IRES (Fig. 4). Therefore, the binding of the IRES likely excludes eIF3 from binding to the 40S by occupying the same space above the ribosome rather than competing for the identical contact point on the subunit (Supplementary Fig. 6a,b), resulting in displacement of eIF3 as observed in the cryo-EM structure of the 40S in complex with CSFV IRES and eIF3<sup>9</sup>. In this structure, the CSFV IRES adopts a very similar overall conformation as the HCV IRES in our structure, except for the missing domain II in the CSFV construct, the non-homologous extension of domain IIIId and domain IIIb, which is mostly disordered in our structure, but ordered in the CSFV structure probably due to the interactions with eIF3.

**Domain IIIf properly positions the start codon.** HCV IRES also interacts extensively with ribosomal protein eS28 at the exit of the mRNA channel via the pseudoknot in domain IIIf (Fig. 5a,b, Supplementary Fig. 4g). The contacts involve hydrophobic interactions with L18 and I43, while the phosphate backbone of the IRES is complemented through charge-charge interactions with side chains of K16 and R31 and R66 (Supplementary Fig. 7a). These rather unspecific interactions are consistent with the function of domain IIIf to insert the viral mRNA into the mRNA channel and to properly position the start codon<sup>3,13,27</sup>. After switching to the elongation stage of translation, the mRNA exits the mRNA channel probably by displacing domain IIIf from its binding site at the exit of the channel.

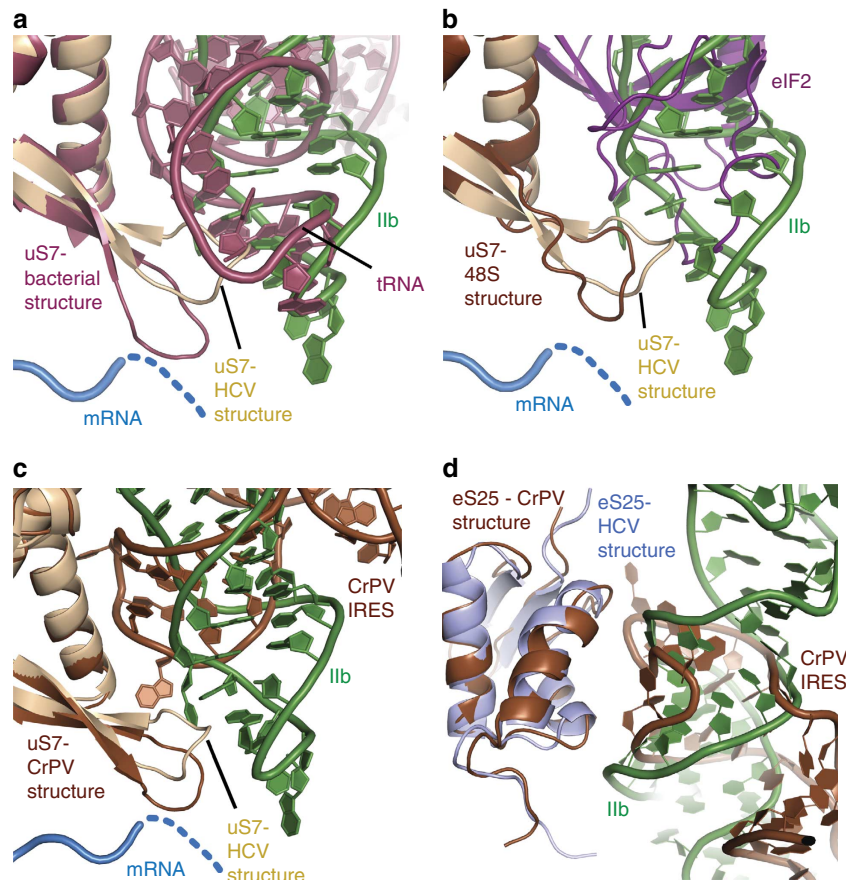


**Figure 5 | Interactions between ribosomal proteins and domain IIIf and domain II of the HCV IRES.** (a) Overview over HCV binding sites. Protein uS7 is shown in ochre, uS11 in pink, eS25 in light blue and eS28 in light green. Other ribosomal proteins are coloured in yellow, 18S rRNA in grey and domain II and IIIef in green and blue, respectively. Interacting residues are marked by spheres. (b-d) Detailed views of binding sites of domain IIIf to eS28 (b), the apical domain of domain II to uS7 and uS11 (c) and a minor interaction site between domain II and uS7 and eS25 (d). Structures are depicted as cartoon. The dashed blue line indicates the approximate path of the mRNA.

