

Kinetic Analysis of Mammalian Translation Initiation

Dissertation

for the award of the degree

“Doctor rerum naturalium” (Dr. rer. nat.)

of the Georg-August-Universität Göttingen

within the doctoral program IMPRS Molecular Biology
of the Georg-August University School of Science (GAUSS)



submitted by

Sung-Hui Yi

易松輝

from Taitung, Taiwan

Göttingen, 2020

Members of the Thesis Committee and the Examination Board

Prof. Dr. Marina V. Rodnina (1st Referee)
Department of Physical Biochemistry
Max Planck Institute for Biophysical Chemistry
Göttingen, Germany

Dr. Sarah Adio (2nd Referee)
Department of Molecular Structural Biology
Georg-August University Göttingen
Göttingen, Germany

Prof. Dr. Markus T. Bohnsack (3rd Referee)
Department of Molecular Biology
University Medical Center Göttingen
Göttingen, Germany

Prof. Dr. Kai Tittmann
Department of Molecular Enzymology
Georg-August University Göttingen
Göttingen, Germany

Further members of the Examination Board

Prof. Dr. Holger Stark
Department of Structural Dynamics
Max Planck Institute for Biophysical Chemistry
Göttingen, Germany

Prof. Dr. Henning Urlaub
Bioanalytical Mass Spectrometry Group
Max Planck Institute for Biophysical Chemistry
Göttingen, Germany

Date of the oral examination: December 15th, 2020

Affidavit

I hereby declare that the presented dissertation entitled "Kinetic Analysis of Mammalian Translation Initiation" has been written independently and with no other sources and aids than quoted.

Göttingen, November 16th, 2020

Sung-Hui Yi

Contents

ABSTRACT	1
1 INTRODUCTION	3
1.1 Mammalian translation cycle	3
1.2 Translation initiation	5
1.3 Cap-dependent translation initiation	7
1.4 43S preinitiation complex formation	8
1.4.1 Formation of the ternary complex, eIF2·GTP·Met-tRNA _i ^{Met}	10
1.4.2 Binding of individual initiation factors to the 40S subunit	11
1.5 mRNA recruitment.....	18
1.5.1 Activation of mRNA by the cap-binding complex, eIF4F	18
1.5.2 Unwinding of secondary structures in 5' UTR by RNA helicases.....	19
1.5.3 mRNA recruitment to the 43S PIC.....	20
1.6 mRNA scanning.....	21
1.6.1 Structures of mammalian 5' untranslated region	21
1.6.2 Scanning.....	25
1.6.3 eIF1 and eIF1A stabilize the scanning complex.....	26
1.6.4 Secondary structures and RNA helicases	27
1.6.5 Leaky scanning and uORF	28
1.7 Non-canonical initiation pathway	29
1.7.1 Cap-dependent scanning-free initiation	31
1.7.2 Cap-independent translation enhancers (CITEs).....	31
1.7.3 Internal ribosome entry sites (IRES) -mediated initiation.....	33
1.8 Start codon recognition and 48S initiation complex formation	35
1.8.1 Relocation of eIF1 from the P site upon start codon recognition.....	35
1.8.2 Movement of N- and C-terminal tails of eIF1A	37
1.8.3 Closed conformation of the 40S subunit.....	39
1.8.4 eIF5-induced GTP hydrolysis of ternary complex.....	42
1.8.5 Non-AUG initiation	43
1.9 80S initiation complex formation	45

1.10	Scope of the thesis.....	47
2	RESULTS	49
2.1	Reconstitution of mammalian translation system.....	49
2.1.1	Components for the reconstituted mammalian translation system.....	50
2.1.2	Toeprinting assay.....	55
2.1.3	Peptide bond formation	68
2.2	Studies of mRNA binding to eIF3 and the 40S subunit by steady-state fluorescence assays	72
2.2.1	mRNA binding to eIF3.....	72
2.2.2	mRNA binding to the 40S subunit	73
2.3	Studies of 80S IC formation by equilibrium fluorescence assays.....	77
2.3.1	Binding affinity of eIF1A variants to the 40S subunit.....	77
2.3.2	Fluorescence changes of labeled eIF1A upon 48S IC and 80S IC formation	80
2.4	Kinetic analysis of mammalian translation initiation	85
2.4.1	Kinetic analysis of eIF1A binding to the 40S subunit.....	85
2.4.2	Identification of discrete eIF1A-dissociation states at different stages of initiation	95
2.4.3	Verification of kinetic checkpoints during initiation by fluorescent eIF1A variants S74C and T120C.....	114
2.5	Cryo-EM structure of human 48S IC.....	118
3	DISCUSSION.....	123
3.1	Reconstitution of <i>in vitro</i> mammalian translation system.....	123
3.1.1	Purification of the human initiation factors.....	123
3.1.2	Assembly of the mammalian initiation complex.....	124
3.1.3	Start-codon selection	125
3.2	Fluorescence-labeled eIF1A as a reporter	127
3.3	Kinetic analysis of eIF1A binding to the 40S subunit	128
3.3.1	Two-population binding-equilibrium model of eIF1A binding to the 40S subunit	128
3.3.2	Binding rates of eIF1A to the 40S subunit in cells	130
3.3.3	Binding affinity of eIF1A to the 40S subunit.....	131
3.4	Kinetic analysis of 48S IC assembly.....	132

3.4.1	Scanning and the dissociation rate of eIF1A	133
3.4.2	Comparison of the eIF1A-dissociation states of the initiation complexes in lower and higher eukaryotes	134
3.4.3	Roles of individual initiation factors during initiation complex assembly	136
3.5	The biochemical pathway of translation initiation in cells.....	139
3.5.1	Summary of the dynamic of the initiation complexes	141
3.6	Perspectives	144
4	MATERIALS AND METHODS	147
4.1	Reconstitution of mammalian translation system.....	147
4.1.1	Purification of the native human factors	147
4.1.2	Purification of the recombinant proteins.....	154
4.1.3	<i>in vitro</i> Transcription	157
4.1.4	Preparation of the fluorescence-labeled eIF1A variants.....	159
4.2	Reconstitution of the Initiation complexes	160
4.2.1	48S IC formation	160
4.2.2	80S IC formation	161
4.3	Functional assays	161
4.3.2	Peptides formation assay	163
4.4	Steady-state fluorescence measurements	164
4.4.1	Gel shift assay	164
4.4.2	Binding affinity of eIF1A to the 40S subunit.....	164
4.4.3	Competition assay of unlabeled WT eIF1A to the 40S-eIF1A-N4C-Alexa555 complex	165
4.4.4	Fluorescence changes of eIF1A-Alexa555 during 48S and 80S IC formation...	165
4.5	Rapid kinetic assays	166
4.5.1	Association kinetic assay	166
4.5.2	Dissociation kinetic assay	167
4.6	Human 48S IC for the cryo-EM structure study	167
5	REFERENCES	169
6	APPENDIX.....	190
6.1	List of figures	190

6.2	List of tables	192
ACKNOWLEDGEMENTS.....		193
CURRICULUM VITAE		Error! Bookmark not defined.

ABSTRACT

During translation initiation in mammalian cells, 80S ribosomes are assembled at the translation start site on mRNA via a cap-dependent scanning mechanism. The formation of the 80S initiation complex (IC) is essential for proper selection of the start codon and requires concerted action of at least 10 eukaryotic initiation factors (eIFs). This process can be divided into five stages: binding of eIFs to the 40S subunit to assemble the 43S pre-IC (PIC), mRNA recruitment to the 43S PIC, mRNA scanning, 48S IC formation upon start codon recognition, and finally, 80S IC formation after 60S subunit joining. Genetic and thermodynamic studies of translation initiation in lower and higher eukaryotes have revealed the major functions of individual eIFs. However, due to the lack of kinetic analysis, the precise order of events and the regulatory mechanisms underlying IC formation in mammals remain unknown.

Here, we show the assembly of human 48S IC in real time using fluorescence-based rapid kinetic techniques in an *in vitro* reconstituted mammalian translation system. We find that eIF1A binds rapidly to two distinct states of the 40S subunit and stabilizes one of them by a two-population binding-equilibrium model. We also observe that the dissociation of eIF1A from the 40S subunit is slower in the presence of eIF1. By monitoring eIF1A dissociation rates, we are able to systematically analyze the kinetic stability of the partial initiation complexes *en route* to the 43S PIC or 48S IC. Depending on the combinations of initiation factors bound to the 40S subunit, the complexes take on one or two kinetically distinct states. The 43S PIC and scanning 48S PIC are single-population, thermodynamically and kinetically stable complexes. Upon start codon recognition, the 48S IC resumes its dynamic behavior with two kinetically distinct populations: one dissociating rapidly and the other slowly. Our results uncover key kinetic checkpoints that drive the thermodynamic equilibrium toward complete assembly of the 48S IC by the individual and combined action of initiation factors. This mechanistic understanding of the kinetically favored routes guiding initiation complex formation sets a milestone in deciphering the regulation of protein synthesis in mammalian cells.

1 INTRODUCTION

Proteins are the elemental functional units in all living cells. In most organisms, genetic information is stored in deoxyribonucleic acid (DNA). A functional segment of DNA, a gene, is expressed in two steps, transcription and translation. RNA polymerase first transcribes DNA into ribonucleic acid (RNA). Subsequently, the ribosome translates the RNA into functional proteins. Protein biosynthesis is a conserved biochemical process in the cells from prokaryotes to eukaryotes. In this study, I will focus on the molecular mechanism of translation in mammals.

1.1 Mammalian translation cycle

The ribosome is a macromolecular ribonucleoprotein complex that catalyzes protein biosynthesis in all kingdoms of life. The ribosome is composed of ribosomal RNAs (rRNAs) and ribosomal proteins (r-proteins) and comprises two major subunits, the large ribosomal subunits (LSU) and the small ribosomal subunits (SSU). The size and nomenclature of the ribosome are conventionally described by its sedimentation coefficient (Svedberg unit, S) during ultracentrifugation. The intact mammalian ribosome is the 80S ribosome with around 4.2 MDa in molecular weight. The 80S ribosome consists of the 60S subunit (LSU) and the 40S subunit (SSU). The 60S subunit is around 2.8 MDa with 28S rRNA, 5.8S rRNA, 5S rRNA and 49 r-proteins, while the 40S subunit is around 1.4 MDa with 18S rRNA and 33 r-proteins. The LSU harbors the peptidyl transferase center (PTC), which catalyzes the peptide bond formation between amino acids, while the SSU houses the mRNA binding channel and the decoding center, where the codon of the mRNA basepairs with the anticodon of the tRNA. The interface between the LSU and SSU harbors the binding sites for translation factors and tRNAs during the translation process. According to the preference to the types of tRNAs, the tRNA binding sites are designated as A (aminoacyl, where aminoacyl-tRNA binds), P (peptidyl, holding the peptidyl-tRNA and deacylated tRNA after peptide bond formation) and E (exit, specific for deacylated tRNA) sites. Together with translation factors and tRNAs, ribosomes translate the genetic information from RNA into functional proteins. In an active growing mammalian cell,

protein biosynthesis is highly efficient as the mRNA transcripts are effectively translated by the ribosomes for several rounds.

Mammalian translation entails four stages: initiation, elongation, termination and ribosome recycling. At each stage, several translation factors and tRNAs coordinate at different spatial and temporal points to facilitate the function of ribosomes. At the initiation phase, at least 10 eukaryotic initiation factors (eIFs) work together with the 40S and 60S subunits to recruit the target mRNA and form the 80S ribosome. The 80S IC is assembled on the initiation codon of an open reading frame (ORF), which defines the starting point of the peptide formation. At the elongation phase, 80S ribosome catalyzes repetitive cycles of peptide bond formation between amino acids together with elongation factor (eEF) 1A, eEF2 and eIF5A. eEF1A brings the aminoacyl-tRNA charged with its specific amino acid to the ribosome, while eEF2 promotes translocation of the ribosome after each round of the peptide bond formation. The peptide bond formation is catalyzed mostly by the ribosome itself in the PTC. In some special cases, eIF5A promotes the formation of peptide bonds between consecutive proline residues, thereby preventing the ribosome from stalling. Next, the translation of an ORF is terminated when the ribosome encounters one of the three stop codons UAG, UAA or UGA. Release factor (eRF) 1 and eRF3 recognize the stop codon and trigger the hydrolysis of the ester bond between tRNA and the nascent peptide, releasing the newly synthesized polypeptide from the ribosome. In the last step of translation, the recycling factor ABCE1 splits the 80S ribosome into the 40S and 60S subunits, which are ready to re-enter further translation cycles. A successful round of the translation cycle generates a full-length polypeptide, which can be further modified and folded into a functional protein.

The elongation step is evolutionarily conserved and follows very similar pathways in eukaryotes, bacteria and archaea. In contrast, initiation mechanisms vary significantly between prokaryotes and eukaryotes, potentially because initiation in eukaryotes is a highly regulated step. The mechanism of translation initiation in mammals is not fully understood yet. In the following sections, I will provide an overview of the current understanding of the molecular mechanism during translation initiation. The majority of the discussion will focus on studies from mammals, mostly from human (*Homo sapiens*) and rabbits (*Oryctolagus*

caniculus). Supplemental information from the lower eukaryotes, e.g. budding yeast (*Saccharomyces cerevisiae*), will also be included where essential.

1.2 Translation initiation

At the initiation phase, ribosomes recruit target mRNAs and select the initiation codon of an open reading frame (ORF). The initiation step largely determines the composition of the cellular proteome and maintains the homeostasis of the protein influx to the cells, i.e. ensures that the spectrum of the synthesized proteins meets the cellular need. Furthermore, the selection of the initiation codons provides a second layer of regulation by defining the translation start site on the desired ORF or by adjusting the expression levels of the main ORF *versus* the potential upstream ORFs (uORF). Start-codon selection warrants the synthesis of full-length protein and avoids the production of aberrant peptide from the incorrect reading frames, which could be toxic and energy consuming for the cells. It is also widely believed that the initiation phase is the rate-determining step of the entire translation process. A computational estimation of yeast protein biosynthesis rates suggests that the duration of the initiation events ranges from 4 s to 233 s with a median value of 40 s. (Shah et al., 2013). In the same study, the authors also suggested that the overall rate of protein production in cells depends primarily on the availability of free ribosomes to enter a translational cycle. Therefore, regulation of ribosome assembly on the start codon of the target mRNA during initiation is a pivotal step to control the overall protein biosynthesis.

Most of the translation initiation factors are highly conserved among eukaryotic organisms. Using *in vitro* reconstituted mammalian translation system and *in vivo* genetic studies (mostly from yeast), the core canonical initiation factors in higher eukaryotes, e.g. human, are identified as eIF1, eIF1A, eIF2, eIF3, eIF4F, eIF4A, eIF4B, eIF4H, eIF5 and eIF5B. Many of these mammalian initiation factors are isolated and characterized as functional proteins in the initiation process early in the 1970s by Staehelin, Hershey, Anderson, and Merrick laboratories (Adams et al., 1975; Benne and Hershey, 1978; Safer et al., 1976; Schreier et al., 1977; Schreier and Staehelin, 1973). However, the search for further components of the

initiation machinery did not stop until the last core initiation factor eIF5B, which is essential for the large subunit joining, was identified in early 2000s (Pestova et al., 2000). Several other regulatory proteins participating in the initiation process were discovered only recently, e.g. DHX29 (Pisareva et al., 2008). These regulatory proteins provide a wider range of regulation during initiation and fine-tune specific protein synthesis in different cellular conditions. Together, all these initiation factors and regulatory proteins cooperatively ensure highly efficient and accurate translation initiation in cells.

The eukaryotic mRNAs have distinct features that increase the complexity of translational controls at the initiation stage. Unlike prokaryotic mRNAs, most eukaryotic mRNAs are post-transcriptionally modified with a 5' 7-methylguanylate (m^7G) cap (**Figure 1**) and 3' polyadenylate (poly-A) tail. In eukaryotic cells, the 5' m^7G cap is modified in a two-step reaction in the nucleus (Cowling, 2009). First, the guanosine-5'-triphosphate (GTP) is attached to 5' end of the mRNA with a 5'-5' triphosphate linkage by mRNA guanylyltransferase. Second, the 7-nitrogen of the guanine is methylated by mRNA cap guanine- N^7 methyltransferase to obtain an m^7G cap-0 structure, where the 0 refers to no subsequent methylation on the next ribonucleotide. Once the mRNA is transported to the cytoplasm, the 5' m^7G cap is recognized and activated by the eukaryotic cap binding complex, eIF4F. The activated mRNA can then be recruited to the ribosome with the help of other eIFs, which is the starting point of the cap-dependent scanning mechanism. In the following sections, I will focus on the canonical mechanisms of cap-dependent and/or scanning-dependent initiation, since these pathways entail a number of similar initiation complexes.

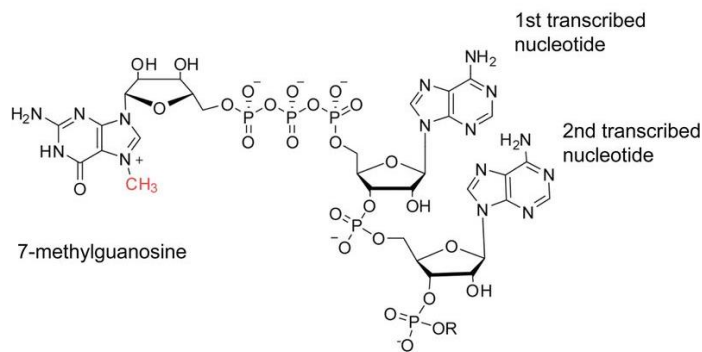


Figure 1. The 5' m⁷G cap structure in eukaryotic mRNA

The m⁷G cap-0 structure. The methyl group of the 7-methylguanosine is indicated in red. The m⁷G cap is covalently bonded to the mRNA with a 5'-5' triphosphate linkage. This figure is adapted from (Cowling, 2009).

1.3 Cap-dependent translation initiation

The canonical mechanism of eukaryotic initiation is the cap-dependent scanning initiation proposed by Kozak in early 1980s (Kozak, 1980a). During cap-dependent initiation, the ribosome binds to the 5' end of an mRNA and scans base-by-base through the 5' untranslated region (5' UTR) until it reaches the start codon. The process of cap-dependent initiation can be divided into five main phases: (1) binding of initiation factors (eIFs) to the 40S subunit to form the 43S preinitiation complex (PIC); (2) recruitment of the eIF4F-bound mRNA to the 43S PIC; (3) scanning; (4) recognition of the start codon to form the 48S initiation complex (IC); and (5) joining of the 60S subunit for the assembly of the 80S IC (**Figure 2**). Although this simplified pathway implies the order of the initiation events, the *bona fide* biochemical process in mammalian cells is still unclear. Studies from lower eukaryotes shed light on the molecular mechanism through a combination of detailed *in vivo* and *in vitro* experiments, including rapid kinetics and structural studies (see reviews (Hinnebusch, 2014, 2017; Hinnebusch and Lorsch, 2012)). However, in higher eukaryotes, many crucial questions remain open due to the lack of a direct assay to monitor the mammalian translation in real

time. In the following section, I will summarize the current understanding of the molecular mechanism in each stage of the mammalian translation initiation.

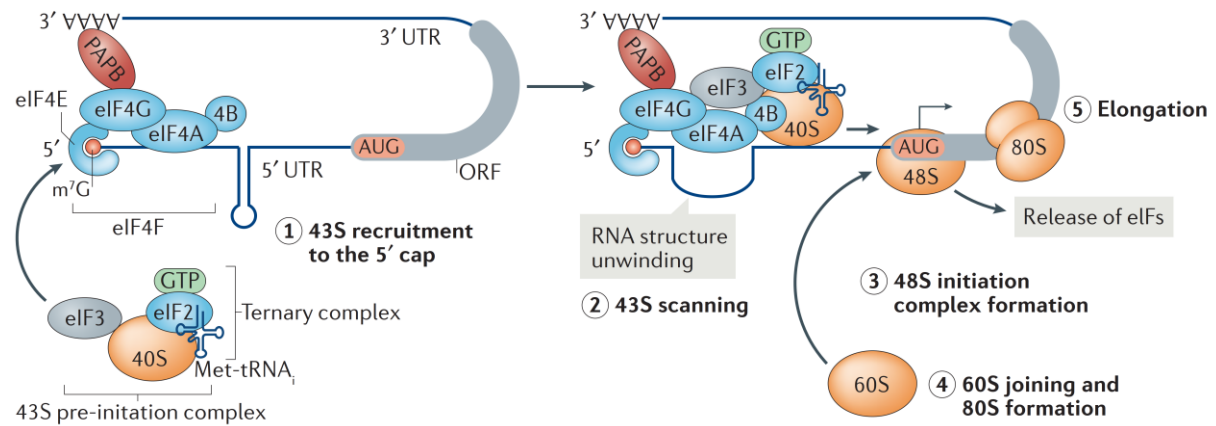


Figure 2. Overview of canonical cap-dependent translation initiation pathway

The key steps of canonical eukaryotic translation initiation: (1) 43S PIC assembly and mRNA recruitment to the 43S PIC; (2) 43S PIC scanning; (3) 48S IC formation upon start codon recognition; (4) 60S joining and 80S IC formation; and subsequent (5) elongation. This figure is adapted from (Leppek et al., 2018).

1.4 43S preinitiation complex formation

43S preinitiation complex (PIC) refers biochemically to the 40S subunit bound with the core initiation factors, eIF1, eIF1A, eIF2·Met-tRNA_i^{Met}·GTP (also called ternary complex, TC), eIF3, eIF4A, eIF4B and eIF5, in the absence of mRNA (**Table 1**). This complex is stable enough to be isolated from the eukaryotic cells or prepared by *in vitro* reconstitution, e.g. by sucrose gradient ultracentrifugation. Due to the lack of kinetic analysis, it is still unknown whether the initiation factors always bind to the 40S subunit prior to the mRNA recruitment in the mammalian cells. In prokaryotes, mRNA can be recruited to the 30S subunit regardless of the presence of either IF1, IF2 or IF3 (Milon et al., 2012), which tempts to speculate that in the eukaryotic cells, mRNA can also bind to the 40S subunit at several intermediate stages. The steady-state thermodynamic framework of translation initiation in yeast and mammalian system has shown that the combined action of the initiation factors on the 40S subunit

increases the affinity of the ribosome to the mRNA. This implies a crosstalk between the initiation factors among each other on the 40S subunit and with the 40S subunit to activate the translation initiation.

Table 1. Number of human initiation factors and ribosomes per HeLa cell

Protein	Mol. weight (kDa)	# per cell (x 10 ⁴)
Cytoplasmic		
Ribosomes	~4200	1064
40S subunit	~1400	-
60S subunit	~2800	-
eIF1	12.7	198
eIF1A	16.5	101
eIF2	~125	210
eIF2B	26	75
eIF2A	65	72
eIF3	~800	164
eIF4A	46.2	801
eIF4B	69.2	131
eIF4E	25	41
eIF4G (1+2)	175.5; 102.3	74
eIF4H	27.4	103
eIF5	49.2	91
eIF5B	138.8	76
PABP	70.7	252
DDX3	73.2	138
DHX29	155.2	12
4E-BP1	12.6	33
4E-BP2	12.9	6

This table is adapted from (Merrick and Pavitt, 2018) based on the mass spectrometry analysis (Kulak et al., 2014). The values for the multi-subunit factors and ribosomes are given from the most abundant subunit. As 1 nM concentration corresponds to around 1000 molecules/cell in the HeLa cells (Moran et al., 2010), the ribosome concentration is about 10 μ M (Duncan and Hershey, 1983) and most of the initiation factors are in the 1 – 8 μ M range.

1.4.1 Formation of the ternary complex, eIF2·GTP·Met-tRNA_i^{Met}

eIF2 is the main carrier of the initiator Met-tRNA_i^{Met} to the 40S subunit in the form of ternary complex (TC), eIF2·GTP·Met-tRNA_i^{Met}. eIF2 consists of three subunits, 2 α , 2 β and 2 γ . eIF2 α and 2 γ subunits bind to the Met-tRNA_i^{Met} (Naveau et al., 2013; Schmitt et al., 2012; Shin et al., 2011), while eIF2 γ subunit harbors the GTPase activity and can also bind to GDP or GTP. eIF2 β is the regulatory subunit that interacts with eIF5 and triggers the GTPase activity of eIF2 γ upon start codon recognition (Jennings et al., 2016). Thermodynamic analysis of nucleotide binding to yeast eIF2 shows that eIF2 alone has a higher affinity (K_d) for GDP (20 nM) than GTP (1700 nM). However, eIF2 binds to Met-tRNA_i^{Met} with 20-fold higher K_d in the presence of GTP than GDP (Jennings et al., 2017; Kapp and Lorsch, 2004). Hence, the exchange of the guanine nucleotide is the rate-limiting step before eIF2 can carry Met-tRNA_i^{Met} for a new round of translation. In eukaryotic cells, eIF2B is the guanine nucleotide exchange factor (GEF) for eIF2 (Pavitt et al., 1998; Rowlands et al., 1988). eIF2B only binds to the non-phosphorylated form of eIF2 and facilitates the exchange of GDP to GTP. If the Ser51 of mammalian eIF2 α is phosphorylated, eIF2 becomes an unfavored substrate for eIF2B (Pavitt et al., 1998; Rowlands et al., 1988), which leads to a global downregulation of translation in cells. During the GTPase cycle of TC, eIF5 acts both as GTPase activating protein (GAP) and GDP dissociation inhibitor (GDI) to eIF2 (Algire et al., 2005; Jennings et al., 2016). eIF5 promotes GTP hydrolysis of TC upon start codon recognition, stabilizes eIF2·GDP and prevents GDP release from eIF2 after eIF2·GDP dissociates from the 48S IC. This implies that eIF5 can form a stable complex with eIF2·GDP in the cytosol (Jennings and Pavitt, 2010). Displacement of eIF5 by eIF2B from the eIF5·eIF2·GDP complex promotes the exchange of GTP, which allows the TC to enter a new round of initiation.

The binding of TC to the ribosome is stabilized by eIF1, eIF1A and eIF3 (Aitken et al., 2016; Maag et al., 2005; Majumdar et al., 2003; Martin-Marcos et al., 2013; Passmore et al., 2007; Saini et al., 2010; Sokabe and Fraser, 2014). eIF1 and the C-terminal tail (CTT) of eIF1A induce a conformation of the 40S subunit that promotes the recruitment of TC (Martin-Marcos et al., 2013; Saini et al., 2010; Sokabe and Fraser, 2014). Together with eIF1 and eIF1A, eIF3 also stabilizes TC in the 43S PIC (Aitken et al., 2016; Majumdar et al., 2003; Sokabe and Fraser,

2014; Valasek, 2012). Specifically, mutations within yeast eIF3b/i/g subunits destabilize TC binding to the 43S PIC (Aitken et al., 2016). After binding to the platform on the 40S subunit, TC places the anticodon of the initiator tRNA into the P site where it can detect the start codon during scanning. The conformations of the TC are classified as P_{out} and P_{in} states depending on the orientations of the initiator tRNA. During scanning, the anticodon of the Met-tRNA_i^{Met} is not fully accommodated into the P site, in a so-called P_{out} state. The full accommodation of the initiator tRNA is prevented by eIF1 and eIF1A-CTT until the start codon is recognized. The base-pairing between the codon and anti-codon triggers the accommodation of the initiator tRNA deeper into the P site, resulting in the P_{in} state.

1.4.2 Binding of individual initiation factors to the 40S subunit

Most initiation factors recruited to the 43S PIC have nanomolar-range affinities to the 40S subunit and stabilize each other on the 40S subunit (Maag and Lorsch, 2003; Sokabe and Fraser, 2014). This means that in the presence of one initiation factor, the 40S subunit can recruit next factors with higher affinity (or in a lower concentration). Considering the cellular concentrations of the 43S PIC factors are around 1 - 8 μ M (**Table 1**) (Duncan and Hershey, 1983; Kulak et al., 2014; Merrick and Pavitt, 2018), the thermodynamic coupling effect would ensure that free 40S subunit is saturated with initiation factors at equilibrium and forms 43S PIC for efficient translation initiation.

Each initiation factor has its defined binding site on the 40S subunit, but can also change its exact location or conformation at different stages. eIF1 (12.7 kDa) and eIF1A (17.5 kDa) are the two smallest initiation factors, and they bind to the 40S subunit at the platform around the decoding center (**Figure 3**). eIF1 binds near the P site, whereas eIF1A binds adjacent to eIF1 at the A site (Passmore et al., 2007; Weissner et al., 2013). eIF1A alone has high affinity, 19 - 62 nM, to the 40S subunit (Fraser et al., 2007; Sokabe and Fraser, 2014) and is not part of the multi-factor complex, MFC (Sokabe et al., 2012) (see below). It is believed that eIF1A is recruited to the 40S subunit early during the recycling step, probably together with eIF1 and eIF3 (Pisarev et al., 2007a).

eukaryotic-specific functions in translation initiation. Even though the NTT and CTT tails have no intrinsic structure in solution, they are partially immobilized and become more rigid on the 40S subunit. Both bacterial IF1 and eIF1A bind to the A site on the small ribosomal subunit with their OB domain, while the NTT and CTT of eukaryotic eIF1A penetrate into the P site at different stages of translation initiation (Yu et al., 2009). The CTT and NTT of eIF1A have distinct functions during scanning and start codon recognition, respectively (Fekete et al., 2005; Fekete et al., 2007). At the late stage of initiation, eIF1A also assists eIF5B to promote the 60S subunit joining (Marintchev et al., 2003) and finally dissociates from the 80S IC after the GTP hydrolysis of eIF5B (Fringer et al., 2007). Participation of eIF1A at almost every step of initiation indicates its key importance in the initiation complex assembly.

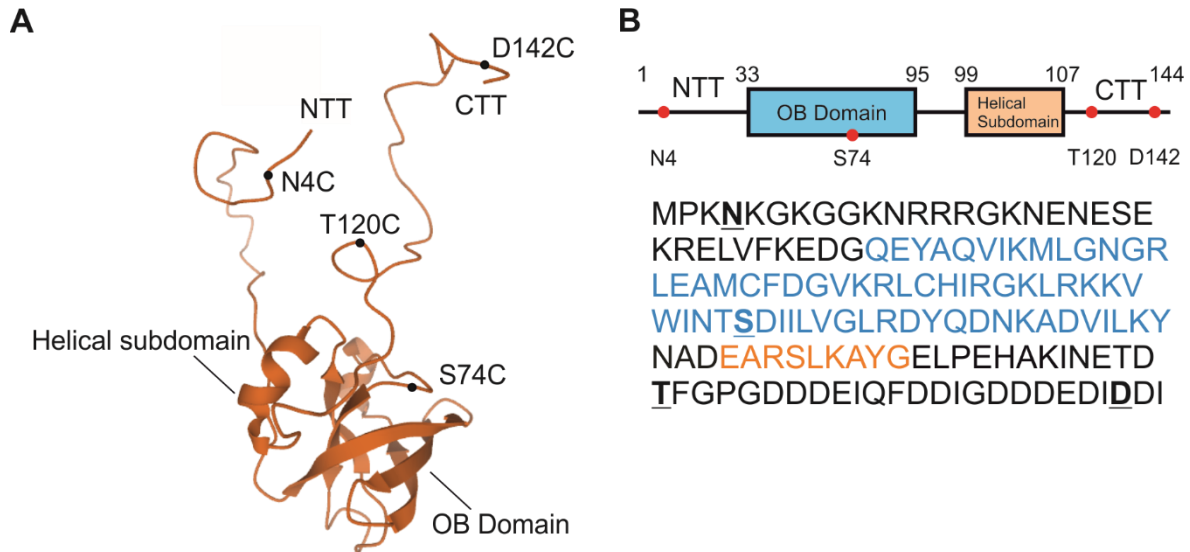


Figure 4. eIF1A structure

(A) eIF1A structure solved by NMR (PDB file: 1D7Q) (Battiste et al., 2000). NTT, CTT and OB domain indicate the N-terminal tail, C-terminal tail and oligonucleotide-binding domain of eIF1A. Positions N4, T120, S74 and D124 are marked with dots. These positions are mutated to single cysteine for fluorescent labeling in this study. **(B)** Scheme and protein sequence of eIF1A with total 144 residues. Numbers in the scheme represent the position of the residues from 1 to 144. The OB domain is colored in blue and the C-terminal helical subdomain in yellow. Positions N4, T120, S74 and D124 are marked in the scheme and underlined in the sequence.

eIF1 alone also binds to the 40S subunit with high affinity, 49 nM in the mammalian system (Sokabe and Fraser, 2014). In both yeast and mammals, eIF1 and eIF1A cooperatively bind to the 40S subunit (Maag and Lorsch, 2003; Sokabe and Fraser, 2014), which increases their affinity to the 40S subunit in mammals (49 to 12 nM for eIF1 and 62 to 21 nM for eIF1A) (Sokabe and Fraser, 2014). Hence, eIF1 could be recruited to the 40S subunit directly (probably in the presence of eIF1A during recycling) as an alternative pathway than in the context of MFC. Bindings of eIF1 and eIF1A induce an open conformation of the 40S subunit (**Figure 5**) (Passmore et al., 2007), where the latch of the mRNA binding channel is opened to allow mRNA docking into the mRNA binding channel (Passmore et al., 2007). Furthermore, in the presence of eIF1 and eIF1A on the 40S subunit, a new connection between the head and shoulder on the solvent side of the 40S subunit is observed, which is not seen in the empty 40S subunit (Passmore et al., 2007). Therefore, the stabilization of the open conformation by eIF1 and eIF1A further promotes subsequent mRNA recruitment and scanning (Pestova et al., 1998a). Both factors also stimulate TC binding to the 40S subunit (Martin-Marcos et al., 2013), most likely also due to the stabilization of the open conformation.

