Molecular Cell, Volume *79*

Supplemental Information

Selective Translation Complex Profiling Reveals Staged Initiation and Co-translational Assembly of Initiation Factor Complexes

Susan Wagner, Anna Herrmannová, Vladislava Hronová, Stanislava Gunišová, Neelam D. Sen, Ross D. Hannan, Alan G. Hinnebusch, Nikolay E. Shirokikh, Thomas Preiss, and Leoš Shivaya Valášek

SUPPLEMENTAL TABLES

Table S1. Libraries overview. Related to STAR Methods.

(smaller: Indicates a smaller insert size cut off (~20 to 50 nt). These yeast libraries were only used for assessment of the co-translational assembly.)

	repli-cate	TCP-seq 80S	TCP-seq 40S	Sel-TCP-seq 40S				80S		
				eIF3b	eIF3a	eIF3c	eIF2β	eIF3b	eIF3a	eIF3c
Human HEK293T	1	x	x	x				x		
	2	x	x							
	3	x	x	x				x		
Yeast *S.cere-visiae*	1-1	x	x		x				x	
	1-2	x	x			x				x
	1-3	x	x				x			
	2-1				x				x	
	2-2					x				x
	2-3						x			
smaller	3	x							x	
smaller	4	x							x	

Table S2. Yeast strains used in this study. Related to STAR Methods.

Strain	Genotype	Reference
SY182	*MAT*a *leu2-3, -112 ura3-52 trp1Δ gcn2Δ tif32Δ URA3::GCN2 ura3* (pRS315-a/TIF32-FLAG-L)	This study
SY183	*MAT*a *leu2-3, -112 ura3-52 trp1Δ gcn2Δ tif32Δ URA3::GCN2 ura3* (YEp181-a/TIF32-L)	This study
LMY61 (SY194)	*mat@ leu2-3, -112 ura3-53 ino1 sui3Δ gcn2Δ p921* (SUI3-FLAG-L)	Alan Hinnebusch
H25	*ura3-52 trp1-63 leu2-3, -112 his4-303(AUU) SUI1 nip1Δ* (YCplac22-c/NIP1-FLAG-T)	Leoš Valášek

SUPPLEMENTAL FIGURES AND FIGURE LEGENDS

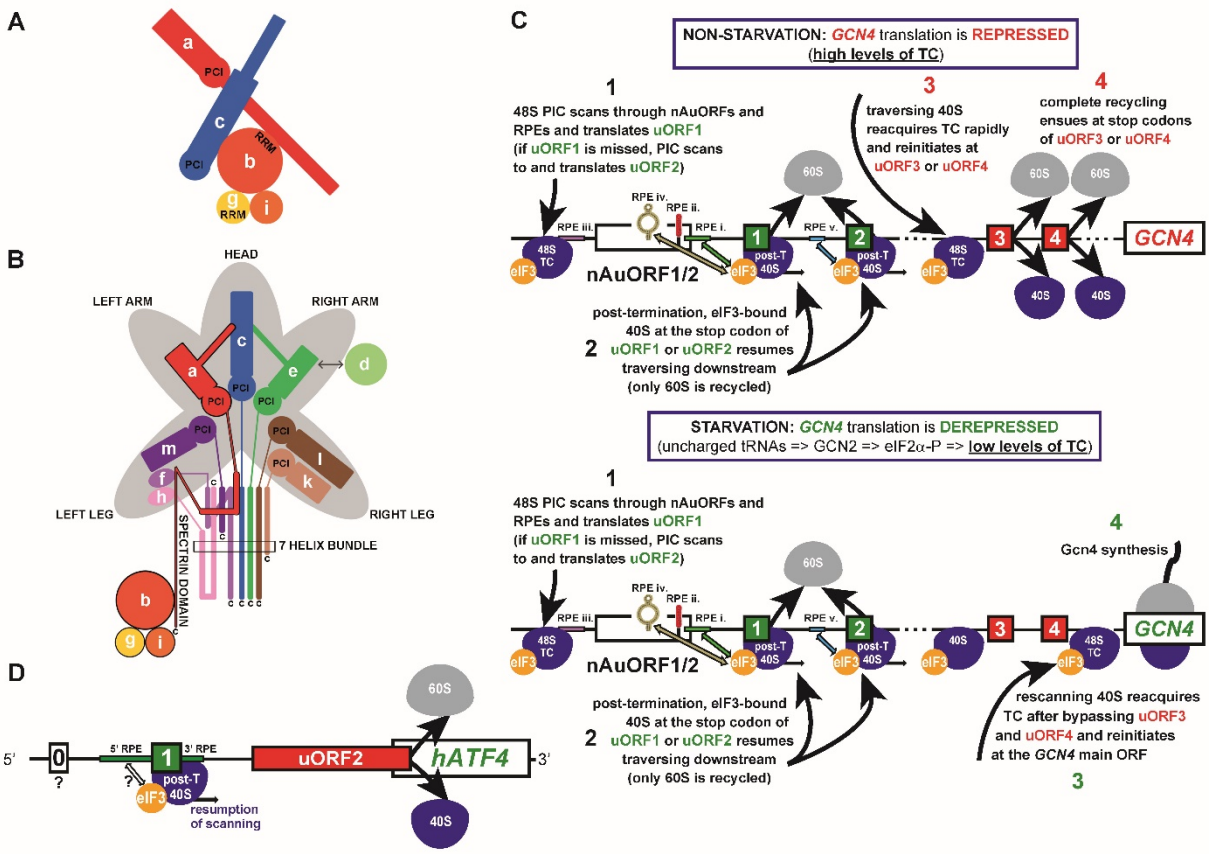


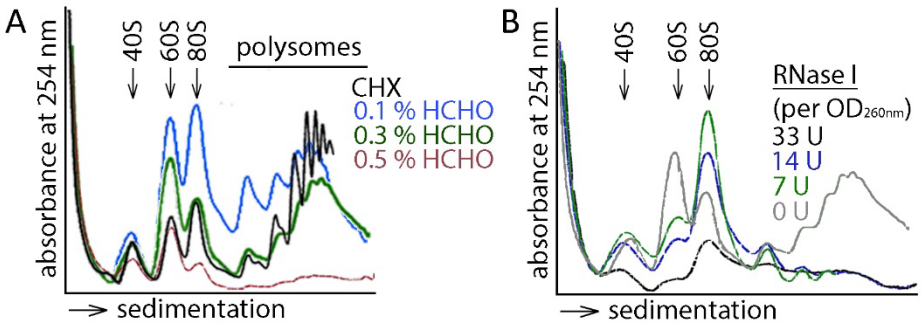
Figure S1. Schematics of eIF3 and translational regulation of *GCN4/ATF4*. Related to Figures 3 to 7.

