Cell Reports, Volume *31*

Supplemental Information

Structural Insights into the Mammalian Late-Stage Initiation Complexes

Angelita Simonetti, Ewelina Guca, Anthony Bochler, Lauriane Kuhn, and Yaser Hashem

SUPPLEMENTAL INFORMATION

SUPPLEMENTAL FIGURES

Figure S1, related to Figure 1. Particle classification outputs, average and local resolutions of the different β-globin and H4 LS48S ICs cryo-EM reconstructions.

Figure S2, related to Figure 2 and Figure 4. Structural differences between β-globin and H4 LS48S ICs.

Figure S3, related to Figure 5. Differences between mammalian and yeast mRNA trajectories in the mRNA channel and the $tRNA_i^{Met}$ interactions.

Figure S4, related to Figure 6. ABCE1 sequence comparison between eukaryotes and a sterical incompatibility with eIF2D.

Figure S5, related to Figure 3. Atomic model of uS7 in mammalian LS48S IC.

Figure S6, related to Figure 7. Comparison between eIF3 of ICs from near-native conditions and *in vitro* assembly.

Figure S1, related to Fig. 1

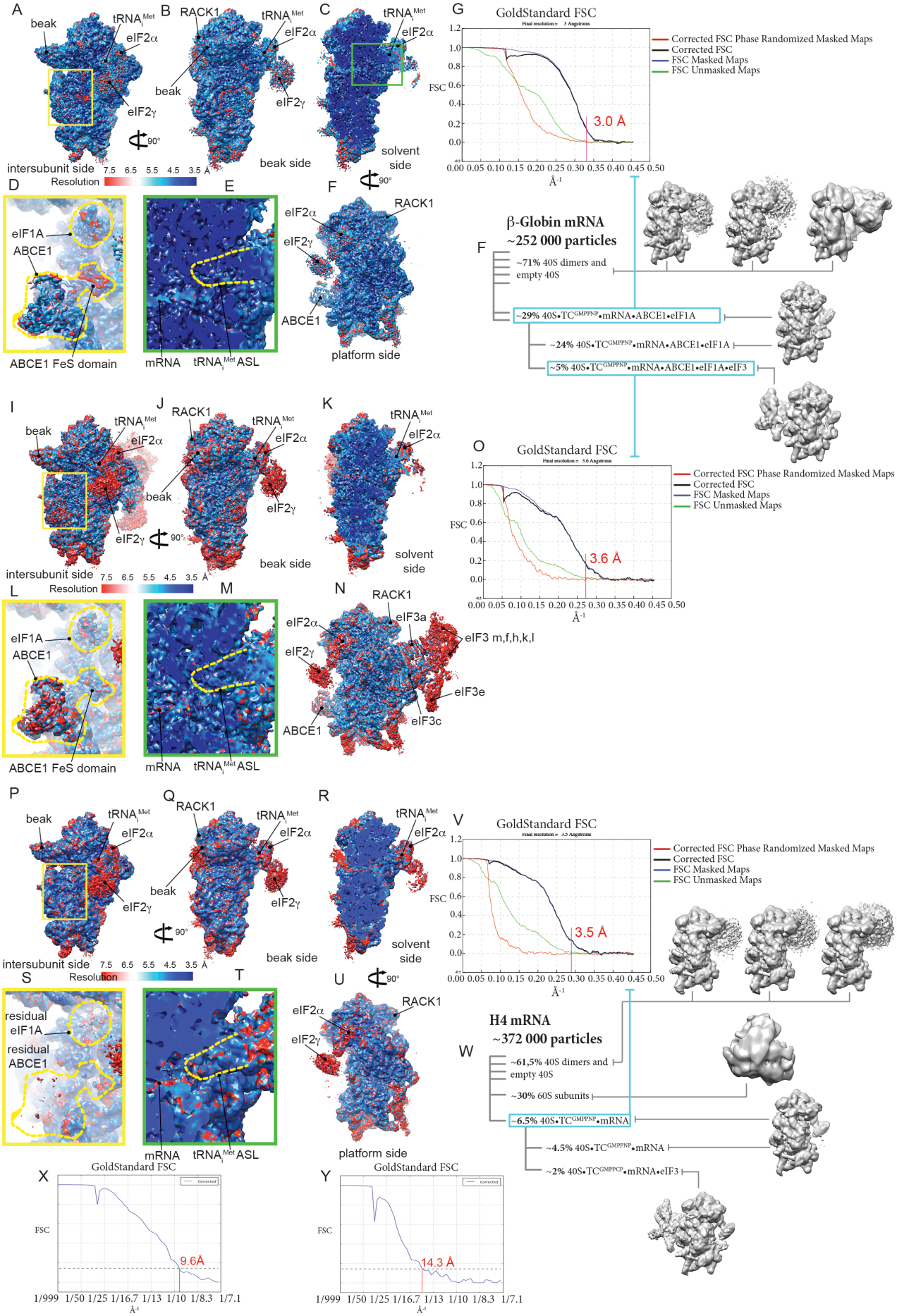

Figure S1, related to Figure 1. Particle classification outputs, average and local resolutions of the different β-globin and H4 LS48S ICs cryo-EM reconstructions. (A-F) Local resolution of the β-globin late-stage IC reconstruction representing a class lacking eIF3 seen from the intersubunit side (A), beak side (B), solvent side (C), platform side (F). (D-E) Insets showing local resolution of eIF1A and ABCE1 (yellow frame) and of mRNA and $tRNA_i^{Met}$ anticodon stem-loop (ASL) (green frame). (G) Average resolution of the reconstruction of the class without eIF3. (H) Particle classification output of the β-globin LS48S IC. (I-N) Local resolution of the β-globin LS48S IC reconstruction representing a class showing eIF3 octamer seen from intersubunit side (I), beak side (J), solvent side (K), platform side (N). (L-M) insets showing local resolution for eIF1A and ABCE1 (yellow frame) and for mRNA and $tRNA_i^{Met}$ ASL (green frame). (O) Average resolution of the reconstruction of the class of β-globin LS48S IC with eIF3. (P-U) Local resolution of the H4 LS48S IC reconstruction representing a class lacking eIF3 seen from the intersubunit side (P), beak side (Q), solvent side (R), platform side (U). The local resolution of eIF3 is not shown, as its average resolution is relatively low. (S-T) Insets showing local resolution of eIF1A and ABCE1 (yellow frame) and of mRNA and $tRNA_i^{Met}$ ASL (green frame). (V) Average resolution of the reconstruction of the class. (W) Particle classification output of the H4 LS48S IC. (X-Y) Average resolutions of the reconstruction of β-globin+ATP LS48S IC reconstructions after their particle sorting (X). Similarly to the counterpart complex with ATP, two major classes stand out, with and without eIF3 (Y).