**Domain II binds in the E-site above the mRNA channel.** Domain II is not required for 40S binding<sup>3</sup> and can adopt a variety of conformations on the ribosome<sup>18</sup>. It has been shown to be important for activation of GTP hydrolysis by eIF2 and subsequent eIF2 release<sup>11</sup>. In addition, domain II is able to manipulate the head–body conformation in the 40S<sup>7,10</sup>. In our structure, domain II adopts a bent conformation reaching into the E-site as observed in previous structures<sup>7,8</sup>. HCV IRES domain II binds to the intersubunit side of the 40S ribosomal subunit through two spatially well-separated interactions (Fig. 5a). The major binding site is formed by the apical loop of domain II, which reaches into the E-site of the 40S close to the mRNA channel. Here it contacts proteins u7 and u11 and is placed in proximity to bases G961 and A962 of the 18S rRNA (Fig. 5c, Supplementary Fig. 4h). The IRES is bound to a  $\beta$ -hairpin structure of u7 via R130 and R135 as well as T133, and to u11 by R66 and Y72 (Supplementary Fig. 7b,c). This hairpin structure of u7 was shown to be involved in contacting both the E-site tRNA and the mRNA codon in the E-site in the structure of a bacterial ribosome<sup>28</sup>. A detailed comparison with an E-site tRNA in the bacterial ribosome<sup>29</sup> (PDB ID 4RBH) shows that both, domain II of the IRES and the tRNA, bind to the hairpin in a similar fashion using the minor groove of the apical loop of domain II and the anticodon loop of the tRNA, respectively

(Fig. 6a). Mutations in the apical loop of domain II, which is conserved in several HCV-like IRESs, including CSFV, have been shown to affect the configuration of the mRNA in the mRNA channel<sup>30</sup> and severely affect translation efficiency<sup>5,31</sup>. It is therefore possible that the placement of domain II in the E-site and interaction with u7 leads to a 40S conformation that is necessary for proper positioning of the mRNA in the mRNA channel. Furthermore, a recent cryo-EM structure of a partial 48S preinitiation complex places eIF2 in direct contact with u7 (Fig. 6b)<sup>32</sup>. This position is incompatible with binding of domain II of the HCV IRES in the E-site. Therefore, eIF2–GTP–tRNA<sub>i</sub> might displace domain II and the competition between domain II and eIF2 may play a role in evicting eIF2–GDP after start codon recognition and progression towards 80S assembly<sup>5</sup>. The conformational flexibility of domain II was also demonstrated by a recent medium-resolution structure of the HCV IRES in complex with the 40S subunit, tRNA and eIF5B<sup>10</sup>. Compared with our structure, the main difference can be observed in the position of domain II, which adopts a more extended conformation in the eIF5B complex, leading to a displacement of the apical loop of domain II of  $>30 \text{ \AA}$  (Supplementary Fig. 8).