(A, B) Similarities and differences between budding yeast and mammalian eIF3. The complex is formed by 5 subunits, a, b, c, g, and i in *S. cerevisiae* (A), while in mammals it consists of 12 subunits, a, b, c, d, e, f, g, h, i, k, l, and m (B). Notably, the proteasome, COP9, initiation factor 3 (PCI) structural domain is shared by several eIF3 subunits; adopted from (Zeman et al., 2019). RRM: RNA recognition motif.

(C) *GCN4* mRNA 5'UTR and regulation. The reinitiation (REI)-permissive uORF1 is translated under both nutrient-replete and depleted conditions. After its translation, the 60S ribosomal subunit is released, whereas the 40S subunit remains bound to the *GCN4* mRNA – due to a specific interaction between REI-promoting elements (RPEs) of uORF1 and eIF3 (Mohammad et al., 2017; Munzarová et al., 2011) – to eventually resume traversing downstream. It cannot start "scanning" per se for the next AUG until it re-acquires an active eIF2 ternary complex (eIF2-TC), the levels of which are reduced under starvation/stress conditions. uORF2, functionally mimicking uORF1, was proposed to serve as its backup to capture 40Ses that 'leaky-scanned' past uORF1, thereby maximizing the REI potential as a fail-safe mechanism (Gunisova and Valasek, 2014). In non-starved cells, where eIF2-TC levels are high, nearly all of the traversing 40Ses rebind the eIF2-TC before reaching the REI-non-permissive uORF3 or uORF4. Upon their translation, terminating ribosomes undergo

full ribosomal recycling, thus preventing REI at the main CDS of *GCN4*. It is noteworthy that uORF4 and not the preceding uORF3 has been thought to represent the main negative element of this system (Hinnebusch, 2005) (see further below). When eIF2-TC levels are low, a large proportion of the 40Ses will bypass uORFs 3 and 4 and – upon eventual acquisition of the eIF2-TC – reinitiate at the *GCN4* start codon. Hence, whereas the stress response shuts down general translation initiation, as eIF2-TCs are required for translation of most mRNAs, it at the same time stimulates *GCN4* translation to trigger stress adaptation programs (reviewed in (Gunisova et al., 2018; Hinnebusch, 2005).

(D) *ATF4* mRNA 5'UTR and regulation. The REI-permissive uORF1 is translated under both non-stress and stress conditions. Upon termination, the 40S remains bound to the *ATF4* mRNA to eventually resume traversing downstream. Under high levels of the eIF2-TC, 40Ses reinitiate at uORF2 whose translation prevents reinitiation at the *ATF4* start codon. Under low eIF2-TC levels, the majority of 40Ses skip uORF2 and reinitiate at *ATF4* instead (not shown) (Lu et al., 2004; Vattem and Wek, 2004). The role of uORF0 is unknown.



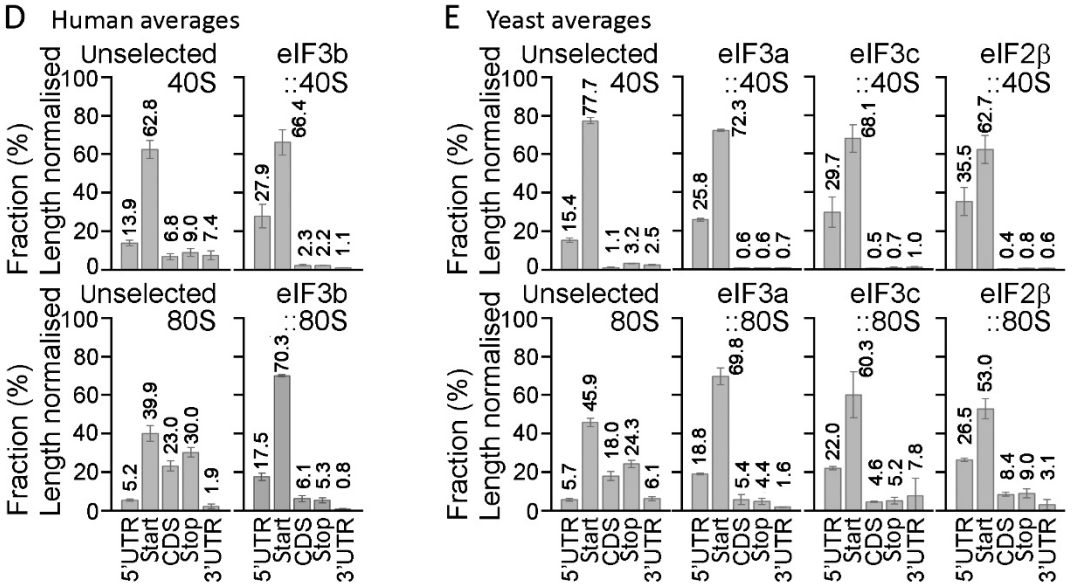
C Pearson correlation of FP count per gene between replicates

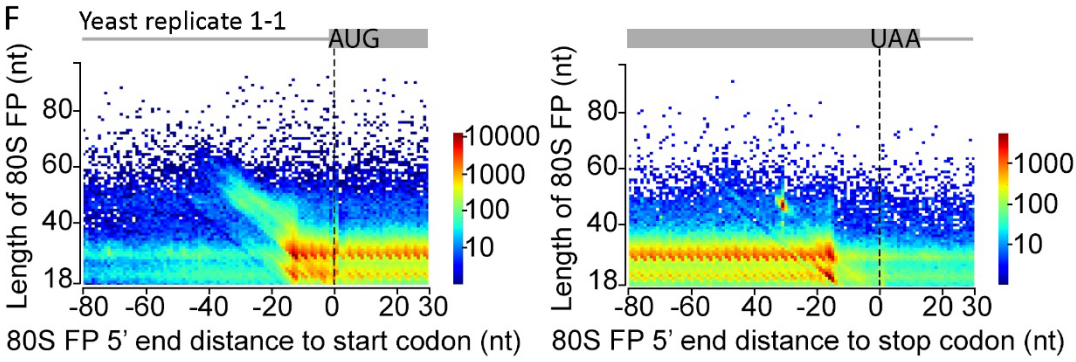
HEK 40S	1	2
repl. 1		
repl. 2	0.89	
repl. 3	0.92	0.73

