nature structural & molecular biology

Article

https://doi.org/10.1038/s41594-023-01196-0

The structure of a human translation initiation complex reveals two independent roles for the helicase eIF4A

Received: 10 January 2023

Accepted: 30 November 2023

Published online: 29 January 2024

Check for updates

Jailson Brito Querido[1,3,4], Masaaki Sokabe[2,4], Irene Díaz-López[1], Yuliya Gordiyenko[1], Christopher S. Fraser[2] ✉ & V. Ramakrishnan[1] ✉

Eukaryotic translation initiation involves recruitment of the 43S pre-initiation complex to the 5′ end of mRNA by the cap-binding complex eIF4F, forming the 48S translation initiation complex (48S), which then scans along the mRNA until the start codon is recognized. We have previously shown that eIF4F binds near the mRNA exit channel of the 43S, leaving open the question of how mRNA secondary structure is removed as it enters the mRNA channel on the other side of the 40S subunit. Here we report the structure of a human 48S that shows that, in addition to the eIF4A that is part of eIF4F, there is a second eIF4A helicase bound at the mRNA entry site, which could unwind RNA secondary structures as they enter the 48S. The structure also reveals conserved interactions between eIF4F and the 43S, probaby explaining how eIF4F can promote mRNA recruitment in all eukaryotes.

Initiation of translation in eukaryotes involves over 20 different eukaryotic initiation factors (eIFs). The process starts with the assembly of the 43S pre-initiation complex (43S) consisting of the 40S ribosomal subunit bound to initiation factors eIF1, eIF1A, eIF3 and eIF5 and a ternary complex (TC) of eIF2, guanosine 5′-triphosphate (GTP), and methionyl initiator transfer RNA ($tRNA_i^{Met}$). In parallel, the cap-binding complex eIF4F bound to the 5′ end of messenger RNA recruits the 43S to form the 48S translation initiation complex, which then scans along the mRNA until it encounters a start codon[1–8].

The cap-binding complex eIF4F consists of a scaffold protein eIF4G, an m^7G cap-binding protein eIF4E, and a DEAD-box helicase eIF4A[9]. Additionally, eIF4G interacts with poly(A)-binding protein PABP located at the 3′ untranslated region (UTR) of mRNA, which further brings the 3′ and 5′ ends close together. In metazoans, the attachment of the 43S to the mRNA–eIF4F–PABP complex is mediated by a direct interaction between eIF3 and eIF4G[3–5,8,10,11]. Recently, we determined the structure of a human 48S complex during scanning, which revealed the binding site of eIF4F[4]. While the resolution was modest, the structure showed that eIF4F interacts with eIF3e and eIF3k/l located upstream of the 43S, near the mRNA channel exit site. This finding was consistent with a slotting mechanism of mRNA recruitment and explained the presence of a blind spot between the m^7G cap structure and the recognition of an initiation codon. Nevertheless, the location of eIF4F in the 48S left an unanswered question about how eIF4A that is part of eIF4F behind the scanning 40S subunit could act as a helicase during scanning and actively unwind secondary structure before it enters the mRNA binding channel.

The helicase activity of eIF4A is stimulated by eIF4G and eIF4B[12–14]. In addition, *Saccharomyces cerevisiae* eIF4B probably induces conformational changes in the 40S subunit to facilitate the attachment of the 43S to mRNA. Yet, it is unclear how eIF4B interacts with eIF4A and/or eIF4F on the surface of the 40S subunit. Interestingly, the ATPase activity of yeast eIF4A does not appear to require eIF4G and eIF4E[15], whereas in mammals recruitment and scanning require the cap-binding complex eIF4F[13]. Thus, it remained uncertain how eIF4A, eIF4B and the eIF4F complex could work together to unwind secondary structure during mRNA recruitment and scanning.

In our prior structure, we used an mRNA that lacked an AUG codon and was shorter than the footprint of the 48S predicted by the structure[4]. In this Article, we capture a later-stage initiation complex of

[1]MRC Laboratory of Molecular Biology, Cambridge, UK. [2]Department of Molecular and Cellular Biology, College of Biological Sciences, University of California, Davis, CA, USA. [3]Present address: Department of Biological Chemistry and Life Sciences Institute, University of Michigan, Ann Arbor, MI, USA. [4]These authors contributed equally: Jailson Brito Querido, Masaaki Sokabe. ✉e-mail: csfraser@ucdavis.edu; ramak@mrc-lmb.cam.ac.uk

the 48S complex positioned over a start codon, using a longer, more physiological mRNA with a long and structured 5′ UTR, AUG codon and a 3′ UTR ending in a poly(A) tail. Unexpectedly, a second eIF4A helicase is bound at the mRNA entry channel of the 48S complex. This eIF4A helicase interacts with universally conserved eIF3 subunits and is entirely separate from the eIF4F complex positioned at the mRNA exit site on the other side of the 40S subunit. The structure identifies the interaction of eIF4B with the entry site bound eIF4A, providing insight into its role in regulating this second eIF4A rather than eIF4F during scanning. The discovery of a second eIF4A molecule positioned at the point of entry of mRNA into the mRNA binding channel of the 40S subunit resolves many seemingly contradictory data, including the unanswered question of how secondary structure is unwound by the helicase activity of eIF4A during scanning.

Structure of a human 48S translation initiation complex

To reconstitute a 48S complex as complete and physiological as possible for cryo-electron microscopy (cryo-EM) analysis, we use a capped mRNA containing a long 5′ UTR (105 nucleotides), an AUG codon, followed by a short coding sequence (9 nucleotides) and a poly(A) tail (~90 nucleotides) (Extended Data Fig. 1). To promote the proposed interaction of the poly(A) tail with the 5′ end of mRNA during complex formation, we included PABP in addition to a full complement of initiation factors. Inclusion of a start codon enables us to generate a late-stage initiation intermediate representing the 48S with the start codon in the P site. To stabilize the complex further, we used Rocaglamide A (RocA), which clamps eIF4A onto polypurine sequences[16–19], while also including three polypurine motifs $(GA)_6$ in the 5′ UTR of the mRNA.

Single-particle reconstruction using cryo-EM reveals density corresponding to the cap-binding complex eIF4F in the same location as seen in our previous structure (Fig. 1 and Extended Data Fig. 2)[4]. Masked classification on the mRNA channel entry site followed by additional masked classification on the region we previously saw for eIF4F yielded a cryo-EM map at 3.1 Å resolution (Extended Data Fig. 3a). Further masked classification on TC yielded a map with a reduced overall resolution of 3.5 Å (Extended Data Fig. 3b) but improved the density for TC.

We observe additional density extending out from the eIF3bgi module at the mRNA channel entry site (Fig. 1). This density was not present in any previous structures of 48S complexes. The local resolution and detailed shape indicate that this additional density is due to a second eIF4A helicase bound to mRNA, in addition to the eIF4A that is part of eIF4F on the other side of the 40S subunit (Fig. 1). To test whether the presence of this second, entry-site eIF4A is a result of its being trapped by RocA, we determined the structure of a complex without RocA and found that both eIF4F and the second eIF4A are still present (Extended Data Fig. 4). To assemble this complex, we used a capped mRNA with the 5′ UTR of β-globin (55 nucleotides) followed by a coding sequence (124 nucleotides) containing a GC-rich region located 27 nucleotides downstream of the AUG. This GC-rich region forms a downstream loop (DLP) in the mRNA and has been used to trap eIF4A in the 48S[20].

The local resolution of 5.9 Å for the cap-binding complex eIF4F after focus refinement (Extended Data Fig. 3c,d) is greatly improved compared to our previous structure[4], and allows us to see direct interactions not only to additional subunits of eIF3 but unexpectedly to the 40S small ribosomal subunit. The improved density shows secondary structure elements and enables us to accurately model the middle domain of eIF4G (HEAT-1 domain) and eIF4A (Fig. 1). Although the mRNA has an m^7G cap and a poly(A) tail, we did not see any additional density that could be assigned to eIF4E or PABP.

The overall conformation of the structure is a post-scanning intermediate—the initiator tRNA is inserted completely into the P site (a P_{in} state) where it bases pairs with the start codon, and the 40S subunit mRNA binding channel is in the closed conformation. Additionally, the overall resolution allowed us to identify density corresponding to the 40S, eIF1A, eIF2a, eIF2b, eIF2g, $tRNA_i^{Met}$, the octameric structural core of eIF3 and its peripheral subunits (b, d, g and i; Fig. 1 and Table 1).

A second eIF4A helicase in the 48S

Compared to our previous 48S structure, we observed additional density at the entry site of the mRNA channel (Fig. 2a,b). Much of this density can be accounted for by eIF4A, which after rigid-body fitting of the known crystal structure of human eIF4A[16] shows close agreement with the density, especially the eIF4A-CTD, for which almost the entire secondary structure was resolved (Fig. 2b). The density for eIF4A-NTD (where 'NTD' is the N-terminal domain) was weaker, presumably because the domain was more disordered. The assignment of this density to eIF4-NTD was dictated by the placement of eIF4A-CTD, the architecture of eIF4A and the orientation of the mRNA (from 5′ to 3′) relative to the ribosome.

eIF4A binds to a pocket formed by eIF3bgi, eIF3a-CTD and ribosomal RNA (rRNA) h16 (Fig. 2c,d). This position of eIF4A agrees well with previous biochemical, cross-link mass spectrometry and proximity labeling data[4,15,21]. Moreover, recent single-molecule data revealed FRET (fluorescence resonance energy transfer) between eIF4A and ribosomal uS19 (ref. 22); the latter is located within FRET distance (~97 Å) of this second eIF4A in our structure.

The two RecA domains of the entry site eIF4A adopt an open conformation that differs from that described for the yeast eIF4A–eIF4G complex[23] (Extended Data Fig. 5). Instead, the conformation of this entry site eIF4A resembles that of the crystallographic structure of the eIF4A–Pdcd4 (Programmed cell death 4) complex[24] (Extended Data Fig. 5). This conformation allows the NTD and CTD to contact the mRNA while the nucleotide-binding domain adopts a more open conformation. The structure shows that eIF4A could remain bound to mRNA as it cycles through the open and closed conformations during scanning, which may help explain the factor-dependent processivity of human eIF4A described previously[14].

Unlike previous 48S structures in which mRNA outside the 40S subunit was not seen, there is clear and continuous density for the mRNA, which allowed us to trace its path for the first time from the cap-binding complex eIF4F upstream of the 43S, through the 40S subunit mRNA channel all the way to the second molecule of eIF4A downstream of the 43S (~40 nt upstream and ~20 nt downstream of the AUG; Fig. 2a,b).

It has been shown that yeast 43S stimulates the ATPase activity of eIF4A[15]. Interestingly, this activity of yeast eIF4A does not require eIF4G or eIF4E[15]. Instead, eIF3, especially eIF3i and eIF3g, are required to stimulate the ATPase activity of eIF4A. Because the cap-binding complex eIF4F binds upstream of the 43S, it was unclear how eIF3 subunits apparently downstream of the 43S could affect the activity of eIF4A. However, it has been shown that the eIF3bgi complex is dynamic, binding to both the mRNA entry channel and the subunit interface[25]. Thus, it was not clear whether the eIF3bgi domain stimulates the eIF4A helicase directly or indirectly on the 40S subunit surface. This is now solved by revealing the direct interaction between the second molecule of eIF4A and the eIF3bgi module (Fig. 2). Indeed, the structure reveals that eIF3i is the main binding partner of this second eIF4A.

eIF3a binds to the solvent-exposed site of the 40S, near the mRNA channel exit site. However, its CTD extends toward the solvent side of the 40S, where it interacts with eIF3bgi at the mRNA channel entry site[26]. Our previous cross-linking mass spectrometry indicated that eIF3a-CTD (Lys632) is in close proximity to eIF4A-CTD (Lys291)[4]. Consistent with this observation, the structure shows a close proximity between $eIF3a_{630–646}$, a disordered region in the eIF3a-CTD and Lys291 of eIF4A-CTD (Fig. 2c,d). Indeed, the structure also explains prior biochemical data indicating that eIF3a-CTD and eIF3bgi play important role during recruitment and scanning[27]. This extensive

https://doi.org/10.1038/s41594-023-01196-0

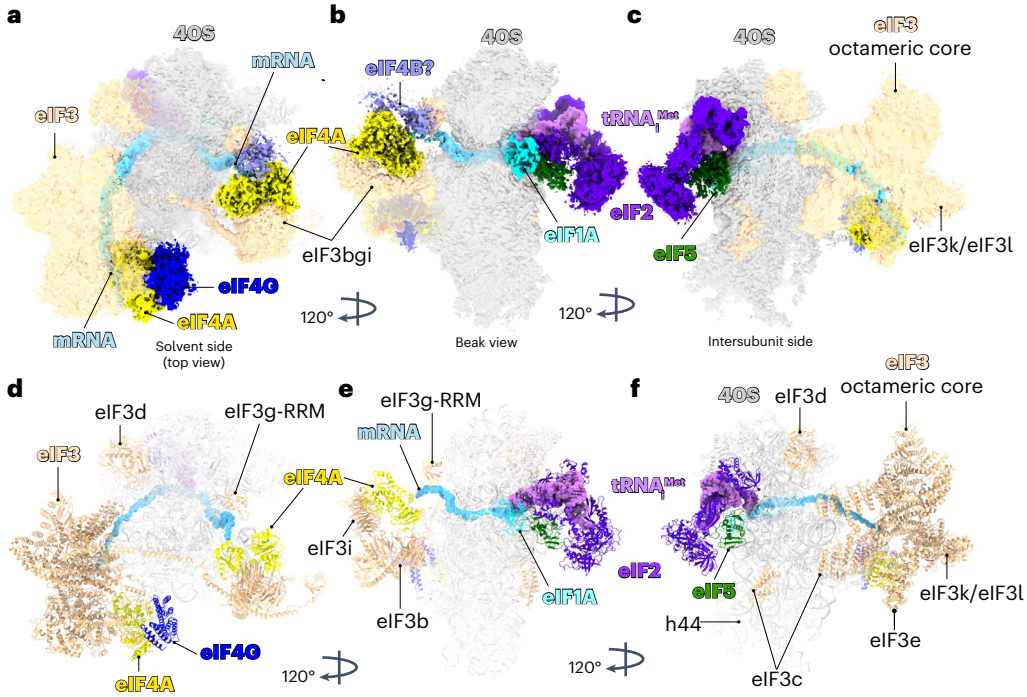


Fig. 1 | Cryo-EM 3D reconstruction and the structure of a human 48S complex. **a**–**c**, Overview of the cryo-EM density map shown in different orientations. **d**–**f**, Overview of the molecular model of human 48S shown in different orientations. The map contains densities for 40S small ribosomal subunit, eIF3, eIF1A, eIF5-NTD, eIF2, tRNA$_i^{Met}$, mRNA, eIF4F (eIF4G and eIF4A) and a second molecule of eIF4A bound to eIF3 at the mRNA channel entry site. The map also contains a possible density for eIF4B in contact with the second molecule of eIF4A at the entry site.

interaction network between eIF4A-CTD and eIF3 can also explain the local resolution disparity between the better-resolved eIF4A-CTD and the NTD. While these interactions stabilize the binding eIF4A-CTD at the entry site, it still allows the highly dynamic transitions between open and closed conformation, which would require a movement of eIF4A-NTD with respect to eIF4A-CTD. This is consistent with a prior single-molecule study suggesting a dynamic transition between open and closed conformation involves a movement of eIF4A-NTD relative to eIF4A-CTD[28].