The second, less-specific contact point of domain II with the 40S subunit involves the minor groove of domain II in the vicinity of base pair C67–G100 and ribosomal proteins u7 and e25



**Figure 6 | HCV IRES domain II occupies similar binding sites as E-site tRNA, eIF2 and CrPV IRES.** (a) The apical loop of domain II binds to the  $\beta$ -hairpin of u7 similar to E-site tRNA. Superposition of the HCV IRES bound to 40S with a bacterial ribosome containing an E-site tRNA<sup>29</sup> (PDB ID 4RBH) based on u7. The structure of the bacterial ribosome is shown in red, in the HCV structure domain II is coloured green, mRNA in blue and u7 in ochre. (b) Superposition of the HCV IRES–80S structure and a partial 48S preinitiation complex<sup>32</sup> (PDB ID 3J81) on u7 illustrating that eIF2 binding and binding of domain II at the E-site are mutually exclusive. (c) Comparison of u7 interactions of the HCV IRES bound to 80S with the CrPV IRES bound to 80S (coloured in brown). (d) Comparison of e25 interactions of the HCV IRES bound to 80S with CrPV IRES bound to 80S (coloured in brown). The dashed blue line indicates the approximate path of the mRNA.

(Fig. 5d, Supplementary Fig. 7d). Interestingly, even though eS25 is not essential for canonical translation initiation, eS25 was demonstrated to be very important for translation initiation of HCV and several other IRESs such as the CrPV IRES<sup>33</sup>. Comparison of the CrPV IRES in complex with the 80S<sup>34</sup> with the HCV IRES structure reveals that in spite of very different RNA folds and modes of interaction with the ribosome, both IRESs form similar contacts with the N-terminal helix–turn–helix motif of eS25 as well as the  $\beta$ -hairpin structure and C-terminal region of uS7 (Fig. 6c,d). Therefore, uS7 and eS25 may represent common interaction sites for IRESs necessary for cap-independent translation initiation.

## Discussion

The results presented here reveal the molecular basis of HCV IRES binding to the human 40S subunit to initiate translation of its viral genome in infected cells. The main interaction site between the HCV IRES and the ribosome is located at the tip of ES7 of the 18S rRNA. Here the apical loop of domain III d and the loop of domain III e sandwich the tip of ES7 providing an extensive binding interface. In agreement with biochemical and genetic evidence that suggest an important role for this contact, these interactions probably serve as the main anchor point for the IRES on the 40S ribosomal subunit (Fig. 1a,b). The binding of domain III ac to protein eS27 orients the IRES, prevents binding of eIF3 to the 40S, and positions domain III b for eIF3 binding. The interaction of domain III f with protein eS28 attaches the IRES at the mRNA channel such that the start codon of the mRNA is properly positioned in the P-site. Finally, the structure reveals two contact areas between the flexible domain II of the IRES and the 40S, the first one comprising eS25 protein, which plays a role in translation initiation by several other IRESs, and the second one between the apical loop of domain II and the E-site of the 40S subunit, where the E-site tRNA would reside during the elongation stage of translation.

Hepatitis C is a very widespread disease, for which no satisfactory treatment has been developed so far. Several studies have attempted to inhibit translation of the HCV genome with an RNA structure that competes with the IRES for 40S binding<sup>35</sup> or small molecules that bind to and conformationally arrest the HCV IRES<sup>36</sup>. The detailed structural information of the interactions between the IRES and the ribosome presented here will help in further development of anti-HCV drugs and provide fundamental insights in the mechanisms of IRES initiation.

## Methods

**Cultivation of HEK293-6E cells.** The HEK293-6E cell line<sup>37</sup> was grown as suspension culture in batch and subsequent perfusion mode in an autoclavable stirred 6-l bioreactor. The bioreactor was equipped with a double-membrane stirrer using Accurel S6/2 tubing (Membrana, Wuppertal, Germany) for aeration as well as for perfusion with internal cell retention using the same membrane in hydrophilized mode<sup>38</sup>. Cells were cultivated at 37 °C using 20-l FreeStyle F17 medium (Life technologies) supplemented with 7.5 mM Glutamine. Aeration was controlled at a dissolved oxygen concentration of 40 and 5% CO<sub>2</sub>. Cells were grown until a cell density of  $2.5 \times 10^7$  cells per ml. The biomass of 6 l culture was collected using a Sorvall RC12BP centrifuge for 30 min at 1,000 r.p.m. The pellets were rinsed with standard PBS buffer, fractionated and transferred to 50-ml Falcon tubes. All steps from collecting were performed at 4 °C. After centrifugation in a Megafuge 40R for 30 min at 1,000 r.p.m. Pellets (15-g wet cells weight per tube) were shock frozen in liquid nitrogen and stored at –80 °C until further use.