HEK 80S	1	2
repl. 1		
repl. 2	0.97	
repl. 3	0.98	0.96

yeast 40S	1-1	1-2
repl. 1-1		
repl. 1-2	0.98	
repl. 1-3	0.99	0.99

yeast 80S	1-1	1-2
repl. 1-1		
repl. 1-2	0.97	
repl. 1-3	0.99	0.99





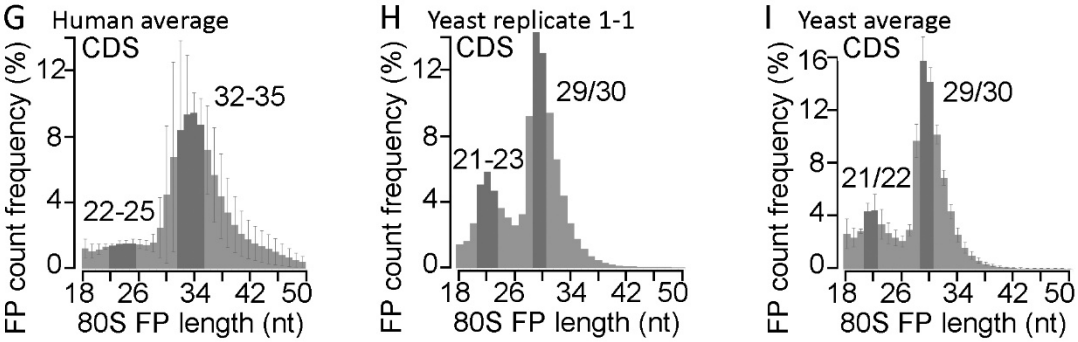


Figure S2. TCP-seq in yeast and human cells – experimental conditions, replicate correlation, and selected characteristics of ribosomal complex footprints. Related to Figure 1.

(A) Polysome profiles of extracts of HEK293T cells that were either treated with cycloheximide (CHX) or crosslinked with different concentrations of formaldehyde (HCHO) as indicated.

(B) Polysome profiles of extracts after crosslinking HEK293T cells with 0.3% formaldehyde that were either left undigested (0U; grey; corresponds to the green profile in (A)) or digested with different RNase I concentrations as indicated.

(C) Pearson correlation coefficients of the normalised footprint (FP) count per gene when comparing replicates of unselected 40S or 80S FP libraries, human (HEK) or yeast (see Table 1 for a full listing of replicates).

(D) Human 40S (top) and 80S (bottom) FP density in different mRNA regions (indicated below the x-axis). FP counts per region were normalised to average region length. Averages of replicates 1, 2 and 3 are shown.

(E) As (D) for all yeast FP densities. Averages of replicates 1-1, 1-2, 1-3 are shown.

(F) Yeast 80S FP length versus 5' end position relative to the first nucleotide (position 0) of start (left; 6,639 sites) or stop codons (right; 6,514 sites). Colour scale represents FP count as indicated to the right. Panel shows representative data from biological replicate 1-1.

(G) Length distribution of human 80S FPs within 9,241 coding sequences (CDS; excluding start and stop codon associated FPs). Averages of replicates 1, 2 and 3 are shown.

(H) As (G) but for 5,818 yeast CDSes using data from replicate 1-1.

(I) As (H) but the average of replicates 1-1, 1-2 and 1-3 are shown.

Error bars in (D), (E), (G), and (I) indicate +/- standard deviation.

See STAR Methods for experimental details, sequencing data treatment as well as selection of mRNAs and appropriate regions for inclusion into plots throughout the figures.