Mammalian eIF4G has two eIF4A-binding domains, one located in the middle domain (HEAT-1) and a second one located at the C-terminal (HEAT-2)[29,30]. The eIF4G HEAT-1 is highly conserved among eukaryotes and is the core of the eIF4F complex, while the eIF4G HEAT-2, not present in *S. cerevisiae*, is poorly conserved and is not essential[29,30]. It has been proposed that HEAT-2 plays a stimulatory role[29]. To test whether HEAT-2 is required for binding of the second molecule of eIF4A, we analyzed a 48S complex with a truncated eIF4G lacking HEAT-2. The second molecule of eIF4A is still present in this complex (Extended Data Fig. 6a–c), suggesting that its binding at the mRNA entry site does not require eIF4G, consistent with previous biochemical data showing that *S. cerevisiae* 43S stimulates the ATPase activity of eIF4A in the absence of eIF4G[15].

A previous biochemical study proposed that eIF4F binds at the mRNA entry site[8]. However, at the entry site, the domain of eIF4A that interacts with eIF4G in the eIF4F complex is involved in interactions with eIF3i (Fig. 3a,b). The structure suggests that the position of eIF4A at the mRNA channel entry site is incompatible with its interaction with the eIF4G HEAT-1 seen in the cap-binding complex eIF4F (Fig. 3a,b). We, therefore, propose that this second molecule of eIF4A acts independently of eIF4F.

A possible location for eIF4B

By itself, eIF4A has low helicase activity. Even though its activity is modulated by the 43S[15], the helicase activity of eIF4A in scanning the 5′ UTR of mRNA with stable secondary structure is enhanced by the cofactor eIF4B or eIF4H[12,13]. The precise binding site of eIF4B on the 48S remained elusive. Our cryo-EM map contains an unassigned density adjacent to the eIF4A-NTD, in close contact with the eIF3g-RNA recognition motif (RRM) and ribosomal protein uS10 (Fig. 3c,d). Although the resolution of the additional density is low, its size and shape are consistent with the eIF4B-RRM[31] (Fig. 3c,d). Consistent with this interpretation, the cryo-EM map of a complex formed without eIF4B does not contain this additional density (Extended Data Fig. 6d–h). This location of eIF4B agrees well with previous biochemical data indicating that eIF4B or eIF4H interact with the same region of the eIF4A-NTD[28,32]. Furthermore, previous yeast two-hybrid analysis identified the ribosomal protein uS10, located in close proximity to the unassigned density, to be the main interaction partner of eIF4B[12]. Our data are also consistent with previous cross-linking mass spectrometry data, which revealed close proximity between eIF4B and ribosomal protein uS3[33]. This possible location, which suggests a direct interaction between eIF4B and the entry-site eIF4A, agrees well with the role of eIF4B in increasing the directionality of eIF4A translocation[14].

In addition to its role as a subunit of eIF4F, we propose that a second molecule of eIF4A at the mRNA channel entry site probably promotes mRNA recruitment and scanning. A second eIF4A in the 48S is also consistent with the finding that eIF4A is in molar excess over other components of eIF4F[34,35]. Nevertheless, even a second eIF4A in the 48S does not account for the substantial excess eIF4A compared to eIF4F in the cell[34,35]. Thus, it is possible that eIF4A could have additional functions apart from its role in the 48S complex. Consistent with this idea, the eIF4A–mRNA interaction has a long lifetime[36], and eIF4A is known to be able to melt RNA secondary structure in the absence of the 43S[34].

eIF4F interacts with eIF3 and ribosomal protein eS7

The recruitment of the 43S to the 5′ UTR of mRNA is a critical step in the translation of eukaryotic mRNAs. Our previous structure of a

Table 1 | Cryo-EM data collection, refinement and validation statistics

	#48S IC, (EMDB-17297) (PDB 8OZ0)
Data collection and processing	
Magnification	105,000
Voltage (kV)	300
Electron exposure (e^{-} Å^{-2})	1.197 (LMB) and 1.2391 (eBIC)
Defocus range (μm)	−1.2 to −3.0
Pixel size (Å)	0.826 (LMB) and 0.829 (eBIC)
Symmetry imposed	C1
Initial particle images (no.)	1,914,949 (LMB) and 422,381 (eBIC)
Final particle images (no.)	241,389
Map resolution (Å)	3.5
FSC threshold	0.143
Map resolution range (Å)	2.5 to 12.5
Refinement	
Initial model used (PDB code)	6ZMW
Model resolution (Å)	3.7
FSC threshold	0.5
Model resolution range (Å)	2.6 to 3.7
Map sharpening *B* factor (Å^2)	−10
Model composition	
Non-hydrogen atoms	121,610
Protein residues	11,916
Ligands	91
B factors (Å^2)	
Protein	48.27
Ligands	5,181
R.m.s.d.	
Bond lengths (Å)	0.011
Bond angles (°)	1.014
Validation	
MolProbity score	1.55
Clashscore	3.45
Poor rotamers (%)	0.13
Ramachandran plot	
Favored (%)	93.80
Allowed (%)	5.99
Disallowed (%)	0.21

FSC, Fourier Shell Correlation.

human 48S complex allowed us to identify the location of eIF4F[4], but the low local resolution and the adjacent highly flexible region of the 43S meant that we could only infer interactions with non-core subunits of eIF3 (eIF3e, eIF3k and eIF3l). Since these subunits do not exist in *S. cerevisiae*, it was also not clear how eIF4F could interact with the yeast 43S complex. The improved local resolution in this region allows us to build eIF4F more accurately and identify its interactions with the 43S in far greater detail, including additional ones that are probably universally conserved.

The density shows almost the entire secondary structure of the middle domain of eIF4G (HEAT-1) and eIF4A (Fig. 4a–c and Extended Data Fig. 7a). Rigid-body fitting of a crystal structure of a human eIF4A (Protein Data Bank (PDB): 5ZC9)[16] allowed us to assign its corresponding density (Fig. 4d–f). A model of the middle domain of eIF4G based on an AlphaFold prediction[37] was used for rigid-body fitting into the density (Fig. 4d–f and Extended Data Fig. 7a). The predicted model agreed well with the density, allowing us to locate the domain of eIF4G in the structure.

eIF4F is positioned on the 5′ side of mRNA relative to the 40S subunit, near the mRNA channel exit site. For the first time, mRNA interacting with eIF4A is clearly visible (Fig. 4a–c), and eIF4A adopts a structure intermediate between the extended conformation, found in a crystal structure of a yeast eIF4A–eIF4G complex without mRNA in which the two RecA domains are rotated away from each other[23], and the closed conformation of an eIF4A–mRNA dimeric complex[16]. This conformation allows the two RecA domains to interact with the mRNA while maintaining their interactions with eIF4G (Fig. 4a–f and Extended Data Fig. 7b,c).

In metazoans, eIF3 plays a critical role during the recruitment of 43S to the eIF4F–mRNA[3–5,8,10]. The structure reveals an unexpectedly large interaction network between eIF3 and eIF4F (Fig. 4g–i). Previously, because of high flexibility, the structure of eIF3l-NTD remained elusive. The improved local resolution reveals a near-complete structure of eIF3l-NTD and its interaction with eIF4G. Interestingly, the interaction between eIF3l-NTD and eIF4G occurs through a domain of eIF4G located between residues Met775 and Ala787, which was unexpected (Fig. 4g). Additionally, the cryo-EM map reveals an additional density connecting eIF4G with eIF3c and eIF3h (Fig. 4). Although the local resolution does not allow us to assign all the side chains in the density, it probably belongs to the C-terminal tail of eIF3c (Fig. 4c). The interaction between eIF3c and eIF3-binding domain of eIF4G was previously predicted by site-specific cross-linking[3]. The structure shows that this interaction occurs through the same eIF3-binding domain of eIF4G that is involved in the interaction with eIF3l (Fig. 4c,g).

Previously, eIF4F was observed to interact with the 43S complex through the eIF3 subunits -e, -k and -l[4]. However, these subunits are not present in *S. cerevisiae* eIF3, and eIF3k and eIF3l are dispensable in *Neurospora crassa* and *Caenorhabditis elegans*[38,39]. Thus, these previously identified interactions could not occur in all eukaryotes, raising the question of how eIF4F makes an interaction with the 43S in those species. In this higher-resolution structure, we see several additional interactions between eIF4F and the 43S complex that are likely to be universally conserved. One of these could be the eIF4G–eIF3c interaction, given that the eIF3c C-terminal tail includes residues highly conserved among eukaryotes, including in *S. cerevisiae* (Extended Data Fig. 8a), which also explains why a knockdown of eIF3c reduces the recruitment of the 43S[40].

The eIF4A component of eIF4F binds to a pocket formed by eIF3c, eIF3e, eIF3h and eIF3l (Fig. 4). We previously described interactions between eIF4A and eIF3e, and eIF3k/eIF3l[4]. The improved local resolution of this post-scanning complex allowed us to further characterize the interaction between eIF4A and the octameric structural core of eIF3. Of the two RecA domains of eIF4A, the C-terminal domain (CTD) is in close proximity to eIF3l, while the NTD makes contact with eIF3c, eIF3e and eIF3h.

The interactions between eIF4A and eIF3h raise many questions. eIF3h is known to interact with METTL3 and promote mRNA circularization during m^{6}A-dependent translation[41]. The two helices from eIF3h extend toward the solvent site of the 43S, but the role of eIF3h in initiation was not clear. The core of the mammalian eIF3 octameric structural core is formed by a seven-helix bundle consisting of two helices from eIF3h and one helix from eIF3c, eIF3e, eIF3f, eIF3l and eIF3k[4,42–45]. Here we see a direct interaction between eIF3h and eIF4A-NTD (Fig. 4i and Extended Data Fig. 8b,c) through which eIF3h may play a regulatory role in initiation (Fig. 4). Consistent with this idea, phosphorylation at residue Ser183 of eIF3h, located just above the domain that interacts with eIF4A (Extended Data Fig. 8), has been implicated in the process of reinitiation in plants[46].

https://doi.org/10.1038/s41594-023-01196-0

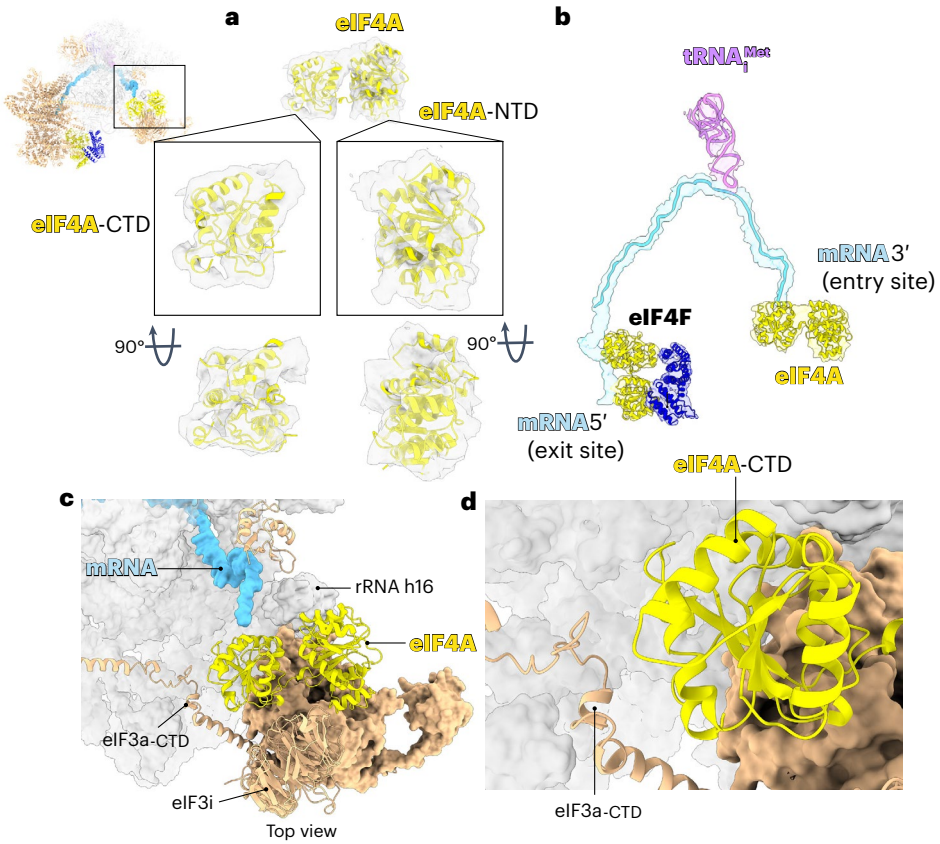


Fig. 2 | Structure of human 48S reveals a second molecule of eIF4A bound at the mRNA channel entry site. a, Atomic model of human eIF4A fitted (correlation 0.9024) into the cryo-EM map filtered to local resolution (6 to 10 Å). **b**, Atomic model of eIF4A, eIF4F, mRNA and tRNA$_i^{Met}$ fitted into a low-pass filter cryo-EM map. **c**, Interactions of eIF4A at the entry site. **d**, The interface between eIF4A-CTD and eIF3a-CTD.