**Sample preparation and data acquisition.** Human ribosomes were prepared similar to previously described methods<sup>39</sup>. In brief, frozen HEK-293 cells were thawed in threefold excess of lysis buffer (50 mM HEPES pH 7.6, 0.5% NP40, 6 mM MgCl<sub>2</sub>, 300 mM NaCl, 50 nM E-64, 20  $\mu$ M leupeptin, 20  $\mu$ M bestatin, 5  $\mu$ M pepstatin A, 20  $\mu$ M phenanthroline, 1  $\mu$ M PMSF and 2 mM DTT) and incubated on ice for 30 min. The lysate was cleared for 10 min at 12,000 g and the supernatant loaded onto a 60% (w/v) sucrose cushion containing 50 mM HEPES pH 7.6, 50 mM KCl, 10 mM MgCl<sub>2</sub> and 5 mM EDTA. Ribosomes were pelleted by centrifugation at

50,000 r.p.m. in a Ti70 rotor (Beckman Coulter) for 20 h. The supernatant was discarded and the pellet resuspended in 1 ml resuspension buffer (50 mM HEPES pH 7.6, 6 mM MgCl<sub>2</sub>, 300 mM KCl, 1 mM DTT, 6.8% sucrose). The resuspended pellet was centrifuged again for 20 min at 20,000 g and the supernatant treated with 1 mM puromycin for 30 min on ice. The ribosomal sample was then loaded onto a 10–40% (w/v) sucrose gradient in 50 mM HEPES pH 7.6, 2 mM MgCl<sub>2</sub>, 150 mM KCl and 1 mM DTT and centrifuged at 24,000 r.p.m. for 17 h in a SW32 rotor. The gradient was fractionated and fractions containing 80S ribosomes were pooled and concentrated as well as buffer exchanged into 80S buffer (20 mM HEPES pH 7.6, 100 mM KCl, 5 mM MgCl<sub>2</sub>) using 100,000 kDa MWCO tabletop-centrifuge concentrators (Sartorius).

HCV IRES was produced by *in vitro* transcription of a linearized plasmid containing the HCV IRES sequence, followed by denaturing polyacrylamide electrophoresis and buffer exchange into water using tabletop-centrifuge concentrators. The IRES was then diluted to 1 mg ml<sup>–1</sup> in folding buffer (20 mM HEPES pH 7.6, 100 mM K-acetate, 2.5 mM MgCl<sub>2</sub>, 0.25 mM spermidine) and folded by heating to 95 °C for one minute and cooling on ice.

The 80S–HCV IRES complex was assembled by adding 300 nM HCV IRES to 100 nM 80S in 80S buffer and incubation for 10 min at 37 °C. The sample was then applied to Quantifoil R2/2 holey carbon grids (Quantifoil Micro Tools) coated with a thin carbon film (prepared using a Balzers BAE 120 thin-film coating system). Immediately, the grids were blotted with damp filter paper from the front and manually plunged into a 1:2 mixture of ethane and propane at liquid nitrogen temperature<sup>40</sup>. Data collection was performed using a Titan Krios cryo-transmission electron microscope (FEI Company) operated at 300 kV equipped with a Falcon II direct electron detector (FEI Company) at a magnification of  $100,719 \times$  (accuracy of magnification calibration approximately 0.5%) and defocus values of –1.5 to –3.4  $\mu$ m. For automated data collection the EPU software (FEI Company) was used. Images were acquired in movie mode using seven movie frames (combined dose 20 electrons per  $\text{\AA}^2$ ) per exposure (combined exposure time 0.77 s) after discarding the first frame (55 ms). Subsequently, the movie frames were aligned using the DOSEGPU DRIFTCORR software<sup>41</sup> to correct for beam-induced specimen motion.