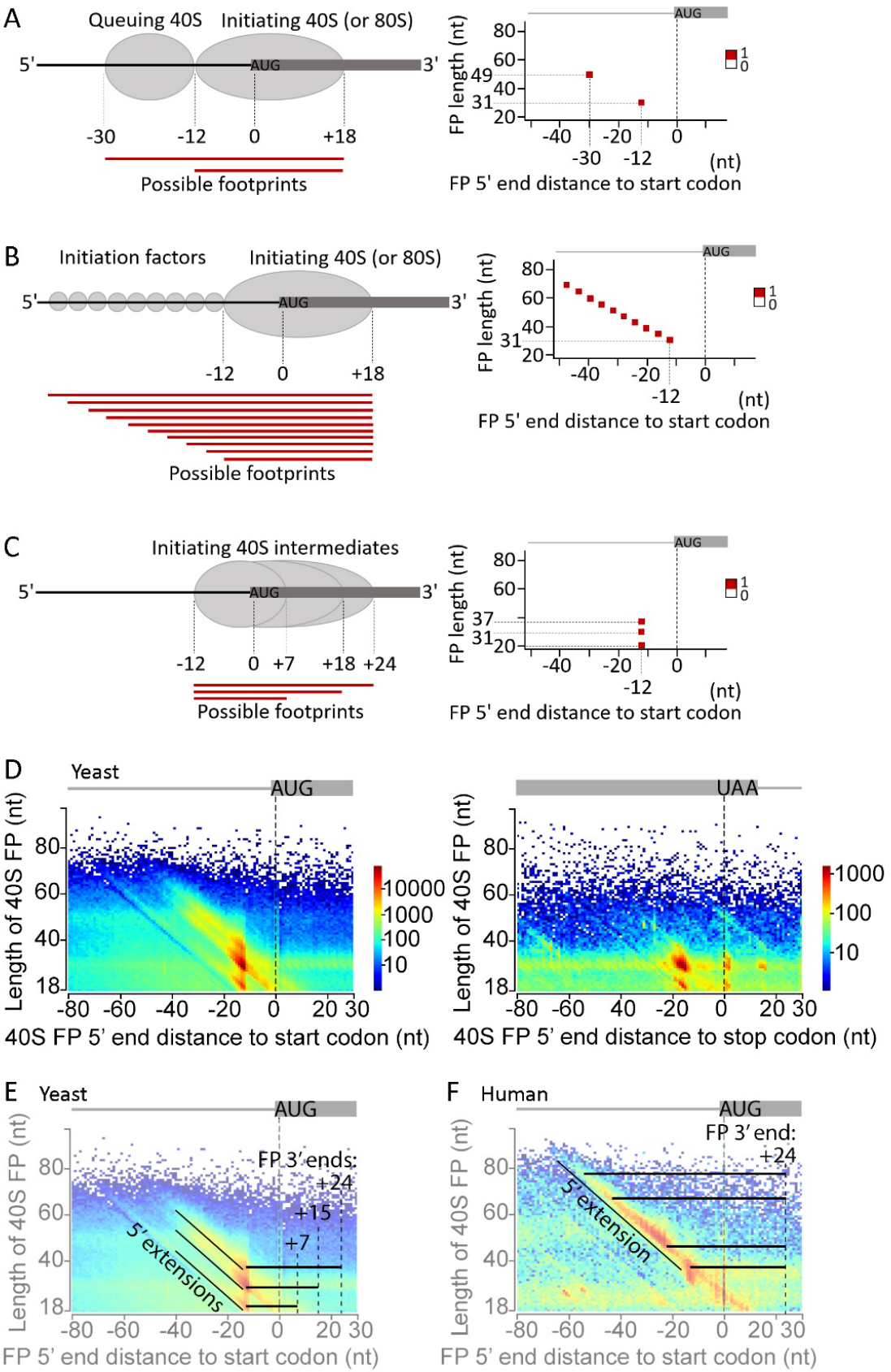

A
Queuing 40S
Initiating 40S (or 80S)
5'
AUG
3'
-30
-12
0
+18
Possible footprints
FP length (nt)
80
60
49
40
31
20
-40
-20
0
-30
-12
(nt)
FP 5' end distance to start codon
B
Initiation factors
Initiating 40S (or 80S)
5'
AUG
3'
-12
0
+18
Possible footprints
FP length (nt)
80
60
40
31
20
-40
-20
0
-12
(nt)
FP 5' end distance to start codon
C
Initiating 40S intermediates
5'
AUG
3'
-12
0
+7
+18
+24
Possible footprints
FP length (nt)
80
60
37
31
20
-40
-20
0
-12
(nt)
FP 5' end distance to start codon
D
Yeast
AUG
Length of 40S FP (nt)
80
60
40
18
10000
1000
100
10
-80
-60
-40
-20
0
20
30
40S FP 5' end distance to start codon (nt)
UAA
1000
100
10
40S FP 5' end distance to stop codon (nt)
E
Yeast
AUG
FP 3' ends:
+24
+15
+7
5' extensions
Length of 40S FP (nt)
FP 5' end distance to start codon (nt)
F
Human
AUG
FP 3' end:
+24
5' extension
Length of 40S FP (nt)
FP 5' end distance to start codon (nt)

Figure S3. Interpretive schematics of 40S and 80S footprint morphology at start codons and TCP-seq in yeast – small ribosomal subunit. Related to Figures 1 and 2.

(A-C) Plausible explanations for the three main types of FP 5’ and 3’ end variations observed at start codons (left panels) and corresponding patterns seen in metagene plots (right panels).

(A) An initiating 40S/80S and a scanning 40S queuing upstream produce 5’ extended FPs with a 5’ end position at around -30 relative to the first nucleotide (position 0) of start codons.

(B) An initiating 40S/80S and variable numbers (initiation) factors upstream produce 5’ extended FPs with a plurality of 5’ end positions.

(C) An initiating 40S complex undergoes conformational/compositional changes on the entry channel side during start codon recognition producing a small number of distinct 3’ extended FPs.

(D) Yeast 40S footprint (FP) length versus 5’ end position relative to the first nucleotide (position 0) of start (left; 6,639 sites) or stop codons (right; 6,514 sites). Colour scale represents FP count as indicated to the right. Data from yeast biological replicate 1-1 are shown.

(E) Overlaid bars to assist in relating actual yeast FP morphology from (D) to conceptual diagrams in (A-C).

(F) Overlaid bars to assist in relating actual human FP morphology from Figure 2A to conceptual diagrams in (A-C).