In humans, the same phosphorylation has been associated with a possible oncogenic role of eIF3h[47]. It is possible that this post-translational modification of eIF3h affects its interaction with eIF4F. Furthermore, this domain of eIF3h also interacts with mRNA (Extended Data Fig. 8).

Unexpectedly, the structure reveals a direct interaction between eIF4A-NTD and ribosomal protein eS7 (Fig. 4j). The structure rationalizes previous studies on translational control whereby eS7 is monoubiquitinated by the human E3 ubiquitin (Ub) ligase CNOT4 or its yeast ortholog Not4 (refs. 48–51). Deubiquitination is required to allow the cap-binding complex eIF4F to bind to the 43S[52]. Here we see a direct interaction between eIF4A-NTD and eS7 (Fig. 4j,k). Superimposing our structure with a structure of Ub[53] at the predicted binding site in eS7 (refs. 52,53) places the Ub in a position where it could clash with eIF4A (Fig. 4k), thereby preventing the binding of the cap-binding complex eIF4F, which suggests why deubiquitination of eS7 may be required for the binding of eIF4F[52]. However, translation can still initiate through eIF4F-independent mechanisms[53].

The cryo-EM map also reveals some additional density between the mRNA and eIF3l (Fig. 4a,b). A portion of it resembles the shape of an A-form double helix, suggesting that a portion of the 5′ UTR forms a stem loop structure, but other parts resemble a protein tail and at this resolution, it was not possible to interpret it unambiguously. It is worth noting that eIF4F was present in the same location in a 48S complex assembled on an mRNA with β-globin 5′ UTR, (Extended Data Fig. 4), and also in our previous structure of a 48S complex using an mRNA with an unstructured 5′ UTR[4]. Thus, the location of eIF4F is not dependent on the unassigned density present here. It is possible that this yet uncharacterized structure may confer additional stability to eIF4F beyond that provided by its interactions with eIF3 and eS7 as well as affect its precise orientation.

Upon start-codon recognition, human eIF5-NTD replaces eIF1

Around the site normally occupied by eIF1, there is additional density that cannot be accounted for by the factor (Fig. 1, Extended Data Fig. 9). Both previous biochemical data and a structure of a yeast initiation complex showed that the yeast NTD of eIF5 (eIF5-NTD) replaces eIF1 upon start codon recognition[54,55]. Rigid-body fitting of the known structure of human eIF5-NTD (PDB: 2E9H) accounted for the entire density present at the platform of the 40S (Extended Data Fig. 9). Moreover, we could fit the zinc-binding domain of eIF5-NTD into the density (Extended Data Fig. 9). This domain is not present in eIF1, confirming that this density arises from bound eIF5-NTD that has displaced eIF1 upon start-codon recognition just as was previously proposed for the yeast system[54,55].

The structure therefore reveals an evolutionarily conserved set of events that occur on start-codon recognition, whose fidelity is enhanced by the factors eIF1, eIF1A and eIF5 (ref. 56). First, eIF1A and eIF1 bind to the 40S platform near the A and P sites, respectively, and promote the open conformation of the 40S that facilitates the binding to the tRNA$_i^{Met}$. During scanning, the tRNA$_i^{Met}$ is not fully inserted into the P site of the ribosome[4], which allows codon–anticodon sampling until recognition of the AUG start codon[4,57]. After start-codon recognition (Fig. 5a), the tRNA$_i^{Met}$ is inserted fully into the P site of the ribosome[42,43,57–61], resulting in a clash with eIF1. Thus, upon start-codon recognition, eIF1 dissociates from the complex and is replaced by eIF5-NTD (Fig. 5). Like eIF1, eIF5-NTD monitors the codon–anticodon interaction (Fig. 5b,c). However, while eIF1 residue Asn39 clashes with the anticodon stem-loop, the eIF5-NTD residue Asn30 adopts a different conformation (Fig. 5d,e), which allows it to monitor codon–anticodon interaction without clashing with the tRNA$_i^{Met}$ (Fig. 5c).

https://doi.org/10.1038/s41594-023-01196-0

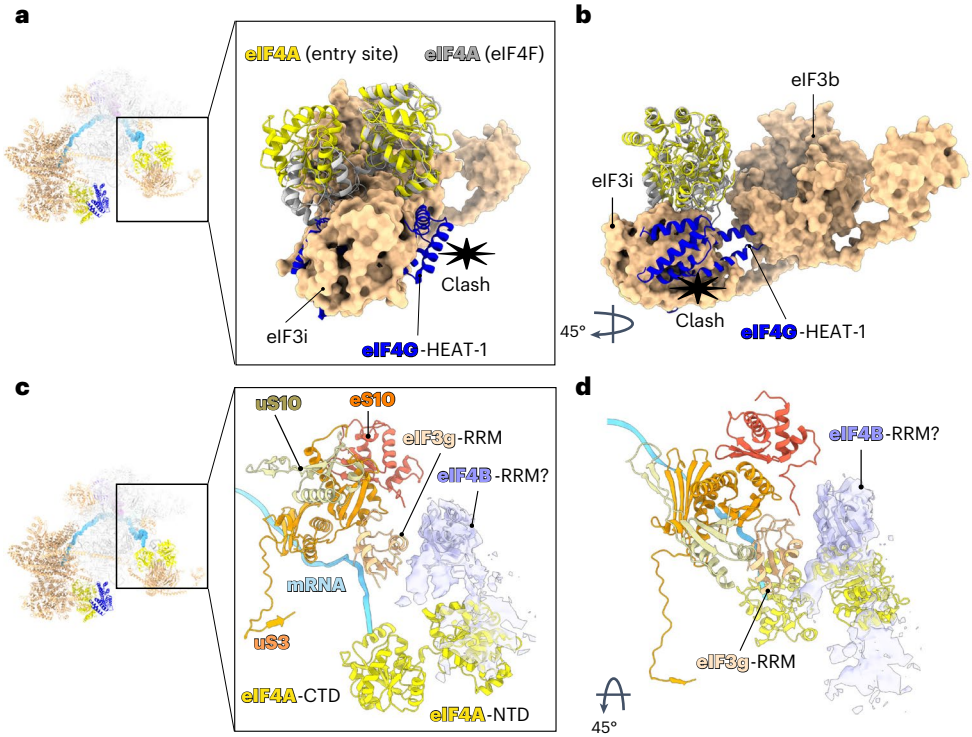


Fig. 3 | Conformational differences between the two molecules of eIF4A and a possible location of eIF4B in the 48S. a,**b**, Superposition of eIF4F with entry site eIF4A to highlight conformational differences shown in in different orientations. eIF4G-HEAT-1 would clash with eIF3i and eIF3b. **c**,**d**, Atomic model of human eIF4B-RRM (PDB: 2J76)[31] fit into the unassigned cryo-EM density map at the entry site shown in in different orientations. The possible location of eIF4B places it in close proximity to eIF4A-NTD, eIF3g-RRM, uS3, eS10 and uS10.

During scanning, eIF1 Loop 2 interacts with the $tRNA_i^{Met}$ D loop, which prevents the accommodation of the $tRNAi^{Met}$ into the P_{IN} state. Upon start codon selection, Loop 2 undergoes conformational changes to allow the accommodation of the $tRNA_i^{Met}$ in the P_{IN} state[57,62]. The destabilization of eIF1 and its replacement with eIF5-NTD avoids a clash with the $tRNA_i^{Met}$ and stabilizes it in the P_{IN} state because eIF5-NTD Loop 2 is both shorter and adopts a different conformation than that of eIF1 (Fig. 5d).

Discussion

In our structure, the cap-binding complex eIF4F binds upstream of the 43S complex, while a second molecule of eIF4A binds downstream, at the entry site of the mRNA binding channel. These locations suggest a model where eIF4A plays eIF4F-dependent and eIF4F-independent roles during mRNA recruitment[15] and scanning (Fig. 6a–c).

The cap-binding complex eIF4F at the 5′ end of mRNA plays a critical role in recruiting the 43S downstream of it, in a process that involves the mRNA being slotted into its channel in the 40S subunit (Fig. 6c). The higher resolution of the current 48S complex reveals how eIF4F makes universally conserved interactions with the 48S complex, including eIF3c and ribosomal protein eS7.

Canonical eIF4F-dependent initiation is enhanced by a topologically closed loop of the mRNA, often mediated by an interaction between eIF4G at the 5′ end and PABP at the 3′ end[63,64]. Such a topology would require a slotting mechanism for recruitment. Similarly, during m^6A-dependent translation, it has been shown that the mRNA is circularized through an interaction between METTL3 located at the 3′ of mRNA and eIF3h[41]. Other modes of initiation also require the slotting mechanism, including that via internal ribosome entry sites, eIF3d-dependent or circular mRNA. Some mRNAs with unusually short 5′ UTR (less than 40 nucleotides in mammals) may use recruitment pathways other than slotting. A recent study proposed that these mRNAs may use the slotting mechanism as well[65], which would require a backward movement of the 43S (3′ to 5′) rather than the small oscillations that are known to exist[22]. Thus, it is more likely that those mRNAs use an alternative recruitment pathway.

In addition to eIF4F, a second eIF4A helicase is coordinated by the 40S subunit and eIF3 so that it binds to mRNA before it enters the mRNA binding channel (Fig. 6c). The function of this entry-site-bound eIF4A may promote the accommodation of the mRNA into the entry channel, especially when secondary structure is present in the 5′ UTR[15,66]. Consistent with this idea, the cryo-EM 3D classification reveals a class of particles with eIF4A bound to the 43S but without the mRNA accommodated into the channel at the entry site (Extended Data Fig. 10a–d), suggesting that the binding of the second eIF4A to the 43S precedes the binding of the mRNA. Given the possible location of the eIF4B-RRM in our structure, it is possible that eIF4B may also play a role in the mRNA accommodation process[67].

The position of eIF4A at the entry site of the mRNA binding channel provides strong evidence that this conserved RNA helicase probably functions to unwind the mRNA secondary structure as it enters the mRNA binding channel during scanning. Thus, we propose a model whereby after recruitment, both eIF4A and eIF4F work in concert to ensure the directionality of scanning (Fig. 6d). In this model, eIF4F would prevent the reverse movement of the 43S complex[4] while the second molecule of eIF4A at the leading edge of the ribosome would use the energy from ATP hydrolysis to unwind downstream RNA secondary structure and translocate along the 5′ UTR of mRNA during scanning (Fig. 6d). It is possible that ATP hydrolysis by eIF4F might further help the movement of the 43S and help overcome energy barriers encountered by the entry channel eIF4A.

Compared to many other helicases, eIF4A has low helicase activity. Even though its ATPase activity is stimulated by the 43S and its cofactor eIF4B, scanning of 5′ UTR of mRNA with highly stable secondary structure may require additional helicases, such as mammalian DHX29, which has been seen to bind at the mRNA channel entry site[44,68]. In our