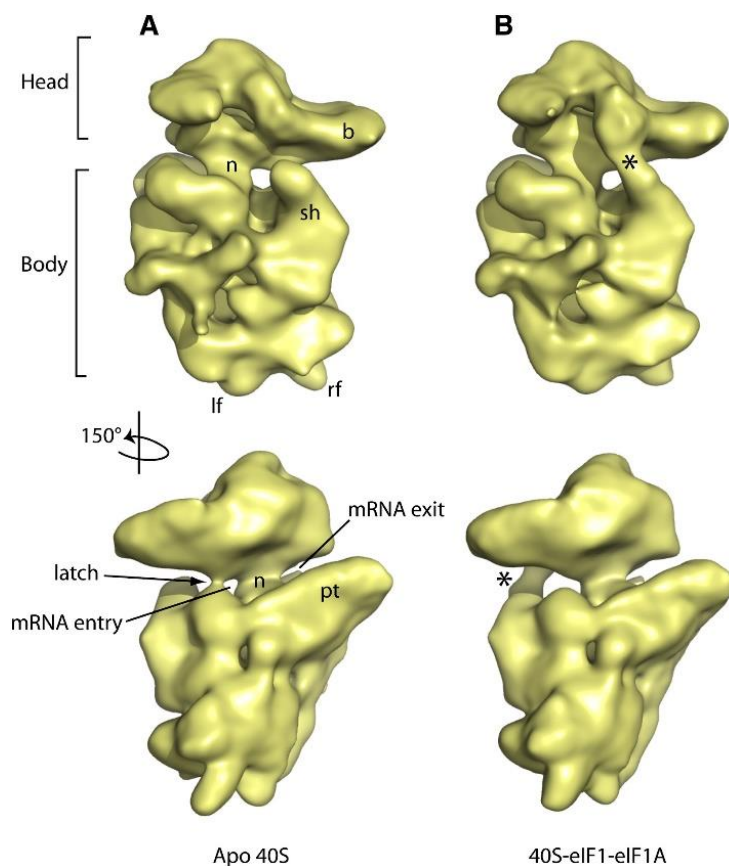


Figure 5. Structures of 40S subunit and 40S-eIF1-eIF1A complex

(A) Structure of the 40S subunit. **(B)** Structure of 40S-eIF1-eIF1A complex. The open conformation of the 40S subunit is stabilized in the presence of eIF1 and eIF1A, where the latch at the mRNA entry channel is opened for the mRNA binding. Abbreviations: b, beak; n, neck; sh, shoulder; pt, platform; lf, left foot; and rf, right foot. The connection between shoulder and head induced by the binding of eIF1 and eIF1A is highlighted by an asterisk. The figure is adapted from (Passmore et al., 2007).

eIF3 is a pivotal initiation factor in mammalian translation with a total molecular mass of approximately 800 kDa. This multi-subunit factor has evolved from 6 subunits in yeast to 13 subunits in mammalian cells (Browning et al., 2001), namely subunit 3a – 3m for human eIF3. It is extremely dynamic in solution and probably exists in several structural forms, which makes it difficult to solve the intact eIF3 structure. eIF3 alone can bind to the 40S subunit in the absence of other factors and promote the recruitment of TC and mRNA (Benne and Hershey, 1978; Peterson et al., 1979; Trachsel and Staehelin, 1979). Upon binding to the 40S subunit, most of eIF3 remains flexible, but parts of the subunits become more rigid and can be captured

in low resolution by cryogenic electron microscopy (cryo-EM) (for review, see (Cate, 2017)). This relatively rigid part comprises a 400 kDa octamer core of eIF3, which is composed of subunit a, c, e, f, h, k, l and m. The architecture of the octamer is similar to the 26S proteasomal regulatory lid and COP9 signalosome, therefore it is also called PCN/MPN domain (proteasome, CSN, eIF3/Mpn1, Pad1-N-terminal) (Scheel and Hofmann, 2005). Early and recent high resolution structures have shown that the octamer binds to the solvent site of the 40S subunit (**Figure 3**) (Brito Querido et al., 2020; Srivastava et al., 1992), while the other 5 peripheral subunits, b, d, g, i, and j, wrap around the 40S subunit body ((Brito Querido et al., 2020) Simonetti 2016) and extend mRNA binding channel at the mRNA exit site in the context of 48S IC (Pisarev et al., 2008; Simonetti et al., 2016). In the 43S PIC, eIF3 acts as a scaffold protein and interacts with almost all initiation factors, e.g. eIF1 and eIF1A (Sokabe et al., 2012), or eIF4G (Villa et al., 2013). eIF3 alone can also recognize specific mRNA elements in the 5' UTR (Lee et al., 2015) and regulate the translation of downstream ORF (for review, see (Shatsky et al., 2018)). eIF3 participates in almost all initiation process and also involves later in re-initiation and ribosome recycling. The essential role of mammalian eIF3 is one of the key differences between lower and higher eukaryotes.

eIF4A is an ATP-dependent DEAD-box RNA helicase (Grifo et al., 1984; Linder and Slonimski, 1989) and forms the cap-binding complex eIF4F together with eIF4E and eIF4G. eIF4B, at the same time, is the helper protein of eIF4A and stimulates its helicase activity. eIF4A and eIF4B together unwind the mRNA secondary structure and promote loading of the 40S subunit onto the mRNA (Ozes et al., 2011; Sokabe and Fraser, 2017). eIF4A and eIF4B can function directly on mRNA with and without the 40S subunit. However, structural studies show that they also interact with the 40S subunit and other initiation factors. The electron density of eIF4B is found at the mRNA entry channel on the 40S subunit (Eliseev et al., 2018), while parts of the eIF4A and eIF4G densities are found on the eIF3 octamer core at the mRNA exit channel (**Figure 3**) (Brito Querido et al., 2020). Furthermore, phosphorylation of eIF4B increases its affinity to eIF3 and 43S PIC (Holz et al., 2005). The rigid density of eIF4A was not clearly captured probably because of its rapid turnover during the unwinding activity and, thus, rapid association with and dissociation from the 43S PIC. In a different model, eIF4B and eIF4A

bind to the mRNA exit channel and act as a pawl of a Brownian ratchet to prevent backward sliding of the 40S subunit (Spirin, 2009). With these two possible binding positions of eIF4A and eIF4B on the 40S subunit, both factors might conduct the unwinding activity and promote mRNA loading at the same time. Notably, eIF4A is the most abundant initiation factor in mammalian cells (**Table 1**) (Kulak et al., 2014; Merrick and Pavitt, 2018), while eIF4B is less abundant. Therefore, it is also possible that multiple sets of eIF4A and eIF4B interacts with one ribosome during initiation (Lu et al., 2014).

The kinetically favored route of 43S PIC formation is still unknown, therefore, the order of factors binding to the 40S subunit remains unclear. The association and dissociation rates of individual initiation factors to the mammalian 40S subunit have to be determined to obtain a sophisticated assembly landscape of the 43S PIC as described for the bacterial system (Milon et al., 2012). Besides, in cells, the recycled free 40S subunit is most likely associated with a few initiation factors, e.g. eIF1A and eIF3, and this partially assembled 43S PIC is then waiting for other essential factors to join. The binding of these remaining factors would serve as a checkpoint for the complete maturation of the 43S PIC.

Alternatively, initiation factors could also join the 40S subunit as a multi-factor complex (MFC), which consists of eIF1, eIF2, eIF3 and eIF5 in eukaryotic cells. Upon their arrival, the initiation factors are accommodated by the 40S subunit and fulfill their functions for the following initiation process. MFC was first discovered in yeast cells (Asano et al., 2000) and later also found in plants (Dennis et al., 2009) and mammalian cells (Sokabe et al., 2012). Mammalian MFC can be isolated from both HeLa cell lysate and rabbit reticulocyte lysates (RRL) by electrophoresis in 4% native polyacrylamide gel with detectable trace of Met-tRNA_i (Sokabe et al., 2012). Human MFC can also be reconstituted *in vitro* from purified human eIF1, eIF2, eIF3 and eIF5 at a physiologically relevant concentration of 300 nM (Sokabe et al., 2012). In the MFC, eIF3 serves as a scaffold protein that provides binding platform for other initiation factors. The interaction between eIF3 and eIF2 determines the stability of the human MFC, whereas in yeast, the key interaction comes from eIF3 and eIF5 (Asano et al., 2000). The affinity of Met-tRNA_i^{Met} to the eIF2·GDP is dramatically increased if eIF2 is in the context of

MFC. This finding suggests that MFC could also facilitate the delivery of Met-tRNA_i^{Met} to the 40S subunit *in vivo*.

1.5 mRNA recruitment

One of the major tasks for the ribosome during initiation is to select a suitable mRNA transcript for translation. The selection of mRNA in the cap-dependent initiation pathway occurs mainly through the activation of mRNA by the cap-binding complex, eIF4F. Binding of eIF4F at the 5' cap of eukaryotic mRNA activates the mRNA. This activated mRNA is then recruited to the ribosome with the assistance of initiation factors. Since cellular mRNAs are in excess over the ribosomes, the competition between mRNA transcripts for free ribosomes determines the translation efficiency. As mentioned above, there are alternative kinetic pathways to assemble the 43S PIC. It is unknown which routes of cap-dependent initiation are kinetically favored in cells and whether they are mRNA-specific. Other initiation mechanisms, differing mainly in the mRNA recruitment step, will be briefly discussed later.

1.5.1 Activation of mRNA by the cap-binding complex, eIF4F

The cap-binding complex, eIF4F, binds to the 5' m⁷G cap of an mRNA and activates it for the subsequent ribosome recruitment (Grifo et al., 1983; Tahara et al., 1981). eIF4F consists of three proteins, eIF4E which binds to the cap structure, eIF4A which is an RNA helicase required for the unwinding of secondary structures proximal to the 5' cap, and eIF4G which is a scaffold protein. eIF4E is the major m⁷G recognition factor in the eIF4F (Gross et al., 2003b) and it alone has a high affinity (~ 100 nM) to the m⁷G (Carberry et al., 1992; Niedzwiecka et al., 2002a; Niedzwiecka et al., 2002b). The cap binding ability of eIF4E is dramatically increased (~20 nM) in the presence eIF4G and decreased in the presence of a secondary structure close to the 5' proximal region (Gross et al., 2003b; O'Leary et al., 2013; Yanagiya et al., 2009). The interaction framework between eIF4F factors has also been elucidated with purified factors from yeast (Mitchell et al., 2010). In yeast, the affinity of eIF4A to eIF4G is increased 10-fold

(from 200 nM to less than 30 nM) in the presence of eIF4E, while eIF4E alone also binds to eIF4G with high affinity (less than 30 nM). The thermodynamic coupling of eIF4A and eIF4E with eIF4G would ensure the stability of the eIF4F complex prior to its binding to the mRNA.

Presumably, the excess free mRNAs must compete for eIF4F due to the limiting amount of eIF4F in cells, especially eIF4E (**Table 1**) (Gingras et al., 1999a; Gingras et al., 1999b; Kulak et al., 2014; von der Haar and McCarthy, 2002). The features that affect the selection of the mRNA by eIF4F remain unclear, but are likely determined by the *cis* elements on the mRNA. The binding of eIF4F to the 5' cap structure is one of the translational control hubs that can adjust the translation efficiency globally. The access of eIF4E to eIF4G is regulated by the eIF4E binding proteins (4E-BPs) (Gross et al., 2003a; Gruner et al., 2016; Richter and Sonenberg, 2005). There are three isoforms of 4E-BPs in mammalian cells, 4E-BP1, 4E-BP2 and 4E-BP3. All 4E-BPs and eIF4G bind to eIF4E with their YXXXXL Φ motif, where X is any amino acid residue and Φ is any hydrophobic residue (Mader et al., 1995; Marcotrigiano et al., 1999). The binding ability of 4E-BPs to eIF4E is regulated through a hierarchical phosphorylation pathway by the mammalian target of rapamycin (mTOR) kinase (Gingras et al., 2001). Once 4E-BPs are phosphorylated, their affinity to eIF4E is reduced 4000-fold and eIF4E is released from the (inactive) 4E-BP-eIF4E complexes for eIF4G to form a complete eIF4F complex. Therefore, the phosphorylation of 4E-BPs leads to an increase of global translation efficiency.

1.5.2 Unwinding of secondary structures in 5' UTR by RNA helicases

Before the 40S subunit attaches to the 5' end of an mRNA, the RNA helicase eIF4A unwinds the 5' proximal region of the mRNA to generate a single-strand RNA region that can accommodate into the mRNA binding channel of the 40S subunit. By shifting the position of a hairpin structure within the first 10 nt of a mRNA, the translation efficiency is dramatically reduced by 50-fold (Babendure et al., 2006). eIF4A alone is a non-processive RNA helicase, but in the presence of eIF4B, eIF4H and eIF4G, eIF4A is shown to be processive, where it remains on the track for 11 base pairs on average before falling off in a single-molecule assay (Garcia-Garcia et al., 2015). The unwinding activity of eIF4A is also strongly enhanced in the context of

eIF4F (Abramson et al., 1987; Feoktistova et al., 2013; Merrick, 2015; Rozen et al., 1990; Sokabe and Fraser, 2017) and in the presence of eIF4B and eIF4H (Harms et al., 2014; Nielsen et al., 2011; Ozes et al., 2011; Rogers et al., 2001a; Rogers et al., 2001b; Rozen et al., 1990). In the eIF4A-knockdown MCF7 breast cancer cells, eIF4A-sensitive mRNAs are transcripts with GC-rich motif or highly structured 5' UTR (Modelska et al., 2015), suggesting the essential unwinding function of eIF4A in translation initiation. Besides unwinding activity, eIF4A also promotes the loading of the mRNA to the 40S subunit in an ATP-dependent manner (Kumar et al., 2016). Since eIF4A is the most abundant initiation factor in cells, an actively translated mRNA could recruit several copies of eIF4A, either in the context of eIF4F or working together with eIF4B and eIF4H, to conduct efficient unwinding activity and promote mRNA binding to the 40S subunit.

1.5.3 mRNA recruitment to the 43S PIC

The interaction between mammalian eIF3 and eIF4F brings the 43S PIC to the 5' end of the mRNA. Crosslinking experiments show that eIF3c, 3d and 3e subunits on the 43S PIC are in close proximity to the eIF4G in the context of eIF4F complex (Hinton et al., 2007; Villa et al., 2013). In contrast, in yeast, the C-terminal domain (CTD) of eIF5 stabilizes the interaction between eIF4G and 43S PIC (Singh et al., 2012; Yamamoto et al., 2005). This difference in the recruitment mechanism between lower and higher eukaryotes also reflects the fact that eIF3 is an essential component for mammalian 48S IC formation but dispensable for the *in vitro* yeast 48S IC reconstitution.

mRNA recruitment requires the coordinated action of initiation factors to induce the open conformation of the ribosome and stabilize the mRNA in the binding channel. The mRNA binding channel is located between the 40S head and body and passes through the decoding center (**Figure 3**). The 40S subunit can bind mRNA with high affinity (~11 nM) (Sokabe and Fraser, 2017). The 43S PIC is also capable of binding to the 5' end of a model mRNA in the absence of eIF4F if the mRNA is unstructured, e.g. containing CAA triplet repeats in its 5' UTR (Pestova and Kolupaeva, 2002). However, natural cellular mRNAs usually bear certain degrees

of secondary structures that would require eIF4A in eIF4F to unwind the 5' proximal region. Other initiation factors also promote mRNA recruitment by stabilizing the binding conformation of the 40S subunit. eIF1 and eIF1A induce the open conformation of the 40S subunit to open the latch at the mRNA entry channel (**Figure 5**) (Passmore et al., 2007). The thermodynamic coupling between the initiation factors, especially eIF3 and TC, on the 43S PIC further stabilizes the mRNA in the mRNA binding channel and increases the affinity of mRNA to the 40S subunit. Only eIF3j protein alone has a negatively regulatory effect on mRNA binding to the 40S subunit (Fraser et al., 2007; Sokabe and Fraser, 2017). eIF3j first binds to the A site near the mRNA entry site and ejects the previous mRNA from the 40S subunit after recycling step (Fraser et al., 2007). eIF3j also impairs mRNA recruitment to the 43S PIC lacking the TC akin to a gate to the mRNA channel (Fraser et al., 2007; Mitchell et al., 2010). However, in the presence of the complete set of 43S PIC components, eIF3j is ejected from the A site to allow the accommodation of the mRNA into the mRNA binding channel (Aylett et al., 2015; Fraser et al., 2007; Sharifulin et al., 2016; Unbehauen et al., 2004). The 43S PIC can sense the completeness of the essential factors early in the initiation process by eIF3j-dependent gating of the mRNA entry channel. Once the mRNA is accommodated in the mRNA binding channel of the 40S subunit, subsequent scanning process happens to search for the start codon.

1.6 mRNA scanning

1.6.1 Structures of mammalian 5' untranslated region

5' untranslated region (5' UTR) contains *cis* elements to regulate the translation process, especially during scanning. 5' UTR usually starts from an m⁷G cap and extends to the start codon of an open reading frame (ORF). The 5' and 3' UTRs encode *cis* elements that can regulate almost every step of the translation (Mayr, 2017; Mazumder et al., 2003). These elements include post-transcriptional modification on the nucleotides, such as N⁶-Methyladenosine (m⁶A) (Meyer et al., 2015), secondary structures (Lee et al., 2015), short upstream open reading frames (uORFs) (Johnstone et al., 2016) and signal sequence motifs (Sinvani et al., 2015). The GC content and the negative folding free energy (ΔG) are often used

as parameters to predict the degree of the secondary structures on an mRNA. mRNAs with higher GC content and lower ΔG tend to form stable secondary structures. However, ΔG alone does not fully reflect the real energy barrier for the ribosome to unwind a structure in cells, because translation process usually requires only local melting of the secondary structure instead of a complete unwound linear form of RNA.

Another parameter to describe the 5' UTR is the length, in terms of nucleotides (nt). Eukaryotic 5' UTRs have a median length of approximately 53 - 238 nt (**Figure 6**) (Pesole et al., 2001), whereas the 3' UTR can be much longer. Higher eukaryotes also have longer 5' UTR than lower eukaryotes with a median length of 218 nt in human mRNAs (Mignone et al., 2002), in comparison to 53 nt in yeast (Nagalakshmi et al., 2008). Longer 5' UTR implies a higher potential to form elements that are involved in complex regulatory networks during translation initiation. However, the 5' UTR length of individual human mRNA transcripts can still vary from less than 10 nt up to more than 1000 nt (Leppek et al., 2018). The mRNA footprint of a free 40S subunit is around 30 nt (Kozak and Shatkin, 1978), whereas the eIF4F-43S PIC protects around 60 nt on the mRNA with extra 30 nt protected by eIF3 at the mRNA exit channel (Pisarev et al., 2008). Considering that the minimum 5' UTR length for the full accommodation of a 43S PIC on the AUG is 45 nt (calculated as 15 nt-long ribosomal footprint of the 5' UTR preceding the AUG codon, plus the 5' extension from eIF3, 30 nt), most of the human mRNA transcripts would fit into the group of mRNAs that can form canonical initiation complexes.

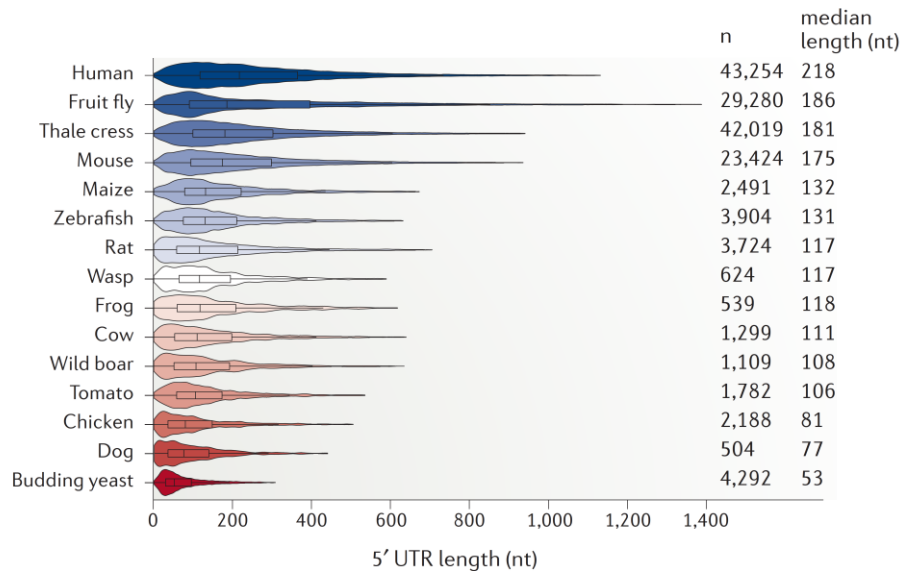


Figure 6. Interspecies variation in mRNA 5' UTR length

The lengths of 5' untranslated regions in eukaryotes. The median lengths range between 53 - 218 nts. This figure is adapted from (Leppek et al., 2018).

The *cis* elements on the mRNA can either interact with the ribosome directly or recruit other *trans* regulatory proteins to modulate the translation efficiency. For example, N⁶-methyladenosine (m⁶A) in the 5' UTR is recognized by eIF3 and then the eIF3 mediates the ribosome recruitment to stimulate the translation of the downstream ORF via scanning mechanism (Meyer et al., 2015; Zhou et al., 2015). Secondary structures in the 5' UTR can also bind regulatory proteins, thereby regulating translation. For example, iron responsive element (IRE) is a 30 nt hairpin structure located close to the 5' proximal region that is bound by IRE-binding protein in the low iron conditions and thereby, prevents the association of 43S PIC (Ma et al., 2012; Muckenthaler et al., 1998). By regulating the structures and modifications of the mRNAs in the cytoplasm, the cells may change the cytosolic proteome to adjust to the fast-changing environments. A great example is from a study in macrophages, where the adjustment of the translation of 90 pre-existing mRNAs after immune stimulation is dependent on the AU-rich elements of the mRNA transcripts (Schott et al., 2014).

In mammals, the presence of upstream ORFs (uORF) and the Kozak sequence are also key features to regulate translation initiation. uORF is a short open reading frame, usually less

than 30 codons, upstream of the main ORF. Since re-initiation on the downstream start sites is not efficient with mammalian ribosome, translation of uORF would likely lead to the down-regulation of the translation of the downstream main ORF. Mammalian cells often use this strategy to fine-tune the expression levels of target genes to meet the cellular needs. Further details will be discussed in the section below.

The Kozak sequence is a consensus sequence, CC A/G CC **AUG** G, flanking the start codon, AUG, on the mRNA transcripts of higher eukaryotes (Kozak, 1981, 1984, 1987a, b). When the positions at -3 and +4 (where A from the AUG is designated as +1) are a purine (A or G) nucleotide and a G, respectively, it is the optimal context for the ribosome to recognize the start codon. If the AUG is surrounded by other combinations of the nucleotides, it is a weak signal for the ribosomes to select the AUG and often leads to leaky scanning. The Kozak sequence and the proceeding coding sequence interact with ribosomal proteins, 18S rRNA and initiation factors, which stabilize the scanning ribosome once it reaches the start codon (**Figure 7**). With all the *cis* elements and *trans* regulatory proteins acting at the 5' UTR, the ribosome scans the mRNA for the start codons in a precisely coordinated manner.

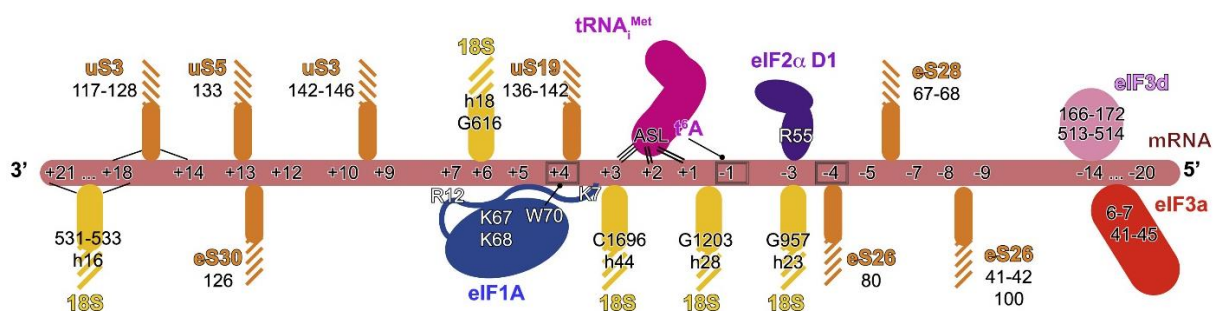


Figure 7. mRNA interactions with mammalian 48S IC

Interactions of the ribosomal proteins, 18S rRNA and initiation factors with the mRNA at the Kozak and coding sequences. A from AUG is designated as +1. The critical interactions are R55 of eIF2 α to position -3 and W70 of eIF1A to position +4. The ribosomal proteins are colored in orange and 18S rRNA in yellow. The figure is adapted from (Simonetti et al., 2020).

1.6.2 Scanning

Scanning refers to the process where the ribosome moves from the 5' end of an mRNA to the start codon. Current models propose that the ribosome scans the 5' UTR base-by-base until it reaches the first AUG in the optimal Kozak sequence context. The scanning mechanism has been proposed early in the 1980s by Kozak (Kozak, 1980a), but remains the least studied step in the translation initiation process due to the difficulty of capturing the scanning intermediates and the lack of a direct assay to measure the scanning process. Measurements of scanning rates with reporter mRNA constructed with different 5' UTR sequences suggest that the scanning rate is 10 nt/s in yeast (Berthelot et al., 2004) and 6 nt/s in human K562 cell free translation system (Vassilenko et al., 2011). Considering that the median length of 5' UTR in human mRNA transcripts is 218 nt, the median scanning time would be around 22 sec. However, it is important to note that the translation efficiency is independent of the 5' UTR length and, in some case, the translation efficiency may even increase with longer 5' UTR (Berthelot et al., 2004; Kozak, 1991b), implying that the speed of scanning in the cell may be higher than that measured so far and may depend on the sequence of the 5' UTR.

The scanning rate is also affected by the local secondary structures and the sequence context. The ribosome tends to select a near-cognate codon or an AUG in a weak Kozak context when followed by a hairpin structure (Kozak, 1990). From crosslinking and structural analysis, the Kozak sequence is in close proximity to many ribosomal proteins, eIF1A, eIF2 α , eIF3d and eIF3a, implying intensive interactions between the initiation complexes to the Kozak sequence (**Figure 7**) (Pisarev et al., 2008; Simonetti et al., 2020). This finding suggests that once the ribosome reaches the Kozak sequence, the scanning rate is slowed down by the interaction between the mRNA and the proteins in the decoding center, which further promotes the selection of the start codon surrounding by the Kozak sequence. In general, the scanning rate of the ribosome determines the dwell time of the ribosome on the local region in the 5' UTR, hence influencing the possibility of the start codon selection by the ribosome.

1.6.3 eIF1 and eIF1A stabilize the scanning complex

The scanning complex is stabilized by interactions with initiation factors. eIF1 and eIF1A-CTT together induce the open conformation of the scanning complex. During scanning, the initiator tRNA is in the P_{out} state, leaving space between the codon and anticodon for inspection. Both eIF1 and eIF1A-CTT play pivotal roles for the fidelity of start codon selection by preventing premature basepairing of the initiator tRNA on near-cognate codons and therefore increasing the stringency of AUG selection (Fekete et al., 2007; Lomakin and Steitz, 2013; Weissner et al., 2013). In contrast, eIF5 stabilizes the P_{in} state, where the initiator tRNA is fully accommodated in the P site, and triggers the start-codon recognition by activating the GTPase activity of eIF2. Notably, translation of eIF1 and eIF5 mRNAs is autoregulated by their protein products, which provides a good example of a self-regulatory feedback loop. The AUG in eIF1 mRNA is in a weak context and is only selected by the ribosome under low cellular eIF1 level (Ivanov et al., 2010). In contrast, the eIF5 mRNA contains an inhibitory uORF with an AUG in weak context to reduce the expression of eIF5 protein in high cellular eIF5 condition. The inhibitory uORF of eIF5 mRNA is only bypassed by the scanning ribosome at high level of eIF1 and low level of cellular eIF5 (Loughran et al., 2012). The NTT of eIF1A also has similar functions as eIF5 in stabilizing the P_{in} state (Saini and Fischer, 2007). During scanning, eIF1A-NTT is tethered by the eIF5-CTD on the scanning complexes (Das et al., 2001; Das and Maitra, 2001; Paulin et al., 2001) and prevents the interaction between eIF5 and eIF2 β (Luna et al., 2013; Obayashi et al., 2017). When the start codon is recognized, eIF1A-NTT is released from eIF5 and reaches into the decoding center to stabilize basepairing, while eIF5-NTD interacts with eIF2 β -NTT to trigger GTP hydrolysis. Other initiation factors also contribute to the stabilization of the scanning complex. For example, eIF3 not only stabilizes TC and other initiation factors on the ribosome, but also interacts with the mRNA outside the mRNA binding channel and possibly also inside the decoding center. On the scanning ribosome, eIF3a and 3d contact the 5' UTR and form an extension of the mRNA binding channel from the exit channel (Pisarev et al., 2008), which also affects the efficiency of the mRNA recruitment (Aitken et al., 2016). Overall, the communications between all the factors and the ribosome increase the efficiency of the mRNA recruitment and scanning.

1.6.4 Secondary structures and RNA helicases

The secondary structures along the road during scanning must be unwound before the mRNA is threaded into the mRNA channel of the 40S subunit. ATP-dependent DEAD-box RNA helicase, eIF4A, is the major roadblock cleaner for the ribosome. Scanning is an ATP-dependent process (Kozak, 1980b). Only when an unstructured mRNA is used in the *in vitro* reconstitution system, the migration of the 43S PIC along the 5' UTR is ATP-independent (Pestova and Kolupaeva, 2002), but still requires the presence of eIF1 and eIF1A to stabilize the open scanning conformation (Passmore et al., 2007). The ATP hydrolysis is probably used by eIF4A to unwind the secondary structure and/or prevent the back tracking of the ribosome (Spirin, 2009). The level of necessity of eIF4A and ATP is also proportional to the degree of the secondary structures on the mRNA (Jackson, 1991; Svitkin et al., 2001). Translation of mRNAs with stable secondary structures (Rubio et al., 2014) or G-quadruplex structures (Wolfe et al., 2014) is highly dependent on eIF4A in mammals. Even in the case of cap-independent initiation, eIF4A still promotes the sliding of the ribosome along the 5' UTR. The mRNA of tobacco mosaic virus containing CAA triplet repeats in the Ω leader sequence, which results in an unstructured form of the mRNA, can bypass the requirement of m⁷G cap and eIF4F. However, eIF4A significantly accelerates the translation process on these unstructured mRNAs (Agalarov et al., 2014).