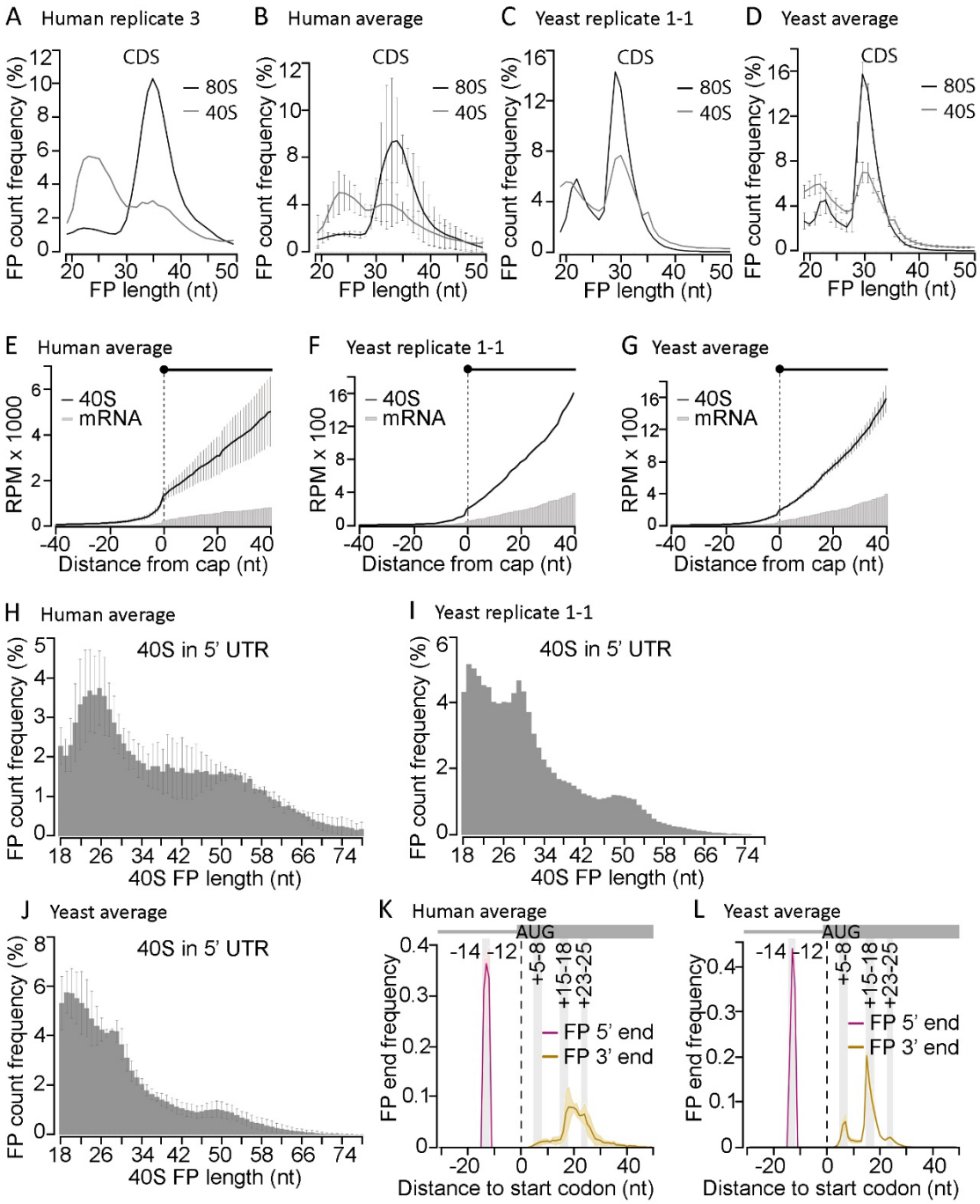


Figure S4. TCP-seq in yeast and human cells – small ribosomal subunit. Related to Figure 2.

(A) Length distribution of human 40S FPs assigned to annotated CDS of 9,241 mRNAs (excluding start and stop codon assigned FPs). Data from human replicate 3 are shown.

(B) As (A) but averages of human replicates 1, 2 and 3 are shown.

(C) As (A) but data (5,818 mRNAs) from yeast replicate 1-1 are shown.

(D) As (A) but averages of yeast replicates 1-1, 1-2 and 1-3 are shown.

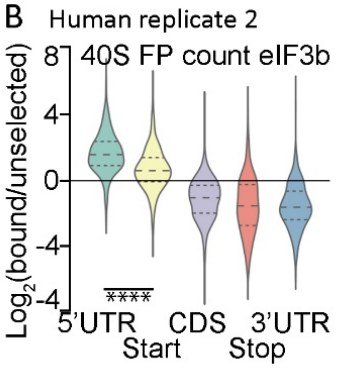
(E) Human 40S FP coverage and mRNA coverage (from ‘regular’ RNA-seq) versus distance from transcription start sites (cap; 1,046 mRNAs, excluding start codon-assigned FPs). RPM, reads per million. Related to Figure 2B.
(F) As (E) but data (198 mRNAs) from yeast replicate 1-1 are shown.
(G) As (E) but averaged data (198 mRNAs) from yeast replicates 1-1, 1-2 and 1-3 are shown.
(H) Length distribution of human 40S FPs assigned to the 5’UTR of 9,241 mRNAs (excluding start codon-assigned FPs). Averages human replicates 1, 2 and 3 are shown. Related to Figure 2C.
(I) As (H) but data (5,818 mRNAs) from yeast replicate 1-1 are shown.
(J) As (H) averages of yeast replicates 1-1, 1-2 and 1-3 are shown.
(K) Human start codon-aligned 40S FP 5’ and 3’ end positions (11,174 sites). FPs with 5’ ends between -14 to -12 positions were chosen to display their 3’ end distribution. Grey vertical grey bars shown yeast 3’ end positions (as in Figures 2D, S4L). Averages human replicates 1, 2 and 3 are shown; shading around line indicates +/- standard deviation. Related to Figure 2E.
(L) As (K) but averaged data (5,994 mRNAs) from yeast replicates 1-1, 1-2 and 1-3 are shown. Related to Figure 2D.
Error bars in (B, D, E, G, H, and J) indicate +/- standard deviation.

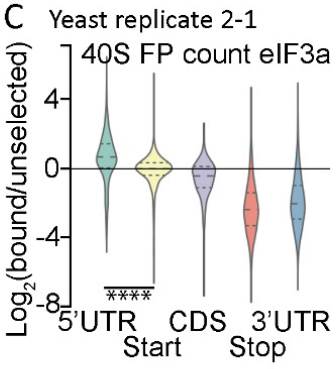
A Pearson correlation of unselected vs. selected 40S libraries 5'UTR+start

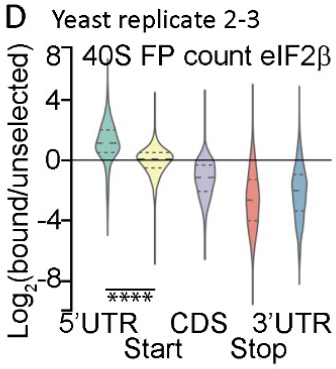
Yeast eIF3a replicate 2	0.92
Yeast eIF2β replicate 2	0.87
Yeast eIF3c replicate 1	0.95
Yeast eIF3c replicate 2	0.81
Human eIF3b replicate 1	0.90

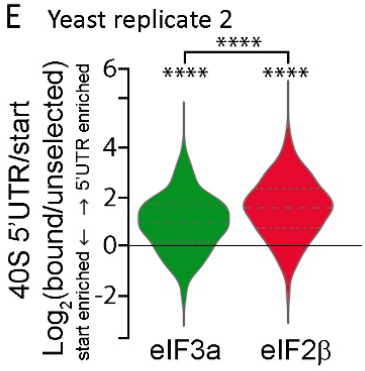
Pearson correlation of FP count per gene between replicates

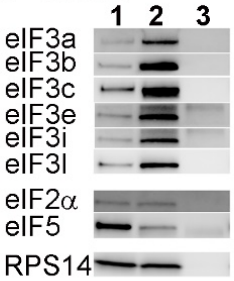
Yeast eIF3a repl. 1 vs. 2	0.97
Yeast eIF2β repl. 1 vs. 2	0.91
Yeast eIF3c repl. 1 vs. 2	0.89
Human eIF3b repl. 1 vs. 3	0.97



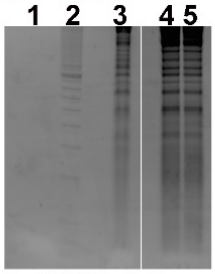




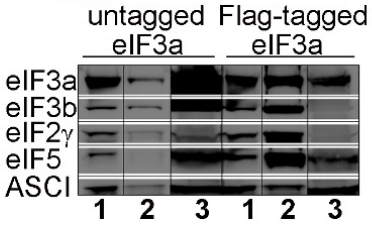




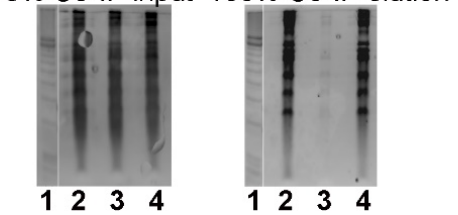
1 4% input
2 30% elution
3 30% elution (Co-IP control)