https://doi.org/10.1038/s41594-023-01196-0

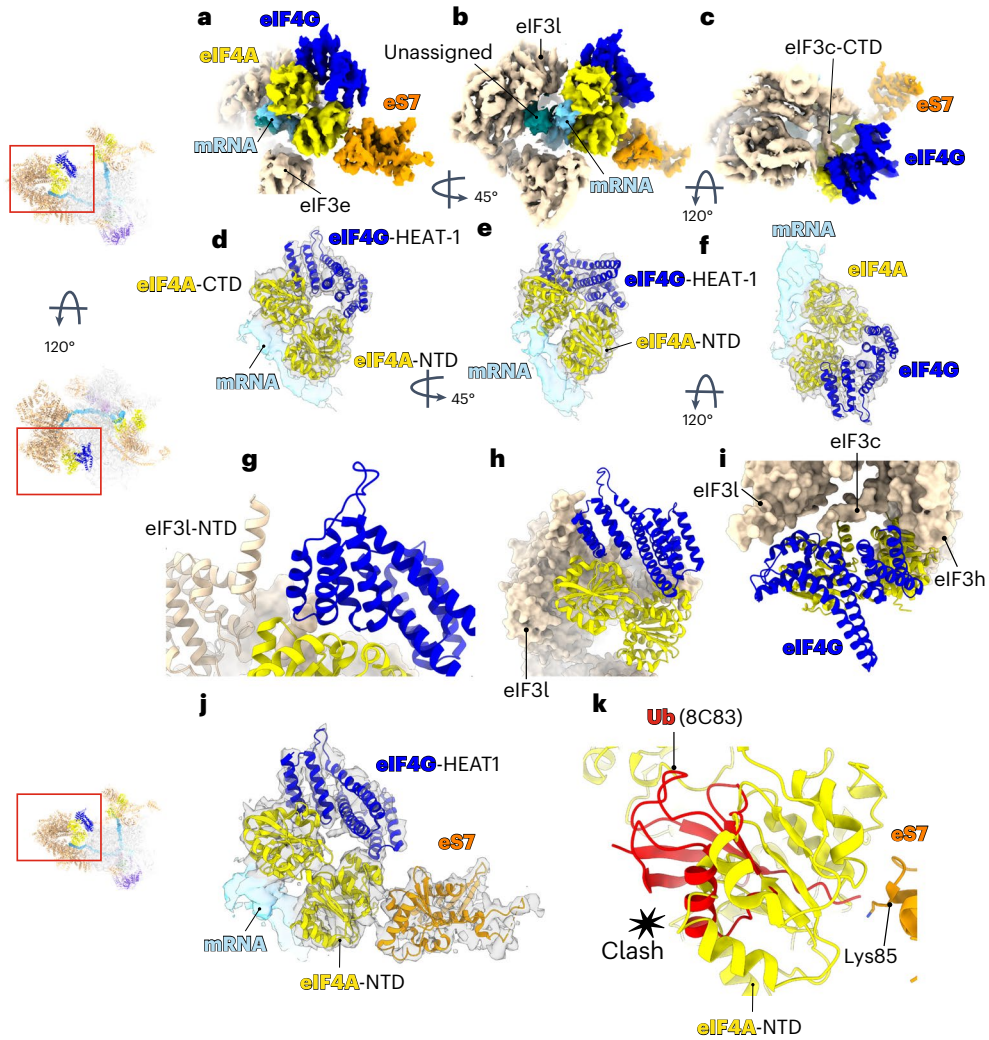


Fig. 4 | Structure of eIF4F and its interactions with eIF3 and ribosomal protein eS7. a–**c**, Cryo-EM map to highlight eIF4F and its interaction network shown in in different orientations. **d**–**f**, Atomic model of human eIF4A and eIF4G fitted into the cryo-EM map shown in in different orientations. **g**–**i**, The insets highlight the interaction network of eIF4F with eIF3. **j**, Close-up to highlight the interaction between eIF4F and ribosomal protein eS7. **k**, Superposition of ubiquitin (PDB:8C83)[53] with the structure of human 48S to highlight the possible clash between eIF4A and ubiquitin at its predicted binding site (Lys85).

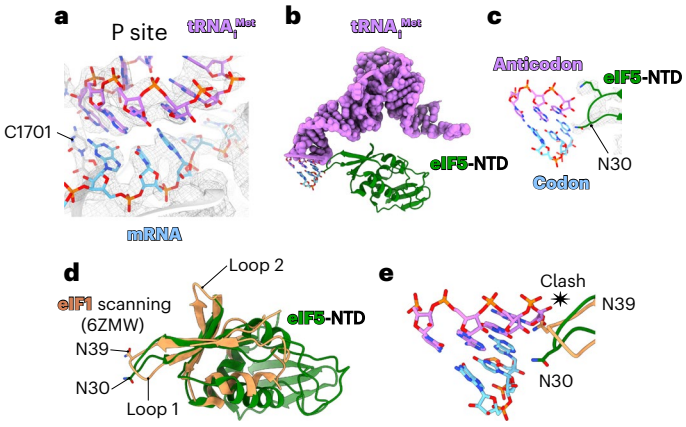


Fig. 5 | Codon–anticodon base pairing in the P site. a, Close-up of the P site to highlight the codon–anticodon base pairing. **b**, eIF5-NTD replaces eIF1 after AUG start codon recognition. **c**, Close-up view illustrating the interaction between eIF5-NTD, tRNA, and mRNA in the P site. **d**, Superposition of eIF5-NTD with the structure of human eIF1 during scanning (PDB: 6ZMW)[4]. **e**, The conformation of eIF1 during scanning would clash with the anticodon stem loop of the tRNA$_i^{Met}$ after start codon recognition.

structure, eIF4A–eIF4B interacts with the 43S at the exact location that DHX29 does[44,68] (Extended Data Fig. 10e–g). Thus, it is possible that, for initiation on specific mRNAs with high secondary structure, other helicases such as DHX29 are recruited to the 43S complex instead of a second eIF4A molecule. Such a method would allow additional control of translation initiation of such mRNAs.

eIF4E stabilizes the binding of mRNA to eIF4G–eIF4A[36] and stimulates the helicase activity of the eIF4F bound eIF4A[69]. Selective 40S ribosome profiling has indicated that scanning is cap-tethered in most human cells[70] (Fig. 6e). In the prior structure of 48S there is a low resolution density that we tentatively assigned to eIF4E[4]. That structure used a very short mRNA, so no meaningful scanning had occurred. In the current structure, scanning has spanned 105 nucleotides before the start codon was reached, thus representing a later stage of initiation. Although local resolution was improved, we did not see such additional density that we could assign to eIF4E, suggesting that it, along with the 5′ cap, is released from eIF4F and the rest of the 48S during scanning, a possibility that has been previously suggested[71,72] (Fig. 6f).

In conclusion, our structural data reveal how two eIF4A helicase proteins coordinate the binding of mRNA to the entry and exit sites of 40S subunit mRNA binding channel. Importantly, the entry-channel-bound eIF4A helicase is precisely positioned to enable its

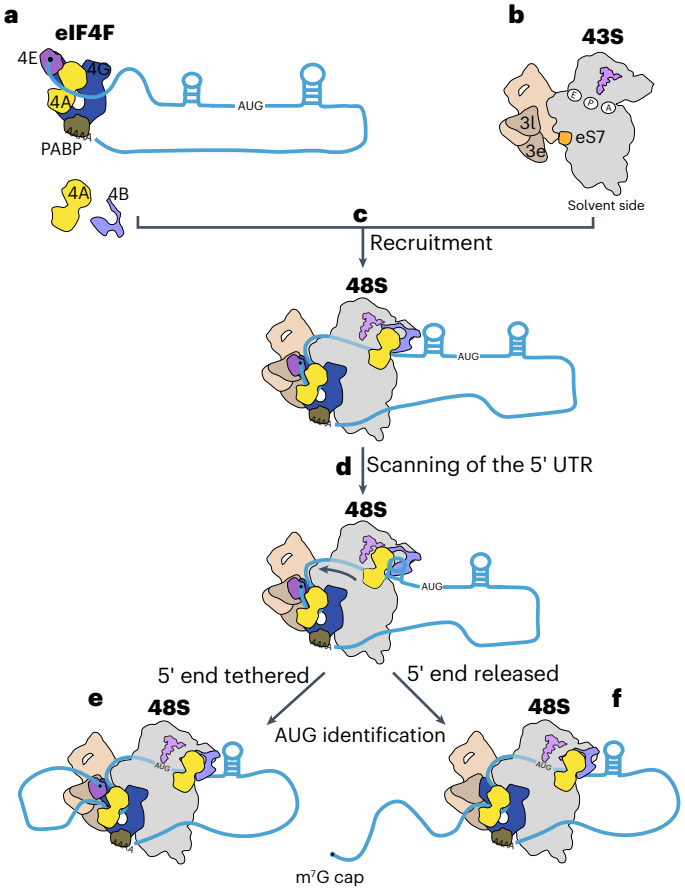


Fig. 6 | Model for translational regulation by eIF4F and eIF4A. a–**c**, The cap-binding complex eIF4F (**a**) binds at the 5′ end of mRNA and recruits the 43S (**b**) to form the 48S complex (**c**). The fuzzy line represents the mRNA into the channel in the 40S. eIF4F binds to eIF3 and ribosomal protein eS7 located upstream of the 43S. The position of eIF4F is compatible with a model whereby the mRNA is slotted into the mRNA channel in the 40S small ribosomal subunit. In addition to eIF4F, a second molecule of eIF4A at the mRNA entry site is likely to play a role in this process by facilitating the accommodation of the mRNA into the channel. The ATPase activity of eIF4A is stimulated by eIF4B and eIF3bgi module 15 located at the entry site. **d**, Thus, eIF4A at the entry site is likely to use the energy from ATP hydrolysis to unwind the RNA secondary structure downstream of the 43S. **e**,**f**, It remains unclear whether eIF4E remains attached to the rest of eIF4F (**e**) or is released (**f**) during the scanning process, but the structure suggests the latter.

ATPase activity to directly unwind secondary structure located downstream of the scanning 48S complex. Our suggested location of eIF4B adjacent to this second eIF4A molecule would explain its role in the helicase activity of eIF4A. The eIF4A helicase has become an important therapeutic target, with eIF4A binding natural products showing promising anti-tumor activity in preclinical studies. The discovery that two copies of eIF4A are bound to the 48S complex may therefore help guide strategies for cancer therapy.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41594-023-01196-0.

References

1. Merrick, W. C. & Pavitt, G. D. Protein synthesis initiation in eukaryotic cells. *Cold Spring Harb. Perspect. Biol.* **10**, a033092 (2018).
2. Marintchev, A. et al. Topology and regulation of the human eIF4A/4G/4H helicase complex in translation initiation. *Cell* **136**, 447–460 (2009).
3. Villa, N., Do, A., Hershey, J. W. B. & Fraser, C. S. Human eukaryotic initiation factor 4G (eIF4G) protein binds to eIF3c, -d, and -e to promote mRNA recruitment to the ribosome. *J. Biol. Chem.* **288**, 32932–32940 (2013).
4. Brito Querido, J. et al. Structure of a human 48S translational initiation complex. *Science* **369**, 1220–1227 (2020).
5. LeFebvre, A. K. et al. Translation initiation factor eIF4G-1 binds to eIF3 through the eIF3e subunit. *J. Biol. Chem.* **281**, 22917–22932 (2006).
6. Harris, T. E. et al. mTOR-dependent stimulation of the association of eIF4G and eIF3 by insulin. *EMBO J.* **25**, 1659–1668 (2006).
7. Gross, J. D. et al. Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E. *Cell* **115**, 739–750 (2003).
8. Kumar, P., Hellen, C. U. T. & Pestova, T. V. Toward the mechanism of eIF4F-mediated ribosomal attachment to mammalian capped mRNAs. *Genes Dev.* **30**, 1573–1588 (2016).
9. Grifo, J. A., Tahara, S. M., Morgan, M. A., Shatkin, A. J. & Merrick, W. C. New initiation factor activity required for globin mRNA translation. *J. Biol. Chem.* **258**, 5804–5810 (1983).
10. Hinton, T. M., Coldwell, M. J., Carpenter, G. A., Morley, S. J. & Pain, V. M. Functional analysis of individual binding activities of the scaffold protein eIF4G. *J. Biol. Chem.* **282**, 1695–1708 (2007).
11. Lamphear, B. J., Kirchweger, R., Skern, T. & Rhoads, R. E. Mapping of functional domains in eukaryotic protein synthesis initiation factor 4G (eIF4G) with picornaviral proteases. Implications for cap-dependent and cap-independent translational initiation. *J. Biol. Chem.* **270**, 21975–21983 (1995).
12. Walker, S. E. et al. Yeast eIF4B binds to the head of the 40S ribosomal subunit and promotes mRNA recruitment through its N-terminal and internal repeat domains. *RNA* **19**, 191–207 (2013).
13. Pestova, T. V. & Kolupaeva, V. G. The roles of individual eukaryotic translation initiation factors in ribosomal scanning and initiation codon selection. *Genes Dev.* **16**, 2906–2922 (2002).
14. García-García, C., Frieda, K. L., Feoktistova, K., Fraser, C. S. & Block, S. M. Factor-dependent processivity in human eIF4A DEAD-box helicase. *Science* **348**, 1486–1488 (2015).
15. Yourik, P. et al. Yeast eIF4A enhances recruitment of mRNAs regardless of their structural complexity. *eLife* **6**, e31476 (2018).
16. Iwasaki, S. et al. The translation inhibitor rocaglamide targets a bimolecular cavity between eIF4A and polypurine RNA. *Mol. Cell* **73**, 738–748.e9 (2019).
17. Iwasaki, S., Floor, S. N. & Ingolia, N. T. Rocaglates convert DEAD-box protein eIF4A into a sequence-selective translational repressor. *Nature* **534**, 558–561 (2016).
18. Wolfe, A. L. et al. RNA G-quadruplexes cause eIF4A-dependent oncogene translation in cancer. *Nature* **513**, 65–70 (2014).
19. Bordeleau, M.-E. et al. Therapeutic suppression of translation initiation modulates chemosensitivity in a mouse lymphoma model. *J. Clin. Invest.* **118**, 2651–2660 (2008).
20. Toribio, R., Díaz-López, I., Boskovic, J. & Ventoso, I. Translation initiation of alphavirus mRNA reveals new insights into the topology of the 48S initiation complex. *Nucleic Acids Res.* **46**, 4176–4187 (2018).
21. Padrón, A., Iwasaki, S. & Ingolia, N. T. Proximity RNA labeling by APEX-seq reveals the organization of translation initiation complexes and repressive RNA granules. *Mol. Cell* **75**, 875–887.e5 (2019).
22. Wang, J. et al. Rapid 40S scanning and its regulation by mRNA structure during eukaryotic translation initiation. *Cell* https://doi.org/10.1016/j.cell.2022.10.005 (2022).