More RNA helicases also help the ribosome during scanning, e.g. DHX29 and DDX3. In the absence of eIF4A, mammalian DHX29 rescues the unwinding activity and helps to load the ribosome onto a structured mRNA both *in vitro* (Pisareva et al., 2008) and *in vivo* (Parsyan et al., 2009). While the binding site of eIF4A on the 40S subunit remains unclear, the structural study shows that DHX29 binds directly to the 40S subunit at the mRNA entry channel in association with eIF3a and 3b (Hashem et al., 2013a; Pisareva and Pisarev, 2016). Since eIF4B also binds to the mRNA entry channel (Eliseev et al., 2018), both DHX29 and eIF4A might function at the entry channel in an mRNA-specific manner, where DHX29 specifically enhances the scanning on the highly structured mRNA (Pisareva and Pisarev, 2016; Pisareva et al., 2008). Unwinding of mRNAs with secondary structures close to the m⁷G cap is facilitated by DDX3 in association with eIF4G (Soto-Rifo et al., 2012), ensuring unwinding activity during the mRNA

loading step. Although the relative amounts of DHX29 and DDX3 in cells are lower than that of eIF4A, these additional RNA helicases help the ribosome to broaden its translation ability on the highly structured mRNAs.

1.6.5 Leaky scanning and uORF

Leaky scanning often happens if the AUG is in a weak Kozak sequence context. Upon leaky scanning, the ribosomes bypass the AUG in a non-optimal sequence context instead of selecting it (Martin-Marcos et al., 2011). The uORFs often utilize leaky scanning mechanism to seize a portion of the scanning ribosomes, allowing only a few ribosomes to bypass the first AUG and reach the downstream AUG in the strong Kozak sequence in the main ORF. When the uORF is in frame with the downstream ORF, the translation from uORF can lead to an N-terminal domain (NTD) extension isoform. The NTDs are often the signals for the subcellular localizations, e.g. mitochondrial isoform (Slusher et al., 1991). The re-initiation efficiency is dependent on the association of eIF3. If the length of the uORF is shorter than 30 nt, higher portions of the ribosomes would still retain eIF3 or (at least) subunit 3h (Park et al., 2001; Roy et al., 2010), thereby increasing the re-initiation efficiency. Besides, highly efficient re-initiation is only found in the eIF4F- and eIF4G-dependent initiation pathway (Poyry et al., 2004), implying the utilization of the cap-dependent scanning mechanism.

The uORF is prevalent in the mammalian mRNAs. In eukaryotes, the mRNA transcripts are mostly monocistronic. However, about half of the mammalian genes encode mRNA transcripts with at least one short uORF (typically less than 30 nt), compared to ~13% of yeast genes (Calvo et al., 2009; Lawless et al., 2009; Resch et al., 2009). The ribosome profiling analysis reveals that a large portion of these uORFs is translated (Ingolia et al., 2009; Johnstone et al., 2016; Lee et al., 2012). Analysis from the mouse embryonic stem cells also shows that CUG and GUG can be used as initiation codon on the short uORFs (Ingolia et al., 2011). The near-cognates codons can be decoded by canonical TC with Met-tRNA^{Met}, or sometimes even by non-canonical imitation factor, eIF2A in complex with Leu-tRNA^{Leu}, especially for the gene

transcripts of the antigenic precursors for loading major histocompatibility complex molecules (Starck et al., 2012).

Finally, the availability of the TC is another determinant of protein expression under stress conditions. The classic example is the expression of mammalian ATF4 mRNA (Lu et al., 2004; Vattem and Wek, 2004; Wek, 2018), a homolog of yeast GCN4. Phosphorylation of eIF2 α inhibits the binding of eIF2B to the eIF2-GDP and therefore decreases the amount of the TC available for the initiation complexes, leading to global downregulation of the translation. The ATF4 mRNA contains two uORFs to regulate the expression of the main ATF4 ORF. The first uORF is efficiently translated by the ribosome and leads to re-scanning after the termination of the first uORF. The re-initiation on the second uORF, which functions as an inhibitory uORF, depends on the concentration of the free TC in the cell. Under rapid growing condition, the TC is in high concentration and the re-initiation complex is loaded with new TC immediately after the termination from the first uORF, which leads to the translation of second inhibitory uORF. Expression of the second inhibitory uORF reduces translation on the downstream main ATF4 ORF. However, under depletion of TC, longer arrival time of TC to the re-initiation complex after translation of the first uORF allows the ribosomes to bypass the second inhibitory uORF and reach the main ATF4 ORF. Presence of the second inhibitory uORF functions as a sensor to the concentration of the cellular TC and ensure the expression of essential transcription factors, ATF4, for cells to survive under the stress condition. In this case, the length between the uORFs balances the rate of scanning and association rate of TC, which provides an example of a sophisticated kinetic regulation applied by the cells.

1.7 Non-canonical initiation pathway

Next to the canonical initiation pathway, alternative initiation routes have been discovered and classified based on the dependence of the m⁷G cap and the scanning mechanism. These include cap-independent scanning mechanism, e.g. cap-independent translation enhancers (CITEs) mediated initiation, cap-dependent scanning free mechanism, e.g. translation initiator of short 5'UTR (TISU) mediated initiation, and cap-independent

scanning free mechanism, e.g. internal ribosome entry sites (IRES) of viral genes (**Figure 8**). Moreover, some viruses, e.g. calicivirus, utilize a viral genome-linked protein (VPg) covalently linked to the 5′ end of their mRNA and recruit translation machinery for the production of its own viral proteins (Goodfellow et al., 2005). Normally, when the canonical pathway is downregulated due to the cellular stress or viral infections, the non-canonical initiation pathways provide alternative routes towards efficient protein biosynthesis.

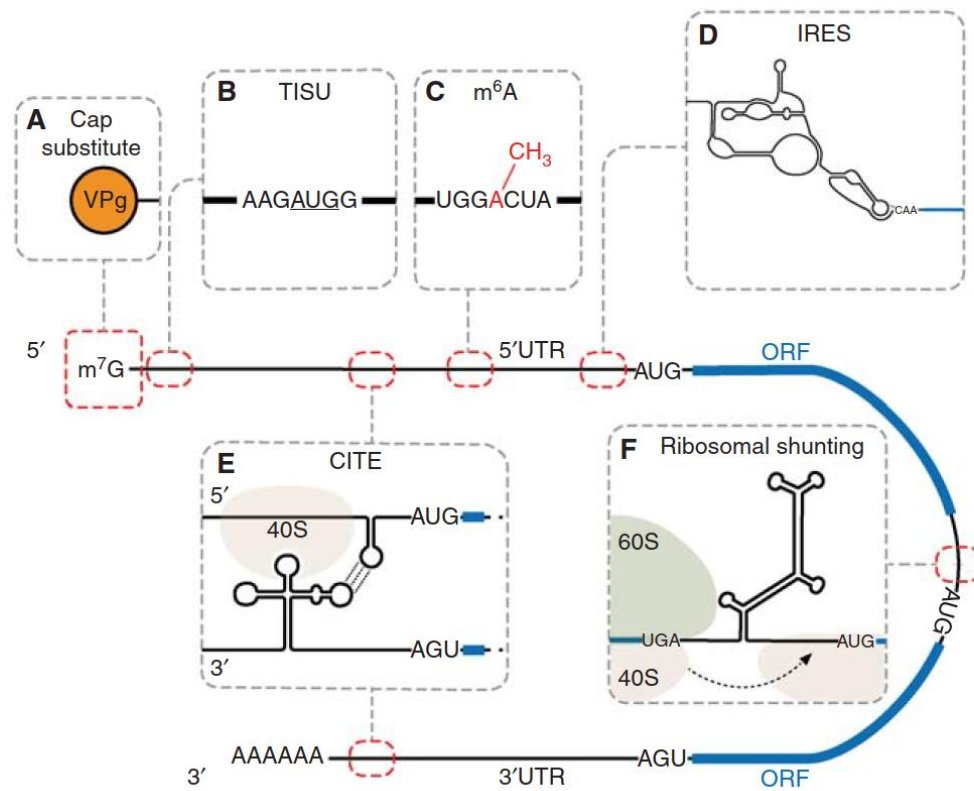


Figure 8. Non-canonical initiation pathway

Possible mechanisms of non-canonical initiation pathways: **(A)** Cap substitute with viral protein linked to the genome (VPg); **(B)** Translation initiator of short 5′ UTR (TISU) elements; **(C)** m⁶A modification in the 5′ UTR; **(D)** Internal ribosomal entry site (IRES) from the viral genes; **(E)** Cap-independent translation element (CITE); and **(F)** Ribosomal shunting. The figure is adapted from (Kwan and Thompson, 2019).

1.7.1 Cap-dependent scanning-free initiation

Around 4% of the mammalian mRNA bears translation initiators of short 5'UTR (TISU) elements and utilize cap-dependent, but scanning-free initiation mechanism (Elfakess and Dikstein, 2008). These mRNAs are also recruited to the ribosome in an eIF4F-dependent pathway, but the subsequent process of start codon selection is different from the canonical scanning mechanism (**Figure 8**). The TISU-mediated genes are mostly housekeeping genes involving respiration, protein metabolism and RNA synthesis, e.g. mitochondrial genes (Sinvani et al., 2015). The TISU-mediated mRNAs contain an AUG very close to the 5' cap with a very short 5' UTR (median 12 nt in length) (Elfakess and Dikstein, 2008). In the canonical scanning mechanism, AUGs located within 20 nt of the 5' cap are usually excluded/bypassed by the ribosome in the presence of eIF1 even if the AUGs are in a strong Kozak sequence (Elfakess and Dikstein, 2008; Kozak, 1991c; Kumar et al., 2016). However, eIF1, together with eIF4GI, mediates efficient translation from the AUGs in the TISU elements close to the 5' cap (Sinvani et al., 2015). Depletion of eIF1, *in vivo* and *in vitro*, diminishes translation directed by the TISU elements but not by an AUG in a strong Kozak sequence (Sinvani et al., 2015). The TISU flanking sequence includes, besides the typical -3A and +4G positions, more elements downstream the AUG to compensate for the short 5' UTR interaction, which ranges from position -2 to +8 (Elfakess et al., 2011). TISU not only regulate translation efficiency but also influence transcription reaction via Ying Yang 1 (YY1) transcription factor (Elfakess and Dikstein, 2008), acting as a link between mammalian transcription and translation.

1.7.2 Cap-independent translation enhancers (CITEs)

Another group of mRNAs possesses cap-independent translation enhancers (CITEs) in the UTR and recruits the ribosomes to the 5' end of the mRNA by factors other than eIF4F, e.g. eIF3 (**Figure 8**). Translation initiation on CITEs-dependent mRNA requires a full canonical initiation complex to scan from the 5' end of the mRNA. Therefore, this new concept of translation initiation, CITEs-mediated translation, is a cap-independent but scanning-dependent process. These *cis* elements can be a post-transcriptional modification of the mRNA,

e.g. N⁶-methyladenosine (m⁶A), or a stem loop structure that is recognized by a mediator protein. Cap-independent initiation permits translation of a subset of cellular and viral mRNAs under conditions where cap-dependent translation is inhibited.

eIF3 is one of the well-studied CITEs mediators. Binding sites of eIF3 on the mRNA transcripts are found not only near the start codon region (Pisarev et al., 2008) but also in the 5' UTR (of around 3% of mRNA transcripts) with its subunits 3a, 3b, 3d and 3g (Lee et al., 2015). eIF3 can be an activator or repressor of translation of the CITE-mediated mRNA (Lee et al., 2015). For example, translation of the cJUN mRNA is activated by eIF3, whereas expression of the B cell translocation gene 1 transcript is repressed after binding to eIF3 (Lee et al., 2015). Further analysis shows that eIF3 recognizes the stem loop structure in the 5' UTR of the cJUN mRNA, while eIF3d binds to the cap structure and prevents eIF4F-mediated initiation (Lee et al., 2016). Binding of the eIF3 to the stem loop in the 5' UTR of cJUN transcript subsequently recruits 43S PIC to the 5' end of the mRNA and activates the translation of the downstream ORF. Besides binding to the mRNA secondary structure, eIF3 (through its eIF3l subunit) can also interact with the m⁷G cap after eIF4E is released from the cap during initiation (Kumar et al., 2016). In these cases, eIF3 not only acts as a scaffold protein to communicate between the initiation factors, but also functions as a mediator to associate the mRNA with the ribosome.

The N⁶-methylation of adenosines (m⁶A) in the 5' UTR also stimulates translation via eIF3, which by itself also recruits 43S PIC, through scanning-dependent, but cap-independent mechanism (Meyer et al., 2015; Zhou et al., 2015). The m⁶A modifications are mainly found near the stop codons and within the 3' UTRs. However, when the modification appears in the 5' UTR within the GAC sequence context, even a single m⁶A can efficiently stimulate the translation (Meyer et al., 2015). The m⁶A-induced ribosome recruitment is especially important under stress conditions, such as heat shock, since the level of methylation is upregulated in the newly synthesized mRNA transcripts (Zhou et al., 2015). Unlike cJUN transcripts, m⁶A-mediated transcripts can also be translated with canonical eIF4F-mediated cap-dependent pathway. The two initiation pathways allow efficient translation of the mRNA under a variety of growth conditions (Coots et al., 2017).

1.7.3 Internal ribosome entry sites (IRES) -mediated initiation

Internal ribosome entry sites (IRESs) are widely used by the viral genes to hijack the host translation system and can achieve high translation efficiency by recruiting truncated/partial version of the canonical initiation complexes in the condition when initiation factors are under limitation (**Figure 9**). For most IRESes, translation initiation occurs through cap-independent and scanning-free mechanism; therefore, the molecular mechanism differs further from the canonical initiation pathway. IRES contains secondary structures that could directly recruit the ribosome to the translation start site and usually position the AUG directly in the P site. These secondary structures either mimic the structures of the codon-anticodon complex, e.g. Cricket Paralysis Virus (CrPV) IRES (Wilson et al., 2000a) or attract initiation factors to recruit the ribosome, e.g. Hepatitis C virus (HCV) IRES (Hashem et al., 2013b). In order to achieve efficient production of the viral protein, only a limited set of initiation factors is required to facilitate the IRES-mediated initiation. For example, HCV IRES is recognized by the eIF3 and only requires additionally TC with the 40S subunit for initiation (Pestova et al., 1998b). In the HCV IRES, eIF3 also binds directly to the RNA secondary structure in a different orientation than in the canonical 48S IC (Hashem et al., 2013b). CrPV IRES requires none of the initiation factor for the 80S ribosome assembly and initiates translation from the A site (Wilson et al., 2000b). To overcome the host immune system, viruses have to replicate in a short infection period. IRESs are probably evolved to ensure rapid and efficient translation of viral proteins at the cost of the host proteins. Different groups IRESs have a variety of specific non-canonical initiation mechanisms, discussion of which would be out of the scope of this work and is nicely depicted a review by (Shatsky et al., 2018). Here, I will continue to describe the mechanism of the cap-dependent and scanning-dependent initiation.

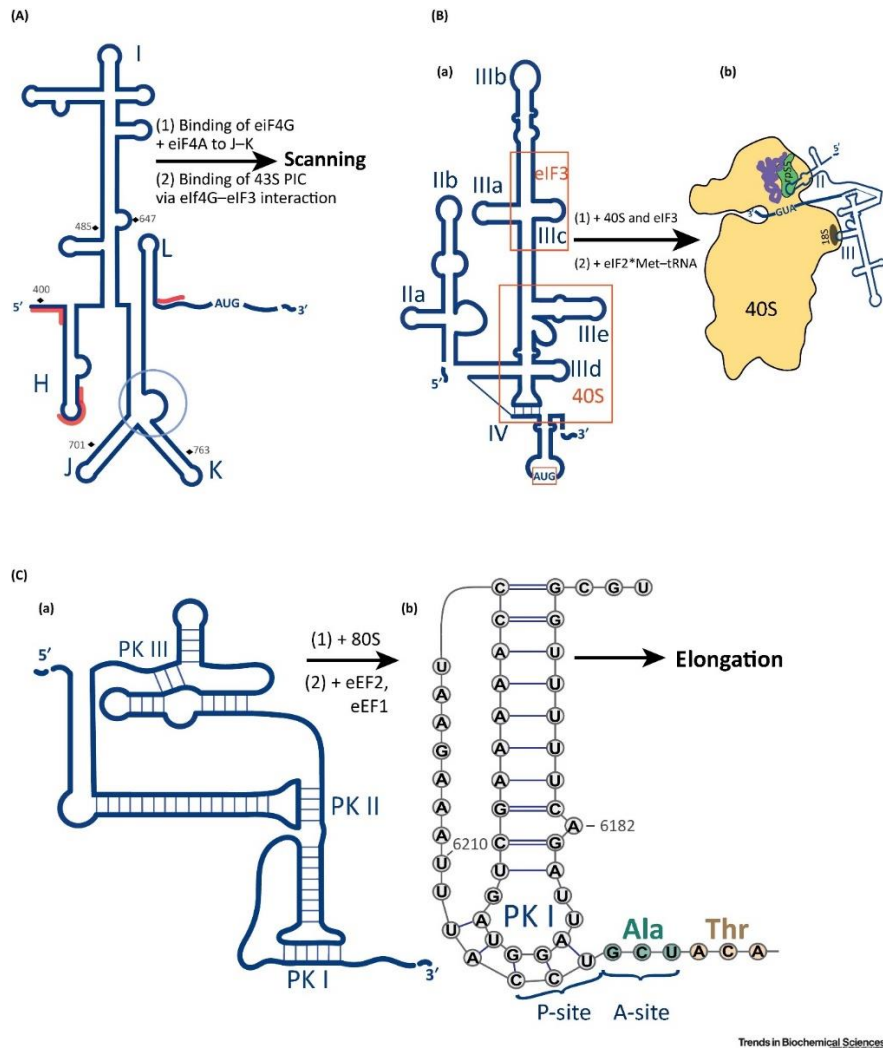


Figure 9. Different types of IRES

(A) IRES elements of the Encephalomyocarditis virus (EMCV). The EMCV IRES requires a full set of initiation factors except for eIF4E and utilizes scanning-dependent initiation pathway. **(B)** IRES elements of the Hepatitis C virus (HCV). The HCV IRES requires mainly eIF3 for the recruitment of the 40S subunit and subsequent TC for the start codon recognition. **(C)** IRES elements of the Cricket Paralysis Virus (CrPV). The CrPV IRES requires only the 80S ribosomes and starts the translation directly at the A site. The figure is adapted from (Shatsky et al., 2018).

1.8 Start codon recognition and 48S initiation complex formation

The start codon recognition is the key step in the initiation process for the ribosome to define the open reading frame and for the formation of a stable 48S IC. The fidelity of the start codon selection is also determined during this step. AUG is the most frequent start codon which is selected with 50-fold higher probability than other near-cognate codons (Fekete et al., 2005). CUG, GUG, UUG are the most commonly used near-cognate codons, but usually in the uORFs (Kearse and Wilusz, 2017). As described above, the sequence context and secondary structures modulate the start codon selection, probably by slowing down the scanning complexes. When an AUG is in an optimal Kozak sequence, CC A/G CC AUG G, it has a higher chance to be selected in cells (Kozak, 1986; Noderer et al., 2014). 40% of human genes utilize AUGs flanking by -3 A/G and +4 G as translation start site, while only 9.6% of human genes use AUGs in neither -3 A/G nor +4 G context (Noderer et al., 2014). Other genes with AUGs in either -3 A/G or +4 G have a moderate efficiency (Kozak, 1991a; Lukaszewicz et al., 2000).

The formation of 48S initiation complex involves concerted factor-interactions and rearrangements of the complex's structures. The major steps include: relocation of eIF1 and eIF1A-tails, stabilization of P_{in} state and eIF5-induced GTP hydrolysis of TC. From structural and steady-state thermodynamic studies, we have a better understanding of the functions of individual factors during the start codon recognition. Mutagenesis *in vivo* studies in yeast have also revealed the individual domains or amino acid residues of the initiation factors that affect the fidelity of the start codon selection. Here, I will introduce the molecular dynamics of the initiation factors and the ribosome during start codon recognition with information obtained mainly from the yeast *in vivo* and *in vitro* studies.

1.8.1 Relocation of eIF1 from the P site upon start codon recognition

eIF1 is the major guard of start codon fidelity. eIF1 enables 43S PIC to discriminate non-AUG triplets and AUGs in the poor Kozak context (Pestova and Morrison, 1998; Pestova and Kolupaeva, 2002; Pisarev et al., 2006). eIF1 also distinguishes the AUGs adjacent to the m⁷G

cap or located within 8 nucleotides from the mRNA 5' cap (Kumar et al., 2016; Pisarev et al., 2006). Only in the absence of eIF1, the AUGs close to the 5' can be translated efficiently *in vitro* (Kumar et al., 2016). The concentration of eIF1 in cells also influences the fidelity and initiation efficiency of weak AUGs, e.g. autoregulation of its own eIF1 mRNA transcript (Ivanov et al., 2010). Moreover, addition of extra eIF1 and eIF1A in the rabbit reticulocyte lysate enhances initiation of AUG in the optimal Kozak sequence on a reporter mRNA, whereas addition of extra eIF5 and eIF5B strengthens the initiation from CUG codon in the uORF (Barth-Baus et al., 2013).

The fidelity of start codon recognition is controlled by the relocation/dissociation of eIF1 from the P site upon start codon recognition (**Figure 10**) (Hussain et al., 2014; Llacer et al., 2015; Llacer et al., 2018; Zhang et al., 2015). Displacement of eIF1 upon AUG recognition relieves the repression of GTP hydrolysis and Pi release from the TC (Algire et al., 2005; Lomakin and Steitz, 2013; Unbehaun et al., 2004). In the scanning complex, eIF1 binds tightly at the P site and blocks Met-tRNA_i^{Met} from full accommodation. eIF1 is then partially displaced (or released) to allow codon-anticodon basepairing upon start codon recognition, probably due to the displacement by eIF5 (Llacer et al., 2018; Nanda et al., 2013). Binding of the eIF5-CTD to the 40S subunit upon start codon recognition reduces the affinity of eIF1 to the 43S-mRNA PIC (Nanda et al., 2013). The eIF5-NTD replaces the vacant position from eIF1 and interacts with eIF2 β to promote the full accommodation of Met-tRNA_i^{Met}, which subsequently triggers GTP hydrolysis (Llacer et al., 2018). This links the dissociation of eIF1 from the 48S IC to the fidelity of start codon selection. The affinity of yeast eIF1 to the 40S-eIF1-eIF1A-TC complexes (43S PIC) ($K_d = 1.4$ nM) is decreased 40-fold in the presence of AUG mRNA ($K_d = 53$ nM) (Maag et al., 2005). Kinetic analysis also shows that eIF1 dissociates faster from the 40S-eIF1-eIF1A-TC complexes (43S PIC) in the presence of AUG mRNA (0.6 s⁻¹) than in the presence of CUC mRNA (0.015 s⁻¹) (Maag et al., 2005). The slower dissociation rate of eIF1 from the 48S-CUC complexes might be due to incomplete basepairing of the mRNA and Met-tRNA_i^{Met}, which would explain why eIF1 stays tightly bound to the P site on the 40S subunit. Notably the dissociation curve of eIF1 from the 48S IC is biphasic; the authors attributed the faster phase to a conformational change of the complex and the slower phase to the actual

dissociation of eIF1 from the complex. The k_{app} values of the fast phases of the 48S·AUG and 48S·CUC complexes are 9 s^{-1} and 2 s^{-1} , respectively.

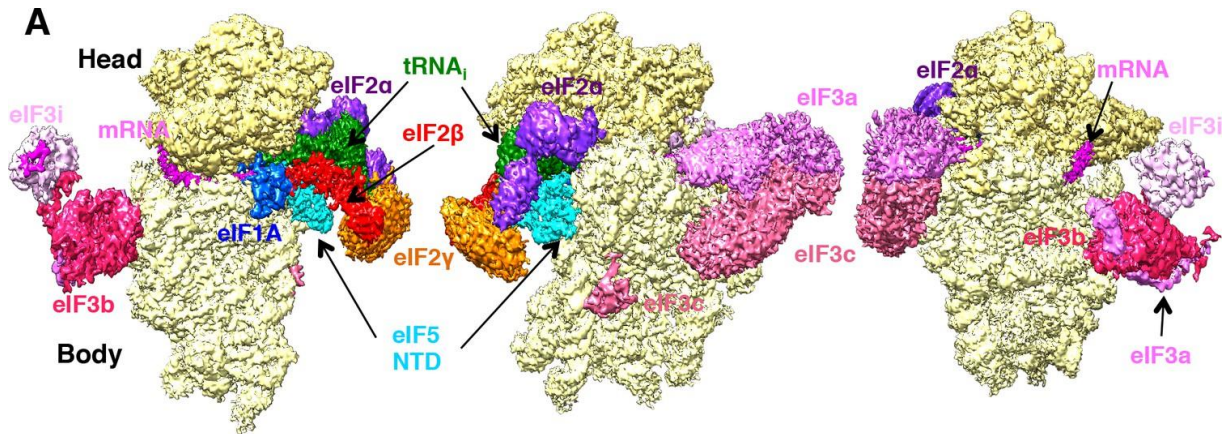


Figure 10. Replacement of eIF1 by eIF5 at the P site in the closed state of yeast 48S PIC after start codon recognition

Yeast 48S PIC is assembled with eIF1, eIF1A, TC, eIF3, eIF5, eIF4A, eIF4B, eIF4E, eIF4G with an unstructured capped AUG mRNA. After start codon recognition, eIF1 is replaced by eIF5-NTD from its original P site to allow full accommodation of Met-tRNA^{Met} into the P_{in} state. This figure is adapted from (Llacer et al., 2018).

1.8.2 Movement of N- and C-terminal tails of eIF1A

Both *in vivo* and *in vitro* studies in yeast have shown that the N- and C-terminal tails (NTT and CTT, respectively) of eIF1A have distinct functions during scanning and start codon recognition (Fekete et al., 2007). During scanning, eIF1A-CTT reaches into the P site and functions together with eIF1 to induce the open scanning conformation of the ribosome and keep the Met-tRNA^{Met} in the P_{out} state (Lomakin and Steitz, 2013; Weissner et al., 2013), while eIF1A-NTT is tethered with eIF5-CTD on the complexes at this stage (Luna et al., 2013; Obayashi et al., 2017). Upon start codon recognition, eIF1A-NTT is released from eIF5-CTD and moves into the A site close to the +4 G and +5 - +7 position (Hussain et al., 2014; Llacer et al., 2015; Lomakin and Steitz, 2013; Simonetti et al., 2020). The codon-anticodon duplex is further stabilized by eIF1A-NTT and thereby, the Met-tRNA^{Met} is fully accommodated in the P_{in} state (Hussain et al., 2014). The eIF1A-CTT moves towards eIF5-NTD upon AUG recognition in a biphasic FRET-signal curve. Notably, the observed rate constant of the slow phase (0.6 s^{-1}) is in

the same range as the release of eIF1 and Pi (0.6 s^{-1} and 0.4 s^{-1} , respectively) (Maag et al., 2006; Nanda et al., 2013). This suggests that the movement of eIF1A-CTT to the eIF5-NTD upon AUG recognition coincides with the release of eIF1 from the 43S-mRNA PIC (Nanda et al., 2013). When the fast- or slow-dissociating yeast eIF1 (G107E/G107K) mutants are used, the FRET signals between eIF1A-CTT and eIF5-NTD also increased and decreased, respectively (Nanda et al., 2013). Mutations in eIF1A-CTT which are slower in the second slow phase also further reduce Pi release but do not affect the dissociation rate of eIF1 (Nanda et al., 2013), implying that the displacement of eIF1 alone is not sufficient to trigger Pi release from TC. Together, the interaction between eIF1A and eIF5, and the relocation of eIF1 from the P site gate the GTP hydrolysis and Pi release upon start codon recognition (Maag et al., 2006; Nanda et al., 2013).

Unlike eIF1, the affinity of eIF1A to the 43S-mRNA (AUG) complexes is very high and is difficult to measure accurately at equilibrium using fluorescence spectroscopy (Maag et al., 2006). To overcome this limitation, Lorsch et al. developed kinetic approaches to measure the stability of eIF1A binding to the ribosome using changes in the anisotropy of the fluorescence reporter on eIF1A. In the *in vitro* yeast reconstituted translation system, the dissociation rates of eIF1A from the 40S-eIF1-eIF1A complex and 43S-mRNA (AUG) complex are 0.27 s^{-1} and 0.015 s^{-1} , respectively (**Table 2**) (Maag et al., 2005), indicating a tighter interaction of eIF1A to the 40S subunit in the presence of TC and AUG mRNA. Further analysis shows that the dissociation of eIF1A from the 48S (AUG) PIC possesses a biphasic kinetic pattern, with a fast phase (0.015 s^{-1}) presumably referring to the open conformation and a slow phase (0.0038 s^{-1}) reflecting the closed conformation (Llacer et al., 2018; Maag et al., 2006); the amplitudes are interpreted as the fraction of the complex in the respective conformation. The biphasic pattern is also seen upon the dissociation of eIF1A from the 43S PIC in the absence of mRNA and eIF5 (with 0.094 s^{-1} and 0.0023 s^{-1}), implying the existence of two distinct populations with two defined conformations of the 40S subunits (Llacer et al., 2018; Maag et al., 2006). Addition of eIF5 to the 48S (AUG) PIC further stabilizes eIF1A on the complexes with two even slower dissociation rates (0.0049 s^{-1} and 0.00053 s^{-1}). eIF5 also shifts the equilibrium of the populations towards the slower phase from 37% to 87% upon AUG recognition (Maag et al., 2006). In conclusion,

the 43S PIC may switch back and forth between open and closed states even in the absence of AUG mRNA. Therefore, during scanning, the ribosome dynamically samples the codon-anticodon interaction on the mRNA, while initiation factors stabilize the conformation upon stable base-pairing.

Table 2. Summary of dissociation rates of eIF1A from yeast 43S PIC and 48S IC

Complex	Composition	A ₁	k _{off1} (s ⁻¹)	A ₂	k _{off2} (s ⁻¹)	
40S·eIF1·eIF1A	40S·eIF1·eIF1A		0.270			*
43S·mRNA (AUG)	40S·eIF1·eIF1A·TC·mRNA		0.015			*
43S PIC	40S·eIF1·eIF1A·TC	0.73	0.094	0.27	0.0023	**
43S PIC+eIF5	40S·eIF1·eIF1A·TC·eIF5	0.57	0.077	0.43	0.00059	**
48S (CUC) PIC	40S·eIF1·eIF1A·TC·mRNA	0.86	0.110	0.14	0.0039	**
48S (CUC) PIC+eIF5	40S·eIF1·eIF1A·TC·eIF5·mRNA	0.67	0.086	0.33	0.0012	**
48S (UUG) PIC	40S·eIF1·eIF1A·TC·mRNA	0.22	0.017	0.78	0.0019	**
48S (UUG) PIC+eIF5	40S·eIF1·eIF1A·TC·eIF5·mRNA	0.2	0.015	0.8	0.0012	**
48S (AUG) IC	40S·eIF1·eIF1A·TC·mRNA	0.63	0.015	0.37	0.0038	**
48S (AUG) IC+eIF5	40S·eIF1·eIF1A·TC·eIF5·mRNA	0.13	0.0049	0.87	0.00053	**
48S (AUG) IC+eIF3+eIF5	40S·eIF1·eIF1A·TC·eIF3·eIF5·mRNA	0.19	0.0068	0.81	0.00081	**

The A₁ and A₂ are the relative portion of the amplitudes to the corresponding k_{offS}. All the complexes are assembled with yeast WT eIFs and native yeast 40S subunit. *The data is adapted from (Maag et al., 2005). **The data is adapted from (Maag et al., 2006).