Figure S2, related to Fig. 2

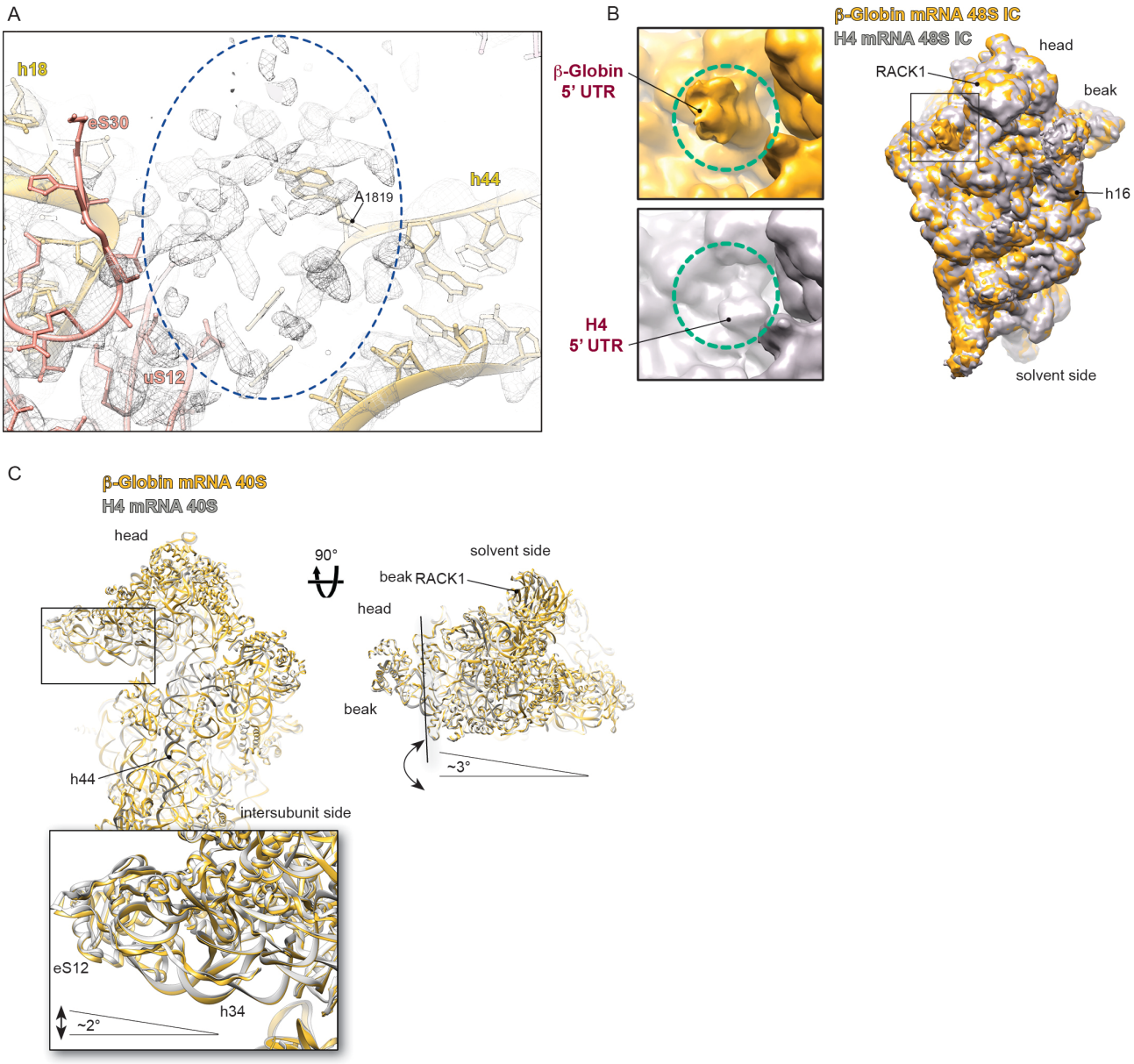

Figure S2, related to Figure 2 and Figure 4. Structural differences between β-globin and H4 LS48S ICs. (A) Residual density for eIF1A in H4 LS48S IC in the A-site, highlighted by dashed blue oval. (B) Comparison of mRNA channel exit between β-globin (orange surface) and H4 (grey surface) 40S reconstructions. (C) Comparison between β-globin (gold ribbons) and H4 (grey ribbons) ICs 40S atomic models. Blowups highlight the conformational changes between both types of complexes at the mRNA channel entrance, the beak and the A, P and E –sites. Dashed coloured ovals indicate the conformational changes in several ribosomal proteins (uS3, eS30 and uS7).