23. Schütz, P. et al. Crystal structure of the yeast eIF4A–eIF4G complex: an RNA-helicase controlled by protein–protein interactions. *Proc. Natl Acad. Sci. USA* **105**, 9564–9569 (2008).
24. Loh, P. G. et al. Structural basis for translational inhibition by the tumour suppressor Pdcd4. *EMBO J.* **28**, 274–285 (2009).
25. Llácer, J. L. et al. Large-scale movement of eIF3 domains during translation initiation modulate start codon selection. *Nucleic Acids Res.* **49**, 11491–11511 (2021).
26. Erzberger, J. P. et al. Molecular architecture of the 40S·eIF1·eIF3 translation initiation complex. *Cell* **158**, 1123–1135 (2014).
27. Aitken, C. E. et al. Eukaryotic translation initiation factor 3 plays distinct roles at the mRNA entry and exit channels of the ribosomal preinitiation complex. *Elife* **5**, e20934 (2016).
28. Sun, Y. et al. Single-molecule kinetics of the eukaryotic initiation factor 4AI upon RNA unwinding. *Structure* **22**, 941–948 (2014).
29. Imataka, H. & Sonenberg, N. Human eukaryotic translation initiation factor 4G (eIF4G) possesses two separate and independent binding sites for eIF4A. *Mol. Cell. Biol.* **17**, 6940–6947 (1997).
30. Korneeva, N. L., Lamphear, B. J., Hennigan, F. L. C., Merrick, W. C. & Rhoads, R. E. Characterization of the two eIF4A-binding sites on human eIF4G-1*. *J. Biol. Chem.* **276**, 2872–2879 (2001).
31. Fleming, K. et al. Solution structure and RNA interactions of the RNA recognition motif from eukaryotic translation initiation factor 4B. *Biochemistry* **42**, 8966–8975 (2003).
32. Rozovsky, N., Butterworth, A. C. & Moore, M. J. Interactions between eIF4AI and its accessory factors eIF4B and eIF4H. *RNA* **14**, 2136–2148 (2008).
33. Eliseev, B. et al. Structure of a human cap-dependent 48S translation pre-initiation complex. *Nucleic Acids Res.* **46**, 2678–2689 (2018).
34. Tauber, D. et al. Modulation of RNA condensation by the DEAD-box protein eIF4A. *Cell* https://doi.org/10.1016/j.cell.2019.12.031 (2020).
35. Itzhak, D. N., Tyanova, S., Cox, J. & Borner, G. H. Global, quantitative and dynamic mapping of protein subcellular localization. *eLife* **5**, e16950 (2016).
36. Izidoro, M. S., Sokabe, M., Villa, N., Merrick, W. C. & Fraser, C. S. Human eukaryotic initiation factor 4E (eIF4E) and the nucleotide-bound state of eIF4A regulate eIF4F binding to RNA. *J. Biol. Chem.* **298**, 102368 (2022).
37. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
38. Cattie, D. J. et al. Mutations in nonessential eIF3k and eIF3l genes confer lifespan extension and enhanced resistance to ER stress in *Caenorhabditis elegans*. *PLoS Genet.* **12**, e1006326 (2016).
39. Smith, M. D. et al. Human-like eukaryotic translation initiation factor 3 from *Neurospora crassa*. *PLoS ONE* **8**, e78715 (2013).
40. Herrmannová, A. et al. Adapted formaldehyde gradient cross-linking protocol implicates human eIF3d and eIF3c, k and l subunits in the 43S and 48S pre-initiation complex assembly, respectively. *Nucleic Acids Res.* **48**, 1969–1984 (2020).
41. Choe, J. et al. mRNA circularization by METTL3-eIF3h enhances translation and promotes oncogenesis. *Nature* **561**, 556–560 (2018).
42. Simonetti, A., Guca, E., Bochler, A., Kuhn, L. & Hashem, Y. Structural insights into the mammalian late-stage initiation complexes. *Cell Rep.* **31**, 107497 (2020).
43. Yi, S.-H. et al. Conformational rearrangements upon start codon recognition in human 48S translation initiation complex. *Nucleic Acids Res.* https://doi.org/10.1093/nar/gkac283 (2022).
44. des Georges, A. et al. Structure of mammalian eIF3 in the context of the 43S preinitiation complex. *Nature* **525**, 491–495 (2015).
45. Kratzat, H. et al. A structural inventory of native ribosomal ABCE1-43S pre-initiation complexes. *EMBO J.* https://doi.org/10.15252/embj.2020105179 (2020).
46. Schepetilnikov, M. et al. TOR and S6K1 promote translation reinitiation of uORF-containing mRNAs via phosphorylation of eIF3h. *EMBO J.* **32**, 1087–1102 (2013).
47. Zhang, L., Smit-McBride, Z., Pan, X., Rheinhardt, J. & Hershey, J. W. B. An oncogenic role for the phosphorylated h-subunit of human translation initiation factor eIF3. *J. Biol. Chem.* **283**, 24047–24060 (2008).
48. Buschauer, R. et al. The Ccr4–Not complex monitors the translating ribosome for codon optimality. *Science* **368**, eaay6912 (2020).
49. Panasenko, O. O. & Collart, M. A. Presence of Not5 and ubiquitinated Rps7A in polysome fractions depends upon the Not4 E3 ligase. *Mol. Microbiol.* **83**, 640–653 (2012).
50. Ikeuchi, K. et al. Collided ribosomes form a unique structural interface to induce Hel2-driven quality control pathways. *EMBO J.* **38**, e100276 (2019).
51. Absmeier, E. et al. Specific recognition and ubiquitination of translating ribosomes by mammalian CCR4-NOT. *Nat. Struct. Mol. Biol.* **30**, 1314–1322 (2023).
52. Takehara, Y. et al. The ubiquitination–deubiquitination cycle on the ribosomal protein eS7A is crucial for efficient translation. *iScience* **24**, 102145 (2021).
53. Ikeuchi, K. et al. Molecular basis for recognition and deubiquitination of 40S ribosomes by Otu2. *Nat. Commun.* **14**, 2730 (2023).
54. Llácer, J. L. et al. Translational initiation factor eIF5 replaces eIF1 on the 40S ribosomal subunit to promote start-codon recognition. *eLife* **7**, e39273 (2018).
55. Nanda, J. S. et al. eIF1 controls multiple steps in start codon recognition during eukaryotic translation initiation. *J. Mol. Biol.* **394**, 268–285 (2009).
56. Hinnebusch, A. G. Structural insights into the mechanism of scanning and start codon recognition in eukaryotic translation initiation. *Trends Biochem. Sci.* **42**, 589–611 (2017).
57. Llácer, J. L. et al. Conformational differences between open and closed states of the eukaryotic translation initiation complex. *Mol. Cell* **59**, 399–412 (2015).
58. Simonetti, A. et al. eIF3 peripheral subunits rearrangement after mRNA binding and start-codon recognition. *Mol. Cell* **63**, 206–217 (2016).
59. Hussain, T. et al. Structural changes enable start codon recognition by the eukaryotic translation initiation complex. *Cell* **159**, 597–607 (2014).
60. Lomakin, I. B. & Steitz, T. A. The initiation of mammalian protein synthesis and mRNA scanning mechanism. *Nature* **500**, 307–311 (2013).
61. Lapointe, C. P. et al. eIF5B and eIF1A reorient initiator tRNA to allow ribosomal subunit joining. *Nature* **607**, 185–190 (2022).
62. Thakur, A. & Hinnebusch, A. G. eIF1 Loop 2 interactions with Met-tRNAi control the accuracy of start codon selection by the scanning preinitiation complex. *Proc. Natl Acad. Sci. USA* **115**, E4159–E4168 (2018).
63. Vicens, Q., Kieft, J. S. & Rissland, O. S. Revisiting the closed-loop model and the nature of mRNA 5′–3′ communication. *Mol. Cell* **72**, 805–812 (2018).
64. Lai, W.-J. C. et al. Intrinsically unstructured sequences in the mRNA 3′ UTR reduce the ability of poly(A) tail to enhance translation. *J. Mol. Biol.* **434**, 167877 (2022).
65. Gu, Y., Mao, Y., Jia, L., Dong, L. & Qian, S.-B. Bi-directional ribosome scanning controls the stringency of start codon selection. *Nat. Commun.* **12**, 6604 (2021).
66. Sokabe, M. & Fraser, C. S. A helicase-independent activity of eIF4A in promoting mRNA recruitment to the human ribosome. *Proc. Natl Acad. Sci. USA* **114**, 6304–6309 (2017).

67. Sen, N. D., Zhou, F., Harris, M. S., Ingolia, N. T. & Hinnebusch, A. G. eIF4B stimulates translation of long mRNAs with structured 5′ UTRs and low closed-loop potential but weak dependence on eIF4G. *Proc. Natl Acad. Sci. USA* **113**, 10464–10472 (2016).
68. Hashem, Y. et al. Structure of the mammalian ribosomal 43S preinitiation complex bound to the scanning factor DHX29. *Cell* **153**, 1108–1119 (2013).
69. Feoktistova, K., Tuvshintogs, E., Do, A. & Fraser, C. S. Human eIF4E promotes mRNA restructuring by stimulating eIF4A helicase activity. *Proc. Natl Acad. Sci. USA* **110**, 13339–13344 (2013).
70. Bohlen, J., Fenzl, K., Kramer, G., Bukau, B. & Teleman, A. A. Selective 40S footprinting reveals cap-tethered ribosome scanning in human cells. *Mol. Cell* https://doi.org/10.1016/j.molcel.2020.06.005 (2020).
71. Filipowicz, W. & Haenni, A. L. Binding of ribosomes to 5′-terminal leader sequences of eukaryotic messenger RNAs. *Proc. Natl Acad. Sci. USA* **76**, 3111–3115 (1979).
72. Sonenberg, N., Shatkin, A. J., Ricciardi, R. P., Rubin, M. & Goodman, R. M. Analysis of terminal structures of RNA from potato virus X. *Nucleic Acids Res.* **5**, 2501–2512 (1978).

Methods

Purification of human eIFs

Ribosome, $tRNA_i^{Met}$ and human eIFs, including $eIF4G_{557–1,105}$, were purified as described previously[61,66,69,73]. Briefly, native eIF2, eIF3, and 40S small ribosomal subunits were purified from HeLa cell extracts. Recombinant eIF1, eIF1A, eIF5, eIF4E and eIF4E were expressed in BL21 (DE3) cells and purified using a combination of affinity and size exclusion liquid chromatography. Recombinant eIF4G was expressed using the baculovirus-insect cell (sf9) expression system. Finally, PABP was purified as described[69]. The protein was expressed in BL21 bacterial cells using maltose-binding protein (MBP) fusion constructs. This construct was designed to include a tobacco etch virus protease cleavage site, enabling the removal of the tag. Subsequently, the recombinant PABP–MBP fusion protein was subjected to incubation with tobacco etch virus protease, facilitating the cleavage of the tag, and underwent additional purification steps to isolate the untagged PABP. Additionally, we used a hydroxyapatite column to remove any remaining nucleic acids bound to it, and the protein was purified further by gel filtration chromatography on a Superdex 200 Increase column in a buffer containing 20 mM HEPES, pH 7.5, 200 mM KCl, 10% glycerol and 1 mM tris(2-carboxyethyl)phosphine.

In vitro transcription and purification of mRNA

We designed a synthetic DNA containing a T7 promoter followed by a 5′ UTR containing three polypurine motifs (GGACAAGAGAGAGAGAGACUCCAACUCCAAGAGAGAGAGAGA

CAACUCCAAGAGAGAGAGAGACAAACCCUCGCUGAGCCGCAGUCAGAUCCUAGCGUCGAGUUG**AUG**CUGUCCGAU). The construct was purchased as a gene block from IDT. After linearization, the DNA plasmid was used as a template for in vitro transcription. The mRNA was purified using a denaturant acrylamide gel (8 M urea) followed by electroelution. The purified mRNA was capped using Vaccinia Capping System (New England Biolabs) and then polyadenylated (~90 nucleotides) (Extended Data Fig. 1) using *Escherichia coli* Poly(A) polymerase (New England Biolabs). After free nucleotide, enzyme and buffer removal, the pure mRNA was stored in water.

We used the same protocol of transcription and purification of DLP mRNA containing β-globin 5′ UTR used for cryo-EM

(GACACUUGCUUUUGACACAACUGUGUUUACUUGCAAUCCCCCAAAACAGACAGA**AUG**AACAACGAGCCACCGCAAACCAGUGACGGCCGCCGGAGGCGCCCGCGCCCGGCGGCCGAGAGAGCAGAACGAGACCACACGGAUCCGAGAAGAUUCAUCCUCCUUCAAUGCCUGGAGGAUA).

In vitro reconstitution of human 48S complexes for cryo-EM

We reconstituted the 48S complex by mixing the 43S with eIF4F, eIF4A, eIF4B, PABP and mRNA in a 25 µl reaction. To reconstitute the 43S, we mixed 0.5 µM 40S with 0.9 µM eIFs and 1.8 µM TC to a final volume of 18 µl in a buffer (20 mM HEPES–KOH pH 7.5, 97 mM KAc, 2.5 mM MgAc, 3% glycerol, 0.1 mM spermidine, 1 mM dithiothreitol and 0.5 mM GMP-PNP). We incubated the 43S reaction mix at 30 °C for 10 min. In parallel, we assembled the cap-binding complex eIF4F by mixing a co-purified eIF4G (residues 165–1,599)–eIF4A complex with eIF4E, mRNA and PABP, to a final concentration of 3.3 µM in 12 µl reaction in a buffer (20 mM HEPES–KOH pH 7.5, 97 mM KAc, 2.5 mM MgAc, 3% glycerol, 0.1 mM spermidine, 1 mM dithiothreitol and 0.5 mM ATP-g-S). Finally, we mixed 1 µM eIF4F, 1 µM eIF4B, 1 µM eIF4A, 43S (0.3 µM 40S and 0.5 µM eIFs) and 0.5 mM RocA in a 25 µl reaction. The reaction mix was incubated at 30 °C for 10 min.

The same protocol was used to assemble the 48S complexes with $eIF4G_{557–1,105}$, without eIF4B and with DLP mRNA. The DLP 48S was assembled without RocA.