1.8.3 Closed conformation of the 40S subunit

Structural studies revealed the conformational changes of the 40S subunit during initiation. Generally, the 40S subunit adopts the open conformation during scanning and the closed conformation after start codon recognition, where the tRNA_i is in the P_{out} and P_{in} state, respectively (Llacer et al., 2015; Llacer et al., 2018). Upon start codon recognition, the head of the 40S subunit moves largely to close the mRNA latch, further restricting the movement of the mRNA and the Met-tRNA_i^{Met} in the decoding center (**Figure 11**) (Llacer et al., 2015). The rRNAs and ribosomal proteins also form interactions with the mRNA around the AUG codons in the decoding center, especially with the Kozak sequence. Helix 44 of 18S rRNA and rpS15

contact the +4 G position (Pisarev et al., 2006; Pisarev et al., 2008), while rpS5 binds to the -3 positions together with eIF2 α (**Figure 7**) (Pisarev et al., 2008). On the other hand, eIF3 octamer core binds to the solvent side of the 40S subunit and the peripherals subunits wrap around the 40S body in the 48S PIC (**Figure 3**) (Brito Querido et al., 2020; Eliseev et al., 2018; Simonetti et al., 2020). The crosslinking data also suggest that eIF3a and eIF3d form an extension of the mRNA channel outside the exit channel and protects at least 18 nt upstream the AUG in the 5' UTR (Pisarev et al., 2008). The interactions of eIF3, especially eIF3a, on the 5' UTR also have strikingly effects on the mRNA recruitment efficiency to the yeast 48S IC (Aitken et al., 2016). In conclusion, eIF3 and the 40S subunit undergo substantial structural rearrangements during start codon recognition.

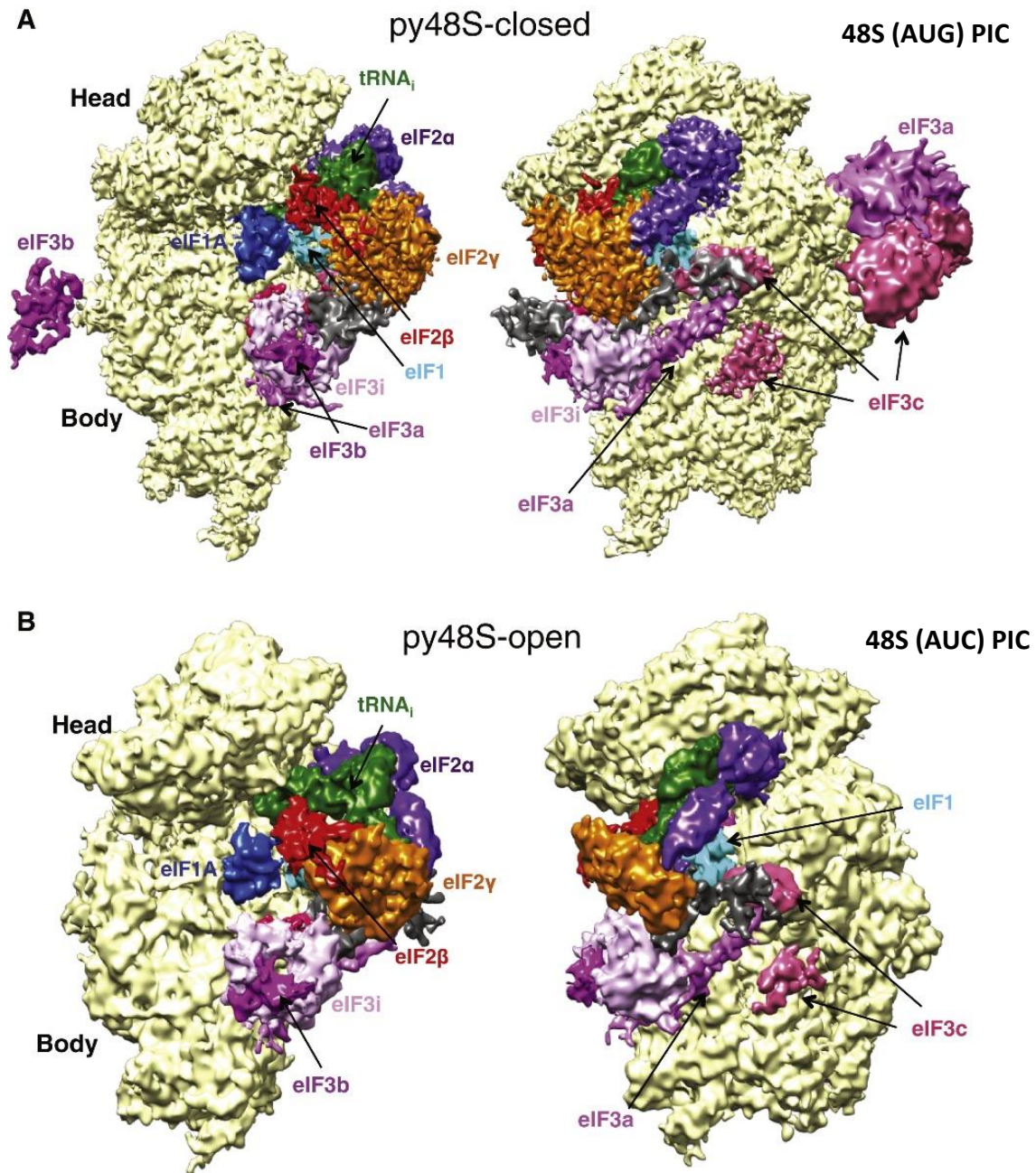


Figure 11. Open and closed conformations of yeast 48S PIC

(A) The closed state of yeast 48S PIC. The yeast 48S PIC is assembled with eIF1, eIF1A, TC, eIF3, eIF5 with an unstructured uncapped AUG mRNA. **(B)** The open state of yeast 48S PIC. The open state of yeast 48S PIC is assembled with eIF1, eIF1A, eIF3, TC on an unstructured uncapped AUC mRNA. After start codon recognition in the 48S (AUG) PIC, 40S head has a large conformational change from an open mRNA latch conformation to a closed one that restricts mRNA binding channel. Initiator tRNA is in the P_{in} state in the closed state of 48S (AUG) PIC. The figure is adapted from (Llacer et al., 2015).

1.8.4 eIF5-induced GTP hydrolysis of ternary complex

GTP hydrolysis and phosphate (Pi) release from the eIF2 ternary complex are the rate determining steps during translation initiation and commit the ribosome to the translation start site. GTP hydrolysis is stimulated by the GTPase activating protein (GAP) function of eIF5, especially by the interaction of eIF5-NTD to the eIF2 β (Das et al., 2001; Paulin et al., 2001). eIF5 is part of the 43S PIC and also an important factor in the multifactor complex. Before the start codon recognition, the eIF5-CTD binds to the 43S PIC and is associated with eIF1A-NTT (Luna et al., 2013; Obayashi et al., 2017). The interaction between eIF5-CTD and eIF1A retains eIF1 in the P site during scanning (Luna et al., 2013). Upon start codon recognition, the full accommodation of the Met-tRNA_i^{Met} in the P_{in} state is achieved following the relocation of eIF1. eIF5-NTD binds to the 40S subunit at the vacant position from eIF1 (Llacer et al., 2018) and stabilizes the codon-anticodon duplex (Saini and Fischer, 2007). The movement of eIF5-NTD is coupled to the eIF1A-CTT (Nanda et al., 2013), which further establishes the essential interaction between eIF5-NTD and eIF2 β (Das et al., 2001; Das et al., 1997; Das and Maitra, 2000). The GAP function of eIF5-NTD induces the GTPase activity of eIF2 γ by providing a classic Arginine finger (Paulin et al., 2001) and derepressing the GTPase activity (Marintchev and Wagner, 2004).

The rate of GTP hydrolysis by eIF2 is controlled via eIF1 and start codon recognition. The key event is the displacement of eIF1 upon AUG recognition, relieving the repression of GTP hydrolysis (Algire et al., 2005; Unbehaun et al., 2004). In the absence of eIF1, the rate of eIF5-stimulated GTP hydrolysis is low and insensitive to the presence of mRNA, i.e., identical in the 43S PIC and 48S IC (Algire et al., 2005; Unbehaun et al., 2004). eIF1 slows down GTP hydrolysis in the 43S PIC, but stimulates GTP hydrolysis in the presence of AUG. The stimulation is higher with the cognate AUG than with the near-cognate codons (Unbehaun et al., 2004), implying the gatekeeper role of eIF1 in controlling GTP hydrolysis. Yeast *in vitro* studies show a link between the dissociation rate of eIF1 and the Pi release rate (Algire et al., 2005). When a rapidly dissociating eIF1 mutant, G107E, was present, the Pi release rate also increased, whereas a slowly dissociating eIF1 mutant, G107K, also slows down the Pi release (Algire et al., 2005; Nanda et al., 2013). Higher levels of non-AUG initiation events are also

observed *in vivo* in the yeast eIF1 mutants with faster dissociation rates (Cheung et al., 2007). In mammals, eIF3 and eIF4F also inhibit eIF5-induced GTP hydrolysis prior the AUG recognition (Majumdar and Maitra, 2005), while eIF4A and eIF4B play an important role, beside their unwinding activity, in promote efficient GTP hydrolysis in the presence of AUG mRNA. After GTP hydrolysis, eIF2·GDP loses its affinity to Met-tRNA_i^{Met} (Jennings et al., 2017; Kapp and Lorsch, 2004), enabling partial release of eIF2·GDP from the 40S subunit probably together with eIF5 (Unbehauen et al., 2004), or by the displacement later from eIF5B (Pisarev et al., 2006). The complete dissociation of eIF2, eIF3 and eIF1 is likely mediated by eIF5B and 60S subunit (Unbehauen et al., 2004), leaving Met-tRNA_i^{Met} in the P site for subsequent large subunit joining and peptide bond formation.

1.8.5 Non-AUG initiation

Although AUG is the predominant start codon, near-cognate codons can also be translation state sites (Peabody, 1989; Zitomer et al., 1984). From the analysis with different translation assays, CUG is generally the most efficient near-cognate codon followed by GUG, ACG and AUU (**Table 3**). By treating the cells with early elongation inhibitors to trap the 80S ribosome at the initiation position before the first translocation cycle, ribosome profiling reveals that the near-cognate codons can act as translational start sites *in vivo*. For example, in the mouse embryonic stem cells, ~60% of the transcripts starts at non-AUG codons, including 75% from the uORFs (Ingolia et al., 2011). The near-cognate codons in the uORFs can downregulate the expression of the main ORF (Ivanov et al., 2008) or generate an N-terminal extended isoform (Liang et al., 2017a; Slusher et al., 1991). One example is the AUU codon in the uORF of polyamine biosynthesis gene which controls polyamine-directed repression of the downstream ORF (Ivanov et al., 2008). The high level of spermidine in the cells decreases the fidelity of the ribosome and also enhances the selection of the AUU in the uORF. Therefore, the downstream ORF that encodes biosynthesis enzymes of polyamines production is inhibited, leading to reduction of the polyamines, e.g. spermidine, in cells. In addition to uORFs, some endogenous and viral genes also initiate from non-AUG start codons (Beerman and Jongens,

2011; Chang and Wang, 2004; Ivanov et al., 2011). In both mouse and human genome, the open reading frame of DAP5, also called eIF4G2, utilizes GUG as the start codon. (Imataka et al., 1997; Imataka and Sonenberg, 1997; Takahashi et al., 2005). DAP5 functions in IRES-mediated translation (Lieberman et al., 2015) and also facilitates cap-independent initiation process that produces proteins involved in stem cell differentiation (Yoffe et al., 2016).

Table 3. Efficiency of Non-AUG start codons in various assays

Biological source	Rabbit reticulocyte lysate (Peabody 1989)	Rabbit reticulocyte lysate (Wei et al. 2013)	Wheat germ extract (Peabody 1989)	<i>Saccharomyces cerevisiae</i> (Clements et al. 1988)	<i>Neurospora crassa</i> (Wei et al. 2013)	HEK293T cells (Ivanov et al. 2010)
Reporter used	Dihydrofolate reductase	Firefly luciferase	Dihydrofolate reductase	β -Galactosidase	Firefly luciferase	Firefly luciferase
AUG	100	100	100	100	100	100
CUG	82	18	36	0.22	10	19
GUG	36	11	8	0.5	6	9
UUG	39	8	10	0.37	2.5	2
ACG	84	5	45	0.39	3.5	7
AUC	47	3	17	0.05	<0.5	2
AUU	67	6	14	0.38	1	3
AAG	14	0.01	3	0.02	<0.5	<0.5
AUA	59	5	30	0.29	3	3
AGG	17	0.02	3	0.04	<0.5	<0.5

The table is adapted from (Kearse and Wilusz, 2017) based on (Clements et al., 1988; Ivanov et al., 2010; Peabody, 1989; Wei et al., 2013).

The non-AUG codons can be decoded by the canonical initiator Met-tRNA_i^{Met}, but in some cases, Leu-tRNA and Val-tRNA are found to decode the CUG/UUG and GUG codons, respectively (Dmitriev et al., 2010; Sendoel et al., 2017; Skabkin et al., 2010; Starck et al., 2012; Starck et al., 2016). Translation initiation on CUG, GUG, ACG, UUG, AGG, AAG, AUA, AUC or AUU near-cognate codons gives rise to full length [³⁵S]-labeled proteins synthesized in an *in vitro* translation system supplemented by [³⁵S]-labeled Met-tRNA_i^{Met} (Peabody, 1989). While Met-tRNA_i^{Met} is usually delivered by eIF2, non-canonical initiator factors, eIF2A and eIF2D, can also facilitate Met-tRNA_i^{Met} binding to the 40S subunit in a GTP-independent or GTP-dependent manner, respectively (Dmitriev et al., 2010; Komar et al., 2005; Skabkin et al., 2013; Starck et al., 2012; Weissner et al., 2017; Zinoviev et al., 2015; Zoll et al., 2002). eIF2A has a low stringency of tRNA selection and can interact with both Met-tRNA_i^{Met} and Leu-tRNA. However, the efficiency of Leu-tRNA deliver to the CUG codons is ~40% higher than that of Met-tRNA_i^{Met} (Komar et al., 2005; Starck et al., 2016). eIF2D, also called ligatin, can deliver both Met-tRNA_i^{Met}

and Val-tRNA. eIF2D promotes initiation on the GUG codon in the Sindbis virus 26S mRNA with Val-tRNA in a GTP-independent manner (Dmitriev et al., 2010; Skabkin et al., 2010). Generally, eIF2A and eIF2D favor the expression of stress-related genes when the canonical translation is downregulated.

The factors that affect the rates of scanning and start codon recognition also determine the efficiency of non-AUG initiation, but their effect is often opposite to that on the canonical AUG pathway. The Kozak sequence has a larger influence on the recognition of non-AUG than AUG codons (Hussain et al., 2014; Pisarev et al., 2006), while the non-AUG codons with a downstream hairpin structure is also used more efficiently *in vitro* and *in vivo* (Kearse et al., 2016; Kozak, 1990; Liang et al., 2017b). In both cases, the scanning complexes are slowed down by the *cis* elements on the mRNA, thereby increasing the usage of a non-AUG codon. Disturbing the functions of initiation factors in fidelity maintenance also changes the balance between AUG and non-AUG initiation. As mentioned above, yeast eIF1 mutants that dissociate faster from the ribosome enhance initiation at non-AUG codons *in vivo* (Cheung et al., 2007). Acceleration of GTPase hydrolysis with mutations by eIF2 and eIF5 results in increasing initiation at non-AUG codons, e.g. UUG (Huang et al., 1997). Similarly, overexpression of eIF5 increases non-AUG translation in human HEK293T cells (Loughran et al., 2012; Nanda et al., 2009), whereas overexpression of eIF1 in mammalian cells inhibits initiation from non-AUG codons (Ivanov et al., 2010). Moreover, start codon fidelity can be further altered via post-translational modification of eIF1 and eIF5. (Homma et al., 2005; Majumdar et al., 2002). Fine tuning of translation efficiency by a broader range of codon usages is a sophisticated post-transcriptional mechanism for cells to adapt to the rapidly changing environments.

1.9 80S initiation complex formation

The 40S subunit and the 60S subunit can spontaneously associate with each other at high Mg^{2+} concentration. In the cellular environment, eIF6 binds to the free 60S subunit and prevents their binding to the 40S subunit in the absence of mRNA (Ceci et al., 2003). The joining of the large subunit and the small subunit after the start codon recognition is mediated by

eIF5B (Pestova et al., 2000), a ribosome-dependent GTPase. eIF5B is homologous to the bacterial IF2 (Pestova et al., 2000) and binds to the same region on the 40S subunit as IF2 at the inter-subunit cleft (Unbehaun et al., 2007).

Before the 60S subunit binds to the interface of the 40S subunit, initiation factors, eIF1, eIF2, eIF3, eIF5 and probably eIF4F dissociate from the complex, in an unknown order, to create room for the 60S subunit. eIF5B also facilitates the displacement of initiation factors in the presence of 60S subunit, likely during 60S subunit joining step (Pisarev et al., 2008; Pisareva and Pisarev, 2014; Unbehaun et al., 2004). At this stage, eIF1 might already relocate from the P site after start codon recognition (Hussain et al., 2014; Llacer et al., 2015) and dissociate completely later during 60S joining (Unbehaun et al., 2004). During or after the recognition event, eIF5 and eIF5B lower the affinity of eIF2 to the 48S IC and partially displace eIF2 (Pisarev et al., 2006; Pisareva and Pisarev, 2014), reflected in a different cross-linking pattern of eIF2 α to the mRNA (Pisarev et al., 2008). The complete dissociation of eIF2·GDP, likely together with eIF5 (Unbehaun et al., 2004), occurs upon 60S subunit joining (Pisarev et al., 2006). eIF3 binds to the 48S IC in a stable manner due to its extra interactions with the mRNA (Pisarev et al., 2008) and the presence of eIF5 and eIF5B do not alter its affinity to the 48S IC (Pisareva and Pisarev, 2014). However, eIF3 is not present in the 80S ribosome (Pisarev et al., 2008) or only parts of the eIF3 subunits, e.g. 3h and 3e, remain on the 80S ribosome (Lin et al., 2020; Mohammad et al., 2017; Roy et al., 2010), suggesting it might (partially) dissociate or relocate its binding position on the ribosome before the joining event. Finally, displacement of the initiation factors by the eIF5B creates the binding site for the 60S subunit on the interface of the 40S subunit.

eIF1A remains bound to the ribosome to facilitate eIF5B for the large subunit joining (Acker et al., 2006; Acker et al., 2009; Marintchev et al., 2003; Olsen et al., 2003). eIF1A-CTT is released from the decoding center after start codon recognition and interacts with eIF5B-CTD, especially at C-terminal domain IV (Nag et al., 2016). This interaction between eIF1A-CTT and eIF5B-CTD is in competition with eIF5 (Lin et al., 2018). GTP hydrolysis of TC is coupled with the displacement of eIF5 and eIF2·GDP by eIF5B (Lin et al., 2018). Domain IV of eIF5B is then released from eIF5 to associate with eIF1A-CTT, which also tightens its interaction to the 40S

subunit (Fringer et al., 2007). The replacement of eIF1A from eIF5 to eIF5B also couples the first GTP hydrolysis of TC after AUG recognition with the second GTP hydrolysis of eIF5B after large subunit joining.

eIF5B is essential for mediating the large subunit joining but its GTPase activity is only required for the release of eIF1A from the A site and the dissociation of eIF5B·GDP from the 80S ribosome (Pisareva et al., 2007). By using GDPNP, eIF5B and eIF1A are trapped on the 80S ribosome. Besides, eIF5B mutant impaired in GTPase activity retains eIF5B and eIF1A on the 80S ribosome (Fernandez et al., 2013; Fringer et al., 2007; Yamamoto et al., 2014). Releasing eIF5B·GDP and eIF1A from the 80S ribosome leads to the stable formation of an elongation competent 80S initiation complex. GTP hydrolysis of eIF5B favors the formation of the 80S IC and commits the 80S ribosome for the elongation cycle, where the vacant A site is ready to receive the second elongator aminoacyl-tRNA for the peptide bond formation in the peptidyl transferase center of the ribosome.

1.10 Scope of the thesis

The canonical scanning mechanism of translation initiation is the most regulated but least studied process throughout the translation cycle. Genetic and steady-state thermodynamic studies in lower and higher eukaryotes have provided insights into the sophisticated networks between initiation factors and the ribosome. However, due to the difficulty to seize the ribosome at intermediate states and the lack of relevant observables to monitor the scanning process, the kinetic analysis on mammalian translation initiation is largely missing. Many questions, therefore, remain unclear: the order of binding and dissociation of each initiation factor to the 40S subunit; the kinetically favored routes of the initiation complex assembly; the checkpoints that regulate the ribosome to overcome the energy barrier from one intermediate state to another one; the kinetics of mRNA selection by the ribosome; the molecular mechanism of the scanning process; the rate-limiting steps during the translation initiation and the signals to trigger the start codon recognition.

In this study, my goal is to develop an experimental approach to study the dynamics of initiation complexes, identify transient intermediates and measure their thermodynamics and kinetics. I will focus at the early stage in translation initiation that leads to the formation of 48S IC. My first aim is to reconstitute the mammalian translation system *in vitro* with purified components and systematically measure the ribosome dynamic during initiation in real time with fluorescent observables. I have chosen eIF1A to generate fluorescence labeled variants at different positions and visualized the dynamics of eIF1A in the initiation complexes using these reporters. Then, I aimed at using the observed fluorescence changes to determine the association and dissociation rates of eIF1A to the 40S subunit with pre-steady state kinetic analysis by stopped-flow apparatus. Finally, I tried to understand the kinetic checkpoints that drive the ribosome *en route* to the formation of 43S PIC and 48S IC. I systematically analyzed the dissociation rates of eIF1A from the partial initiation complexes and identified the discrete binding states of eIF1A at different stages during initiation. With thorough real-time kinetic analysis, this study will shed light on the dynamic of the 40S subunit during the initiation complex assembly for the first time.

2 RESULTS

To elucidate the molecular mechanism of the translation initiation in higher eukaryotes, I reconstituted *in vitro* mammalian translation system with purified factors. The reconstituted translation system allows me to study the mechanisms and the roles of individual factors in translation by dissecting the intermediate steps of the initiation process and quantifying the thermodynamic and kinetic parameters. Here, I utilized the purified native human proteins and ribosomes from HeLa cell cytoplasmic lysate in combination with the recombinant proteins from *E. coli*. The tRNA_i and mRNA are prepared via the T7-polymerase *in vitro* transcription reaction. The activities of the reconstituted initiation complexes are verified with the toeprinting assay and the peptide-bond formation assay. With the introduction of the site-specific fluorescently labeled eIF1A mutants, I am able to monitor conformational rearrangements of the factor at different stages of initiation and thus identify elemental steps during initiation and study them from the kinetic and thermodynamic perspective. For the first time, the results shed light on the assembly landscape of the mammalian translation initiation complexes from a view point of quantitative kinetic analysis.

2.1 Reconstitution of mammalian translation system

To reconstitute the mammalian translation system, the conventional core factors eIF1, eIF1A, eIF2, eIF3, eIF4A, eIF4B, eIF5B are required, as well as the 40S and 60S subunits, Met-tRNA_i^{Met} and mRNA. eIF1, eIF1A, eIF4A and eIF4B are recombinantly expressed and purified from *E. coli*. The other protein components, which are mostly multi-factor complexes, are isolated from the native source, HeLa cell lysate, in order to retain their activities. **Figure 12** illustrates the overview of the respective preparations.

2.1.1 Components for the reconstituted mammalian translation system

2.1.1.1 Purification of the recombinant eIF1, eIF1A, eIF4A and eIF4B

Single-subunit initiation factors, eIF1, eIF1A, eIF4A and eIF4B, are recombinantly expressed and purified from *E. coli* (**Figure 12A**). The human cDNAs of these factors are inserted into expression vector pET24a. eIF1 and eIF1A are expressed in *E. coli* strains BL21 codon-plus RIPL (Agilent Technologies), while eIF4A and eIF4B are expressed in *E. coli* strain Rosetta. It is important to start each purification with at least 6L of cell culture to ensure a decent yield of each recombinant protein, especially for eIF1A and eIF4B since they are usually expressed relatively poorly in the *E. coli*. Furthermore, when introducing the mutations in eIF1 and eIF1A, it is important to retain two endogenous cysteines in both proteins. Loss of endogenous cysteines in human eIF1 will defect the ability of eIF1 binding to the 40S subunit. Furthermore, even though the two cysteines are substituted with serine and phenylalanine in rabbit eIF1A, replacement of either one with Ser and Phe in human eIF1A leads to low expression of the recombinant protein and its accumulation in the inclusion bodies.

The non-tagged recombinant proteins are isolated based on their differences in isoelectric points. They are purified through a series of consecutive anion and cation exchange columns by fast protein liquid chromatography (FPLC), e.g. Hitrap Q HP column and Hitrap SP HP column, respectively. The Hitrap Q HP column and Hitrap SP HP column not only separate the target proteins from the impurities but also effectively remove the RNase contamination from the lysate. RNase-free preparation is crucial for the reconstituted system to protect all the RNA components and the ribosomes. Small amounts of the RNase contamination might reduce the effective amount of the mRNA, the tRNA and even the ribosomes. The Mono Q 5/50 GL column is used to further remove impurities from the target proteins and also partially concentrate the proteins, which allows to avoid the conventional filter-based concentration step that often significantly reduces the yield. For the eIF4A purification, ADP is optionally added in all purification buffers to stabilize the protein and increase the yield.

Finally, the identity of recombinant proteins was confirmed by both SDS-PAGE and mass spectrometry analysis. eIF1 and eIF1A show clear single bands at 13 and 17 kDa,

respectively, while eIF4A and eIF4B have larger sizes of 46 and 69 kDa, respectively (**Figure 12A**). The activity of the recombinant proteins to promote initiation complex formation is tested in the toeprinting assay and the peptide formation assay.

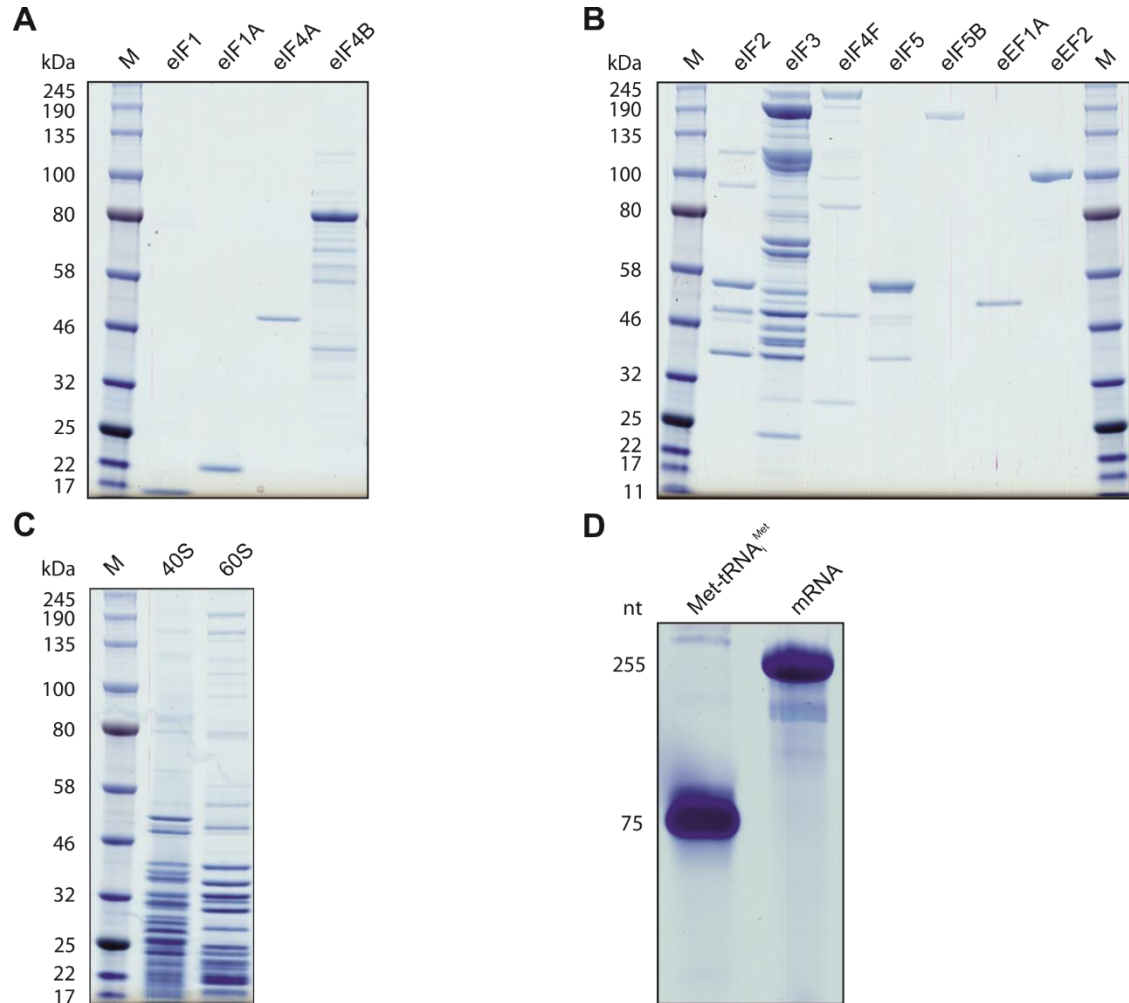


Figure 12. Components for the reconstitution of the mammalian translation systems

(A) Recombinant human proteins, eIF1, eIF1A, eIF4A and eIF4B, from *E. coli*. **(B)** Native initiation factors, eIF2, eIF3, eIF4F, eIF5B, and elongation factors, eEF1A and eEF2, from HeLa cell lysate. Native eIF5 is purified from rabbit reticulocyte lysate. **(C)** Endogenous human ribosomes, 40S and 60S, from HeLa cell lysate. **(D)** *in vitro* transcribed model mRNA and human initiator tRNA^{Met}. The tRNA^{Met} is subsequently aminoacylated with [³H]-labeled methionine.

2.1.1.2 Purification of the native human eIF2, eIF3, eIF4F, and eIF5B

The multisubunit initiation factors and the ribosomes are purified from the HeLa cell cytoplasmic lysate (**Figure 12B**). The nucleus-free HeLa cell lysate is kindly provided by the research group of Prof. Dr. Holger Stark. The polysomes in the cytoplasmic lysate are first isolated by the ultracentrifugation as the starting source for both the initiation factors and the ribosomes. The polysome-bound initiation factors are washed off the ribosomes by elevating the salt concentration in the buffer. The remaining factor-free polysomes are subjected for the ribosome purification. The mixture of the polysome-bound proteins is further precipitated with ammonium sulfate (AS) according to their sizes. The large initiation factors eIF3 and eIF4F precipitate in the 0-40% AS fraction, while eIF2 and eIF5B precipitate in the 40-50% AS fraction. The small native initiation factors, eIF1, eIF1A, eIF4A and eIF4B, remain in the > 50% AS fraction, which can be purified further. The native initiation factors are then isolated from the impurities by a series of anion and cation exchange chromatography steps, i.e. HiTrap Q HP column, the Hitrap SP HP column and the Mono Q 5/50 GL column. Markedly, the eIF3 preparation always carries significant amount of the eIF4F components and other RNA binding proteins due to its strong interactions with them. If the experiment requires contaminant-free eIF3 preparation, subsequent sucrose gradient separation can be applied. However, this polishing step causes dramatic loss of the yield of eIF3. Since the recovery of the eIF3 from the polysome fraction is relatively low, I have tried to purify the eIF3 from the cytosolic polysome-free extract (S100). Even though the yield is improved and eIF3 contains all subunits, it does not stimulate the formation of initiation complexes *in vitro*. The preparation of the cytosolic eIF3 can also contain significant amounts of RNase. Therefore, I only used high quality eIF3 from the polysome-bound fraction in the following experiments. In comparison to eIF3, the yields of eIF2 and eIF5B preparation are relatively high. eIF2 and eIF5B are easily separated from each other on the Mono Q 5/50 GL column. Notably, the eIF2 preparation always co-elutes with ABC50 protein, which also stimulates translation initiation in mammalian cells in both cap-dependent and cap-independent pathways (Paytubi et al., 2009). Even though eIF5B is a single subunit protein, recombinant eIF5B is insoluble and forms inclusion bodies in *E. coli* cells. The recombinantly expressed eIF5B also performs poorly in promoting the 60S subunit

joining in the toeprinting assay. Therefore, I utilized the native eIF5B for the reconstitution of the initiation complexes. Finally, eIF5 is a low abundant protein in the HeLa cell lysate but can be purified with the same procedure from the rabbit reticulocyte lysate (RRL). Since the size of the eIF5 is similar to that of the eIF2 subunits, it is important to identify the correct target fractions during the chromatography purification. The identity of native factors was confirmed by SDS-PAGE and mass spectrometry analysis.