Figure S3, related to Fig. 5

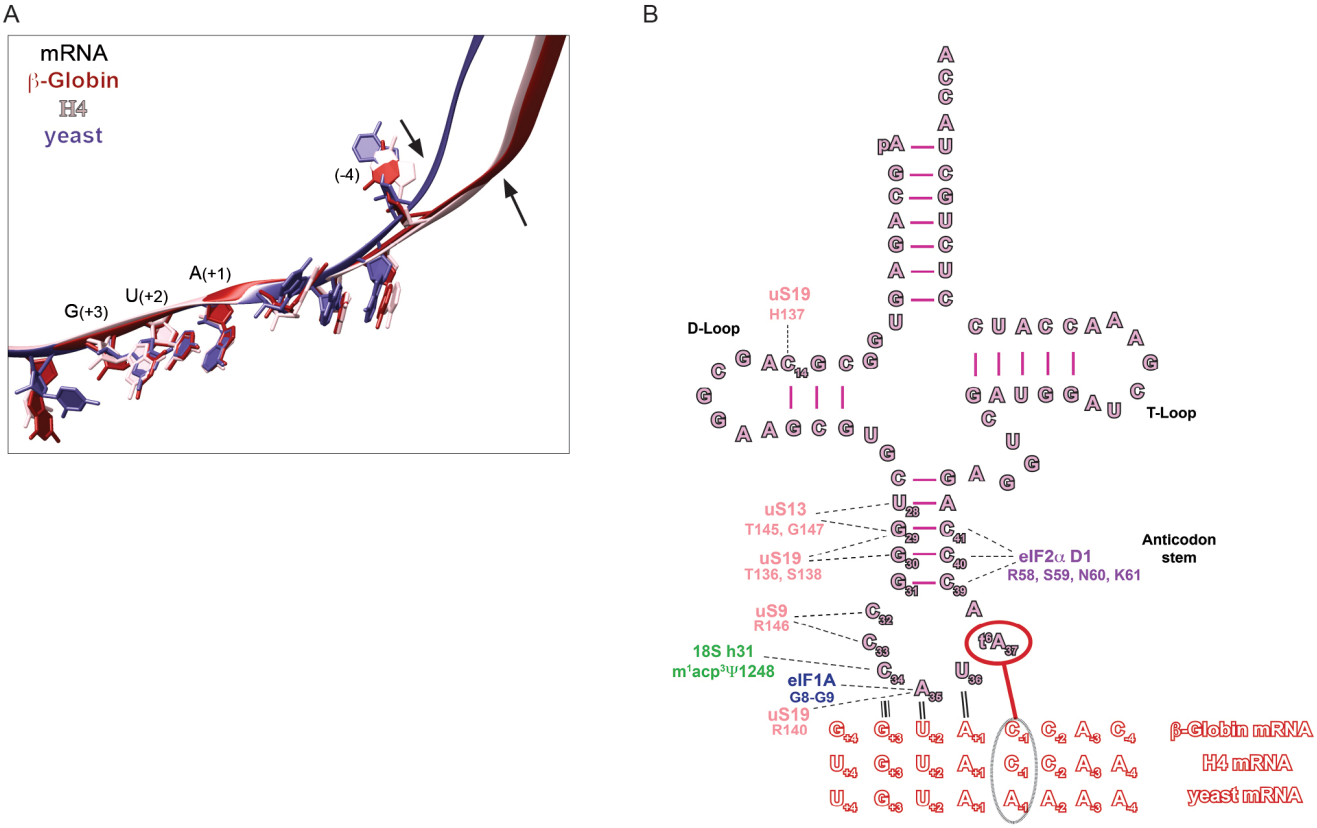

Figure S3, related to Figure 5. Differences between mammalian and yeast mRNA trajectories in the mRNA channel and the $tRNA_i^{Met}$ interactions. (A) Superimposition of yeast optimal mRNA Kozak consensus sequence in py48S-eIF5 IC structure (Llácer et al., 2018) and mammalian mRNAs showing a smoother P/E kink in the case of β-globin and H4 ICs, indicated by arrows. (B) The cloverleaf representation of the $tRNA_i^{Met}$ summarizing the interactions within the LS48S IC described in the Results section. The interaction of $t^6A(37)$ modification with (-1) mRNA position is highlighted by dashed-line oval.