Cryo-EM grid preparation

The assembled 48S complexes were cross-linked using 1.5 mM BS3 (final concentration) on ice for 45 min to prevent dissociation of eIFs during the grid preparation. As described before[4], the cross-link reaction in the presence of spermidine (a polyamine with quenching properties) and on ice has a mild effect—comparable with the complex without BS3. Three microliters of 140 nM 48S complex (based on the concentration of the 40S) was applied onto UltrAuFoil R1.2/1.3 300 mesh gold grids precovered with graphene oxide (Sigma) suspension made in-house. Briefly, the UltrAuFoil gold grids were washed using deionized water and then dried overnight at room temperature. The grids were then glow-discharged for 5 min at 30 mA, followed by incubation with 3 µl of graphene oxide (0.2 mg ml^{-1}) for 1 min. After the incubation and blotting, the grids were washed three times using 20 µl deionized water. The grids were dried for at least 1 h at room temperature and then used.

We prepared the grids using an FEI Vitrobot Mark IV at 4 °C and 100% humidity. The grids were blotted for 7 or 8 s at blotting force −15 and then plunged into liquid ethane at 93 K in a precision cryostat system produced at the MRC Laboratory of Molecular Biology[74].

Cryo-EM data collection

Data were collected on Titan Krios microscopes (ThermoFisher) equipped with a K3 direct electron detector camera (Gatan) at a magnification of 105,000× and at pixel sizes of 0.826 Å per pixel or 0.829 Å per pixel (48S dataset from eBIC/Diamond). The data were collected using EPU software, super-resolution counting mode using a Bio-quantum energy filter (Gatan) (binning 2), faster acquisition mode, and with defocus ranging from −1.2 µm to −3.0 µm.

The following doses/frames were used: 1.197 $e^- Å^{-2}$ (48S dataset from eBIC/Diamond); 1.2391 $e^- Å^{-2}$ (48S dataset from MRC LMB); 0.6889 $e^- Å^{-2}$ (48S without eIF4B and 48S complexes with $eIF4G_{682–1,105}$). DLP 48S was collected using Falcon 4 electron detector, counting mode and 1 $e^- Å^{-2}$ per frame.

Image processing

Motion correction was performed using the implementation in RELION 4 (ref. 75). Movies were aligned using 5 × 5 patches with dose weighting. CTF (contrast transfer function) was estimated using CTFFIND4.1 (ref. 76). After 2D classification, we used the cryo-EM map of a human 48S[4] after low-pass filtering to 60 Å as a reference for 3D classification. After 3D classification and refinement, we performed mask classification to select only particles containing eIF3 octameric structural core. We used Bayesian polishing in RELION to correct beam-induced motion[75]. After polishing, we performed mask classification at the entry site, followed by the eIF4F binding region and TC. Finally, we used multi-body refinement and flexibility analysis in RELION[77] to account for the predominant molecular motions.

Model building, fitting and refinement

We used our previous structure of human 48S[4], previous cryo-EM and crystal structures of eIFs[16,23,25,26,42,44,54,60,78–80], as well as AlphaFold[37], to build and refine the atomic model. Fitting and model building and local refinement were performed in Coot[81]. We used Phenix for real space refinement[82].

Figures

All figures were made using ChimeraX[83].

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The atomic model and cryo-EM maps have been uploaded to the Protein Data Bank (PDB: 8OZ0) and to the Electron Microscopy Data Bank (EMD: 17297). Maps of 3D focus refinement and 3D multi-body have also been uploaded to EMDB under the same accession number (EMD: 17297).

References

73. Sokabe, M. & Fraser, C. S. Human eukaryotic initiation factor 2 (eIF2)–GTP–Met–tRNAi ternary complex and eIF3 stabilize the 43S preinitiation complex. *J. Biol. Chem.* **289**, 31827–31836 (2014).
74. Russo, C. J., Scotcher, S. & Kyte, M. A precision cryostat design for manual and semi-automated cryo-plunge instruments. *Rev. Sci. Instrum.* **87**, 114302 (2016).
75. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).
76. Rohou, A. & Grigorieff, N. CTFFIND4: fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
77. Nakane, T., Kimanius, D., Lindahl, E. & Scheres, S. H. Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION. *eLife* **7**, e36861 (2018).
78. Lee, A. S., Kranzusch, P. J., Doudna, J. A. & Cate, J. H. D. eIF3d is an mRNA cap-binding protein that is required for specialized translation initiation. *Nature* **536**, 96–99 (2016).
79. Schmitt, E. et al. Structure of the ternary initiation complex aIF2-GDPNP-methionylated initiator tRNA. *Nat. Struct. Mol. Biol.* **19**, 450–454 (2012).
80. Wei, Z. et al. Crystal structure of human eIF3k, the first structure of eIF3 subunits. *J. Biol. Chem.* **279**, 34983–34990 (2004).
81. Casañal, A., Lohkamp, B. & Emsley, P. Current developments in Coot for macromolecular model building of electron cryo-microscopy and crystallographic data. *Protein Sci.* https://doi.org/10.1002/pro.3791 (2019).
82. Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D* **74**, 531–544 (2018).
83. Goddard, T. D. et al. UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).

Acknowledgements

We thank the LMB and eBIC (proposal BI23268-52, funded by the Wellcome Trust, MRC and BBSRC) cryo-EM facilities for their support during data collection, and J. Grimmett and T. Darling for computing. We thank N. Desai and N. R. James for preparing the cell extract and 80S ribosomes from which J.B.Q. purified the 40S small ribosomal subunits. We also thank J. Lorsch, T. Dever, A. Hinnebusch and V. Chandrasekaran for feedback on the manuscript. J.B.Q. was supported by an FEBS long-term fellowship; I.D.-L. was supported by an EMBO Postdoctoral Fellowship; V.R. was supported by the UK Medical Research Council (MC_U105184332), a Wellcome Trust Investigator award (WT096570) and the Louis-Jeantet Foundation; C.S.F. was supported by the NIH (grant R01 GM092927) and a seed grant for international collaboration from Global Affairs and the College of Biological Sciences at the University of California, Davis.

Author contributions

M.S., C.S.F. and Y.G. purified eIFs; J.B.Q. purified the mRNA, assembled the complexes, prepared cryo-EM grids, performed the cryo-EM data collection and processing, and built and refined atomic models; I.D.-L. performed the cryo-EM analysis of the DLP 48S complex; J.B.Q., Y.G., I.D.-L. and V.R. interpreted the structure; J.B.Q., M.S., C.S.F. and V.R. wrote the manuscript with input from all authors. V.R. supervised the project.

Competing interests

The authors declare no competing interests.

Additional information

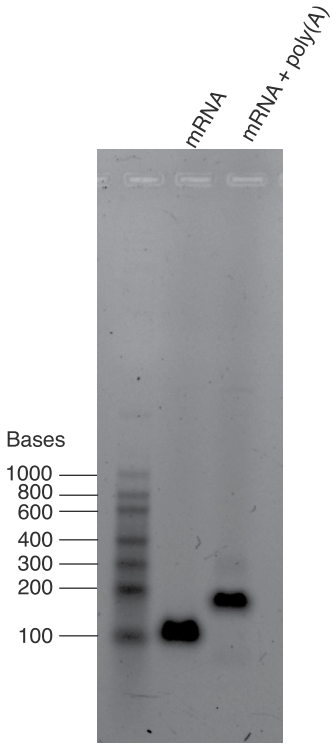
Extended data is available for this paper at https://doi.org/10.1038/s41594-023-01196-0.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41594-023-01196-0.

Correspondence and requests for materials should be addressed to Christopher S. Fraser or V. Ramakrishnan.

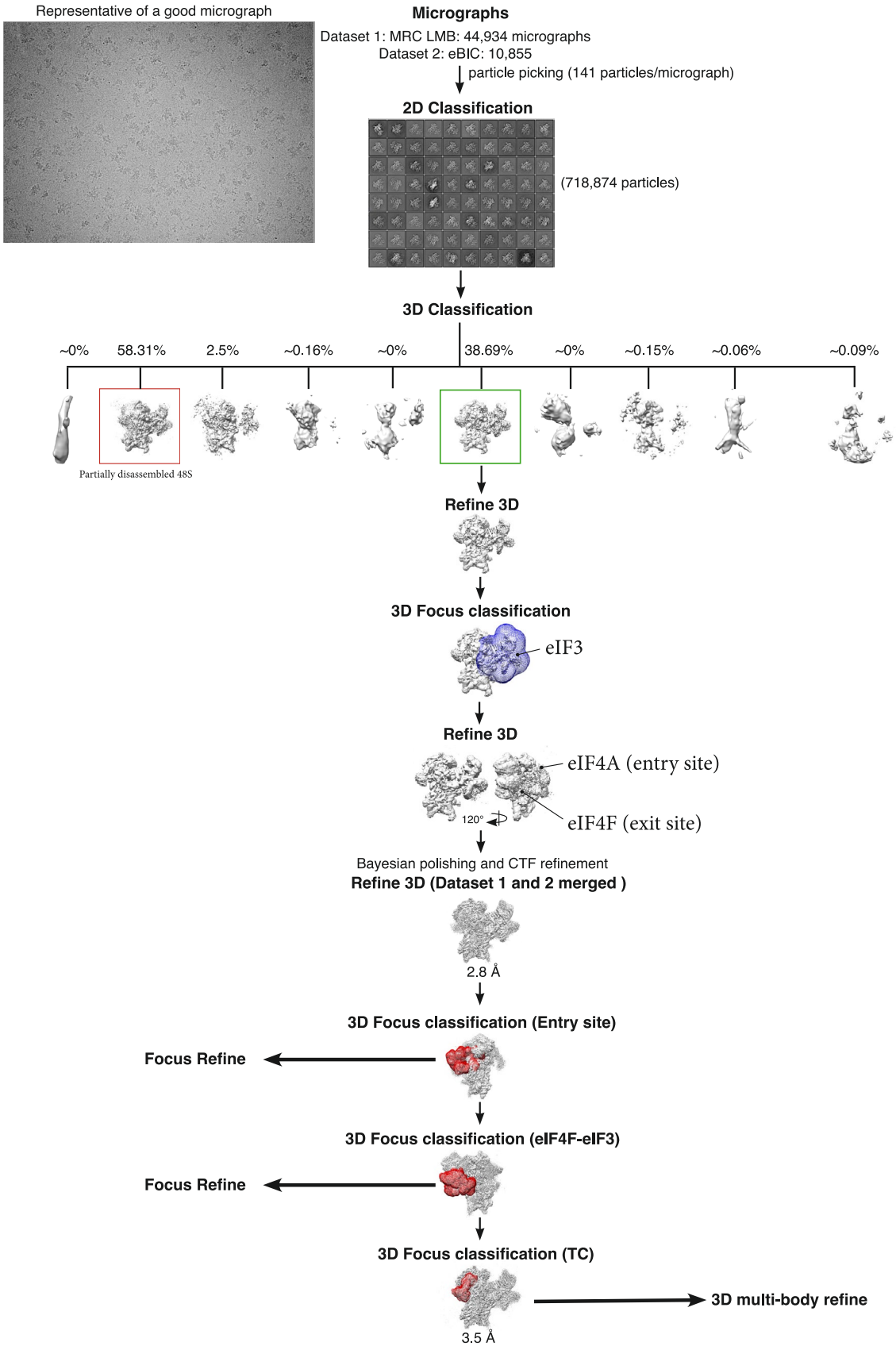
Peer review information *Nature Structural & Molecular Biology* thanks Andrei Korostelev, Nahum Sonenberg and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Sara Osman was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

https://doi.org/10.1038/s41594-023-01196-0



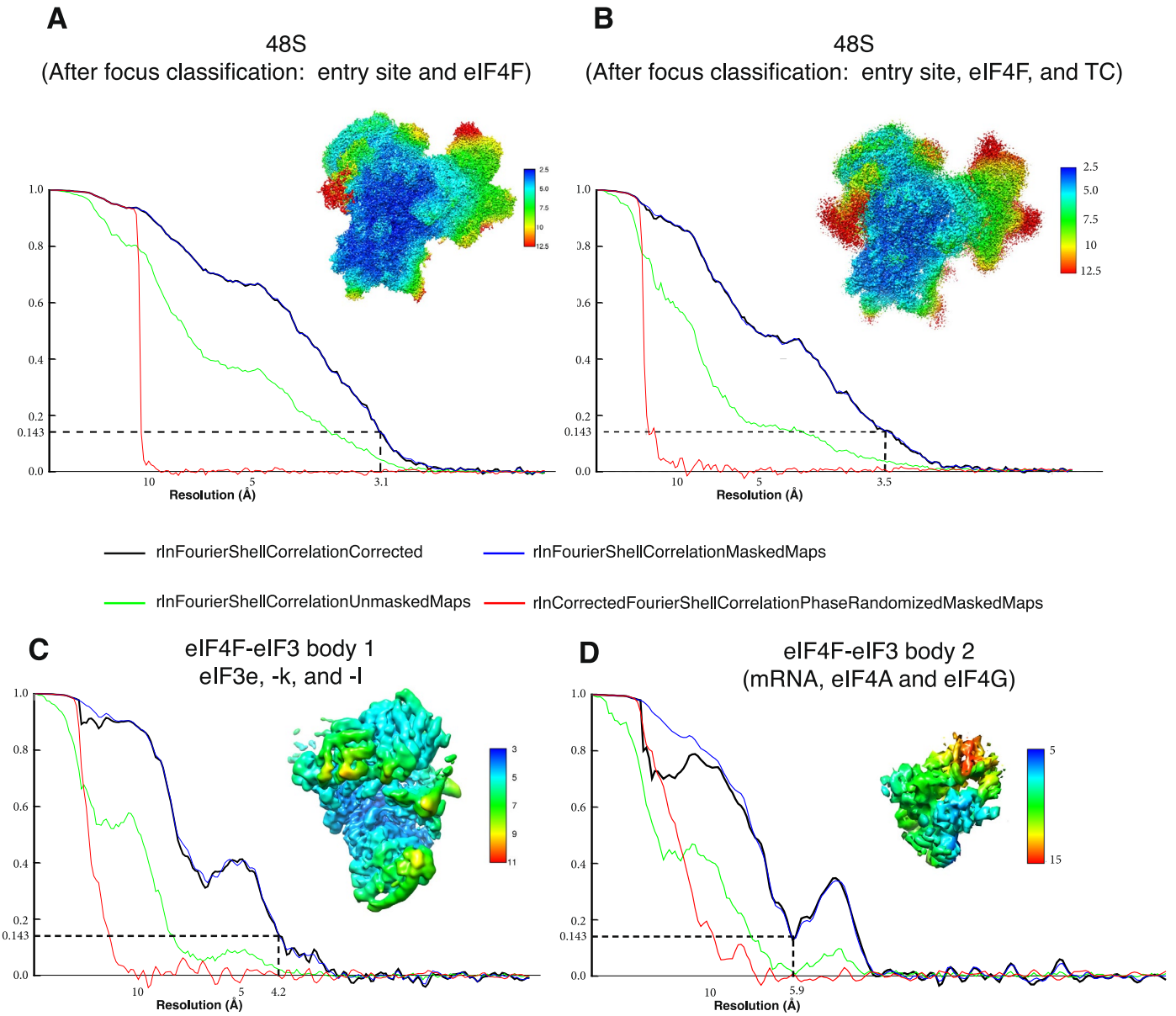
Extended Data Fig. 1 | Analysis of capped mRNA after Poly(A) tailing. The polyadenylated mRNA has additional ~90 nucleotides at the 3′ end. The poly(A) tailing was performed according to the protocol provided by the manufacturer (NEB; M0276). The final mRNA, containing a poly(A) stretch, was stored and used to assemble the complex for cryo-EM.