2.1.1.3 Purification of human ribosome

Human ribosomes are purified from factor-free polysomes (**Figure 12C**). First, puromycin is added to release the pre-existing peptides bound to peptidyl-tRNA. Then the 60S subunit dissociates from the 40S subunit by increasing the KCl concentration to 0.5 M. Small and large subunits are separated by centrifugation on a 10-30% sucrose gradient. The gradient is fractionated and the absorbance of each fraction at 260 nm is measured to identify ribosome subunits via their RNA content. To ensure high quality of the ribosome preparation, only 5 - 6 peak fractions of the 40S and 60S subunits are pooled and concentrated thereafter. The fractions that are too close to the separation boundaries are avoided. Mass spectrometry analysis is used to confirm the presence of ribosome subunits in the preparation.

2.1.1.4 Preparation of the model mRNAs

There are two different mRNAs used in this study, the model mRNA (**Figure 12D**) and short 13XCAA model mRNA. The mRNAs are prepared by *in vitro* transcription using T7-polymerase without the conventional supplement of GMP in the reaction mixture to retain the tri-phosphate group at the 5' end of the mRNA for the potential capping reaction. The sequence of the model mRNA starts with triple Gs to increase the transcription efficiency of the T7-polymerase. The 5' UTR consists of 30 CAA triplet repeats to form unstructured single strand RNA region for higher efficiency of the ribosome recruitment. Four restriction sites inserted between the CAA triplets for cloning purpose results in formation of low-energy

secondary structures in the 5' UTR as markers of the positions. The 4th restriction site is adjacent to the Kozak sequence and contains a near-cognate AUC codon in the palindrome sequence. The Kozak sequence is adapted from the 13 nucleotides upstream of the human β -globin ORF, which utilizes -3 A and +4 G in the critical positions. The 5' UTR of the model mRNA is 132 nt in length. The reporter ORF encodes an oligopeptide, Met-Val-Arg-Phe-Lys-Ala, with UGA as stop codon. Near-cognate codon versions of the model mRNA are generated by replacing AUG with near-cognate codons. Finally, the 3' UTR consists of the first 102 nucleotides of the human β -globin ORF. Towards the end of the 3'UTR is the specific binding site of the fluorescently labeled primer for the toeprinting assay. The total length of the model mRNA is 255 nt.

The short 13XCAA model mRNA consists of similar elements as the model mRNA. The major difference is that there are only 13 CAA triplet repeats in the 5' UTR without the restriction sites present in the model mRNA. The total length of the 5' UTR is thus only 56 nt, allowing it to be a potential mRNA template for the solid-state RNA synthesis. The reporter ORF only encodes a tripeptide, Met-Val-Phe, with UGA as the stop codon. The Met-Val (MV) dipeptide and the Met-Val-Phe (MVF) tripeptide synthesized from the reporter ORF have distinct elution positions from the RP-8 HPLC column, allowing for quantitative analysis of the ribosome activities. The 3' UTR of the short 13XCAA model mRNA also utilizes the first 102 nucleotides of the human β -globin ORF. The total length of the short 13XCAA model mRNA is 170 nt.

2.1.1.5 Preparation of human Met-tRNA_i^{Met}

Human Met-tRNA_i^{Met} is also prepared by *in vitro* transcription and subsequently aminoacylated with [³H]-labeled methionine using S100 extract prepared from *E. coli* (**Figure 12D**). One important detail in the preparation of PCR products for the *in vitro* transcription reaction is using less than 50 μ L of the PCR reaction mixture in each PCR tube to ensure the homogenous temperature during the PCR reaction. Furthermore, using fresh and high specificity of [³H]-Met for the aminoacylation can improve the resolution of the peptide

detection in the fractions collected from the HPLC separation, which can further reduce the required amount of the reconstituted initiation complexes for each sample in the peptide formation experiment.

2.1.2 Toeprinting assay

2.1.2.1 Toeprinting assay of 48S IC, 80S IC and 80S EC

To verify the formation and activity of the reconstituted initiation complexes, I performed a fluorescence-based toeprinting (TP) assay to detect the position of the initiation complexes on the mRNA. The toeprinting assay is a variant of the primer extension assay. Generally, a fluorescence-labeled DNA primer is annealed to the 3' UTR of the mRNA and is extended by reverse transcription (RT) using SuperScript III Reverse Transcriptase. The cDNA fragments are then purified by phenol extraction, separated in the 8% UREA-PAGE and imaged with a fluorescent scanner. The lengths of the fluorescence-labeled cDNA fragments indicate the positions of stable initiation complexes, secondary RNA structures which block the reverse-transcriptase progression, or the full-length cDNAs of the model mRNA.

Since the model mRNA has a long unstructured 5' UTR with CAA repeats, the 48S IC can form in the absence of the cap-binding complex eIF4F (Pestova and Kolupaeva, 2002). This allows me to reconstitute the 48S IC using purified native human proteins, eIF2, eIF3, and the 40S subunit, and the recombinant proteins, eIF1, eIF1A, eIF4A and eIF4B, in the presence of the Met-tRNA_i^{Met} and the model mRNA. To form the 80S IC, I included additional eIF5, eIF5B and the 60S subunits into the 48S IC mixture. For the formation of the 80S elongation complexes (80S EC), purified bacterial tRNAs (Val, Arg, Phe, Lys, and Ala) and native human eEF1A and eEF2, kindly provided by Dr. Dmitry Burakovskiy, are added to the 80S IC sample to promote the peptide formation and translocation.

The toeprinting shows that the 48S IC and 80S IC position at the AUG codon of the reporter ORF, while the 80S EC moves towards the 3' end after 5 rounds of peptide bond formations and locates at the stop codon, UGA. The reverse-transcription (RT) reaction also

reveals the secondary structures of the model mRNA, which result probably from the palindromic restriction sites in the 5' UTR and high GC-content in the 3' UTR (**Figure 13**).

The formation of 48S IC on the AUG codon results in a strong band at around 15 nt downstream the AUG position (**Figure 13**) (Pestova and Kolupaeva, 2002). The footprint of the 48S IC is around 60 nts with 30 nts covered by the eIF3 in the 5' UTR and 30 nts covered by the 40S subunit around the AUG codons. Therefore, the 48S IC band locates around 15 nt downstream the AUG position. The intensity of the 48S IC band also correlated with the efficiency of complex formation (Pisarev et al., 2007b). Even though toeprinting is a qualitative assay due to the non-linear readout of the RT reaction, the intensity ratio between the 48S IC band and the F.I. band can still provide an estimation for the efficiency of the complex formation. In the presence of the 48S IC factors, the secondary structure bands in the 5' UTR are also reduced or disappear either by the action of eIF4A, unwinding activity of the scanning ribosome or simply due to the occlusion by the ribosomes and factors. Last, the 40S subunit alone or the 40S-eIF3 complex cannot produce a clear stable band after the RT reaction (data not shown). Only the base-pairing between the codon and anticodon can trigger the stable 48S IC formation and leads to a strong 48S IC band. Any partial initiation complex before start codon recognition is not stable and might simply fall off from the mRNA when encountered by the reverse-transcriptase.

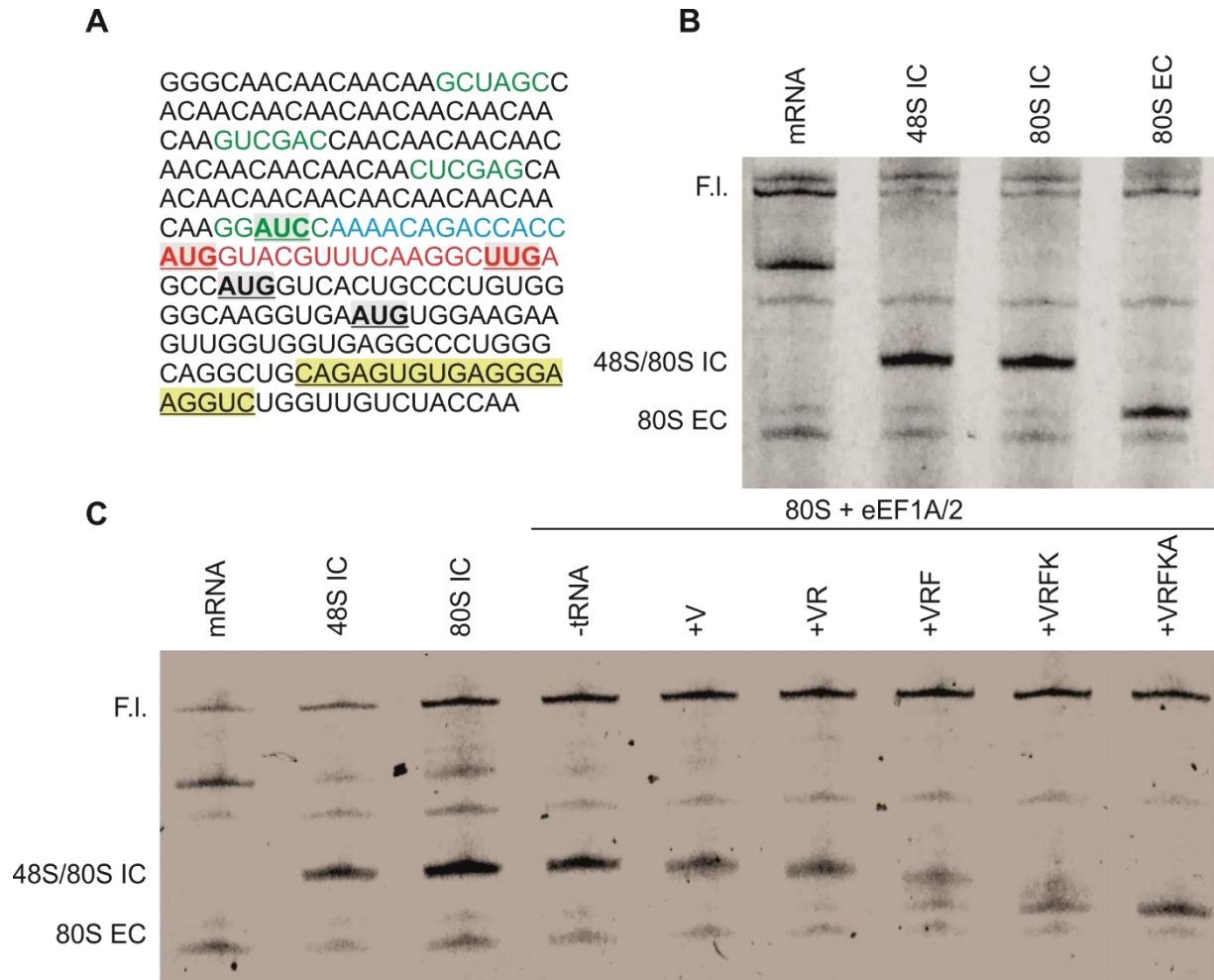


Figure 13. Toeprinting assay of reconstituted the 48S IC, 80S IC and 80S EC

(A) Sequence of the model mRNA. Restriction sites (RS) are green; AUG and near-cognate codons are in bold, underlined and highlighted in grey; Kozak sequence is blue; the reporter open reading frame (ORF) is red; the binding site of the toeprinting primer is underlined and highlighted in yellow. **(B)** The toeprinting assay of the model mRNA without ribosome complex, and with 48S IC, 80S IC and 80S EC. F.I. is the full-length cDNAs from the model mRNA. The cDNA bands resulting from 48S IC, 80S IC and 80S EC are indicated. **(C)** Stepwise elongation of the 80S EC from AUG to the stop codon. The individual charged tRNA is sequentially included into the 80S EC together with eEF1A and eEF2. -tRNA is designated to the 80S IC with eEF1A and eEF2 but without additional aa-tRNAs.

Additions of eIF5, eIF5B and the 60S subunits to the 48S IC sample lead to the formation of the 80S IC. The 80S IC band locates at the same position as the 48S IC band, 15 nts downstream the AUG position (**Figure 13**). The intensity of the 80S IC band is as strong as the

48S IC band, indicating that the large subunit joining does not affect the amount of the initiation complexes on the AUG codon. However, since the positions of the 48S IC and 80S IC are both at the AUG codon, the toeprinting assay cannot distinguish whether the bands are generated from the 48S IC or the 80S IC.

To further verify whether the 80S IC is active *in vitro*, I initiated the elongation reaction by supplying the corresponding aa-tRNAs with native human elongation factors, eEF1A and eEF2. The 80S ribosomes move through the coding region which encodes 5 amino acids, Met-Val-Arg-Phe-Lys-Ala. After the last peptide bond formation between the Lys and the Ala, eEF2 promotes the translocation of the 80S ribosomes and locates the UGA to the A site. Therefore, the cDNA of 80S EC is at around 15 nts downstream the UGA codon (**Figure 13**). The formation of the 80S EC band confirms that all the purified components of the mammalian translation system are active and can be successfully used for the reconstitution of the mammalian 48S IC, 80S IC and 80S EC.

To demonstrate the movement of the 80S EC during elongation, I performed stepwise elongation with the 80S ribosomes by adding the individual aa-tRNA one after another (**Figure 13C**). In the first round of elongation, upon peptide bond formation between Met and Val and the subsequent translocation, the ribosome moves by one codon towards the 3' direction, resulting in a characteristic RT stop. This result further confirms that the movement of the 80S ribosomes on mRNA is due to the actual elongation. The intensity of the cDNA bands from elongation complexes decrease slightly after each round of the peptide bond formation. However, there is only one 80S EC band for each elongation step, indicating that the ribosomes move either to the next codon or drop off from the mRNA.

2.1.2.2 The role of individual eIFs in the initiation complex formation

To identify the minimum essential components of active 48S IC and 80S IC, I tested the initiation complex formation in the absence of individual initiation factors with the toeprinting assay. In the 48S IC positive control lane, the 48S IC band is strong and sharp, indicating a high efficiency of the complex formation when all the factors are present in the reconstitution

reaction (**Figure 14A**). The hairpin structures in the 5' UTR which are visible in the negative control (mRNA only) lane, have also disappeared. In the absence of both eIF1 and eIF1A, two weak bands appear corresponding to the complexes on the AUG codon and possibly also on the AUC codon in the 5' UTR. In the absence of ternary complex (TC) or any component of TC, i.e. eIF2 and Met-tRNA_i^{Met}, the 48S IC band disappears. The mammalian eIF3 is an essential component for the 48S IC formation which is not necessarily required in the yeast reconstituted initiation complexes. Therefore, exclusion of eIF3 disrupts the mammalian 48S IC formation. On the other hand, when I omitted eIF4A and eIF4B in the reconstitution reaction, the 48S IC band is still clearly formed. Since the purified native eIF3 always carries significant amount of eIF4F components, especially eIF4A and eIF4G, it is very difficult to remove eIF4A and eIF4B completely in the reaction. Therefore, the presence of the eIF4A in the eIF3 preparation needs to be taken into consideration when interpreting the result. In conclusion, eIF1, eIF1A, TC, eIF3 and probably eIF4A and eIF4B are essential for the mammalian 48S IC formation. All purified proteins also proved active in the reconstituted system (**Figure 14A**).

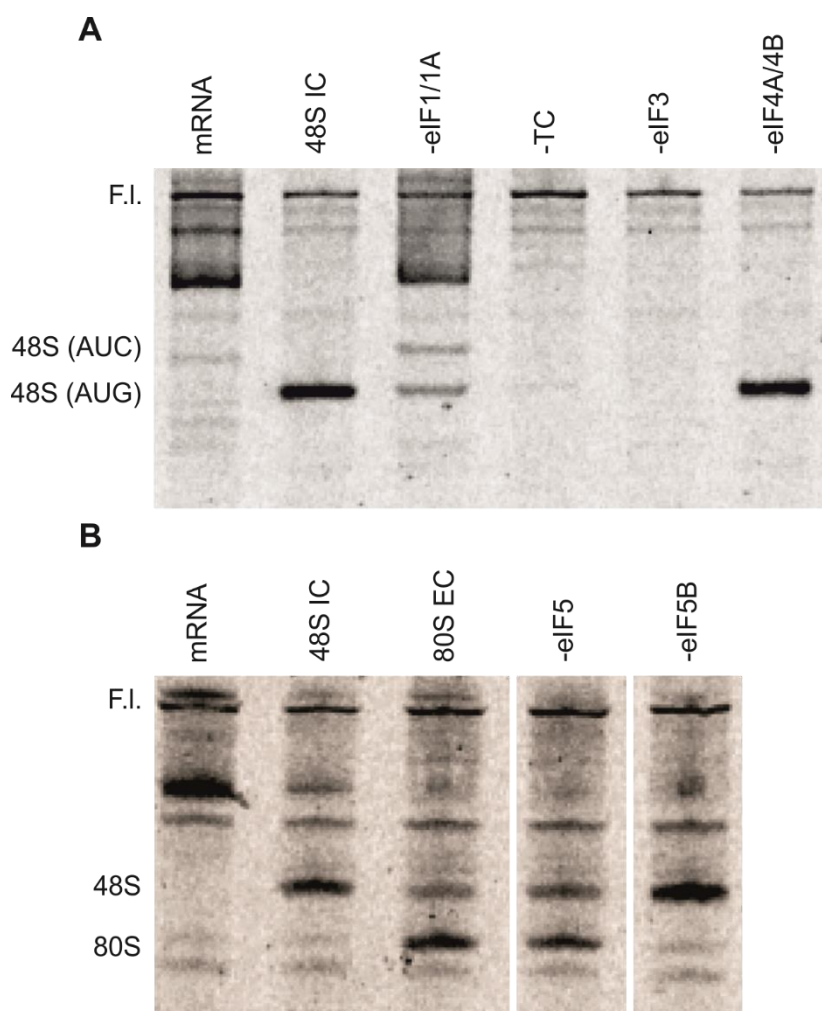


Figure 14. Functions of individual initiation factor in the reconstituted 48S IC and 80S IC

(A) 48S IC formation in the omission of individual initiation factor. F.I. indicates the full-length cDNAs of the model mRNA. 48S (AUC) is the 48S IC located at the AUC in the 5' UTR. 48S (AUG) is the 48S IC located at the AUG in the reporter ORF. **(B)** 48S IC and 80S EC formation in the absence of eIF5 or eIF5B. 80S EC indicates the 80S ribosomes located at the stop codon, UGA.

The functions of eIF5 and eIF5B are tested together with the 60S subunit joining (**Figure 14B**). In this experiment, the ratio of the 60S subunit to the 40S subunit is 1:1. Only part of the 48S IC is converted to the 80S IC and 80S EC. Therefore, both 48S IC and 80S EC bands appear in the same sample and indicate the beginning and end of the reporter ORF. Addition of eIF5 does not increase the portion of the 80S IC/EC bands, indicating that formation of 48S IC and 80S IC is independent of the eIF5 in the equilibrium condition. However, in the absence of

eIF5B, no 80S EC is observed and probably all the ribosomes stay in the 48S IC state on the AUG codon, indicating that eIF5B is essential for the large subunit joining.

In summary, the minimum set of components required for functional 48S IC are eIF1, eIF1A, TC, eIF3, eIF4A, eIF4B, mRNA and the 40S subunit. Addition of eIF5B and the 60S subunit is sufficient to assemble the 80S IC. eIF5 does not influence the 48S IC and the 80S IC formation when studied in the equilibrium condition. eIF1 and eIF1A have functional redundancies during the scanning and start codon recognition. However, exclusion of both factors in the reconstituted system largely decreases the efficiency of the 48S IC formation. Neither eIF2 nor Met-tRNA_i^{Met} alone can promote the formation of the 48S IC. The initiator tRNA must be delivered by the eIF2 in the ternary complexes to the 40S subunit for the initiation complex formation. Mammalian eIF3 is the biggest initiation factor and its composition differs largely from the eIF3 in lower eukaryotes. Even though the yeast 48S IC can be assembled without eIF3, the mammalian 48S IC strictly requires eIF3. eIF4A and eIF4B are important for mRNA recruitment to the ribosomes. In my toeprinting assay, removing recombinant eIF4A and eIF4B has little effect on the efficiency of the 48S IC formation. This is most likely due to the presence of native eIF4A and eIF4B in the eIF3 preparation.

2.1.2.3 48S IC formation on different featured model mRNAs

To better understand the roles of eIF1 and eIF1A in the start codon selection, I compared the 48S IC formation in the presence and absence of eIF1 and eIF1A on mRNAs bearing AUG or CUC in the ORF (**Figure 15**). The CUC codon is a negative control that is not recognized by the ribosomes as translation start site. The model mRNA template utilizes the ORF of human β -globin gene as the 3' UTR sequence, therefore, two extra AUGs are present after the reporter ORF, designated as 2nd and 3rd AUG. The 2nd AUG, which localizes 4 nts after the stop codon of the reporter ORF, is also embedded in the Kozak sequence, upstream -3 G and downstream +4 G. Two near-cognate codons are in the reporter ORF, AUC in 5' UTR and UUG in the ORF. These AUGs and near-cognate codons are recognized differently by 48S ICs that are distinct regarding the presence of eIF1 or eIF1A (**Figure 15B**).

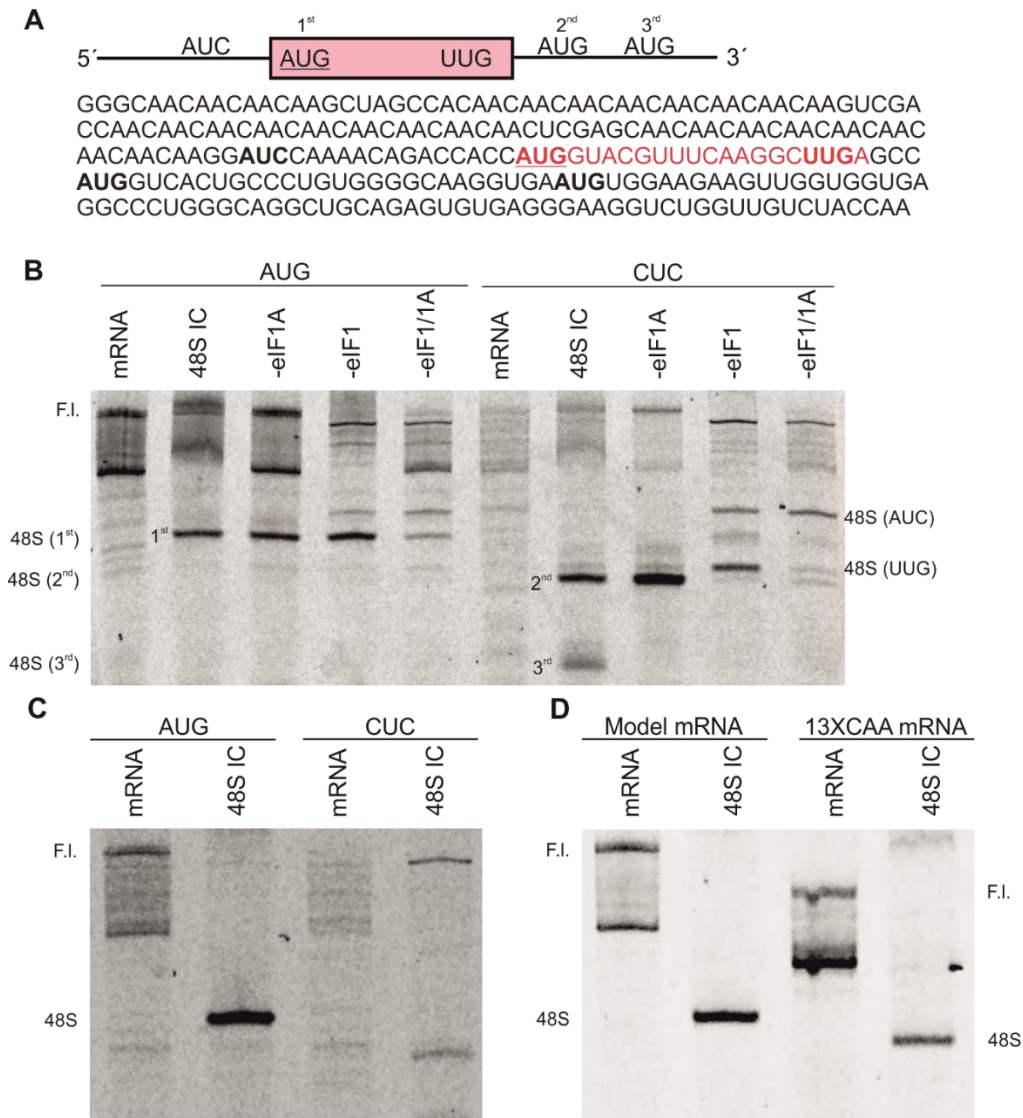


Figure 15. The 48S IC formation on the model AUG/CUC mRNA and short 13XCAA model mRNA

(A) Schematic and sequence of the model mRNA. The ORF is marked in red. The AUG underlined is replaced to CUC for the model CUC mRNA. AUGs and near-cognate codons, AUC and UUG, are bold. **(B)** 48S IC formation on the model AUG/CUC mRNA in the absence of eIF1 and/or eIF1A. 48S IC (1st) indicates the 48S IC assembled on the AUG in the ORF, while 48S IC (2nd) and 48S IC (3rd) indicate the 48S IC that form at the 2nd and 3rd AUG codon in 3' UTR. 48S (AUC) and 48S (UUG) indicate the 48S IC forming on the AUC in the 5' UTR and the UUG in the ORF. **(C)** The 48S IC formation on the model AUG/CUC mRNA without the 2nd and 3rd AUG in the 3' UTR. **(D)** The 48S IC formation on the short 13XCAA model mRNA. The positions of the full length of the short 13XCAA model mRNA and the 48S IC are indicated on the right panel.

eIF1 is the main fidelity regulator and eIF1A promotes the scanning mode of the 48S IC. Since the AUG in the reporter ORF is in the optimal Kozak sequence, the 48S IC easily selects this AUG even in the presence of either eIF1 or eIF1A alone. However, in the absence of eIF1, small portion of the ribosomes is captured at the AUC position. In the absence of both eIF1 and eIF1A, very few 48S IC is formed on the AUG and part of the ribosomes also stops at the AUC in the 5' UTR. Notably, the 2nd and 3rd AUGs are not selected under any condition, supporting the first AUG rule in the scanning model. The 1st AUG on the mRNA is in an optimal Kozak context and therefore, all ribosomes are seized on this start codon as soon as they reach it. On the CUC mRNA, the positions of the 48S IC are largely determined by eIF1 and eIF1A. In the presence of both eIF1 and eIF1A, no 48S IC formation occurs on the CUC codon, but the 2nd and 3rd AUGs are selected with a preference on the 2nd AUG position. In the absence of eIF1A, only the 2nd AUG is selected. Without eIF1, 48S IC forms at AUC and UUG codons located upstream to the 2nd AUG or the 3rd AUG. In the absence of eIF1 and eIF1A, hardly any 48S IC is detected, just a few on the AUC codon. In conclusion, the order of codons, the strength of the Kozak sequence and the presence of eIF1 and eIF1A altogether determine the selection of the start codons by the ribosomes.

To compare the initiation complex formation on the mRNA with only one start codon, I replaced the 2nd and 3rd AUG with CUC and tested the 48S IC formation on the corresponding AUG and CUC mRNAs (**Figure 15C**). The assembly of 48S IC on the AUG mRNA remains highly efficient. On the CUC mRNA, no toeprint of 48S IC is detected. The weak band in the lower part of the lane results probably from RNA secondary structures in the 3' UTR. After mutagenesis, the GC content in the 3' UTR is increased; therefore, the efficiency of the RT reaction on the CUC mRNA is reduced. The results clearly indicate that the 48S IC can select the AUG codon and reject the CUC codon as assembly site on the mRNA in the presence of eIF1 and eIF1A.

I further designed a simpler mRNA with only 13 CAA triplet repeats and optimal Kozak sequence in the 5' UTR following by a tri-peptide (MVF) reporter ORF and the identical 3' UTR as the model mRNA (**Figure 15D**). RT of the shorter 13XCAA mRNA results in a shorter full-length cDNA product compared to the model mRNA. The 48S IC formation is as efficient as on the model mRNAs, indicating that the efficiency of initiation complex formation reaches

optimal on both model mRNAs. This result is also consistent with the recent ribosome profiling finding that the 5' UTR length (which is positively correlated with the possibility of natural mRNA secondary structures in the 5' UTR) does not influence the processivity of 43S PIC but rather the presences of uORFs reduce globally the translation of downstream coding sequence (CDS) (Giess et al., 2020). Therefore, both model mRNAs are suitable for the reconstitution of active initiation complexes.

In order to study the usage of near-cognate codons by the ribosomes, I replaced the 1st AUG codon in the reporter ORF to near-cognate codons differing in the first or third positions to the AUG and tested 48S IC formation on these near-cognate codon mRNAs. Based on the selection frequency of these codons by the ribosomes, near-cognate codons are divided into strong and medium near-cognate codons. Strong near-cognate codons are AUC, UUG and GUG, which are selected by the 48S IC even in the presence of eIF1 (**Figure 16**). Medium near-cognate codons are AUU, AUA and CUG, which are excluded in the presence of eIF1 but can be selected by the ribosomes in the absence of eIF1. The model CUC mRNA, as mentioned above, is the negative control and CUC codon is not used as start site in all conditions. The near-cognate codons in the ORF, which is in an optimal Kozak context, are selected with higher chances by the ribosomes than those in the 5' UTR or 3' UTR. However, the near-cognate codons are not attractive enough to seize all the scanning ribosomes. The 2nd AUG and 3rd AUG are partially selected in almost all the near-cognate mRNAs except the AUU and GUG mRNAs by the 48S IC. In the absence of eIF1A, the 2nd AUG is preferentially selected by the ribosomes over the 3rd AUG, leading to stronger intensities of the 2nd AUG bands on AUC, AUA, CUG and CUC mRNAs. This finding agrees with the function of eIF1A in promoting the scanning mode of the ribosomes. In the absence of eIF1, all the near-cognate codons in the ORF become dominantly selected by the ribosomes with some of the ribosomes seized earlier on the AUC or later on the UUG. Without both eIF1 and eIF1A, the efficiency of complex formation decreases largely, and the ribosomes are detected mostly on the AUC and less on the targeted near-cognate codons in the ORF. Overall, the combination of the strength of the codon, the Kozak sequence and the presence of eIF1 and eIF1A determines the degree of usage of AUG or the near-cognate codons on the model mRNA.