Figure S4, related to Fig. 6

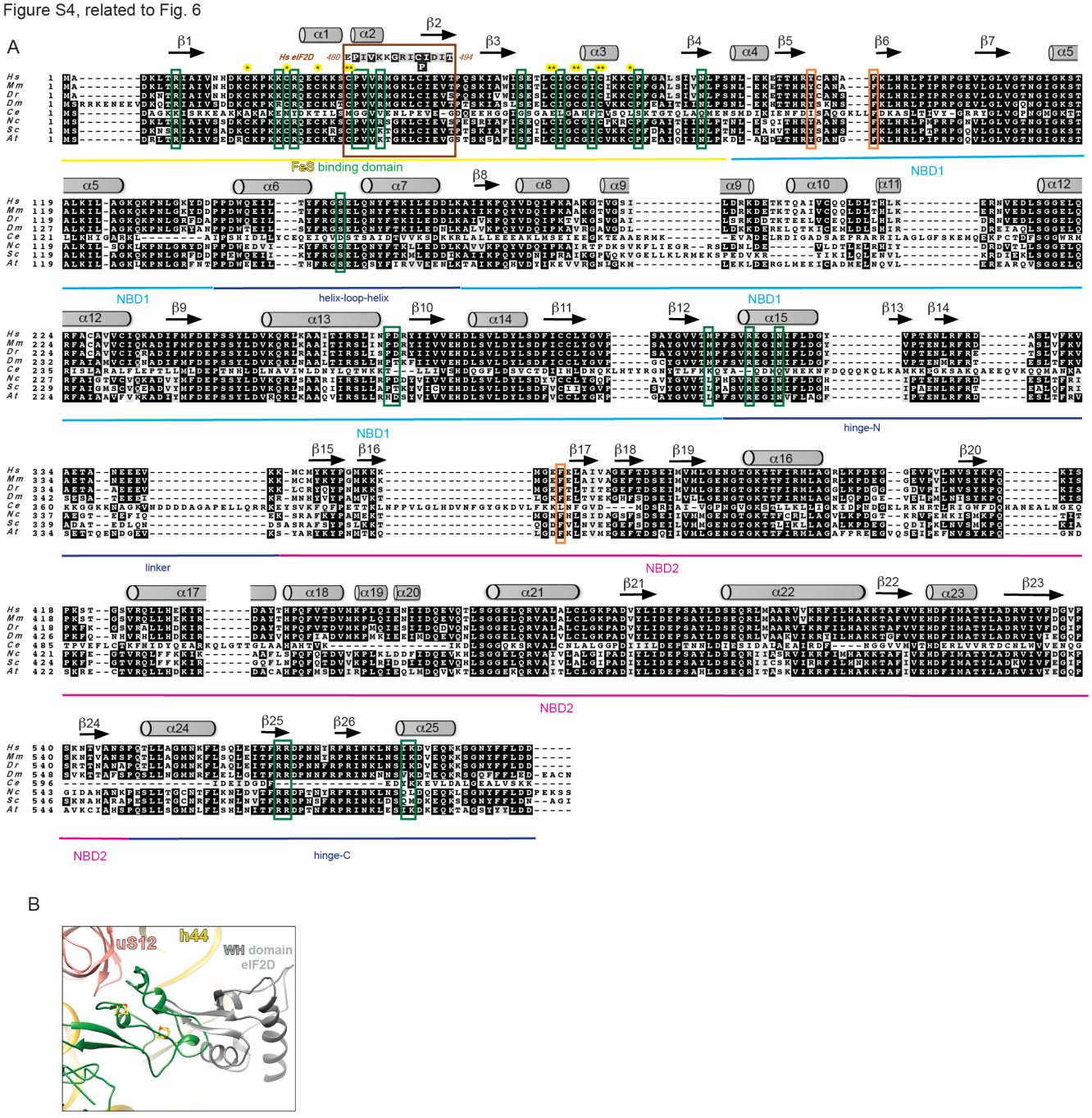

Figure S4, related to Figure 6. ABCE1 sequence comparison between eukaryotes and a sterical incompatibility with eIF2D. (A) Alignment of ABCE1 among eight representative eukaryotic species with secondary structure elements labelled, helices α and β-sheets (according to (Karcher et al., 2008)), and functional domains, indicated by bottom coloured lines. The residues involved in the interactions described in Results section are framed in green. In brown frame, the sequence of Fe-S cluster that showed similarity to the N-terminal part of eIF2D SUI domain (~40% sequence identity). The coordinated cysteines of cluster (I) (*) and (II) (**) are labelled in yellow. In orange frames, the residues responsible for proper ATP/GTP binding in nucleotide binding domains (NBD1 and NBD2) pockets. The accession numbers used for the