**Structure calculation.** To discard poor regions on the grid, overview images of the holes were inspected. CTF parameters of the images were estimated with CTFFIND<sup>42</sup> from within RELION<sup>43</sup> and the resulting power spectra analysed to discard poor quality micrographs. A total of 638,532 particles was selected with Batchboxer (EMAN 1.9)<sup>44</sup>. Using RELION, an initial round of two-dimensional classification was performed with 100 classes at a pixel size of 5.6  $\text{\AA}$  on the object scale (80-pixel frames). Classes containing 80S ribosomes were selected and subjected to three-dimensional classification with 10 classes at the same pixel size using a low-pass filtered reconstruction of the 80S ribosome as an initial reference (Supplementary Fig. 1). Again, only homogeneous 80S classes were selected, resulting in a dataset of 404,357 particles. A test-refinement without a mask using only 44,801 particles in RELION yielded a 4.2  $\text{\AA}$  reconstruction of the 80S–HCV IRES complex, however, the resolution of the 40S and the HCV IRES was significantly lower as the alignment of the particles is dominated by the 60S subunit. Therefore, the full data set was refined using focused alignment in RELION by applying a mask around the 40S and the HCV IRES, as described previously<sup>34</sup>. A 3.9  $\text{\AA}$  resolution (according to a gold standard refinement procedure<sup>45</sup>) reconstruction of the 40S ribosomal subunit was obtained with clear density for the HCV IRES.

**Model building, refinement and validation.** Model building was performed in Coot<sup>46</sup> and O<sup>47</sup> using the structure of the pig 40S ribosomal subunit (pdb code 4W23)<sup>19</sup> as the starting model. As the head and body were in a different conformation in the pig apo-40S ribosomal subunit, they were individually rigid body fitted into the density using Chimera<sup>48</sup>. The sequences of the rRNA and the proteins were mutated to the human sequence. A rearrangement was observed in the structure of expansion segment 7 (ES7), which was completely rebuilt, and many obvious errors in both the 18S rRNA and the ribosomal proteins were corrected using the 3.0  $\text{\AA}$  yeast<sup>49</sup> and the 3.7  $\text{\AA}$  *Tetrahymena thermophila*<sup>50</sup> structures as a guide. While the appearance of bulky side chains in the EM density allowed the correct assignment of the protein register also in areas of lower local resolution, the exact side chain conformations could not be established for all amino acids with certainty in these regions. As a starting point for building the HCV IRES, the previously published model (4UJC)<sup>10</sup> was used as well as the crystal or NMR structures of individual domains (domain II (1P5P)<sup>14</sup>, domain III c (11DV)<sup>17</sup>, domain III abc (1KH6)<sup>16</sup>, domain III d (1F84)<sup>15</sup> and III e (3T4B)<sup>13</sup>).

The resulting model was refined against back-calculated structure factors using PHENIX, applying the mhl target function as described<sup>51,52</sup>. Briefly, after one round of rigid body refinement using individual proteins, 18S rRNA segments and the HCV IRES as groups, the structure was refined to convergence using nine cycles of individual coordinate and grouped B-factor refinement (two groups per residue). To preserve proper geometry also in areas of lower local resolution, automatically detected Ramachandran, secondary structure and base-pair restraints were applied throughout. To prevent over-refinement, an appropriate weighting of the geometry versus the crystallographic term was established empirically, aiming at good model geometry with a low *R* value (Supplementary Table S1). To validate

the structure, the final model was refined against the two half-set maps using a similar procedure after resetting the B-factors and applying a random coordinate shift of 0.5 Å (Supplementary Fig. 2d).

**Creation of figures.** Figures of molecular structures were created using the UCSF Chimera package from the Computer Graphics Laboratory, University of California, San Francisco (supported by NIH P41 RR-01081)<sup>48</sup> and PyMOL (The PyMOL Molecular Graphics System, Version 1.7 Schrödinger, LLC.). Local resolution plots were generated in ResMap<sup>53</sup>.