1 100% elution of Co-IP control
2 ladder
3 100% elution of α-eIF3b::40S
4+5 unselected 40S (10% of Co-IP input)



1 1% input
2 33% elution
3 10% supernatant



1 ladder
2 eIF3a-FLAG (strain SY182)
3 untagged eIF3a (strain SY183)
4 eIF2β-FLAG (strain SY194)

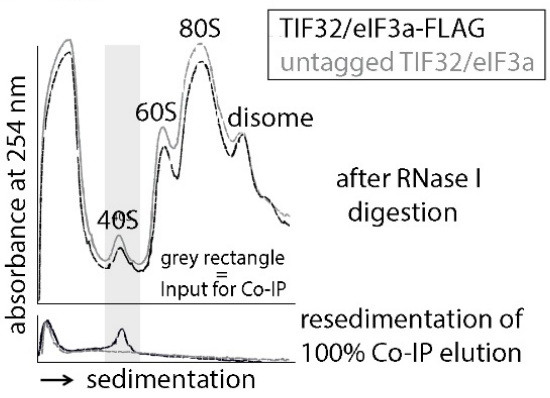


Figure S5. Sel-TCP-seq in yeast and human cells – small ribosomal subunit and controls for immunoprecipitation. Related to Figure 3.

(A) Upper table: Pearson correlation coefficients of the normalised FP count per gene, when comparing FOI-selected and unselected 40S libraries for yeast replicate 2 (eIF3a and eIF2β as FOIs), yeast replicates 1 and 2 (eIF3c as FOI), and human replicate 1 (eIF3b as FOI). Related to Figure 3A-C. Lower table: Pearson correlation coefficients of the normalised FP count per gene, when comparing replicates of the same FOI::40S libraries.

(B) Human eIF3b::40S versus unselected 40S FP count ratios within indicated mRNA regions (n = 1,129; dashed line: median; dotted lines: 1st and 3rd quartiles; ****: p-value << 0.0001, two sample t-test). Data from human replicate 2 are shown. Related to Figure 3G.

(C) As (B) but for yeast eIF3a::40S versus unselected 40S FP count ratios (n = 1,797). Data from yeast replicate 2-1 are shown. Related to Figure 3H.

(D) As (B) but for yeast eIF2β::40S versus unselected 40S FP count ratios (n = 1,749). Data from yeast replicate 2-3 are shown. Related to Figure 3I.

(E) Differences in FOI (indicated below the x axis) persistence upon initiation complex arrival at start codons. The ratio of FOI::40S FP counts in 5'UTR versus start codon regions of 1,070 mRNAs was divided by the equivalent ratio of unselected 40S FP counts. Both FOI::40S complexes redistribute to 5'UTRs relative to the unselected 40S (**** p << 0.0001, one sample t-test), but this is more pronounced for eIF2β than eIF3a (**** p << 0.0001, two sample t-test). Data from yeast replicate 2 are shown. Related to Figure 3L.

(F) Western blot of anti-eIF3b co-immunoprecipitation fractions. Input: human RNase I-digested 40S fraction from formaldehyde-crosslinked cells. Elution: Material eluted from beads either exhibiting the anti-eIF3b antibody or from a control experiment with beads lacking the antibody.

(G) Gel electrophoresis of RNA from eluates after anti-eIF3b co-immunoprecipitation and control as described in (F). Ladder: 10 bp DNA ladder

(H) Western blot of anti-FLAG co-immunoprecipitations with the RNase I-digested 40S fractions derived from formaldehyde-crosslinked yeast cells expressing FLAG-tagged eIF3a, or control cells without tag.

(I) Gel electrophoresis of RNA from eluates after an anti-FLAG co-immunoprecipitation from tagged and control yeast material as described in (H). Ladder: 10 bp DNA ladder

(J) Re-sedimentation of the eluates of an anti-FLAG co-immunoprecipitation from the 40S fractions of formaldehyde-crosslinked and RNase I-digested yeast cells expressing FLAG-tagged eIF3a or untagged eIF3a.

In (F-J) percentages given are relative to the amount of starting material.