alignment: *Hs*: NP_002931.2, *Mm*: NP_056566.2, *Dr*: NP_998216.2, *Dm*: NP_648272.1 (pixie), *Ce*: NP_506192.1 ABC transporter class F, *Nc*: XP_963869.3, *Sc*: AJV19484.1, *At*: OAP04903.1). (B) Superimposition of ABCE1 in β-globin LS48S IC (in green) with eIF2D (PDB ID: 5OA3) (Gouridis et al., 2019), showing the sterical clashing of Fe-S cluster (I) with winged-helix (WH) domain of eIF2D (in grey).

Figure S5, related to Fig. 3

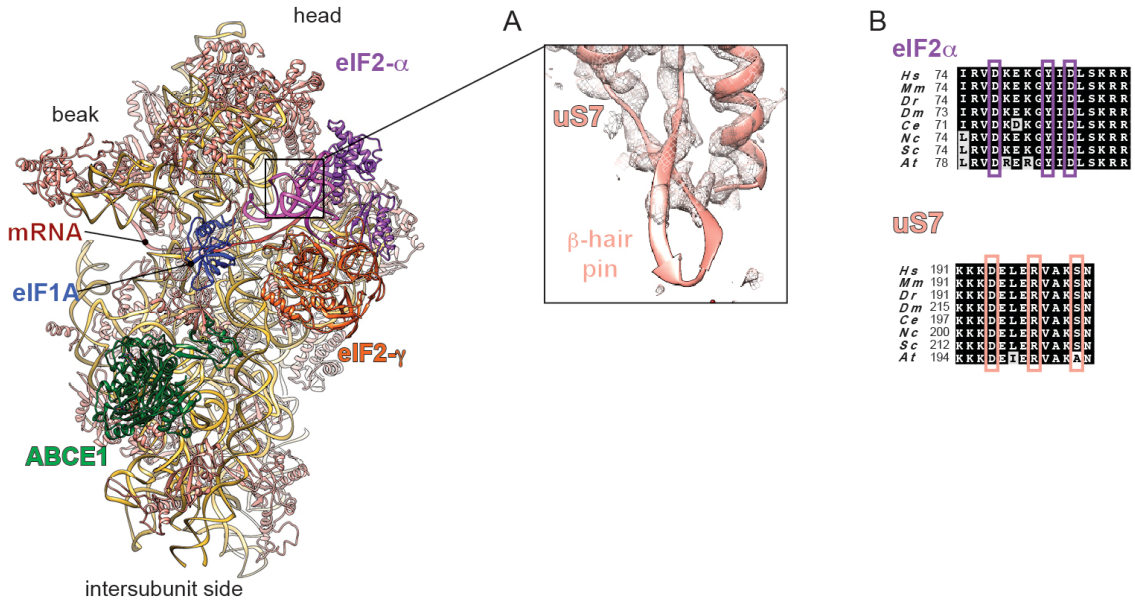

Figure S5, related to Figure 3. Atomic model of uS7 in mammalian LS48S IC. (A) Ribbon representation of uS7 in its electron density, showing the lack of density only in the mRNA-contacting β-hairpin, indicating the flexibility of this part of the protein. (B) Sequence conservation of the interacting residues in eIF2α and uS7, showed by (Visweswaraiah and Hinnebusch, 2017), among eukaryotic species showed in coloured frames.