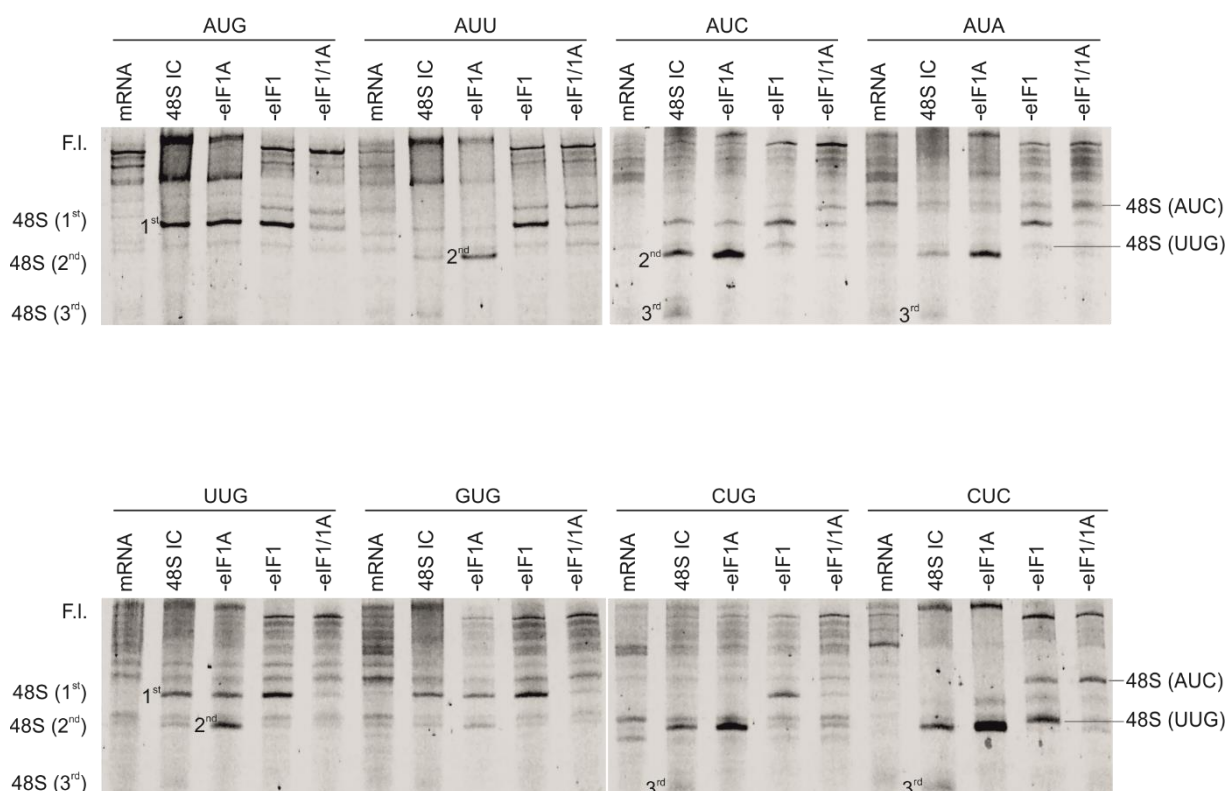


Figure 16. 48S IC formation on mRNAs bearing near-cognate start codons

48S IC is formed on the near-cognate codon mRNAs in the absence of eIF1 and/or eIF1A. The AUG in the ORF is replaced with AUU, AUC, AUA, UUG, GUG, CUG and CUC in the model mRNA. F.I. indicates the full-length cDNAs of the model mRNA. 48S IC (1st) marks the 48S IC assembled on the AUG or near-cognate codons in the reporter ORF, while the 48S (2nd) and 48S (3rd) indicate the 48S IC forming at the 2nd and 3rd AUG in 3' UTR. 48S (AUC) and 48S (UUG) indicate the 48S IC forming on the AUC in 5' UTR and the UUG in the ORF.

2.1.2.4 Functional tests of different translation components used in this study

Toeprinting is also a functional assay to test the activity of individual components. By introducing an unknown factor in the validated reconstitution system, I can test if the unknown factor can promote or affect the 48S IC formation with the toeprinting assay.

To fluorescently label different parts of the eIF1A with maleimide derivative fluorophores, I introduced cysteine residues in the N-terminal tail (NTT), OB domain and C-

terminal tail (CTT) of human eIF1A individually at positions N4, S74, T120, D125 and D142 (**Figure 4**). The eIF1A mutants were labeled with Alexa555 dye and used as observables to monitor the fluorescence changes during the formation of 48S IC and 80S IC in the later fluorescent studies.

To verify the activity of the eIF1A cysteine variants in promoting initiation complex formation, I replaced the WT eIF1A with eIF1A mutants in the absence of eIF1 (**Figure 17**). Since eIF1 and eIF1A have overlapping functions, the activity of eIF1A in promoting 48S IC formation only becomes obvious in the absence of eIF1. Fortunately, all eIF1A mutants support the 48S IC formation and rescue the function of eIF1 (**Figure 17**).

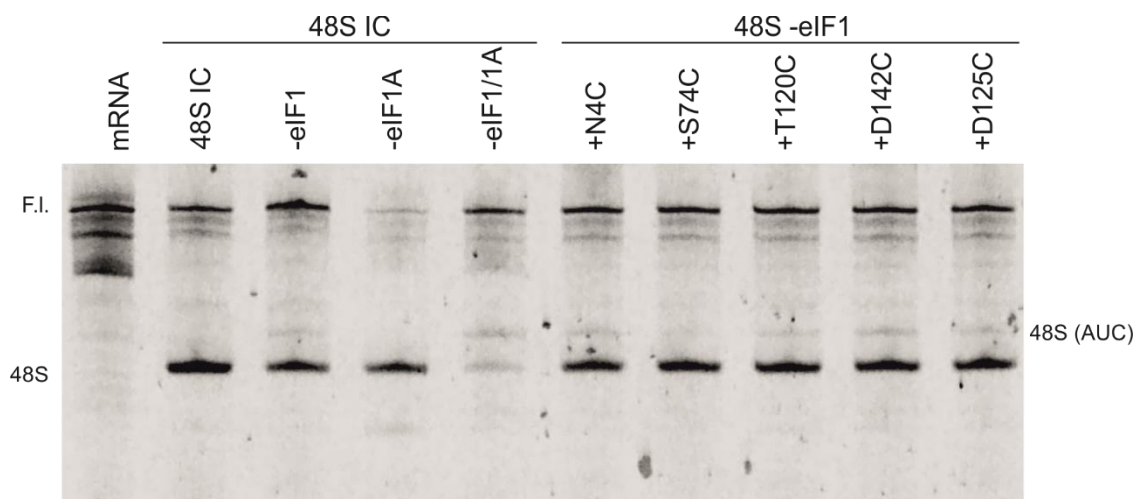


Figure 17. Toeprinting assay using different eIF1A variants

48S IC is reconstituted with eIF1A variants, N4C, S74C, T120C, D142C and D125C, in the absence of eIF1 on the model mRNA. F.I. indicates the full-length cDNAs of the model mRNA. 48S marks the 48S IC forms on the AUG in the ORF. 48S (AUC) indicate the 48S IC form on the AUC codon.

To investigate whether different derivatives of Met-tRNA^{Met} can integrate into 48S IC, I performed toeprinting analysis comparing *in vitro* transcribed human Met-tRNA^{Met} with different variants of bacterial fMet-tRNA^{fMet}, namely fMet-tRNA^{fMet}, Bpy-Met-tRNA^{fMet}, Cy3-Met-tRNA^{fMet}, and Cy5-Met-tRNA^{fMet} (**Figure 18A**). In the absence of any tRNA, only the cDNA of the full-length mRNA is detected, implying that the tRNA is one of the essential components for the 48S IC formation. Interestingly, the human eIF2 can also utilize native bacterial initiator

tRNA for the formation of the human 48S IC independent on the modifications of the fluorescent probes attached to it. All fluorescent analogues of fMet-tRNA^{fMet} lead to stable formation of 48S IC. Cy5-Met-tRNA^{fMet} has a similar excitation wavelength as the fluorescent toeprinting primer; therefore, the lane shows a high fluorescence background.

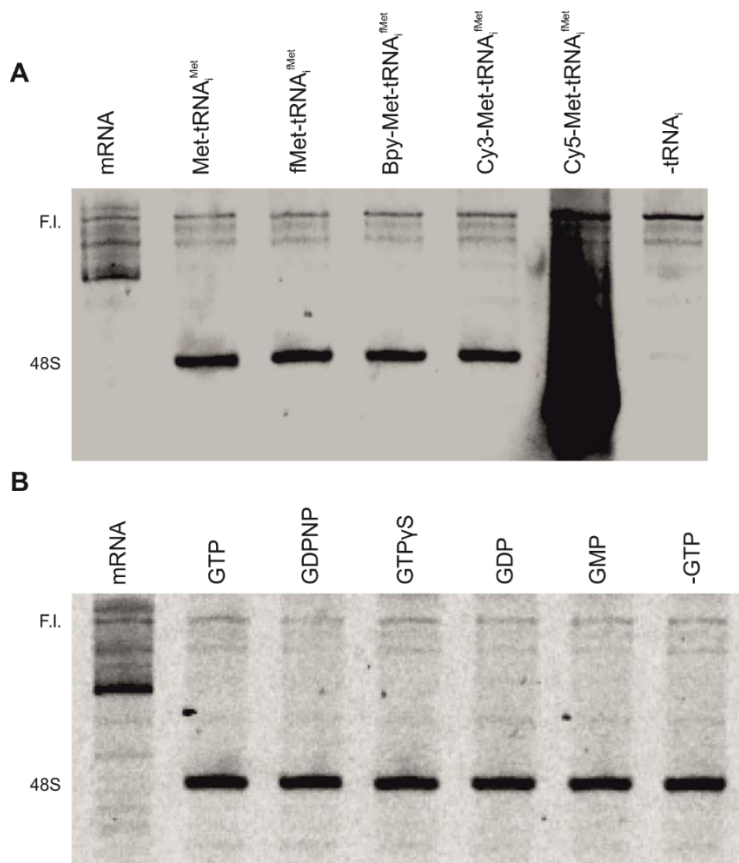


Figure 18. Toeprinting assay using initiator tRNA and GTP analogs

(A) 48S IC formation with variants of the bacterial fMet-tRNA^{fMet} and fluorescently labeled Met-tRNA^{fMet}. The *in vitro* transcribed human initiator tRNA is replaced with native bacterial initiator tRNA and fluorescently labeled initiator tRNA in the 48S IC formation. **(B)** 48S IC formation in the presence of GTP analogs, GDPNP, GTPγS, GDP, GMP. GTP is substituted with GTP analogs or eliminated from the reaction buffer. F.I. indicates the full-length cDNAs of the model mRNA. The 48S marks the 48S IC forms on the AUG in the reporter ORF.

Furthermore, I examined if GTP hydrolysis is essential for the codon-anticodon base pairing and whether the 48S IC can form in the presence of GTP analogs. I replaced the GTP in the reaction buffer with GTP analogs, GDPNP, GTPγS, GDP, and GMP, and allowed the

exchange of eIF2-bound nucleotide in the presence of Met-tRNA_i^{Met} at 37 °C for 10 min prior to the 48S IC formation (**Figure 18B**). In the negative control (-GTP) sample, the GTP was removed completely from the reaction buffer and the complexes were formed in the absence of any additional GTP nucleotide. Surprisingly, RT stop characteristic for the 48S IC is detected even in the absence of additional GTP. Since eIF2 has higher affinity to the GDP than GTP in the absence of tRNA, the native eIF2 could remain in the GDP-bound state when it is purified. Therefore, the native eIF2 might form ternary complex with its pre-bound GDP and still carry the initiator tRNA to the ribosomes. Interestingly, the formation of the 48S IC in the presence of GDP indicates that GTP hydrolysis is not required for 48S IC formation. Similarly, in the presences of GDPNP or GTPγS, the formation of 48S IC occurs to similar extent as with GTP. In conclusion, the 48S IC can be assembled with GTP or its analogs.

2.1.3 Peptide bond formation

The activity test of the purified ribosomes is crucial to verify the quality of each purification process. The toeprinting assay is a fast and easy way to qualitatively test the activity of the ribosomes and other components. However, the direct and quantitative measurement of the ribosome activity relies on the peptide bond formation. According to the stepwise elongation toeprinting results, the amount of the ribosomes on the mRNA is decreased after introducing extra aa-tRNAs and the subsequent translocation steps. In order to quantify the ribosome activity precisely, I measured the dipeptide, Met-Val (MV), formation by the ribosomes in the presence of Val-tRNA and eEF1A but without eEF2, avoiding the translocation step. MV dipeptide is separated from free aa-tRNA by reversed-phase (RP) HPLC. The efficiency of the MV dipeptide formation is around 50% to 80% relative to the amount of 80S ribosomes (**Figure 19A**). I also compared the activity of the initiation complexes constituted with WT eIF1A, eIF1A-N4C and eIF1A-N4C-Alexa555 or without eIF1A, which shows similar efficiency in all tested conditions (**Figure 19B**), in agreement with the toeprinting assay (**Figure 17**). Therefore, eIF1A-N4C-Alexa555 will be a potential observable to monitor the initiation complex formation.

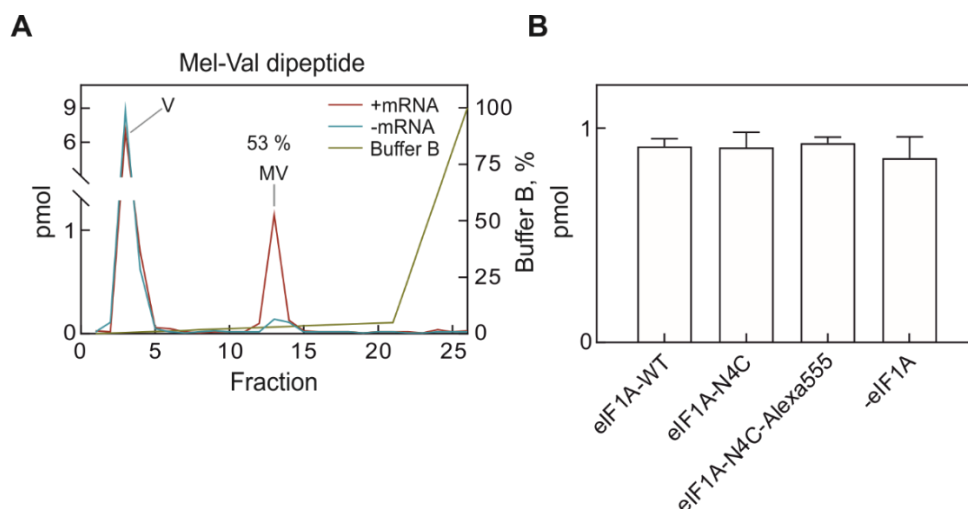


Figure 19. Met-Val dipeptide formation by 80S IC

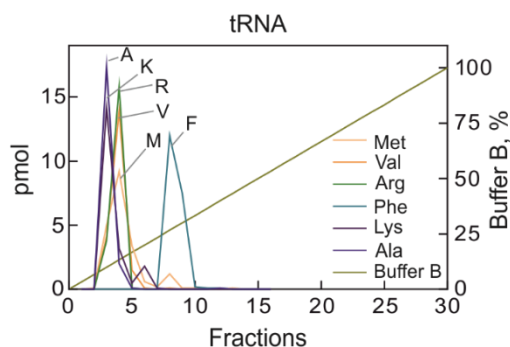
(A) Elution profile of Met-Val dipeptide on the RP-8 column in the HPLC system. The dipeptide is formed by the 80S IC with additions of Val-tRNA and eEF1A into the reconstitution reaction. The experiment without mRNA is a negative control, testing for non-specific peptide bond formation. The efficiency of the dipeptide formation is calculated by dividing the amount of MV by the amount of total 80S ribosomes. The efficiency of peptide bond formation in this experiment is 53%. The gradient of the elution profile is shown as yellow line. **(B)** Formation of dipeptide by 80S IC reconstituted with eIF1A mutants. The 80S ICs are reconstituted with eIF1A mutants or its fluorescently labeled form. The amounts of dipeptides from these 80S ICs are quantified.

To confirm the ability of the ribosomes to promote elongation, I tested the translation of an oligopeptide, Met-Val-Arg-Phe-Lys-Ala, using RP HPLC. The individual tRNAs are charged with [^{14}C]-labeled amino acids. Hence, the elution profiles of individual amino acid and the oligopeptides can be quantified via measuring the ^{14}C radioactivity. The polar single amino acids are eluted from the column in the first ten fractions of the gradient (**Figure 20A**), whereas the oligopeptides are retained longer on the column and eluted usually from fraction 13th to 15th (**Figure 20B**). Around 22% of total 80S ribosomes produces oligopeptide in this assay.

To examine the influence of different buffers and salts on initiation complex formation, I tested oligopeptide formation in Hepes buffer supplemented with KOAc and MgOAc and Tris-HCl buffer with KCl and MgCl_2 (**Figure 20B**). In conclusion, the oligopeptide formation occurs

with similar efficiency in both buffer systems. Therefore, I have chosen Hepes buffer system for further experiments as it is less sensitive to pH changes due to temperature variations.

A



B

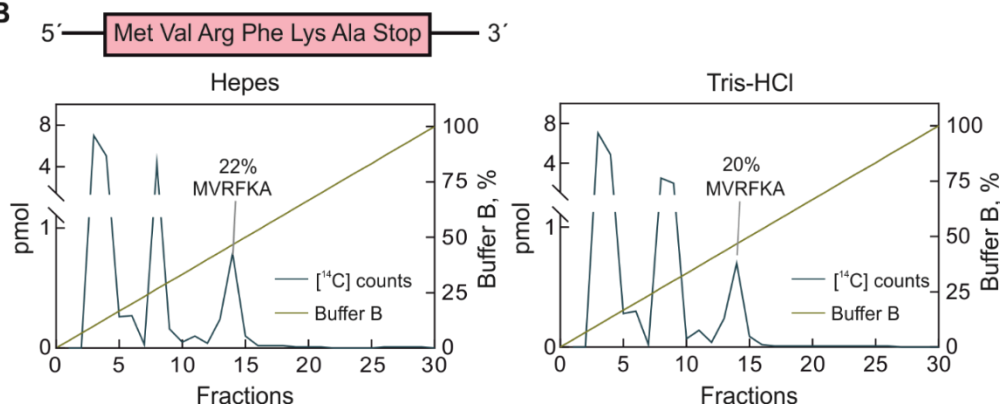


Figure 20. Oligopeptide formation by 80S IC

(A) Elution profile of each amino acid from the RP-8 column in HPLC system. The pmol of each amino acid is calculated from the radioactivity of [³H] or [¹⁴C]-labeled Met, Ala, Lys, Arg, Val, and Phe. The elution gradient is shown in yellow line with the percentage of buffer B. The radioactivity of each 1 mL elution fraction is analyzed. **(B)** Elution profile of the MVRFK oligopeptide from the RP-8 column in the HPLC system. The oligopeptide is extracted from the 80S ribosomes. The amount of oligopeptide is calculated from the radioactivity of the [¹⁴C]-labeled Ala, Lys, Arg, Val, and Phe. The efficiency of the oligopeptide formation is calculated by dividing the amount of detected oligopeptides by the amount of total 80S ribosomes. Schematic of the ORF is shown. Left panel, translation in Hepes with KOAc and Mg(OAc)₂ buffer; right panel, same in Tris-HCl with KCl and MgCl₂ buffer; both at pH 7.5.

I next defined the optimal salt concentration for the IC reconstitution. The MV dipeptide assay was used to measure the tolerance of the initiation complexes to the varying salt concentrations. Potassium is the major monovalent salt that maintains the ionic strength

in the buffer and has direct influence on the stability of the purified proteins. The cellular concentration of K^+ is around 60 – 150 mM in a stress-free condition. It is crucial to store the ribosomes in the buffer containing minimum of 50 mM K^+ while other purified factors are maintained in the buffer with 100 mM K^+ . The reconstituted initiation complexes are relatively stable in the salt range between 65 to 125 mM KOAc with the optimum at 95 mM (**Figure 21A**). It is, therefore, important to keep the temporal elevation of the salt concentration during reconstitution of complexes within the tolerance range to ensure the stability of the ribosomes and translation factors.

Mg^{2+} is the most prevalent divalent cation that influences the ribosome activity due to the ionic stabilization for the negatively charged rRNAs. The total Mg^{2+} concentration in the mammalian cells is around 20 mM (Romani, 2011), while the free Mg^{2+} concentration is in the range of 0.5 to 1 mM (Shenvi et al., 2005). Therefore, I tested the MV dipeptide formation at the free MgOAc concentration ranging from 0 mM to 25 mM. The optimal free MgOAc concentration for the initiation complex formation lies between 2.5 mM and 3.75 mM (**Figure 21B**), well within the physiological range. This finding also proves that the MV peptide formation by the reconstituted IC is not an artificial effect caused by high Mg^{2+} concentration.

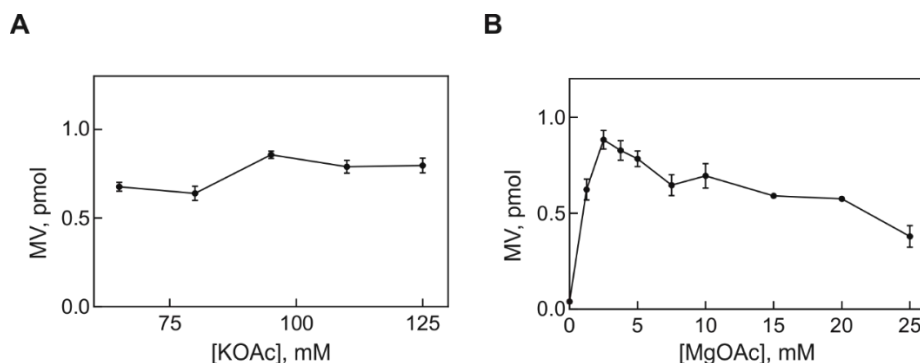


Figure 21. MV dipeptide formation at different concentrations of KOAc and MgOAc

(A) MV dipeptide formation at different KOAc concentrations from 65 to 125 mM. **(B)** MV dipeptide formation at different MgOAc concentrations from 0 to 25 mM.

2.2 Studies of mRNA binding to eIF3 and the 40S subunit by steady-state fluorescence assays

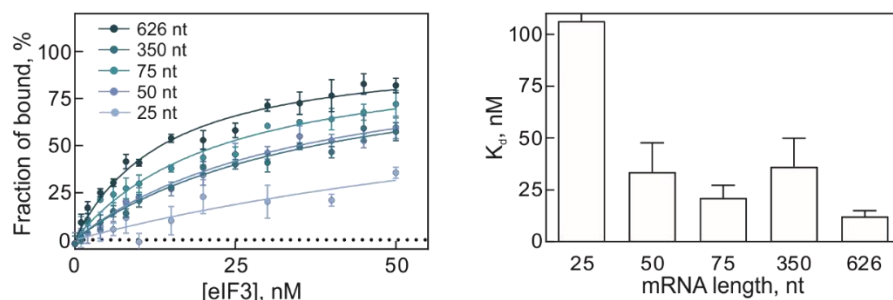
2.2.1 mRNA binding to eIF3

Recently discovered interactions between eIF3 and mRNA reveal new regulatory functions of eIF3 during translation initiation besides its scaffold protein role (Cate, 2017; Lee et al., 2015; Lee et al., 2016). Binding of eIF3 to the mRNA may provide an alternative route to recruit the initiation complexes to certain growth-related mRNAs (Lee et al., 2015). eIF3 in the context of the 48S IC forms an extension outside the mRNA exit channel and stabilizes the 5' UTR during scanning and start-codon recognition (Pisarev et al., 2008). To better understanding these interactions, I measured the affinity of the eIF3 to mRNA transcripts of different lengths. I utilized human β -globin gene as template for *in vitro* transcription and created a library of mRNAs ranging from 25 nt to 626 nt (full-length). The mRNA transcripts were labeled with fluorescein 5-thiosemicarbazide at the 3' end and used as a fluorescence observable to study the interaction between the mRNA and eIF3.

The binding affinity (K_d) of eIF3 to the mRNA fragments is around 25 nM for mRNA lengths between 50 nt to 626 nt, but the K_d decreases dramatically when the mRNA transcript is 25 nt long (**Figure 22A**). It seems that the binding of eIF3 to the mRNA requires a minimum footprint of about 25 - 50 nt. Next, I confirmed the interaction of eIF3 and mRNA using fluorescent-based electrophoretic mobility shift assay (EMSA). Since the native eIF3 is too bulky to migrate in conventional native 6% PAGE system, I utilized the 0.7% agarose gel electrophoresis. Fluorescein-labeled mRNA was incubated with different amounts of eIF3 to form the eIF3-mRNA complexes. Free mRNAs were then separated from bound mRNA by the electrophoresis. Since the detection limit of the labeled mRNA, 50 nM, is already above the estimated binding affinity, the calculation of the binding affinity from the EMSA result would not be accurate. However, the EMSA verifies co-migration of the mRNA with eIF3 in the agarose gel (**Figure 22B**). Surprisingly, there is an extra band of mRNAs which migration on gel is separate but dependent from eIF3. Mass-spectrometry (MS) analysis shows that these mRNAs are bound to single eIF3 subunits, eIF4A, RNA helicases and other RNA binding proteins

which were co-purified together with eIF3 during the preparation. In conclusion, these experiments show that a long mRNA can bind with high affinity to intact eIF3 alone or the subunits of eIF3 as well as to the copurifying translation components.

A



B

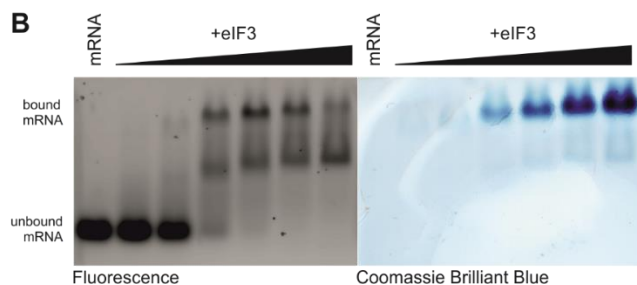


Figure 22. mRNA binding to eIF3

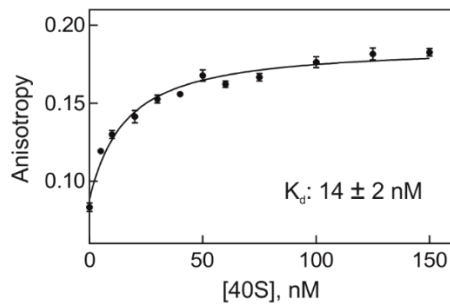
(A) Binding of eIF3 to transcripts of different lengths. The fluorescein-labeled mRNA transcripts with lengths from 626 to 25 nt are generated from the human β -globin ORF by shortening the 3' end. The binding of the mRNA to eIF3 is monitored in the fluorometer with increasing concentrations of eIF3. The fraction of bound mRNA is analyzed with the quadratic equation to obtain the binding affinity (K_d). **(B)** The electrophoretic mobility shift assay (EMSA) of eIF3-bound fluorescein mRNA. Increasing concentrations of eIF3 with the fluorescein mRNA are separated on the 0.7% agarose gel. The localization of the fluorescein labeled mRNA is visualized with the Typhoon fluorescent scanner, while the position of eIF3 is analyzed with the conventional Coomassie brilliant blue staining.

2.2.2 mRNA binding to the 40S subunit

To elucidate the ability of 40S subunit to bind mRNA, I determined the K_d of the 40S subunit to the 3'-fluorescein-labeled mRNAs. I titrated mRNA with increasing 40S subunit

concentrations in a fluorometer and detected anisotropy changes of the reporter. The K_d is around 14 ± 2 nM (**Figure 23A**), similar to the published value of 11 nM (Sokabe and Fraser, 2017). The EMSA also shows the co-migration of the mRNA with the 40S subunit (**Figure 23B**), indicating the direct binding of the mRNA to the 40S subunit in the absence of initiation factors. Even though this interaction is not stable enough to resist the reverse-transcriptase that generates a cDNA band in the toeprinting assay (data not shown), it is stable enough to migrate through the agarose gel during electrophoresis. However, binding of fluorescein-labeled mRNA to the 40S subunit does not lead to any change in the fluorescence intensity, which is, unfortunately, difficult to employ in the fast-kinetic study.

A



B

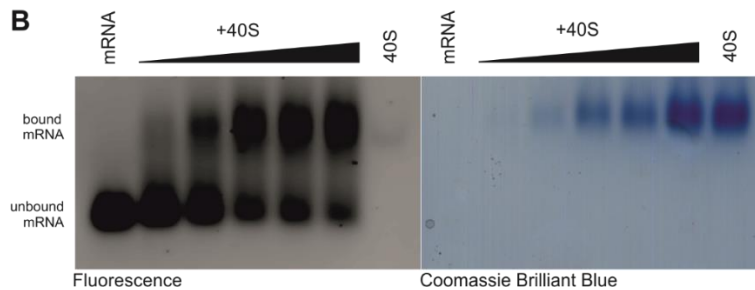


Figure 23. mRNA binding to the 40S subunit

(A) Binding affinity of the 40S subunit to the fluorescein-labeled model mRNA. Anisotropy of the fluorescein-labeled mRNA is monitored with increasing concentrations of the 40S subunit in the fluorometer. The binding affinity is calculated by fitting the resulting curve with the quadratic equation.

(B) Electrophoretic mobility shift assay (EMSA) of the 40S-bound fluorescein mRNA. Increasing amounts of the 40S subunits are incubated with the fluorescein-labeled mRNA and separated in the 0.7% agarose gel. The fluorescein-labeled mRNA is detected with the fluorescence scanner, while the 40S subunit is detected with Coomassie brilliant blue.

To confirm that the measured interaction between mRNA and 40S subunit is not an artefact caused by the labeling position or the chosen fluorescent dye, I purchased a custom synthesized 13XCAA mRNA labeled with Alexa488 dye on the deoxyuridine at the -42nd position upstream the AUG, which is 15 nt downstream the 5' end (mRNA-Alexa488). The anisotropy assay reveals that the mRNA-Alexa488 has a higher affinity (3.2 ± 0.6 nM) to the 40S subunit than that of the fluorescein-labeled mRNA (14 ± 2 nM). The presence of eIF1 and eIF1A slightly increases the affinity to 1.3 ± 0.3 nM (**Figure 24A**). Binding of the mRNA-Alexa488 to 40S subunit leads to a significant decrease in the fluorescence intensity. This allows me to study rapid binding kinetic of the mRNA-Alexa488 to the 40S subunit in a stopped-flow apparatus. The fluorescence decrease upon binding yields a triphasic curve with around 4% amplitude change in total (**Figure 24B**). To measure the dissociation rate of the mRNA from the 40S subunit, I mixed the 40S-mRNA-Alexa488 complexes with excess amount of the unlabeled mRNA. Unfortunately, the fluorescence change is too small to obtain meaningful information (**Figure 24B**).

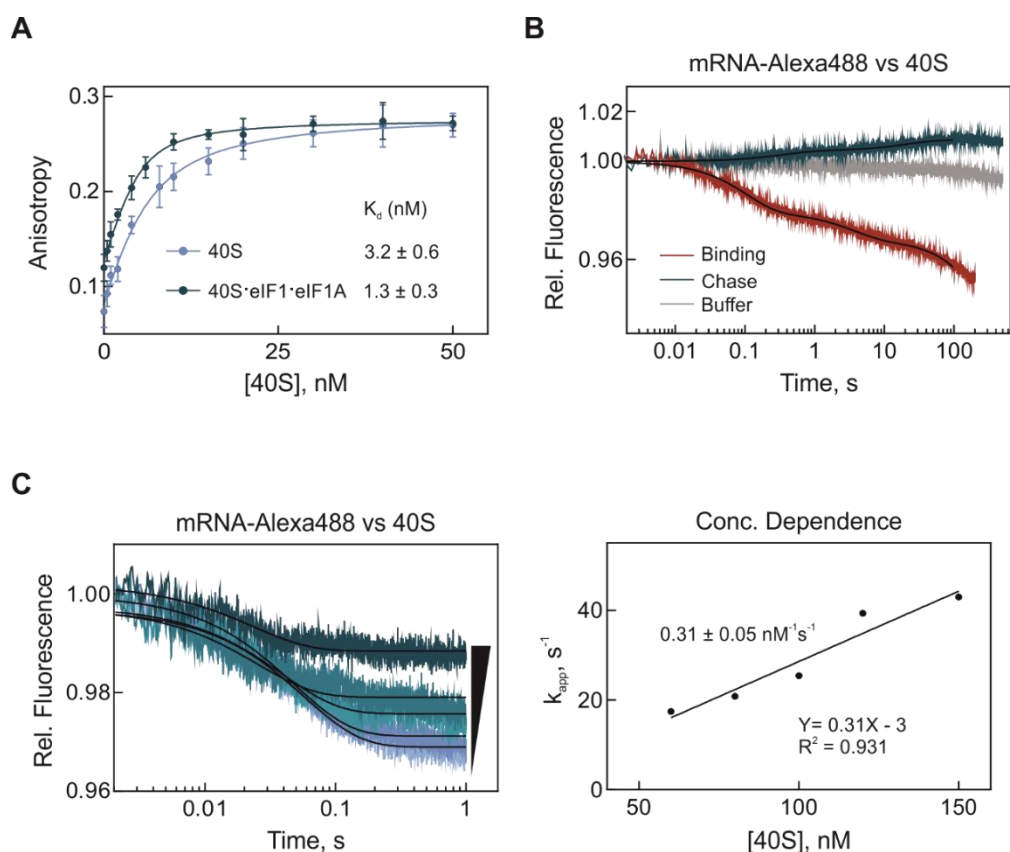


Figure 24. Binding kinetics of the 40S subunit to the Alexa488-labeled mRNA

(A) The binding affinity of the 40S subunit to the Alexa488-labeled mRNA in the absence or presence of eIF1 and eIF1A. **(B)** Real-time monitoring of the binding and dissociation of the 40S subunit to the Alexa488-labeled mRNA in the stopped-flow apparatus. The binding reaction is shown in red and analyzed with three-exponential function. The chase experiment is shown in blue and analyzed with two-exponential function. The buffer control is shown in grey. **(C)** Pseudo-first order titration of the 40S subunits to the Alexa488-labeled mRNA. Increasing amounts of the 40S subunits are mixed with the Alexa488-labeled mRNA in the stopped-flow apparatus and monitored in the fast reaction time range. The binding traces are analyzed with single-exponential function. The apparent rate constants are plotted against its corresponding concentrations of the 40S subunits to obtain the association rate constant (k_{on}), $0.31 \pm 0.05 \text{ nM}^{-1} \cdot \text{s}^{-1}$.