## References

- Gravitz, L. Introduction: a smouldering public-health crisis. *Nature* **474**, S2–S4 (2011).
- Hinnebusch, A. G. The scanning mechanism of eukaryotic translation initiation. *Annu. Rev. Biochem.* **83**, 779–812 (2014).
- Kieft, J. S., Zhou, K., Jubin, R. & Doudna, J. A. Mechanism of ribosome recruitment by hepatitis C IRES RNA. *RNA* **7**, 194–206 (2001).
- Pestova, T. V., Shatsky, I. N., Fletcher, S. P., Jackson, R. J. & Hellen, C. U. A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initiation of hepatitis C and classical swine fever virus RNAs. *Genes Dev.* **12**, 67–83 (1998).
- Otto, G. A. & Puglisi, J. D. The pathway of HCV IRES-mediated translation initiation. *Cell* **119**, 369–380 (2004).
- Brown, E. A., Zhang, H., Ping, L. H. & Lemon, S. M. Secondary structure of the 5' nontranslated regions of hepatitis C virus and pestivirus genomic RNAs. *Nucleic Acids Res.* **20**, 5041–5045 (1992).
- Spahn, C. M. *et al.* Hepatitis C virus IRES RNA-induced changes in the conformation of the 40S ribosomal subunit. *Science* **291**, 1959–1962 (2001).
- Boehringer, D., Thermann, R., Ostareck-Lederer, A., Lewis, J. D. & Stark, H. Structure of the hepatitis C virus IRES bound to the human 80S ribosome: remodeling of the HCV IRES. *Structure* **13**, 1695–1706 (2005).
- Hashem, Y. *et al.* Hepatitis-C-virus-like internal ribosome entry sites displace eIF3 to gain access to the 40S subunit. *Nature* **503**, 539–543 (2013).
- Yamamoto, H. *et al.* Structure of the mammalian 80S initiation complex with initiation factor 5B on HCV-IRES RNA. *Nat. Struct. Mol. Biol.* **21**, 721–727 (2014).
- Locker, N., Easton, L. E. & Lukavsky, P. J. HCV and CSFV IRES domain II mediate eIF2 release during 80S ribosome assembly. *EMBO J* **26**, 795–805 (2007).
- Filbin, M. E., Vollmar, B. S., Shi, D., Gonen, T. & Kieft, J. S. HCV IRES manipulates the ribosome to promote the switch from translation initiation to elongation. *Nat. Struct. Mol. Biol.* **20**, 150–158 (2013).
- Berry, K. E., Waghay, S., Mortimer, S. A., Bai, Y. & Doudna, J. A. Crystal structure of the HCV IRES central domain reveals strategy for start-codon positioning. *Structure* **19**, 1456–1466 (2011).
- Lukavsky, P. J., Kim, I., Otto, G. A. & Puglisi, J. D. Structure of HCV IRES domain II determined by NMR. *Nat. Struct. Mol. Biol.* **10**, 1033–1038 (2003).
- Lukavsky, P. J., Otto, G. A., Lancaster, A. M., Sarnow, P. & Puglisi, J. D. Structures of two RNA domains essential for hepatitis C virus internal ribosome entry site function. *Nat. Struct. Mol. Biol.* **7**, 1105–1110 (2000).
- Kieft, J. S., Zhou, K., Grech, A., Jubin, R. & Doudna, J. A. Crystal structure of an RNA tertiary domain essential to HCV IRES-mediated translation initiation. *Nat. Struct. Mol. Biol.* **9**, 370–374 (2002).
- Rijnbrand, R., Thiviyanathan, V., Kaluarachchi, K., Lemon, S. M. & Gorenstein, D. G. Mutational and structural analysis of stem-loop IIIIC of the hepatitis C virus and GB virus B internal ribosome entry sites. *J. Mol. Biol.* **343**, 805–817 (2004).
- Fuchs, G. *et al.* Kinetic pathway of 40S ribosomal subunit recruitment to hepatitis C virus internal ribosome entry site. *Proc. Natl Acad. Sci. USA* **112**, 319–325 (2015).
- Voorhees, R. M., Fernandez, I. S., Scheres, S. H. & Hegde, R. S. Structure of the mammalian ribosome-Sec61 complex to 3.4 Å resolution. *Cell* **157**, 1632–1643 (2014).
- Hellen, C. U. & de Breyne, S. A distinct group of hepacivirus/pestivirus-like internal ribosomal entry sites in members of diverse picornavirus genera: evidence for modular exchange of functional noncoding RNA elements by recombination. *J. Virol.* **81**, 5850–5863 (2007).
- Matsuda, D. & Mauro, V. P. Base pairing between hepatitis C virus RNA and 18S rRNA is required for IRES-dependent translation initiation in vivo. *Proc. Natl Acad. Sci. USA* **111**, 15385–15389 (2014).
- Luttermann, C. & Meyers, G. The importance of inter- and intramolecular base pairing for translation reinitiation on a eukaryotic bicistronic mRNA. *Genes Dev.* **23**, 331–344 (2009).
- Psaridi, L., Georgopoulou, U., Varaklioti, A. & Mavromara, P. Mutational analysis of a conserved tetraloop in the 5' untranslated region of hepatitis C virus identifies a novel RNA element essential for the internal ribosome entry site function. *FEBS Lett.* **453**, 49–53 (1999).