Figure S6, related to Fig. 7

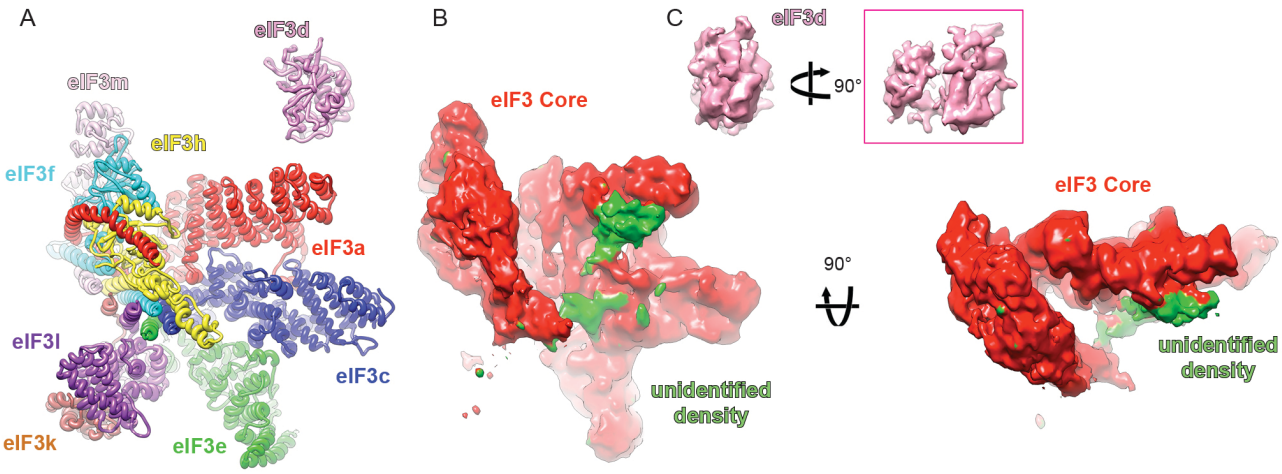

Figure S6, related to Figure 7. Comparison between eIF3 of ICs from near-native conditions and *in vitro* assembly. (A) Atomic model of eIF3 octamer core along with eIF3d subunit (in coloured ribbons). (B) Segmented cryo-EM densities of eIF3 octamer core along with eIF3d (filtered to 8 Å), showing unidentified density (green surface) in contact with eIF3 a and c subunits. (C) Fitting of eIF3d partial crystal structure into its cryo-EM density.

SUPPLEMENTAL TABLES

Table S1, related to Figure 1. Refinement and validation statistics for the three LS48S IC structures from *O.cuniculus*.

Validation[a]	**LS48S IC with β-globin mRNA**	**LS48S IC+eIF3 with β-globin mRNA**	**LS48S IC with H4 mRNA**
PDB ID	**6YAL**	**6YAM**	**6YAN**
Clashscore[b]	11.92 (63rd percentile)	17.51 (40th percentile)	3.67 (97th percentile)
Molprobity score	2.56 (44th percentile)	2.77 (33rd percentile)	1.95 (78th percentile)
Favored rotamers	92.03%	91.49%	93.14%
Allowed rotamers	5.27%	5.41%	5.14%
Poor rotamers	2.70%	3.10%	1.72%
Ramachandran favored	86.02%	85.15%	87.39%
Ramachandran allowed	10.07%	10.83%	9.03%
Ramachandran outliers	3.91%	4.02%	3.58%
Correct sugar puckers	99.68%	99.57%	99.46%
Correct backbone conformation	84.88%	84.52%	83.81%

[a] Compiled using MolProbity (Chen et al., 2010)
[b] Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1,000 atoms.