https://doi.org/10.1038/s41594-023-01196-0



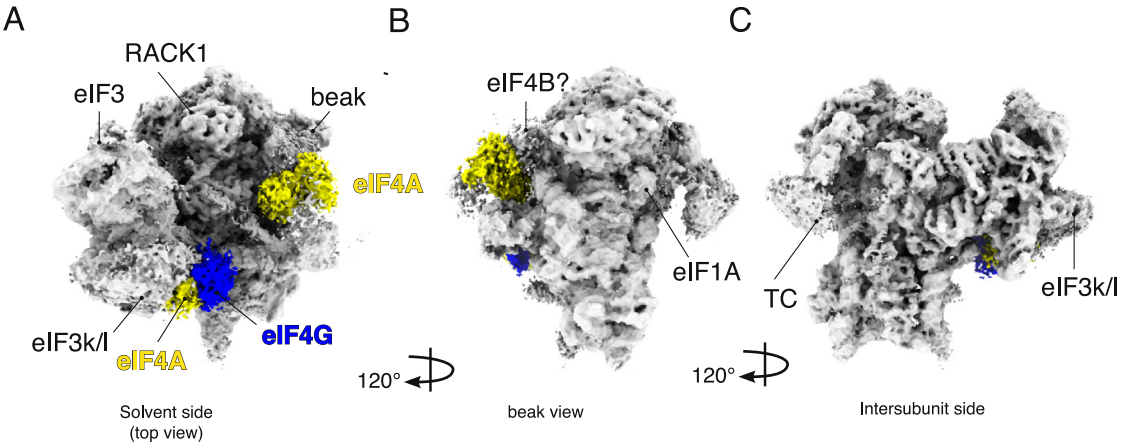
Extended Data Fig. 2 | Cryo-EM data analysis of human 48S. We performed two independent data collections, and each dataset (datasets 1 and 2) was analyzed independently. After polishing and CTF refinement, the two datasets were merged, which yielded the 2.9 Å 3D reconstruction of a human 48S.

https://doi.org/10.1038/s41594-023-01196-0



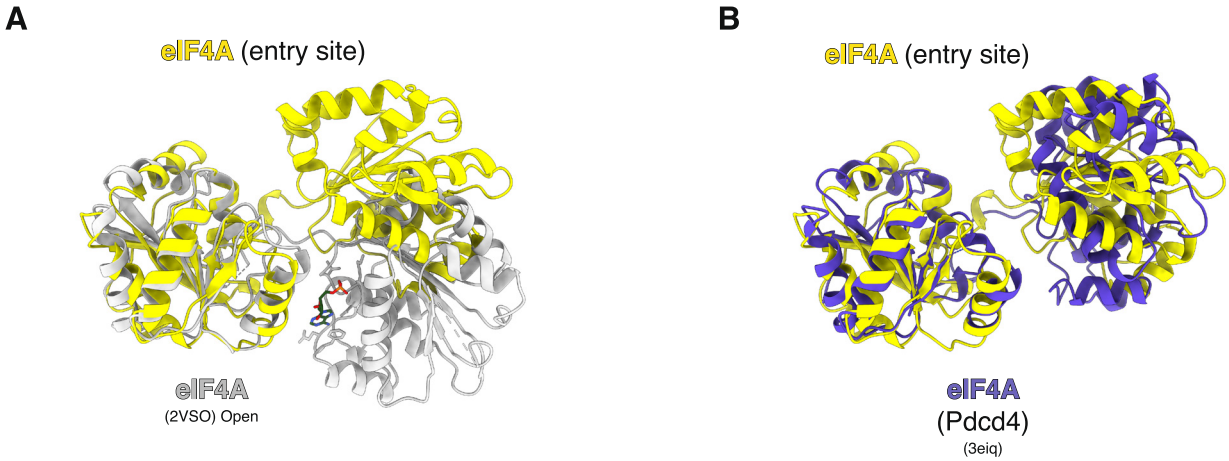
Extended Data Fig. 3 | Overall and local resolution. (**a** and **b**) Fourier shell correlation (FSC) curve and local resolution of the 48S complex after focus 3D focus classification at the entry site, eIF4F, and ternary complex. (**c** and **d**) Overall and local resolution of eIF4F-eIF3 after multi-body refinement.

https://doi.org/10.1038/s41594-023-01196-0



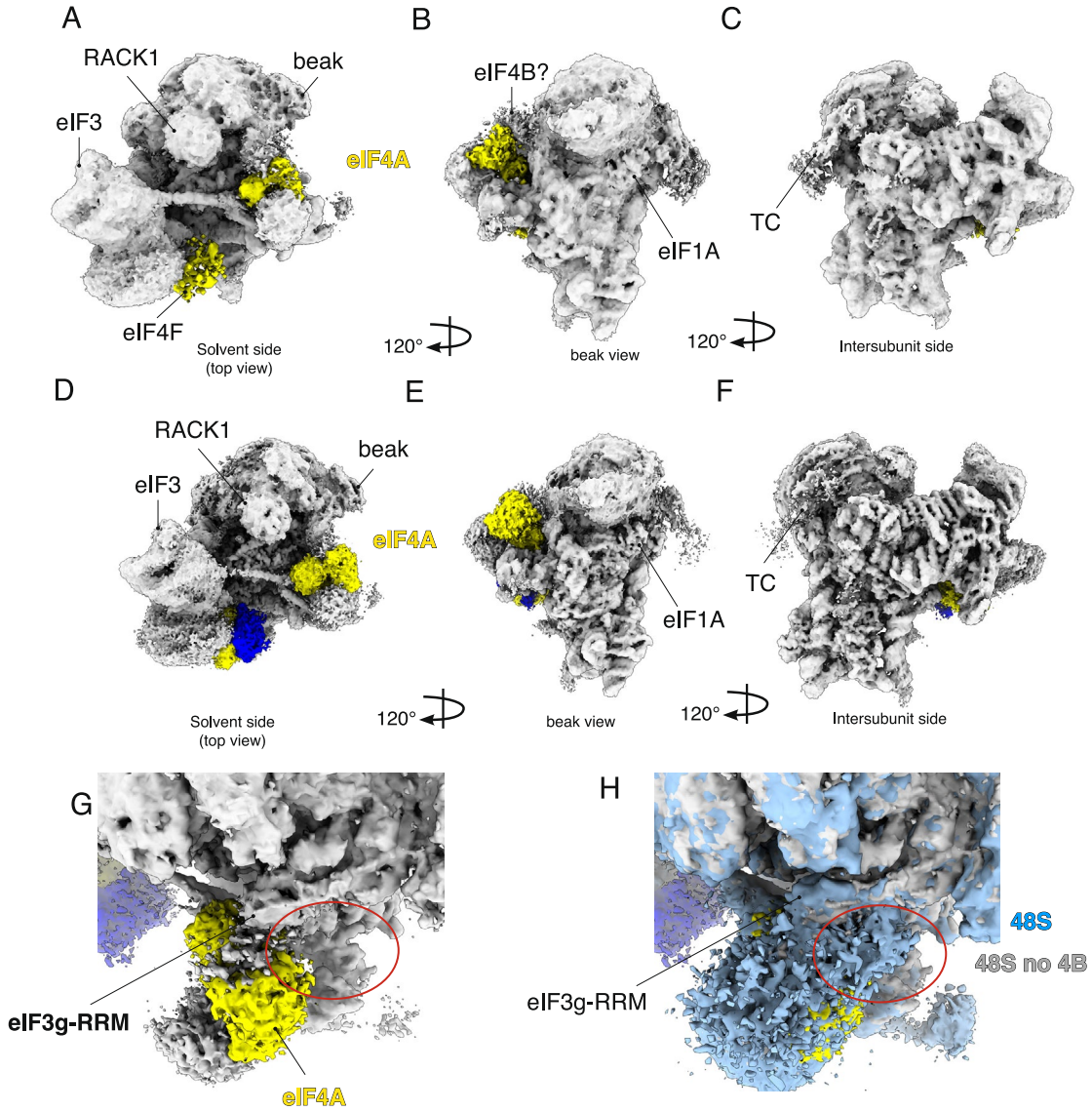
Extended Data Fig. 4 | Cryo-EM 3D reconstruction of a human 48S complex assembled without Rocaglamide (RocA). (**a** to **c**) Overview of the cryo-EM density map. The complex contains both eIF4F and the second molecule of eIF4A bound at the mRNA channel entry site. To assembly this complex, we used an mRNA with highly stable RNA stem-loop structure (DLP) located 83 nucleotides downstream of the cap. This mRNA also has an AUG start codon.

https://doi.org/10.1038/s41594-023-01196-0



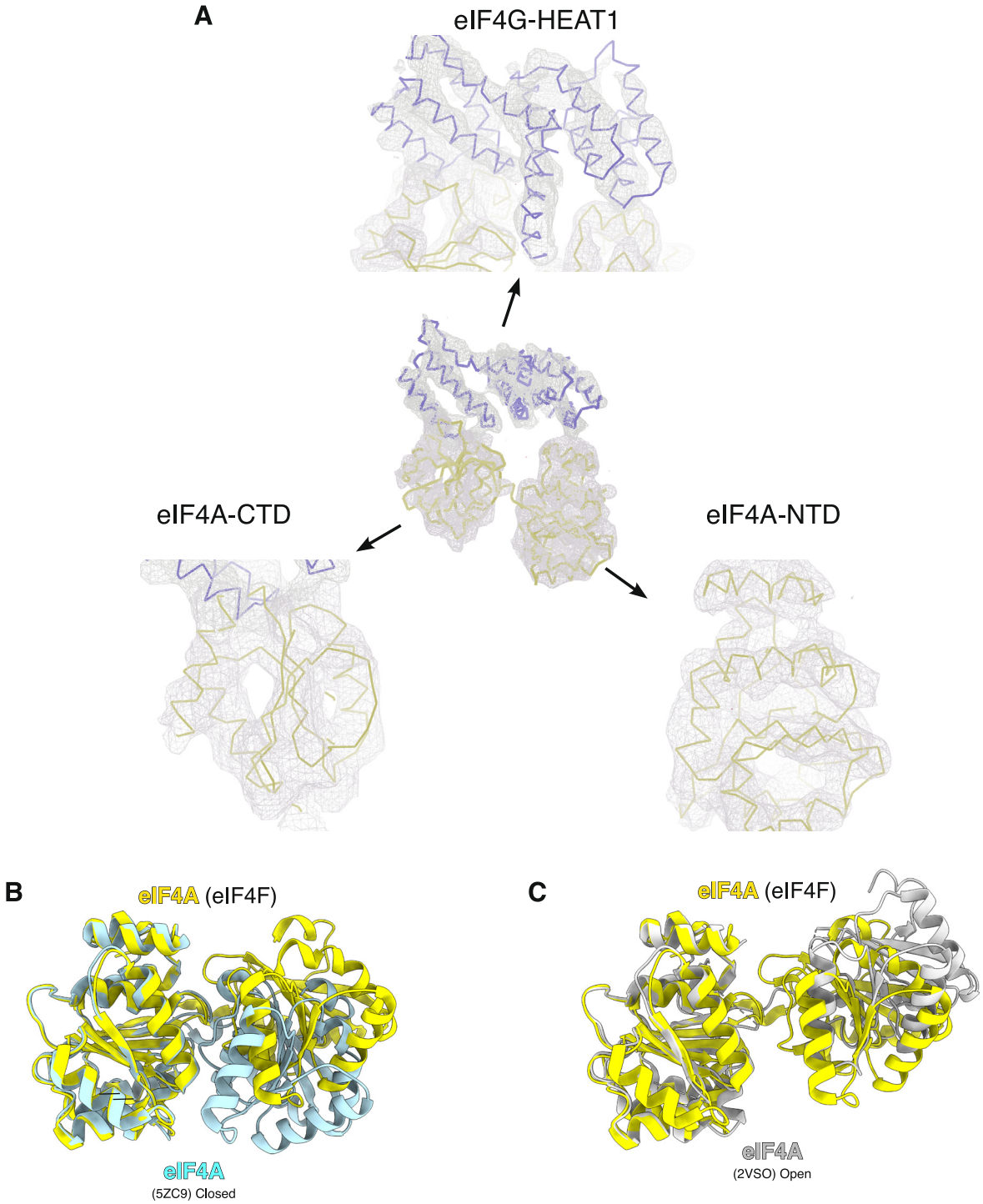
Extended Data Fig. 5 | Conformation of molecule of eIF4A bound at the mRNA entry site. (**a**) Superposition of eIF4A with the structure of a yeast eIF4A in complex with eIF4G (PDB:2VSO)[23]. (**b**) Superposition of eIF4A with the structures of eIF4A in complex with Pdcd4 (PDB:3eiq)[24].