- Easton, L. E., Locker, N. & Lukavsky, P. J. Conserved functional domains and a novel tertiary interaction near the pseudoknot drive translational activity of hepatitis C virus and hepatitis C virus-like internal ribosome entry sites. *Nucleic Acids Res.* **37**, 5537–5549 (2009).
- Melcher, S. E., Wilson, T. J. & Lilley, D. M. The dynamic nature of the four-way junction of the hepatitis C virus IRES. *RNA* **9**, 809–820 (2003).
- Hashem, Y. *et al.* Structure of the mammalian ribosomal 43S preinitiation complex bound to the scanning factor DHX29. *Cell* **153**, 1108–1119 (2013).
- Berry, K. E., Waghay, S. & Doudna, J. A. The HCV IRES pseudoknot positions the initiation codon on the 40S ribosomal subunit. *RNA* **16**, 1559–1569 (2010).
- Yusupov, M. M. *et al.* Crystal structure of the ribosome at 5.5 Å resolution. *Science* **292**, 883–896 (2001).
- Polikanov, Y. S. *et al.* Negamycin interferes with decoding and translocation by simultaneous interaction with rRNA and tRNA. *Mol. Cell.* **56**, 541–550 (2014).
- Filbin, M. E. & Kieft, J. S. HCV IRES domain IIb affects the configuration of coding RNA in the 40S subunit's decoding groove. *RNA* **17**, 1258–1273 (2011).
- Kalliampakou, K. I., Psaridi-Linardaki, L. & Mavromara, P. Mutational analysis of the apical region of domain II of the HCV IRES. *FEBS Lett.* **511**, 79–84 (2002).
- Hussain, T. *et al.* Structural changes enable start codon recognition by the eukaryotic translation initiation complex. *Cell* **159**, 597–607 (2014).
- Landry, D. M., Hertz, M. I. & Thompson, S. R. RPS25 is essential for translation initiation by the Dicistroviridae and hepatitis C viral IRESs. *Genes Dev.* **23**, 2753–2764 (2009).
- Fernandez, I. S., Bai, X. C., Murshudov, G., Scheres, S. H. & Ramakrishnan, V. Initiation of translation by cricket paralysis virus IRES requires its translocation in the ribosome. *Cell* **157**, 823–831 (2014).
- Bhat, P. *et al.* Targeting ribosome assembly on the HCV RNA using a small RNA molecule. *RNA Biol.* **9**, 1110–1119 (2012).
- Dibrov, S. M. *et al.* Structure of a hepatitis C virus RNA domain in complex with a translation inhibitor reveals a binding mode reminiscent of riboswitches. *Proc. Natl Acad. Sci. USA* **109**, 5223–5228 (2012).
- Tom, R., Bisson, L. & Durocher, Y. Culture of HEK293-EBNA1 cells for production of recombinant proteins. *CSH Protoc.* **2008** doi:10.1101/pdb.prot4976 (1 March 2008).
- Blasey, H. D. & Jaeger, V. in *Animal Cell Culture and Production of Biologicals*. (eds Sasaki, R. & Ikura, R.) (Kluwer Academic Publisher, 1991).
- Khatter, H. *et al.* Purification, characterization and crystallization of the human 80S ribosome. *Nucleic Acids Res.* **42**, e49 (2014).
- Tivol, W. F., Briegel, A. & Jensen, G. J. An improved cryogen for plunge freezing. *Microsc. Microanal.* **14**, 375–379 (2008).
- Li, X. *et al.* Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat. Methods* **10**, 584–590 (2013).
- Mindell, J. A. & Grigorieff, N. Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol.* **142**, 334–347 (2003).
- Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).
- Ludtke, S. J., Baldwin, P. R. & Chiu, W. EMAN: semiautomated software for high-resolution single-particle reconstructions. *J. Struct. Biol.* **128**, 82–97 (1999).
- Scheres, S. H. & Chen, S. Prevention of overfitting in cryo-EM structure determination. *Nat. Methods* **9**, 853–854 (2012).
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
- Jones, T. A. Interactive electron-density map interpretation: from INTER to O. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2115–2125 (2004).
- Pettersen, E. F. *et al.* UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
- Ben-Shem, A. *et al.* The structure of the eukaryotic ribosome at 3.0 Å resolution. *Science* **334**, 1524–1529 (2011).
- Weisser, M., Voigts-Hoffmann, F., Rabl, J., Leibundgut, M. & Ban, N. The crystal structure of the eukaryotic 40S ribosomal subunit in complex with eIF1 and eIF1A. *Nat. Struct. Mol. Biol.* **20**, 1015–1017 (2013).
- Greber, B. J. *et al.* The complete structure of the large subunit of the mammalian mitochondrial ribosome. *Nature* **515**, 283–286 (2014).
- Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
- Kucukelbir, A., Sigworth, F. J. & Tagare, H. D. Quantifying the local resolution of cryo-EM density maps. *Nat. Methods* **11**, 63–65 (2014).