To understand the meaning of the first phase in the fluorescence change of the binding reaction between mRNA-Alexa488 and the 40S subunit, I applied the pseudo-first order reaction assumption and monitored the binding with increasing concentrations of the 40S subunit within 1 second range. Higher concentrations of the 40S subunits result in less change

of the fluorescence intensity due to the inner filter effect of the 40S subunit. All obtained curves are analyzed with one-exponential fitting and resulting apparent rate constants (k_{app}) show a linear dependence on the 40S concentration, indicating that the first phase of binding trace on **Figure 24B** represents a bi-molecular association reaction. The slope of the linear regression, which represents the k_{on} , is $0.31 \pm 0.05 \text{ nM}^{-1} \cdot \text{s}^{-1}$ (also $310 \pm 50 \text{ } \mu\text{M}^{-1} \cdot \text{s}^{-1}$) (**Figure 24C**). In conclusion, the experiments with two differently-labeled mRNAs show that mRNA can bind to the 40S subunit rapidly and with high affinity even in the absence of the initiation factors. To analyze the interaction of the 40S subunit with initiation factors, I next employed fluorescence-labeled eIF1A variants and characterized their interactions with the 40S subunit during initiation complex formation with both fluorescence-based equilibrium and rapid kinetic methods.

2.3 Studies of 80S IC formation by equilibrium fluorescence assays

2.3.1 Binding affinity of eIF1A variants to the 40S subunit

I chose eIF1A as a reporter for kinetic studies because, according to the literature, eIF1A stays on the 40S subunit until 80S ribosome forms and it also participates in most steps during initiation. Additionally, eIF1A can be recombinantly expressed in *E. coli*, allowing me to introduce site-specific fluorophore at different regions of the eIF1A protein. Importantly, eIF1A binds to the A site on the 40S subunit, which provides the opportunity to monitor the start-codon recognition event in a close vicinity. Therefore, I introduced site-specific cysteines at positions N4C in the N-terminal tail, S74C in the OB domain, and T120C and D142C in the C-terminal tail and site-specifically labeled them with Alexa555 C2 maleimide dyes.

First, I measured the affinity of eIF1A-N4C-Alexa555 to the 40S subunit by monitoring anisotropy changes at increasing 40S concentrations in the fluorometer. The binding affinity (K_d), $8.2 \pm 0.5 \text{ nM}$, is calculated by fitting the anisotropy changes with the quadratic equation (**Figure 25A**). In the presence of eIF1, the K_d of the eIF1A-N4C-Alexa555 to the 40S subunit is $5 \pm 0.3 \text{ nM}$ (**Figure 25A**), which represents the upper limit to the K_d value, because the

concentration of the labeled eIF1A in this assay is also 5 nM. The anisotropy of the unbound eIF1A-N4C-Alexa555 is 0.28 at 25 °C, indicating that the rotation freedom of the fluorophore is restricted on eIF1A. After binding to the 40S subunit, the anisotropy of the bound eIF1A-N4C-Alexa555 is increased to 0.36 at 25 °C. Addition of >50 nM of 40S subunit do not further increase the anisotropy, indicating that the binding reaction has reached saturation. Also, eIF1A-S74C-Alexa555 and eIF1A-T120C-Alexa555 show high affinity to the 40S subunit with values of 6.8 ± 0.5 and 26 ± 4 nM, respectively (**Figure 25B**). The anisotropies of eIF1A-S74C-Alexa555 and eIF1A-T120C-Alexa555 are both about 0.31 off the ribosomes and 0.36 on the ribosomes (at 25 °C). The binding affinities from this assay are comparable to the published affinities of eIF1A to the 40S subunit (19 - 62 nM (Fraser et al., 2007; Sokabe and Fraser, 2014)).

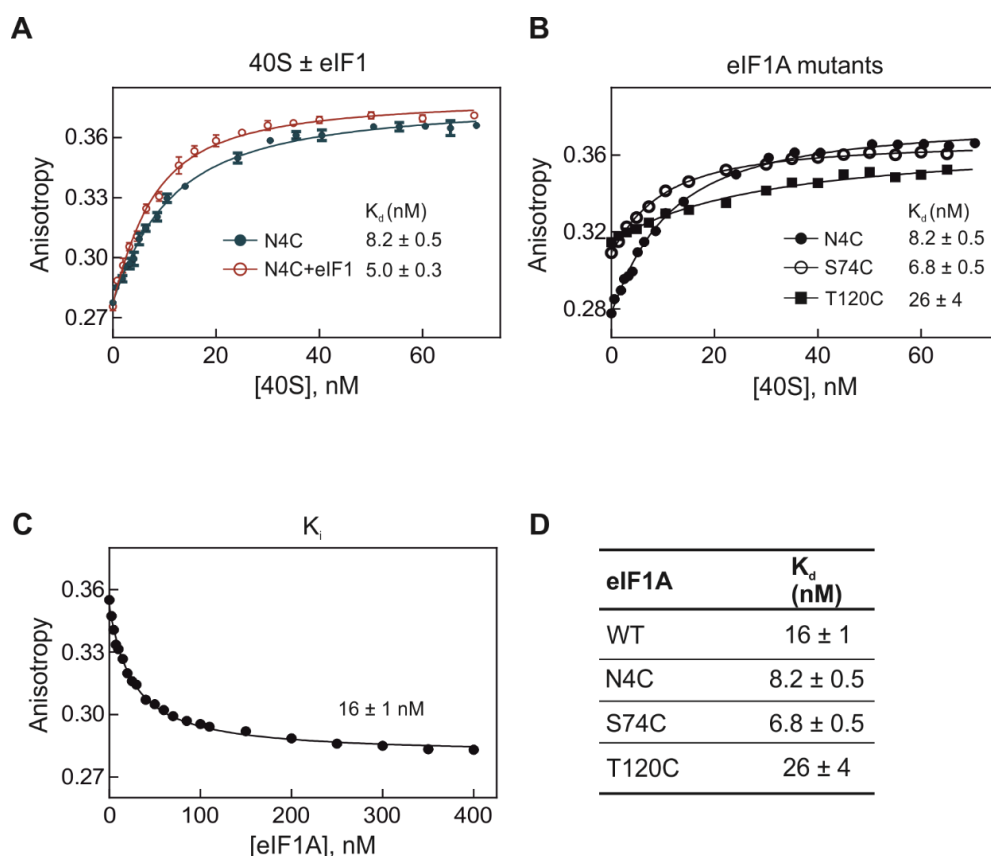


Figure 25. Binding of eIF1A to the 40S subunit

(A) The binding affinity of eIF1A-N4C-Alexa555 to the 40S subunit in the presence/absence of eIF1. The anisotropy of eIF1A-N4C-Alexa555 is measured with increasing concentrations of the 40S subunits in the absence or presence eIF1. The binding curve is analyzed with the quadratic equation to obtain the corresponding binding affinity. **(B)** The binding affinity of eIF1A N4C, S74C and T120C mutants to the 40S subunit. The anisotropies of eIF1A-N4C-Alexa555, eIF1A-S74C-Alexa555, and eIF1A-T120C-Alexa555 are monitored with increasing concentrations of the 40S subunits. The binding curve is analyzed with the quadratic equation to obtain the corresponding binding affinity. **(C)** The competition assay of the WT eIF1A to the 40S-eIF1A-N4C-Alexa555 complexes. Increasing amounts of the unlabeled WT eIF1A are added to compete with the binding of eIF1A-N4C-Alexa555 to the 40S subunit. The dissociation of the eIF1A-N4C-Alexa555 is monitored as decrease of anisotropy. The inhibitory constant, K_i , of the WT eIF1A to the 40S subunit is calculated from the IC_{50} of the competition curve. **(D)** Summary of binding affinity of eIF1A mutants to the 40S subunit. Binding affinity of WT eIF1A to the 40S subunit is from the competition assay with 40S-eIF1A-N4C-Alexa555 complexes.

To directly compare the affinities of the labeled and unlabeled eIF1A, I performed the competition binding assay with the unlabeled WT eIF1A. Pre-formed 40S-eIF1A-N4C-Alexa555 complexes are mixed with increasing amounts of unlabeled WT eIF1A. Because the WT eIF1A competes with eIF1A-N4C-Alexa555 for binding to the 40S subunit, the anisotropy decreases upon addition of WT eIF1A. The hyperbolic concentration dependence yields the affinity constant K_i of 16 ± 1 nM for the WT eIF1A (**Figure 25C**). The fact that K_d and K_i values are similar indicates that the fluorophore and mutation sites do not affect the affinity of eIF1A to the 40S subunit.

2.3.2 Fluorescence changes of labeled eIF1A upon 48S IC and 80S IC formation

To confirm that eIF1A-N4C-Alexa555 is a suitable observable for kinetic studies, I tested whether the fluorescence-labeled eIF1A reports on compositional changes of ribosome complexes during initiation complex formation. I monitored the fluorescence changes of the eIF1A-N4C-Alexa555 in response to the stepwise additions of the components for the 48S IC and 80S IC formation in the fluorometer (**Figure 26**). Binding of eIF1A to the 40S subunit results in the increase of the fluorescence intensity (almost 180%) and also of the anisotropy close to a saturated value of 0.36 (**Figure 26**), implying that the microenvironment of the fluorophore changes after binding and the rotational freedom of the fluorophore is restricted on the 40S subunit. The addition of eIF1 slightly increases the fluorescence and the anisotropy, probably due to the eIF1-induced conformational change of the 40S subunit or the cooperative binding effect of eIF1 and eIF1A that induces higher amount of bound eIF1A on the 40S subunit. Upon addition of eIF3, however, fluorescence intensity and anisotropy decrease. This is unexpected because eIF3 is a scaffold protein that stabilizes binding of almost all initiation factors on the 40S subunit. Hence, the decrease of the fluorescence intensity is likely due to an eIF3-induced conformational change of the 40S subunit. The lower anisotropy also implies that the eIF1A becomes more flexible on the 40S subunit in the presence of eIF3. The addition of the TC partially reverses the effect of eIF3 on fluorescence intensity and also restores the high anisotropy of eIF1A in the complexes. The complete 43S PIC is formed after the addition of

eIF4A and eIF4B, which do not alter the fluorescence intensity and anisotropy. Addition of the mRNA triggers formation of the 48S IC. The accompanying increase of the fluorescence intensity indicates that microenvironment of eIF1A-N4C-Alexa555 changes of eIF1A upon start codon recognition. After the 48S IC formation, additions of eIF5 and eIF5B slightly decrease the fluorescence intensity and anisotropy. Finally, 60S subunit joining leads to a large decrease in both fluorescence intensity and anisotropy due to the dissociation of the eIF1A from the 80S ribosomes. In the absence of the 40S subunit, none of the initiation components causes the changes in the fluorescence and anisotropy (**Figure 26B**). Since the fluorescence of eIF1A-N4C-Alexa555 responds to microenvironment changes only when bound to the 40S subunit, I conclude that eIF1A-N4C-Alexa555 can serve as a reliable reporter of the 48S IC formation.

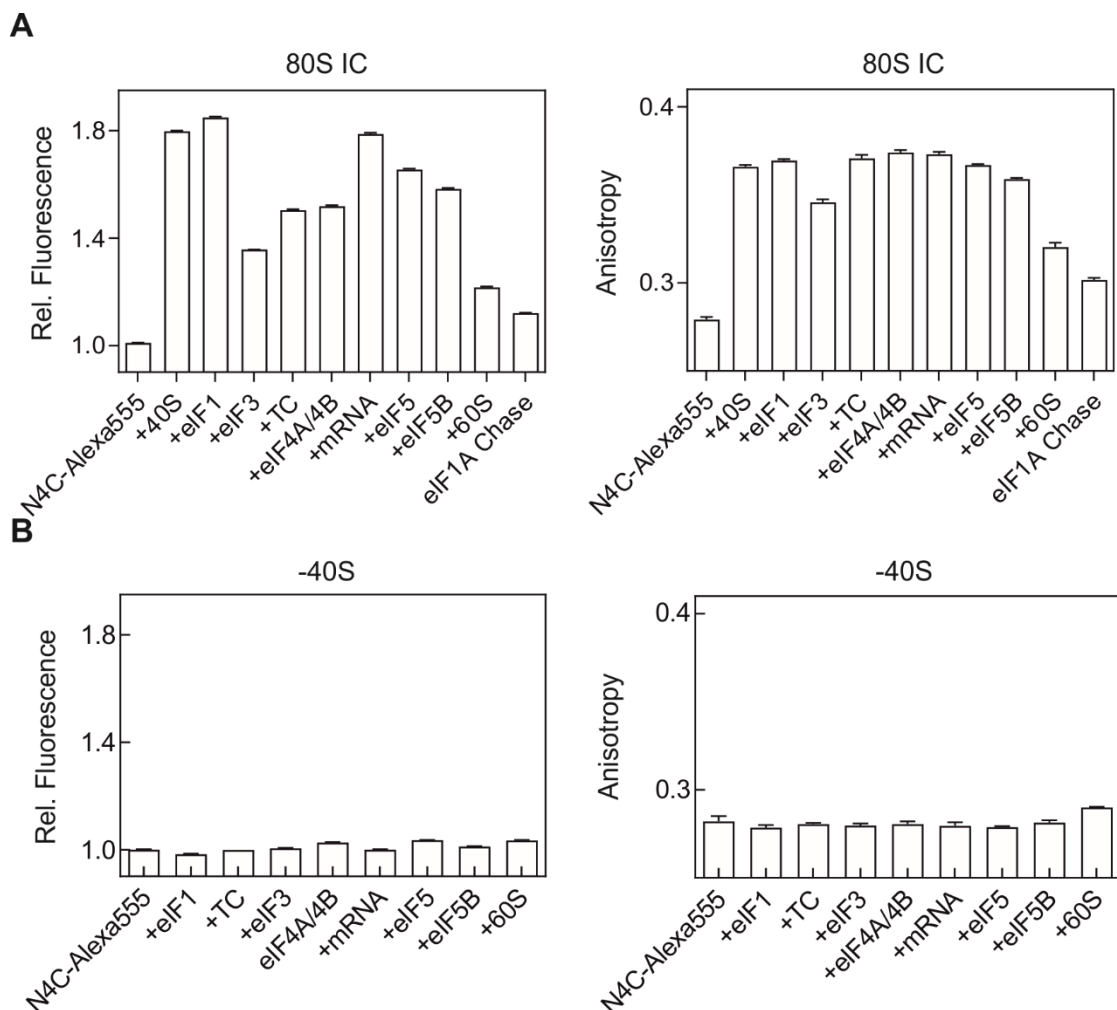


Figure 26. Fluorescence profiles of eIF1A-N4C-Alexa555 during 80S IC formation

(A) Fluorescence intensity and anisotropy changes of eIF1A-N4C-Alexa555 during 80S IC formation. The fluorescence intensity and anisotropy of eIF1A-N4C-Alexa555 are monitored in the spectrofluorometer upon stepwise addition of individual components for the 80S IC formation. The relative fluorescence intensity of each condition is normalized to the basal intensity of free eIF1A-N4C-Alexa555 as 1. The chase means addition of 10-fold excess of unlabeled WT eIF1A to the reaction mixture **(B)** Changes in the fluorescence intensity and anisotropy of eIF1A-N4C-Alexa555 with initiation factors in the absence of the 40S subunit. The fluorescence changes from the eIF1A-N4C-Alexa555 occur exclusively on the 40S subunit.

I next monitored other labeled eIF1A variants and their fluorescence changes during the 48S and 80S IC formation. All of the tested eIF1A versions respond to sequential 48S IC and 80S IC formation, however, in different profiles due to their distinct label positions (**Figure 27**).

The intensity of fluorophore attached to the S74C and T120C increases only slightly when eIF1A binds to the 40S subunit, while the reporter at the position D142C demonstrate even smaller increase in fluorescence. Since also eIF1A-S74C has a high binding affinity to the 40S subunit (**Figure 25**), the differences in the fluorescence intensity can be explained by particular orientations and distinct environments of the individual fluorophores on the 40S subunit. Upon addition of eIF1, fluorescence intensity of eIF1A-S74C-Alexa555 increases dramatically, indicating that a conformational change of the 40S subunit is induced by eIF1. The effect is less pronounced with eIF1A-T120C-Alexa555 and eIF1A-D142C-Alexa555. Addition of eIF3 does not have much effect except for N4C position. The presence of TC induces a large fluorescence change of eIF1A-T120C-Alexa555 and the signal remains high upon subsequent additions of eIF4A, eIF4B and mRNA. Similarly, the fluorescence remains high in the sample of eIF1A-S74C-Alexa555 after the addition of TC and subsequent eIF4A, eIF4B and mRNA, whereas these factors induce gradual decrease of eIF1A-D142C-Alexa555 fluorescence intensity. To validate that labeled eIF1A can dissociate from the 48S IC, I added excess amount of unlabeled WT eIF1A over the complex (chase), which leads to a decrease of fluorescence intensity to the level of unbound labeled eIF1A. This suggests that the fluorescence changes during the 48S IC formation correspond to specific interactions of labeled eIF1A variants with the initiation complexes and the binding of eIF1A to the ribosomes is reversible (**Figure 27**).

80S IC formation is completed when initiation factors are released from the ribosomes. The addition of eIF5 and eIF5B reduces the fluorescence intensity of all labeled eIF1A variants and a further dramatic drop of the signal is induced by 60S subunit joining (**Figure 27**). Addition of excess unlabeled WT eIF1A to the 80S IC mixture did not further affect the fluorescence intensity, implying that fluorescent eIF1A has been dissociated from the ribosomes and obtained its fluorescent state as that of the free eIF1A-Alexa555.

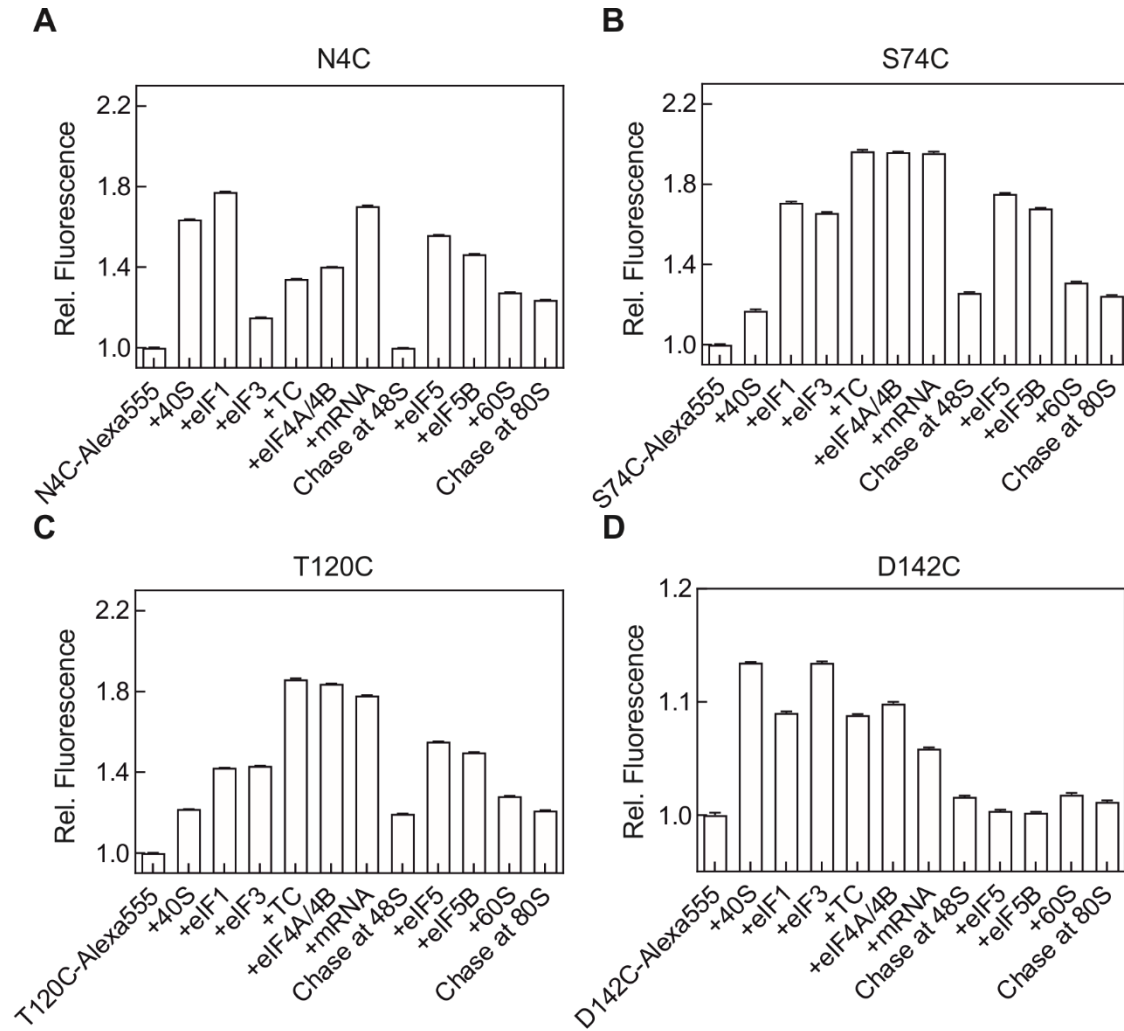


Figure 27. Fluorescence profiles of eIF1A-Alexa555 variants during 80S IC formation

(A-D) Fluorescence intensity changes of Alexa555-labeled eIF1A N4C, S74C, T120C, D142C variants upon stepwise addition of the components that lead to the 80S IC formation. eIF5, eIF5B and 60S subunits are added to the 48S IC in an independent experiment from the chase at 48S. The relative fluorescence intensity of each event is normalized to the basal fluorescence intensity of each free Alexa555-labeled eIF1A. The chase means addition of 10-fold excess of unlabeled WT eIF1A to the reaction mixture.

2.4 Kinetic analysis of mammalian translation initiation

2.4.1 Kinetic analysis of eIF1A binding to the 40S subunit

eIF1A has a high binding affinity to the 40S subunit and promotes 48S IC and 80S IC formation even in the absence of eIF1 (Figure 17). Therefore, I utilized eIF1A-N4C-Alexa555 to follow 43S PIC, 48S IC and 80S IC assembly by analyzing the fluorescence intensity changes that occur at almost every step during initiation complex formation. The finding obtained with eIF1A-N4C-Alexa555 were later confirmed using eIF1A variants labeled at different positions, i.e. eIF1A-S74C-Alexa555 and eIF1A-T120C-Alexa555.

To study the binding of eIF1A to the 40S subunit, I used pre-steady state kinetic approaches to obtain the association rate constant (k_{on}) by rapidly mixing eIF1A-N4C-Alexa555 with increasing concentrations of 40S subunits in a stopped-flow apparatus. The fluorescence intensity of eIF1A-N4C-Alexa555 increases upon binding to the 40S subunit and reaches a plateau after 200 s. The fluorescence intensity change is biphasic, with both phases contributing about equally to the observed fluorescence increase. Two-exponential fitting yields two apparent rate constants, k_{app1} and k_{app2} (Figure 28A). Both k_{app} values show linear dependence of 40S subunit concentration with slopes of $5.1 \pm 0.6 \mu\text{M}^{-1}\cdot\text{s}^{-1}$ and $2.2 \pm 0.1 \mu\text{M}^{-1}\cdot\text{s}^{-1}$ (Figure 28A). This suggests that both reactions are bi-molecular association reactions, i.e. the fluorescence changes reflect the binding of eIF1A to the 40S subunit. In the simplest model, the two slopes would represent the two k_{on} values of eIF1A binding to two kinetically distinct populations of 40S subunits.

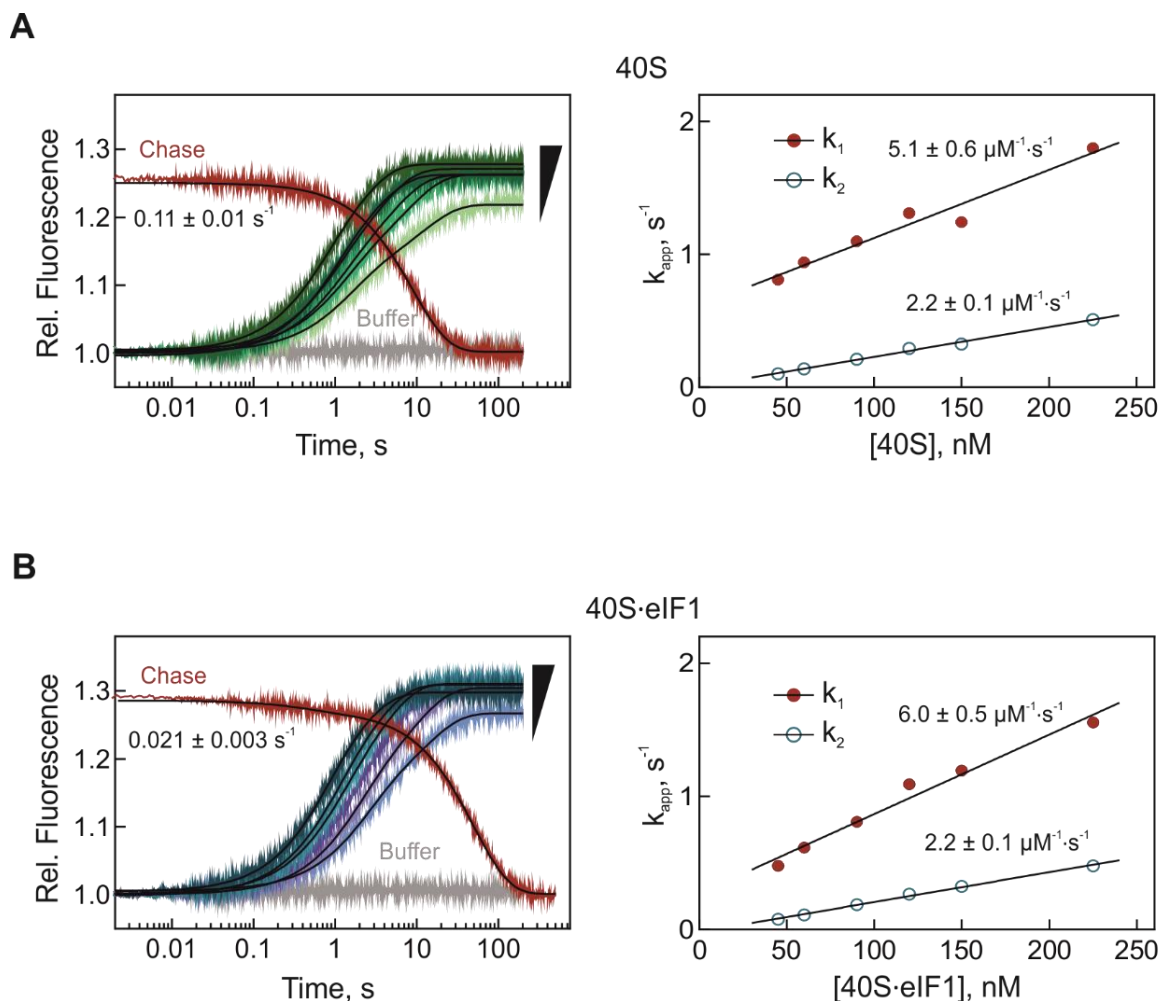


Figure 28. Kinetic analysis of binding and dissociation of the eIF1A-N4C-Alexa555 to the 40S subunit

(A) Left panel: Time courses of the 40S subunit binding (green) and dissociation (red) of the eIF1A-N4C-Alexa555. The eIF1A-N4C-Alexa555 (15 nM) was rapidly mixed with increasing concentrations (45, 60, 90, 120, 150, 225 nM) of the 40S subunits. In the chase experiment, 15-fold excess of unlabeled WT eIF1A is rapidly mixed with the 40S-eIF1A-N4C-Alexa555 complexes. Time courses were evaluated by two-exponential fitting (black line). Right panel: Concentration dependence of the two apparent rates constants (k_{app}). The linear concentration dependence of the k_{app} s yields the association rate constants (k_{on}) $5.1 \pm 0.6 \mu\text{M}^{-1} \cdot \text{s}^{-1}$ and $2.2 \pm 0.1 \mu\text{M}^{-1} \cdot \text{s}^{-1}$. The dissociation curve is fitted with one-exponential function to obtain the dissociation rate constant (k_{off}), $0.11 \pm 0.01 \text{ s}^{-1}$. **(B)** Left panel: Time courses of the 40S-eIF1 complex binding (blue) and dissociation (red) of the eIF1A-N4C-Alexa555. The eIF1A-N4C-Alexa555 (15 nM) was rapidly mixed with increasing concentrations (45, 60, 90, 120, 150, 225 nM) of the 40S subunits in the presence of eIF1 (500 nM). In the chase experiment, 15-fold excess of unlabeled

WT eIF1A is rapidly mixed with the 40S-eIF1-eIF1A-N4C-Alexa555 complexes. Time courses were evaluated by two-exponential fitting (black line). Right panel: Concentration dependence of the two k_{app} s. From the linear concentration dependence of the k_{app} s, the k_{on} values are $6.0 \pm 0.5 \mu\text{M}^{-1}\cdot\text{s}^{-1}$ and $2.2 \pm 0.1 \mu\text{M}^{-1}\cdot\text{s}^{-1}$. The dissociation curve is fitted with one-exponential function to obtain the k_{off} , $0.021 \pm 0.003 \text{ s}^{-1}$.