https://doi.org/10.1038/s41594-023-01196-0



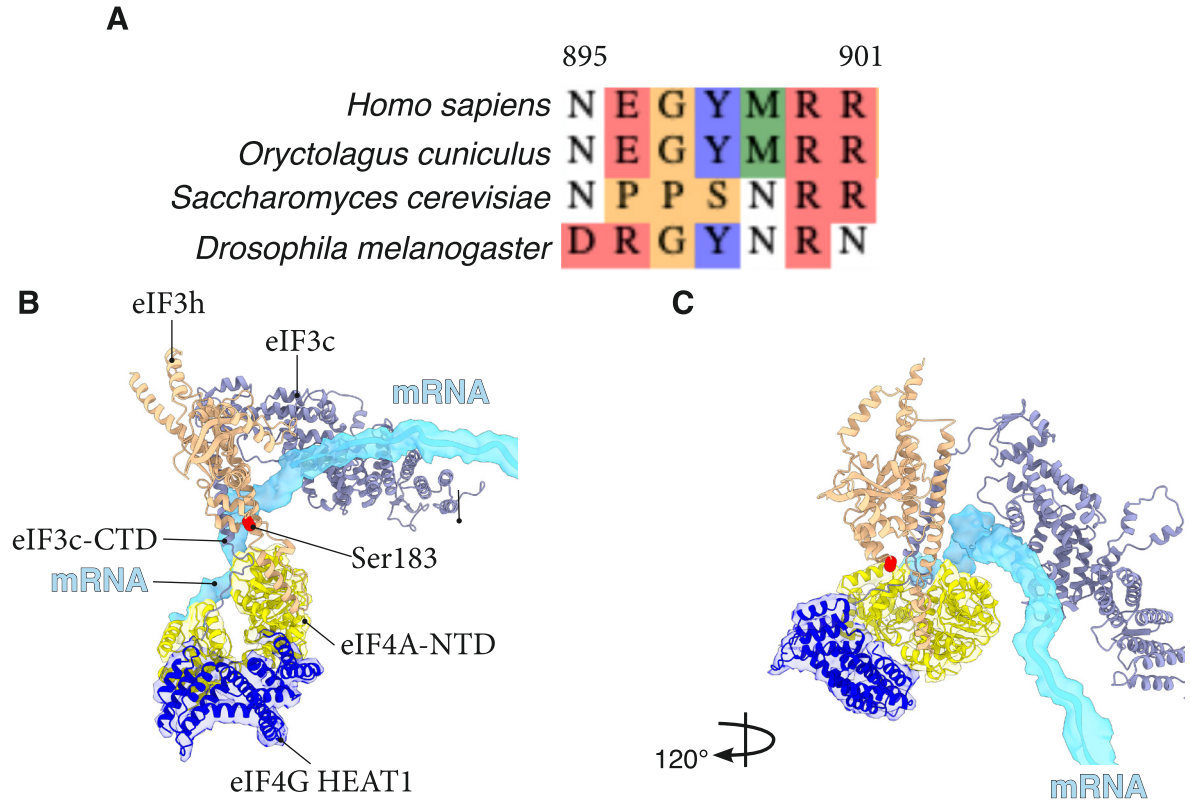
Extended Data Fig. 6 | Cryo-EM 3D reconstruction of a human 48S assembled with a truncated eIF4G (residues 557–1137) or without eIF4B. (**a** to **c**) Overview of the cryo-EM density map of a 48S complex assembled with a truncated eIF4G (residues 557–1137). The eIF4A density is visible even without 3D focus classification. The eIF4F density is visible, but is faint without focus classification. (**d** to **f**) Overview of the cryo-EM density map of a 48S complex assembled without eIF4B. The red oval highlights the absence of the density in the complex without eIF4B (**g**) that can be seen in the structure of 48S assembled eIF4B (**h**).

https://doi.org/10.1038/s41594-023-01196-0



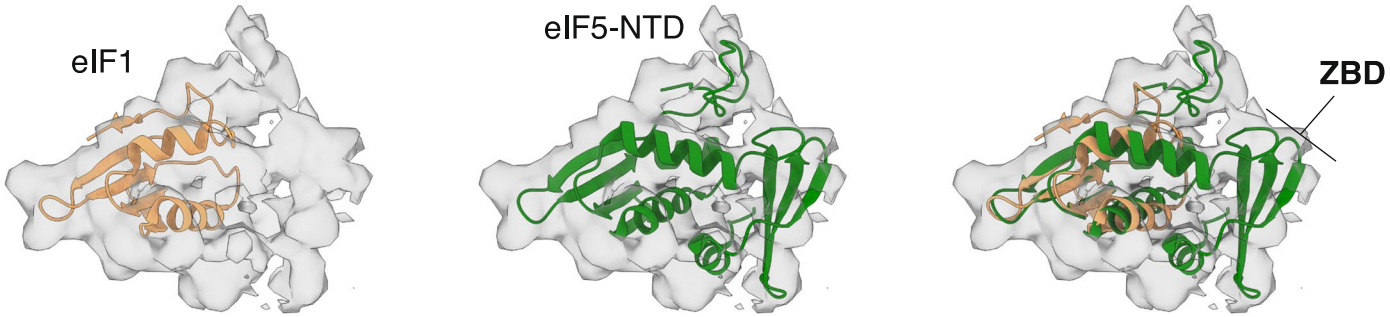
Extended Data Fig. 7 | Human eIF4F. (**a**) Atomic model of human eIF4F (eIF4G and eIF4A) fitted into the cryo-EM map after focus refinement on eIF4F. (**b**) Superposition of eIF4A that is part of the cap-binding complex in the 48S with the structure of human free eIF4A in complex with mRNA and RocA (PDB: 5ZC9)16. (**c**) Superposition of this eIF4A with the structure of a yeast eIF4A in complex with eIF4G (PDB: 2VSO)23.

https://doi.org/10.1038/s41594-023-01196-0



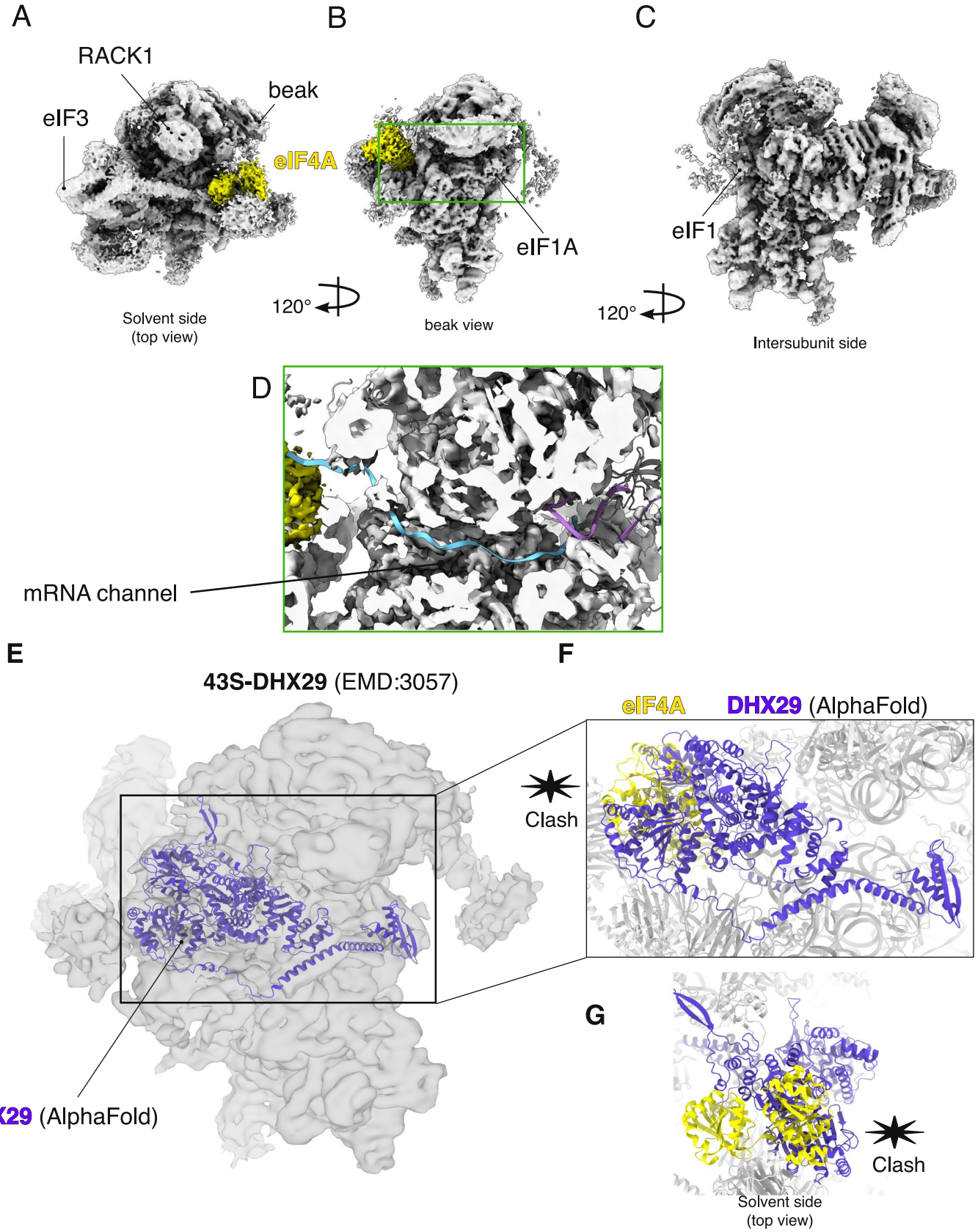
Extended Data Fig. 8 | Multiple sequence alignment and Interaction of eIF3 with eIF4F. (**a**) Multiple sequence alignment of the eIF3C C-terminal tail. The site of interaction with eIF4G is conserved among eukaryotes. (**b**, **c**) Interaction of eIF3h and eIF3c with eIF4F. eIF3h interacts directly with eIF4A-NTD. eIF3h and eIF3c provide clamping and guide the mRNA (cryo-EM map) from eIF4F towards the mRNA channel in the 40S.

https://doi.org/10.1038/s41594-023-01196-0



Extended Data Fig. 9 | Map-to-model fits of human eIF1 and eIF5. Rigid-body fitting of human eIF1 (PD- B:6ZMW)[4] and eIF5-NTD (PDB: 2E9H) into the cryo-EM map. Zinc binding domain (ZBD).

https://doi.org/10.1038/s41594-023-01196-0



Extended Data Fig. 10 | Cryo-EM 3D reconstruction of a class of particles without mRNA into the channel and the position of eIF4A. (**a** to **c**) Overview of the cryo-EM density map of a class of particles without mRNA into the channel. (**d**) The inset shows a lack of density for mRNA using mRNA and tRNA superimposed from the 48S structure. (**e**) Rigid-body fitting of human DHX29 (AlphaFold prediction) into a prior cryo-EM map of a rabbit 43S (EMD:3057)[44]. (**f** and **g**) The inset shows that DHX29 and the entry-site eIF4A would overlap and could not coexist.

nature portfolio

Corresponding author(s): Venki Ramakrishnan and Christopher S. Fraser

Last updated by author(s): Nov 7, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Please do not complete any field with "not applicable" or n/a. Refer to the help text for what text to use if an item is not relevant to your study.
For final submission: please carefully check your responses for accuracy; you will not be able to make changes later.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☒	☐	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Cryo-EM datasets were collected with EPU (Thermo Fisher Scientific, version 2.12.1).
Data analysis	The cryo-EM data was processed using RELION (version 4.0). Coot (version 0.9.8.5) was used for model building. The atomic model was refined using Phenix (version 1.20.1-4487). ChimeraX (version 1.6.1) was used to prepare the figures.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The atomic model, along with cryo-EM maps, has been deposited in the Protein Data Bank (PDB: 8OZ0) and the Electron Microscopy Data Bank (EMD:17297). Additionally, 3D focus refinement and 3D multi-body maps have been included in the EMDB with the same accession number (EMD:17297).

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	N/A
Reporting on race, ethnicity, or other socially relevant groupings	N/A
Population characteristics	N/A
Recruitment	N/A
Ethics oversight	N/A

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	This study involves the in vitro reconstitution of the translation initiation complex, followed by its analysis using cryo-electron microscopy (cryo-EM). Therefore, it does not require the use of any statistical methods to determine the sample size. The number of micrographs collected was based on our previous structural study of translation initiation (reference 4 in the manuscript).
Data exclusions	Individual micrographs with resolution > 5Å and CTF Figure of Merit < 0.2 were excluded.
Replication	Two individual datasets from the sample sample were collected.
Randomization	N/A
Blinding	N/A

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	The Hela cell line was used to purify eIF2, eIF3, and 40S small ribosomal subunits. E. coli BL21 was used for expression and purification of recombinant eIFs.
Authentication	N/A
Mycoplasma contamination	No mycoplasma contamination was detected.
Commonly misidentified lines (See ICLAC register)	N/A

Plants

Seed stocks	N/A
Novel plant genotypes	N/A
Authentication	N/A