## Acknowledgements

We thank the ETH Zürich Scientific Center for Optical and Electron Microscopy (ScopEM) for access to EM equipment and are indebted to P. Tittmann for technical support. N.Q. was supported by a European Molecular Biology Organization (EMBO) long-term fellowship. This work was supported by the Swiss National Science Foundation (SNSF), the National Center of Excellence in Research Structural Biology and RNA and Disease programs of the SNSF, and European Research Council grant 250071 (to N.B.) under the European Community's Seventh Framework Programme.



The large-scale cultivation was performed at the Helmholtz Protein Sample Production Facility, Braunschweig, Germany and supported by Instruct, part of the European Strategy Forum on Research Infrastructures (ESFRI). We would like to thank Basil Greber and Christopher H. Aylett for critically reading our manuscript.

### Author contributions

J.v.d.H. produced the HEK293 cells. N.Q. performed preparation of ribosome samples and cryo-EM grids. N.Q. and D.B. acquired cryo-EM data and calculated cryo-EM reconstructions. N.Q. and M.L. built the atomic model. M.L. performed coordinate refinement of the atomic model. N.Q., M.L., D.B. and N.B. interpreted the structure. All authors contributed to the final version of the paper.

### Additional information

**Accession codes:** The cryo-EM map of the 40S-HCV IRES complex has been deposited in the EMDB with the accession code EMD-3019 and the structure of the 40S-HCV IRES complex has been deposited in the PDB with the accession code 5A2Q.

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

**Competing financial interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**How to cite this article:** Quade, N. *et al.* Cryo-EM structure of Hepatitis C virus IRES bound to the human ribosome at 3.9-Å resolution. *Nat. Commun.* 6:7646 doi: 10.1038/ncomms8646 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>