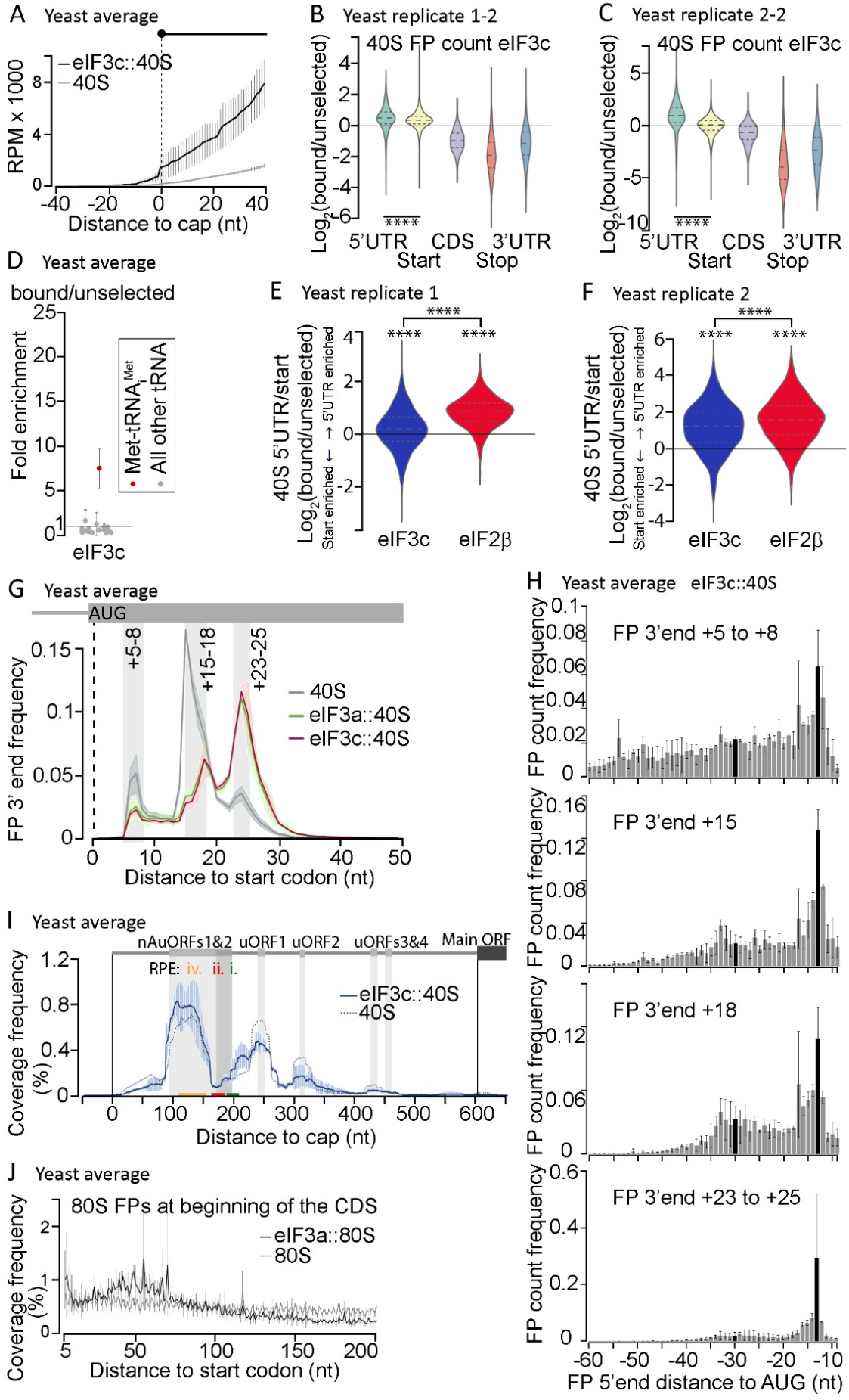

A Yeast average
RPM x 1000
eIF3c::40S
40S
Distance to cap (nt)
B Yeast replicate 1-2
40S FP count eIF3c
Log2(bound/unselected)
5'UTR CDS 3'UTR
Start Stop
****
C Yeast replicate 2-2
40S FP count eIF3c
D Yeast average
bound/unselected
Fold enrichment
Met-tRNAiMet
All other tRNA
eIF3c
E Yeast replicate 1
40S 5'UTR/start
Log2(bound/unselected)
Start enriched ← → 5'UTR enriched
eIF3c eIF2β
F Yeast replicate 2
G Yeast average
AUG
+5-8
+15-18
+23-25
40S
eIF3a::40S
eIF3c::40S
FP 3' end frequency
Distance to start codon (nt)
H Yeast average eIF3c::40S
FP count frequency
FP 3'end +5 to +8
FP 3'end +15
FP 3'end +18
FP 3'end +23 to +25
FP 5'end distance to AUG (nt)
I Yeast average
nAuORFs1&2 uORF1 uORF2 uORFs3&4 Main ORF
RPE: iv. ii. i.
eIF3c::40S
40S
Coverage frequency (%)
Distance to cap (nt)
J Yeast average
80S FPs at beginning of the CDS
eIF3a::80S
80S
Coverage frequency (%)
Distance to start codon (nt)

Figure S6. Yeast eIF3c::40S Sel-TCP-seq. Related to Figures 3-5.

(A) Yeast eIF3c::40S FP and unselected 40S FP coverage aligned to annotated transcription start sites (cap; 198 mRNAs, excluding start codon associated FPs). Averages of replicates 1 and 2 are shown; RPM, reads per million. Related to Figures 3E,F.

(B) Yeast eIF3c::40S versus unselected 40S FP count ratios within indicated mRNA regions (n = 1,499; dashed line: median; dotted lines: 1st and 3rd quartiles; **** p << 0.0001, two sample t-test). Data from yeast replicate 1-2 are shown. Related to Figure 3H,I.

(C) As (B) but data (n = 1,510) for yeast replicate 2-2 are shown. Related to Figure S5C,D.

(D) Enrichment of tRNA read counts in yeast eIF3c::40S versus unselected 40S library. Averages of yeast replicates 1 and 2 are shown. Related to Figure 3K.

(E) Differences in FOI (indicated below the x axis) persistence upon initiation complex arrival at start codons. The ratio of FOI::40S FP counts in 5'UTR versus start codon regions of 1,068 mRNAs was divided by the equivalent ratio of unselected 40S FP counts. Both FOI::40S complexes redistribute to 5'UTRs relative to the unselected 40S (**** p << 0.0001, one sample t-test), but this is more pronounced for eIF2β than eIF3c (**** p << 0.0001, two sample t-test). Data from yeast replicate 1 are shown. Related to Figure 3L.

(F) As (E) but for data (n = 1,065) from yeast replicate 2. Related to Figure S5E.

(G) Yeast start codon-aligned, unselected and FOI::40S, FP 3' end position frequencies (5,994 sites). Grey vertical bars are shown as in Figure 2D. The averages of yeast replicates 1 and 2 are shown; shading around lines indicates +/- standard deviation. Related to Figure 4A.

(H) 5' end position frequencies of eIF3c::40S FPs ending at specific 3' end positions as indicated. Selection matches grey bars in (G) but respecting the partition of the +15 to +18 area into two peaks. For better orientation, bars representing the -30 and -13 positions are highlighted in black. Error bars indicate +/- standard deviation. Averages of yeast replicates 1 and 2 are shown. Related to Figure 4D,E.

(I) Yeast unselected eIF3c::40S coverage frequency along the 5'UTR of *GCN4* mRNA (schematic with regulatory elements shown on top). uORF locations are emphasised by vertical grey shading; positions of reinitiation-promoting elements (RPE) i., ii., and iv. are further indicated by coloured bars on x-axis. Unselected 40S track is indicated as dashed line. Averages of replicates 1 and 2 are shown. Related to Figure 5C,D.

(J) Yeast unselected 80S FP and eIF3a::80S FPs coverage frequency at the beginning of the CDS of 6,638 mRNAs. Averages of yeast replicates 1 and 2 are shown. Related to Figure 3M.

Error bars in (A, D, H-J) represent +/- standard deviation.