Table S2, related to Figures 2-7. Summary of the interaction found in the β-globin LS48S and H4 LS48S ICs, described in this paper. The contacts found in this work are highlighted in blue. Stars indicate the interactions with the nucleotide bases.

protein factors	mRNA position β-globin	mRNA position H4	tRNA position	40S component	function
eIF1A					
Lys7	G(+3)				start-codon recognition by the eIF1A NTT
Gly8-Gly9			A(35)		
Trp70	G(+4)*			18S rRNA A1819	Kozak dependent
Arg12	**C(+7)**				
Lys67	G(+6)				
Lys68	U(+5)	**eIF1A absent**			
Gly9, Lys10				18S rRNA C1696	influence on the stability the cognate codon:anticodon duplex by the eIF1A NTT
Lys16, Asn17				C1327	
Asn44, Arg46				C1705 A1817-A1819,	might depend on the mRNA sequence
Lys64 , Arg 62				G604, C605	
Arg82, Tyr84, Gln58				uS12 Leu91, Gly56	
Asp83				eS30 Arg82	
	G(+6)	**U(+6)**		18S rRNA **G616**	
	A(+1)*, C(-1)*			18S rRNA G1203	stabilizing the ASL in the P-site
			C(34)	$m^1acp^3\Psi$1244	
	G(+4)			uS19	
			A(35)	**Arg140**	
			G(29), G(30)	**Thr136**	
				uS13	
			U(28), G(29)	**Thr145, Gly147**	
				uS9	
			C(33), C(32)	**Arg 146**	
				eS26	reinforcement of start-codon recognition in case of suboptimal Kozak sequence
	C(-4)*	**A(-4)***		**His80**	
	(-8)			Ile41	stabilizing the mRNA channel exit site
	(-9)			Arg42	
	(-8), (-9)			Arg100	
				eS28	
	(-5)			Arg66	
	(-7)			Arg67	

protein factors	mRNA position		tRNA position	40S component	function
	β-globin	H4			
eIF2α Arg55 Arg57, Ser58, Asn60, Lys61	C(-3)*		C(39-41)		stabilizing the ASL in the P-site
	mRNA entry channel (+14) to (+18) (+9), (+10) (+12)*, (+13)* A(+13)			uS3 Arg117 helix α (117-128) β-hairpin (142-146) eS30 Lys126 uS5 Ala133	stabilizing the mRNA
ABCE1 HLH and NBD1: Ser150, Arg306, Asn310 NBD2: Lys584, Ile583, Arg566, Arg567 NBD1: **Pro265, Asp266** Fe-S cluster: Arg7, Lys20, Pro66, Asn74 Pro30, Val31, Arg33, Ile56, Ile60				18S rRNA U478, A455, A454, C453, C452 A455, A454, C453 eS24 **C-terminal helix** **Gly128** 18S rRNA C471 G1718 A1719 G470 uS12 Ile50, Leu52, Leu62, Glu63, Ile75	binding to 40S
eIF3c **Asn388** **Arg340, Asn384** **Gly341, Lys343** **Arg450** **Lys342** **Thr391, Tyr392** **Lys343** eIF3a **Asn10** **Lys13** **Arg14** **Phe18**				eS27 **Glu75** **Thr61** **Gln65** **Cys59** 18S rRNA **G925** **C1112** **U1116** eS1 **Asp77** **Asn76** **Asp191** **Pro190**	binding to the 40S platform site
Gln6, Arg7, Arg41, Gln44, Lys45 eIF3d **S166–E172, Asn513, Lys514**	**5'UTR** **5'UTR**				binding to mRNA 5'UTR