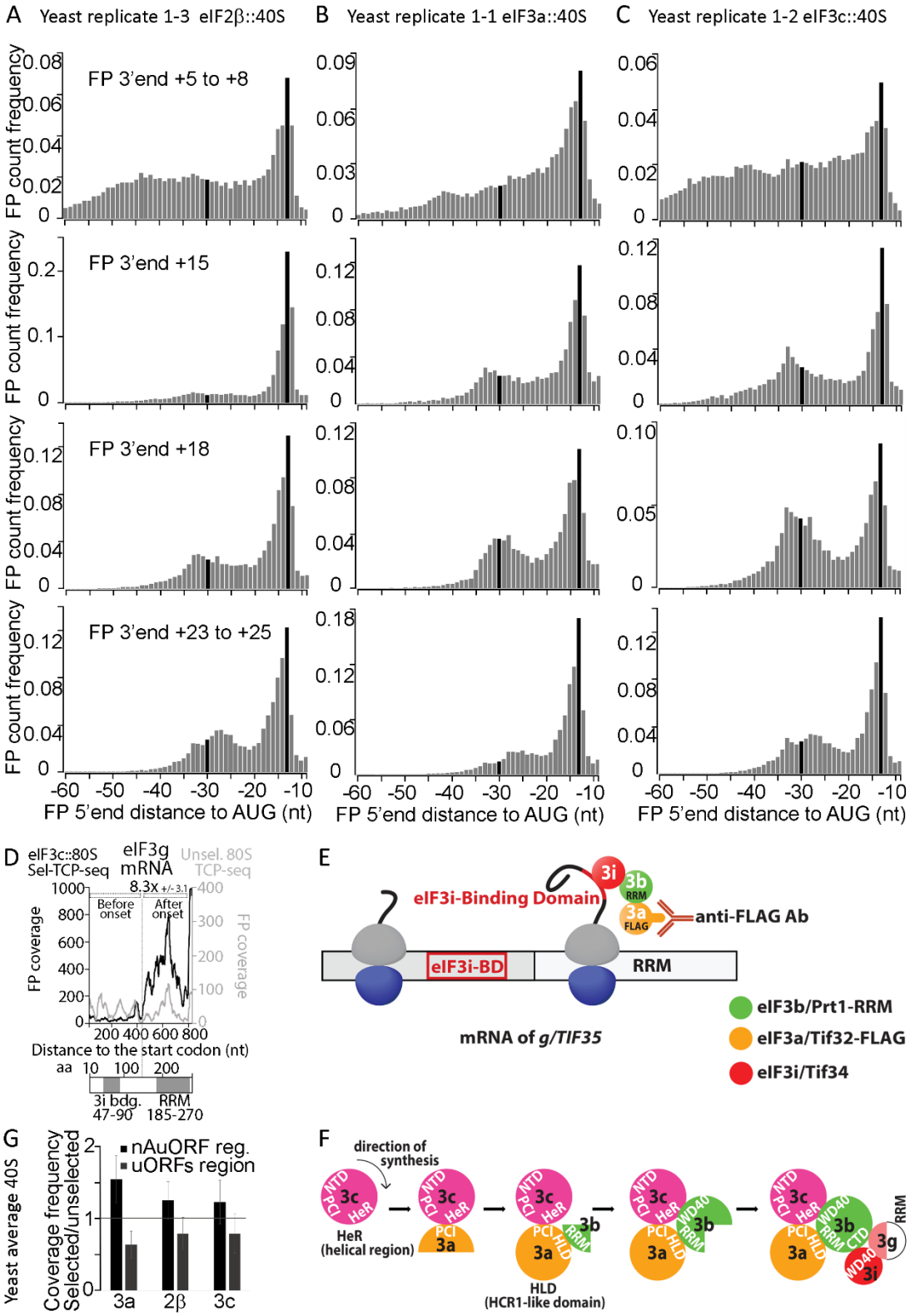


Figure S7. Related to Figures 4, 5, and 7.

(A) 5' end position frequencies of eIF2β::40S FPs ending at specific 3' end positions as indicated to the right of each row. Selection matches the grey bars in Figure 4A but respecting the partition of the +15 to +18 area into two peaks. For orientation, bars representing the -30 and -13 positions are highlighted in black. Data for yeast replicate 1-3 is shown. Related to Figure 4D.

(B) As (A) but for eIF3a::40S FPs using data from yeast replicate 1-1. Related to Figure 4E.
(C) As (A) but for eIF3c::40S FPs using data from yeast replicate 1-2. Related to Figure S6H.
(D) FP coverage tracks for selected (black; eIF3c as 'bait' protein indicated on top of left axis) and unselected 80S FP (grey; right axis) across the mRNA encoding 'target' protein eIF3G (domain structure shown underneath graph). Dotted line indicates selected 80S FP coverage increase. Tracks from replicate 1 are shown and average fold induction from replicates 1 and 2 is given with standard deviation. Fold induction of the 80S Sel-TCP-seq coverage was normalized to that seen with unselected 80S. Related to Figure 7C.
(E) Schematic illustrating the co-translational interactions of the yeast eIF3a-b-i subcomplex with nascent eIF3g, where eIF3i and eIF3b subunits mediate the contact between eIF3a-FLAG and nascent eIF3g, as revealed by FLAG co-immunoprecipitation in a strain expressing FLAG tagged eIF3a. RRM: RNA recognition motif.
(F) Schematic illustrating the co-translational assembly pathway of eIF3 in yeast. The fully synthesized eIF3c co-translationally assembles with eIF3a forming an eIF3a-3c dimer, which subsequently picks up nascent eIF3b. At first, through a direct interaction between eIF3a and the eIF3b-RRM, later by an additional contact mediated by the helical region of eIF3c and the C-terminal end of eIF3b, forming a stable trimeric eIF3a-b-c subcomplex. This is consistent with earlier analysis demonstrating that this subcomplex is stable in cells, binds eIF1 as well as eIF5, and stimulates ribosome binding of mRNA and Met-tRNA$_i^{Met}$ *in vitro* (Phan et al., 2001)). Though, it contrasts with the human eIF3 assembly pathway, where the eIF3a and eIF3b subunits were proposed to form a nucleation core around which all other eIF3 subunits assemble into a 12-subunit complex (Wagner et al., 2016). This trimer then co-assembles with eIF3i but only once it is fully synthetized, contrary to eIF3g that can be picked up co-translationally. RRM: RNA recognition motif. NTD: N-terminal domain. CTD: C-terminal domain. WD40: WD40 domain. PCI: Proteasome-COP9-Initiation factor 3 domain.
(G) The ratio of selected versus unselected 40S FP coverage frequency is shown for two areas of the GCN4 5' UTR. nAuORF region: +90 nt to +165 nt; uORFs region: +225 nt to +500 nt (0 corresponds to the transcript start). Averages from replicates 1 and 2 are plotted with standard deviations. Related to Figures 5C,D and Figure S6I.