The Y-axis intercepts of the linear regressions provide the dissociation rates of eIF1A from the two 40S subunit populations of about 0.56 s^{-1} and close to 0, respectively (**Figure 28A**). In the simple model, if the two subpopulations of 40S subunits are independent of each other and do not interconvert, dissociation of eIF1A from the complex should be also biphasic, with the k_{app} values matching the value predicted from the titration experiment (**Figure 28A**). To verify this, I measured the dissociation rate constant (k_{off}) of the 40S-eIF1A-N4C-Alexa555 complex in a chase experiment. I formed 40S-eIF1A-N4C-Alexa555 complex and rapidly mixed it with 15-fold excess (over the complex) of unlabeled WT eIF1A in the stopped-flow apparatus. The decay of the fluorescence intensity upon dissociation was recorded until the fluorescence reaches a plateau (250 s). Initially, the time course of the chase experiment shows high fluorescence intensity, indicating that most of the eIF1A-N4C-Alexa555 remains bound to the 40S subunit. Upon dissociation of labeled eIF1A, the excess amount of unlabeled WT eIF1A occupies the binding site on the 40S subunit and prevents eIF1A-Alexa555 from rebinding to the 40S subunit. Given the concentrations in the experiment, the final part of the time course indicates the fluorescence intensity of the unbound eIF1A-N4C-Alexa555 in reaction buffer. In contrast to the predictions of a simple model, the time course of the fluorescence change is single-exponential, with a $k_{off} = 0.11 \pm 0.01 \text{ s}^{-1}$. The amplitude of the fluorescence change in the chase experiment is also almost identical to that of the binding curves. This suggests that eIF1A-N4C-Alexa555 dissociation occurs from a single population complex and almost all labeled eIF1A is chased out by the unlabeled WT eIF1A from the complex. A model that can explain this kinetic data is a binding-equilibration model, in which eIF1A binds to two kinetically distinct populations of the 40S subunits and, subsequently, the eIF1A-bound 40S subunits undergo a conformational change towards a dominant and thermodynamically favored complex.

Next, I performed the same experiments in the presence of eIF1. The time courses of binding reactions are also biphasic. The two k_{app} values show linear dependence on the concentrations of 40S-eIF1 complexes with the slopes similar to those without eIF1, $6.0 \pm 0.5 \mu\text{M}^{-1}\cdot\text{s}^{-1}$ and $2.2 \pm 0.1 \mu\text{M}^{-1}\cdot\text{s}^{-1}$ (**Figure 28B**). The dissociation of eIF1A-N4C-Alexa555 from the 40S-eIF1-eIF1A complex shows a monophasic fluorescence change with $k_{off} = 0.021 \pm 0.003 \text{ s}^{-1}$ (**Figure 28B**). This kinetic behavior is also consistent with the binding-equilibration model, i.e. after eIF1A and eIF1 bind to the 40S subunit, the complex equilibrates into a stable state. In conclusion, the association rate constants for eIF1A binding to the 40S subunit remain similar in the presence of eIF1. However, the dissociation rate constant of eIF1A is 5-fold lower from the eIF1-bound 40S subunit. This finding suggests that the binding rate of eIF1A to the 40S subunit is independent on eIF1, but if eIF1 is present, eIF1A stays longer on the 40S subunit due to cooperative interaction with eIF1. The higher affinity of eIF1A to the 40S subunit in the presence of eIF1 is consistent with the steady-state thermodynamic experiments. Kinetic analysis shows that higher affinity results from slow dissociation of eIF1A (**Table 4**).

Table 4. Summary of analytical rate constants of eIF1A binding to 40S and 40S-eIF1 complex

eIF1A	k_{on} ($\mu\text{M}^{-1}\cdot\text{s}^{-1}$)	k_{off} (s^{-1})
40S	5.1 ± 0.6 ; 2.2 ± 0.1	0.11 ± 0.01
40S-eIF1	6.0 ± 0.5 ; 2.2 ± 0.1	0.021 ± 0.003

Since the simple kinetic model does not satisfy the experimental data, I had to test whether all data are consistent with the binding-equilibration model (**Figure 29B**). For this purpose, I applied numerical integration of the data from binding and dissociation experiments using the KinTek Explorer software. The two-population binding-equilibrium model assumes that free 40S subunits fluctuate between two states, designated as B and D, before binding to eIF1A-N4C-Alexa555, designated as A. A can bind to B and D in a reversible manner, resulting in 40S-eIF1A complexes C and E, respectively (**Figure 29B**). Each reversible binding reaction has its association and dissociation rate constants, designated as k_{+1} , k_{-1} , k_{+2} , k_{-2} . The two populations of the 40S subunits, B and D, also interconvert with rate constants k_{+3} and k_{-3} .

Likewise, the eIF1A-bound 40S subunits C and E are in equilibrium with each other with rate constants k_{+4} and k_{-4} . All equilibria in the model form a closed thermodynamic cycle. In this model, the unbound substrate A, and complex C and E are fluorescent with their characteristic fluorescence coefficients. In the chase experiment, the unlabeled WT eIF1A is designated as N. N can bind to B and D with the same rate constants as A, which are k_{+1} , k_{-1} , k_{+2} , k_{-2} , and result in products M and L, respectively. M and L are the unlabeled versions of C and E, but interconvert with the same rate constants, k_{+4} and k_{-4} . In this model, I arbitrarily assumed that the characteristics of the labeled-eIF1A remain independent of the fluorophore, which allows me to link the values for the same rate constants during the fitting. In each experiment, I also allowed the total free 40S subunits reach the equilibrium between B and D before the binding reaction with A. All time courses from binding and chase experiments were simultaneously fitted in the KinTek Explorer software by numerical simulations. The results with the lowest Chi-squared values and the values closest to the analytical results are shown here.

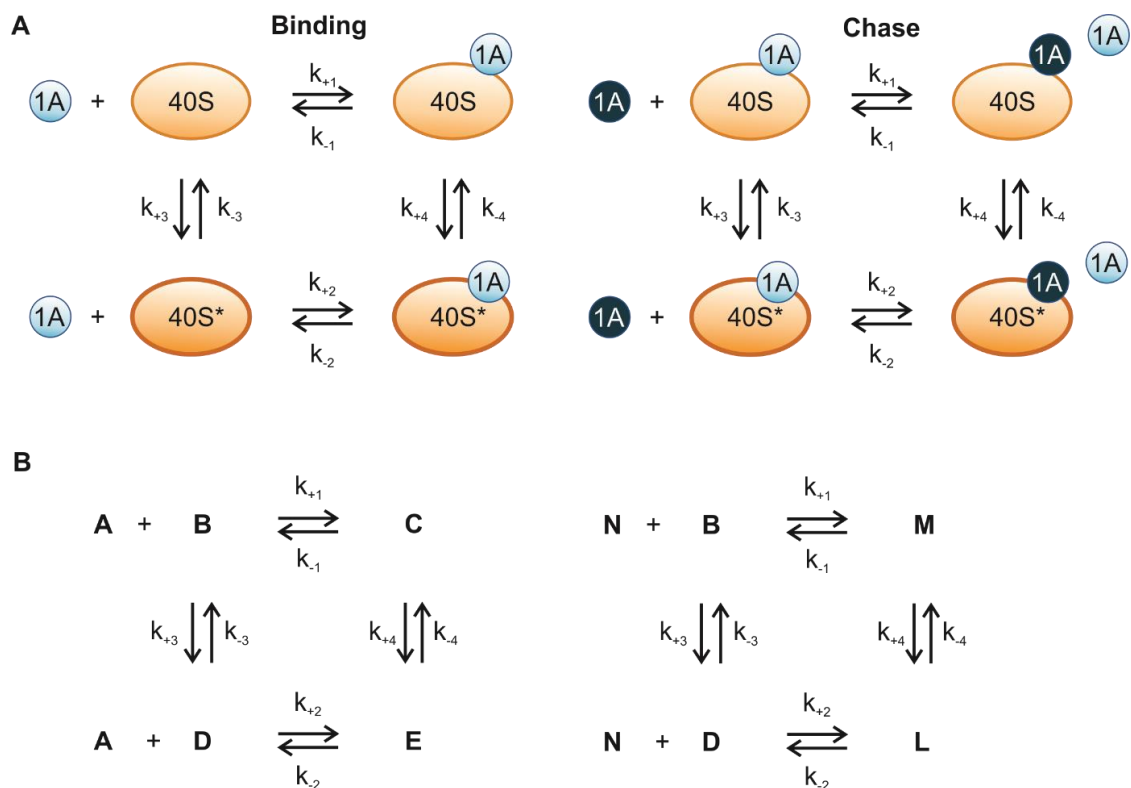


Figure 29. Two-population binding-equilibrium model of eIF1A binding to the 40S subunit

(A) Left schematic: Two bimolecular association reactions of eIF1A binding to the 40S subunit. 40S subunits equilibrate between 40S and 40S* states. Both states can reversibly bind eIF1A. The resulting complexes, eIF1A·40S and eIF1A·40S*, are also in equilibrium; the fluorescence of eIF1A (in light blue circle) is higher in the complexes than free in solution. Right schematic: In the chase experiment, excess amount of unlabeled WT eIF1A (in dark circle) compete with the binding site of labeled eIF1A on the 40S subunit upon dissociation of labeled eIF1A. The fluorescence of labeled eIF1A returns to its original state after being chased by the unlabeled eIF1A. **(B)** Left: Model of the eIF1A binding to the 40S subunit. eIF1A (reactant A) binds to two populations of the 40S subunits, B and D, resulting in complexes C and E, respectively. The binding and dissociation rate constants of A to B and D are shown as k_{+1} , k_{-1} , and k_{+2} , k_{-2} , respectively. States B and D can interconvert with rate constants k_{+3} and k_{-3} . After binding of eIF1A to the 40S, the two complexes C and E are also in equilibrium with rate constants k_{+4} and k_{-4} . A, C, E are fluorescent species; B and D are non-fluorescent. Right: In the chase experiment, unlabeled WT eIF1A is assigned as N. It can bind to the 40S subunits, B and D, with the same rate constants as A. Once the labeled eIF1A dissociates from the 40S subunits, the excess amount of the unlabeled WT eIF1A would bind to the 40S subunits and form complexes M and L. Since WT eIF1A is non-fluorescent, the complexes M and L are also non-fluorescent.

The results of the fit to the binding-equilibration model can satisfactorily explain all pre-steady state kinetic data, signifying that this relatively simple model is suitable to describe the eIF1A-40S interaction. The results of fitting suggest that the association rate constants of eIF1A to 40S subunits are $5 \pm 1 \mu\text{M}^{-1}\cdot\text{s}^{-1}$ and $30 \pm 4 \mu\text{M}^{-1}\cdot\text{s}^{-1}$, for the two 40S sub-populations (**Figure 30A**), followed by a conformational change that favors population E. In the absence of eIF1A, the two 40S subunit conformations are present in similar amounts, but after eIF1A binding, 95% of complexes convert to the E state. The rate constant for eIF1A dissociation from the E state (k_{-2}) is $0.104 \pm 0.001 \text{ s}^{-1}$, which is significantly smaller than the dissociation rate constant from the C state, and agrees well with the dissociation rate measured in the chase experiment ($0.11 \pm 0.01 \text{ s}^{-1}$). The binding-equilibration model provides a new insight into the mechanism of initiation by showing that free 40S subunits are likely to be dynamic in the absence of ligands and that the two conformations captured in our experiments have different kinetic preferences for eIF1A binding. However, once eIF1A is bound, the less favorable 40S conformation (which binds eIF1A less readily and allows it to dissociate faster) is converted to a state that is favorable for stable eIF1A binding. Notably, these conformations are not controlled by mRNA binding or codon recognition, but solely by the 40S interaction with eIF1A.

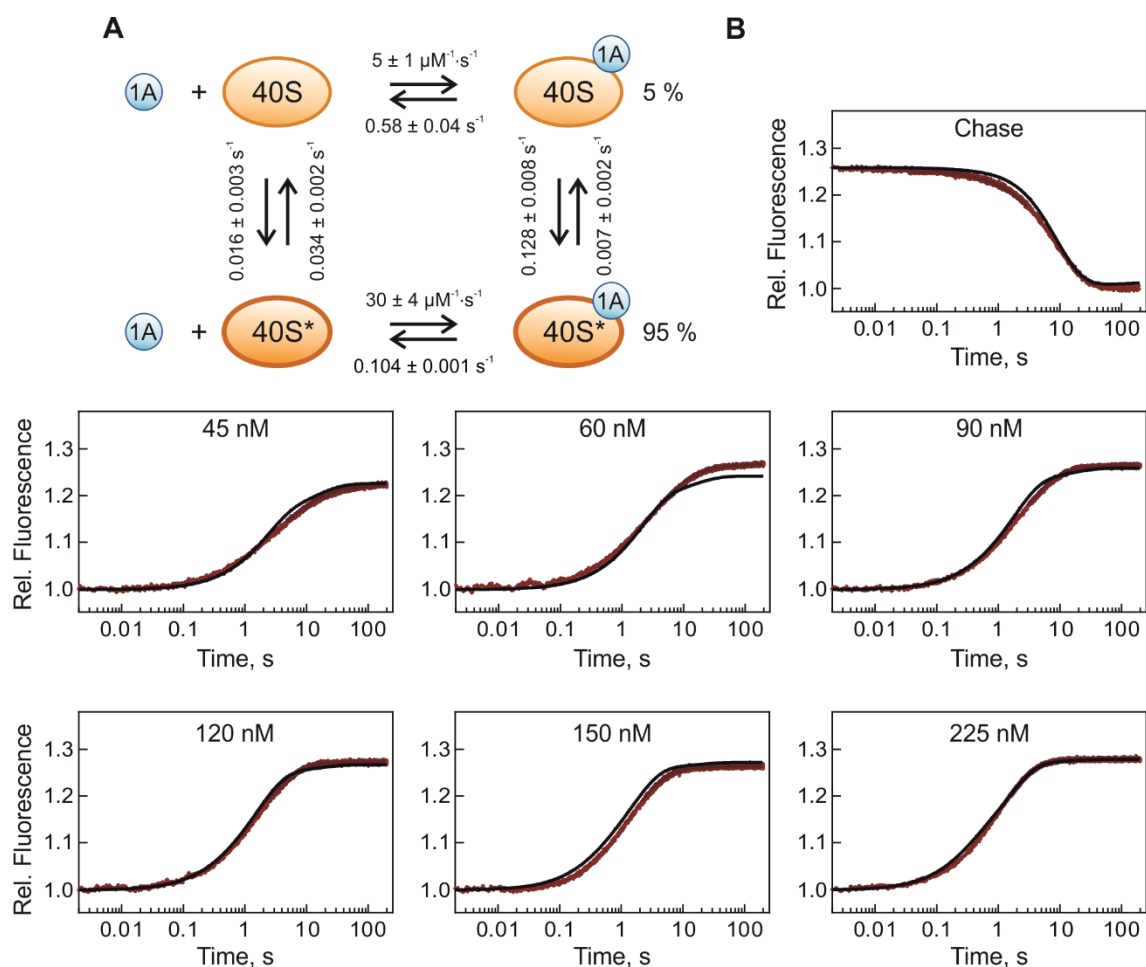


Figure 30. Fitting the data to the binding-equilibration model by numerical integration

(A) The two-population binding-equilibrium model and elemental rate constants resulting from numerical integration. **(B)** Fluorescent changes and associated simulations of the binding and dissociation reactions. The simulation result with the lowest Chi-squared value and closest to the analytical values is shown. The calculated fluorescent constants for A (eIF1A), C (40S·eIF1A) and E (40S*·eIF1A) are 0.066, 0.043 and 0.089, respectively. The Chi-squared of the simulation is 317683 with $N = 27873$ and degree of freedom (DoF) = 27861.

The same binding-equilibration model was then applied to the data from experiments of eIF1A binding to the 40S·eIF1 complex. eIF1 was found to not alter the distribution of 40S states (B and D) (**Figure 31A**). The association rate constants, k_{+1} and k_{+2} , are also largely independent of eIF1 at $5.2 \pm 0.2 \mu\text{M}^{-1}\cdot\text{s}^{-1}$ and $28.6 \pm 0.8 \mu\text{M}^{-1}\cdot\text{s}^{-1}$, respectively. After binding of eIF1A, the equilibrium between complex C and E is shifted even further towards the

population E (98%). Both k_{+4} and k_{-4} are somewhat slower in the presence of eIF1, $0.037 \pm 0.002 \text{ s}^{-1}$ and $0.0006 \pm 0.0002 \text{ s}^{-1}$, respectively, compared to $0.128 \pm 0.008 \text{ s}^{-1}$ and $0.007 \pm 0.002 \text{ s}^{-1}$ in the absence of eIF1. The dissociation rate constant, k_{-1} , from complex C is $0.9 \pm 0.2 \text{ s}^{-1}$, 1.5-fold higher than in the absence of eIF1. However, the dissociation rate constant, k_{-2} , from complex E is $0.027 \pm 0.001 \text{ s}^{-1}$ in the presence of eIF1, almost 4-fold smaller than that without eIF1, $0.104 \pm 0.001 \text{ s}^{-1}$. This agrees well with the results of the chase experiment suggesting a $k_{\text{off}} = 0.021 \pm 0.003 \text{ s}^{-1}$ from the 40S-eIF1-eIF1A complex. In summary, eIF1A binds to two interconverting populations of 40S subunits with two association rate constants independent of eIF1. eIF1 stabilizes eIF1A on the 40S subunits and shifts the equilibrium of the eIF1A-bound 40S subunits even further towards the stable state E (98%). Additionally, eIF1A dissociates even more slowly from the 40S subunits in the presence of eIF1. The consequence is a higher binding affinity of eIF1A to the 40S subunit in the presence of eIF1 ($7 \pm 2 \text{ nM}$ for 40S and $2.6 \pm 0.2 \text{ nM}$ for 40S-eIF1 complex), providing a kinetic evidence for the cooperative binding of eIF1A and eIF1 to the 40S subunits (**Table 5**).

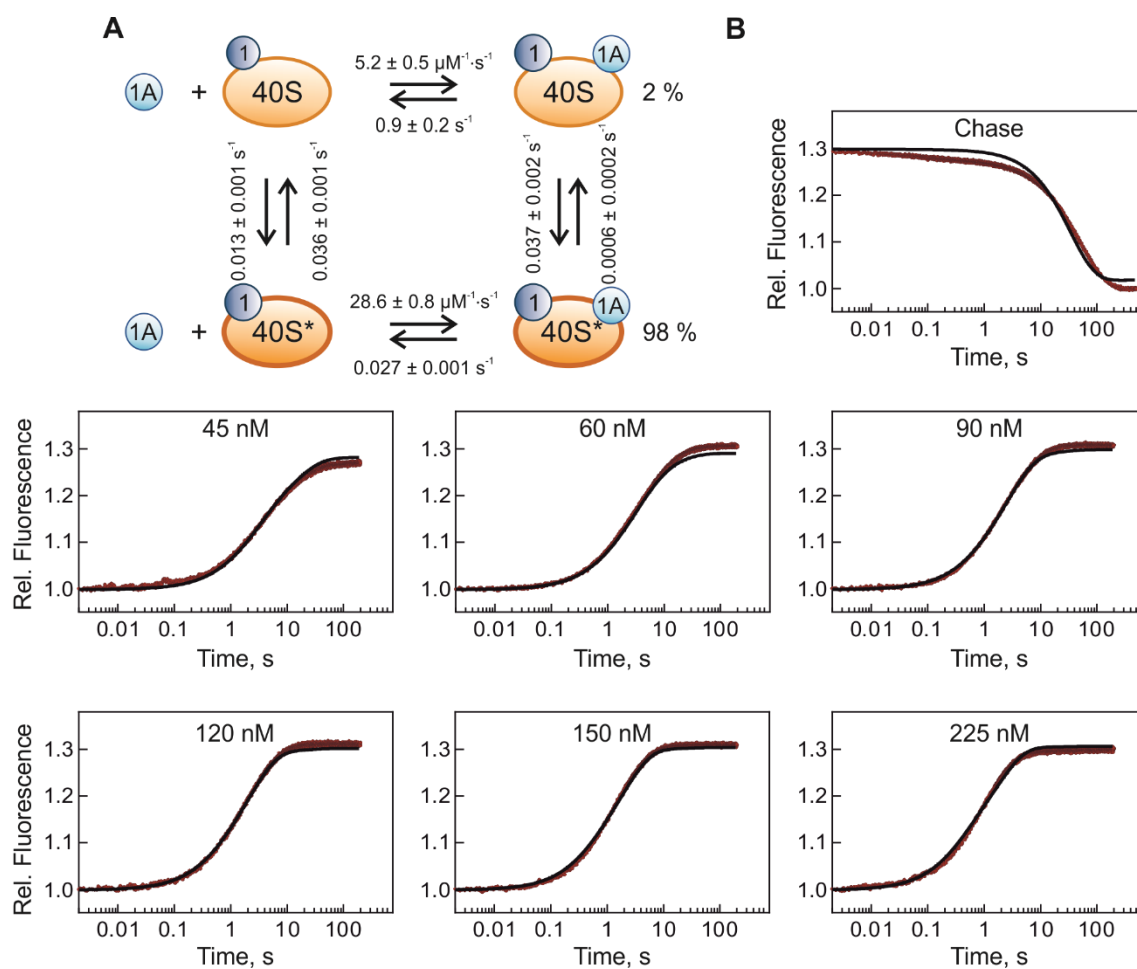


Figure 31. Kinetic model of eIF1A binding to the 40S-eIF1 complex

(A) Elemental rate constants in the two-population binding-equilibration model were obtained by numerical integration. **(B)** Fluorescent changes and associated simulations of the binding and dissociation reactions. The fitting result with the lowest Chi-squared value and closest to the analytical values is shown here. The calculated fluorescent constant for A (eIF1A), C (40S-eIF1-eIF1A) and E (40S*-eIF1-eIF1A) is 0.067, 0.057 and 0.088, respectively. The Chi-squared of the simulation is 298963 with $N = 27878$ and $\text{DoF} = 27866$.

Table 5. Summary of K_d s of eIF1A binding to 40S and 40S-eIF1 complex by kinetic and thermodynamic analysis

Complex	Calculated K_d^* (nM)	Measured K_d^{**} (nM)
40S	7 ± 2	8.2 ± 0.5
40S-eIF1	2.6 ± 0.2	5.0 ± 0.3

*The K_d is calculated from the equation, $K_d = (k_{-3}/k_{+3}) * (k_{-2}/k_{+2})$, with elemental rate constants from the numerical integrations in **Figure 30** and **Figure 31**. The two-tailed p value is < 0.01 between K_d s of both complexes by Student's t-test.

** The values are taken from the steady-state binding assay (**Figure 25**).

2.4.2 Identification of discrete eIF1A-dissociation states at different stages of initiation

The detailed analysis of eIF1A binding to 40S and 40S-eIF1 complexes shows that the k_{off} of eIF1A is sensitive to the conformation and composition of the complexes, which implies that differences in the k_{off} values can be interpreted as changes in the conformations of the complexes. Low k_{off} values suggest that the corresponding state is thermodynamically favored and/or kinetically stable. This prompted me to measure the k_{off} values for eIF1A-40S complex in the presence of other initiation factors to understand which of them favor the conformation that stabilizes eIF1A binding. I formed the complexes with labeled eIF1A in the presence of another individual initiation factor and performed the chase experiments by adding excess amounts of unlabeled WT eIF1A in the stopped-flow apparatus. The time courses were analyzed with either one-exponential or two-exponential functions to obtain the k_{off} values (**Figure 32**).

eIF1 and TC are the two initiation factors that were found to stabilize eIF1A binding on 40S subunit, and they result in a k_{off} value that is about 5-fold lower than that of the 40S-eIF1A complex in the absence of other factors. Unlike with eIF1, the dissociation time courses of

eIF1A from the 40S·TC·eIF1A complex are biphasic, with rate constants $k_{\text{off}1} = 0.1090 \pm 0.0007 \text{ s}^{-1}$, which is similar to eIF1A dissociation from 40S subunit, and a slower reaction with $k_{\text{off}2} = 0.0184 \pm 0.0002 \text{ s}^{-1}$, similar to eIF1A dissociation from 40S·eIF1·eIF1A complex (**Figure 32**). Binding of TC to the 40S·eIF1A complex was monitored by fluorescence changes in the 40S·eIF1A-Alexa555, and TC binding was found to be saturated at 0.3 μM TC (**Figure 33**). This indicates that the biphasic curve observed in the chase experiment of 40S·TC·eIF1A complex is not a result of heterogeneous (i.e. unsaturated) complexes. The biphasic curve would rather suggest that eIF1A dissociates from two distinct populations of the 40S·TC·eIF1A complexes. Unlike eIF1, binding of TC to the 40S subunit does not completely shift the ribosome populations into a stable state; rather, a fraction of ribosomes remains in the state favored by the eIF1A binding alone. eIF3, eIF5B and mRNA do not change the k_{off} of the 40S·eIF1A complex, whereas eIF4A together with eIF4B even slightly increase the k_{off} (**Figure 32**). Moreover, additions of eIF1 and mRNA to the 40S·eIF1A complexes increases the fluorescence amplitudes, whereas bindings of TC, eIF3, eIF4AB and eIF5B to the same complexes decrease the fluorescence intensities. Although some of the complexes have similar eIF1A k_{off} values, the large variations in the fluorescence amplitudes indicate that the initiation factors affect the microenvironment of eIF1A in the corresponding complexes and could also reflect the conformations of the 40S subunit in these complexes (**Figure 32D**). For example, subunits of eIF3 and eIF5B might have direct contacts to eIF1A or bind in the vicinity of eIF1A in such a way that they change the microenvironment of the fluorophore attached to eIF1A. eIF4A and eIF4B bind at the entry channel next to the A site close to eIF1A, which also affects the surrounding of the fluorophore. eIF1 and TC, on the other hand, have less effect on the fluorescence amplitudes of the corresponding complexes, but result in slower k_{off} values. In these cases, eIF1A might have different orientations or binding positions that leads to more stable interactions with the 40S subunit and, thereby, lower k_{off} values.

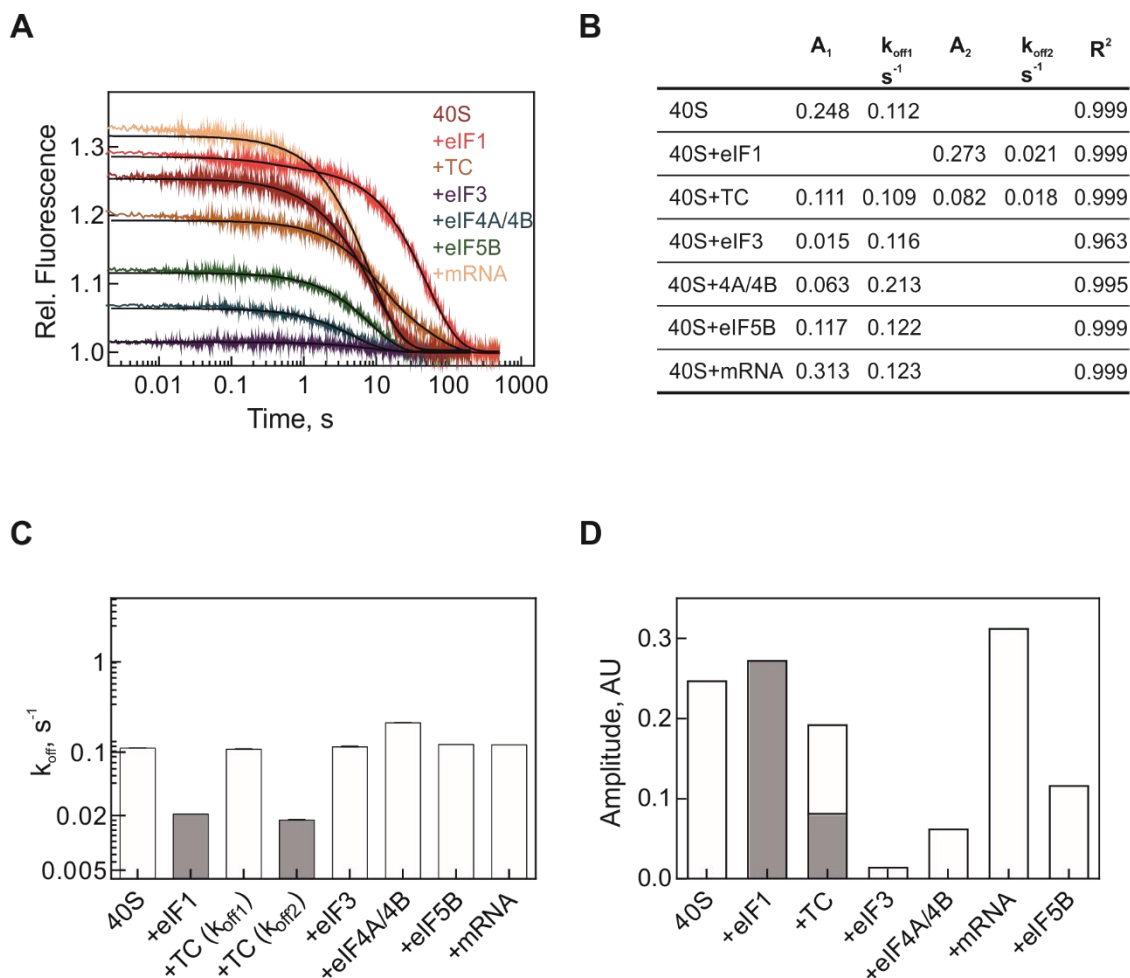


Figure 32. Dissociation of eIF1A from the 40S subunit in the presence of other individual initiation factor

(A) Kinetics of dissociation of eIF1A from the 40S subunit in the chase experiments. The time courses were analyzed by single-exponential fitting except for 40S+TC time course, which is biphasic. **(B)** Summary of the k_{off1} and k_{off2} values and the fluorescence amplitudes A_1 and A_2 , respectively, from exponential fittings. R^2 is the coefficient of determination of the corresponding exponential fitting. **(C, D)** Bar charts of the k_{off} values and the respective fluorescence amplitudes. The states that stabilize eIF1A on the complexes are marked in grey. All SEMs of the k_{off} values are below 0.001 in this figure. AU, arbitrary unit

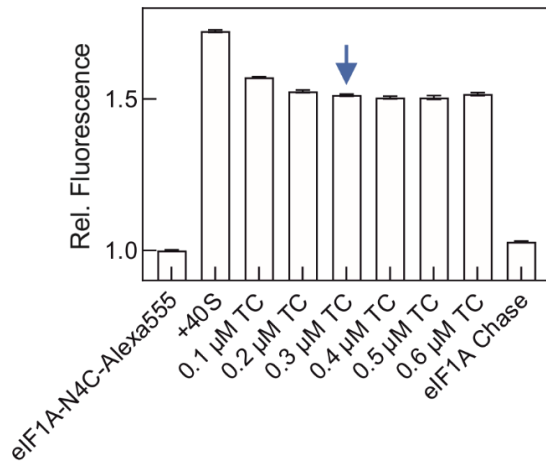


Figure 33. Fluorescence changes during titration of the 40S-eIF1A complex with TC

Increasing concentrations of TC are added to the 40S-eIF1A complex. The fluorescence intensities of the 40S-eIF1A-Alexa555-TC complex remain steady after addition of 0.2 μM TC, indicating saturation of TC binding to the 40S-eIF1A-TC complex. The 40S-eIF1A-TC complex in the chase experiment (**Figure 32**) is formed with TC at the saturating concentration, 0.3 μM , which is indicated by the blue arrow. 15 nM of eIF1A-N4C-Alexa555 and 90 nM of 40S were used in this measurement as the same condition as that in the chase experiment.

Next, I determined the k_{off} of initiation complexes that may occur *en route* to 43S PIC formation, i.e. with different combinations of eIF1, TC and mRNA. Any combination of two or more ligands, such as 40S-eIF1-TC-mRNA-eIF1A complex, stabilizes eIF1A on the 40S subunit, resulting in a lower k_{off} (**Figure 34**). The conclusion of the previous chase experiments showed that two-phase dissociation is observed for the 40S-TC-eIF1A-mRNA complex, implying the presence of two conformations of the complex (E and C states). Upon addition of eIF1, however, only one dissociation phase is observed, indicating that one conformation of the complex is dominant. The two populations observed for the 40S-TC-eIF1A complex are still observed in the presence of mRNA, but disappear upon addition of eIF1. When both eIF1 and TC bind to the 40S subunit, the 40S subunit is found exclusively in the kinetically stable conformation and a single-exponential dissociation is observed in the chase experiment. The lower k_{off} value of eIF1A from the 40S subunit in the presence of eIF1 and TC also supports the finding that binding affinity of eIF1A to the 40S subunit is further increased in the presence of both factors (Fekete et al., 2007). A slow, single-exponential dissociation from the ribosome is

also the case in the presence of eIF1, TC, eIF3 and eIF4A/4B (in the absence of mRNA), i.e. in the 43S PIC (**Figure 35**). In conclusion, addition of components of the 43S PIC promotes a shift in the 40S conformation towards a kinetically stable state.

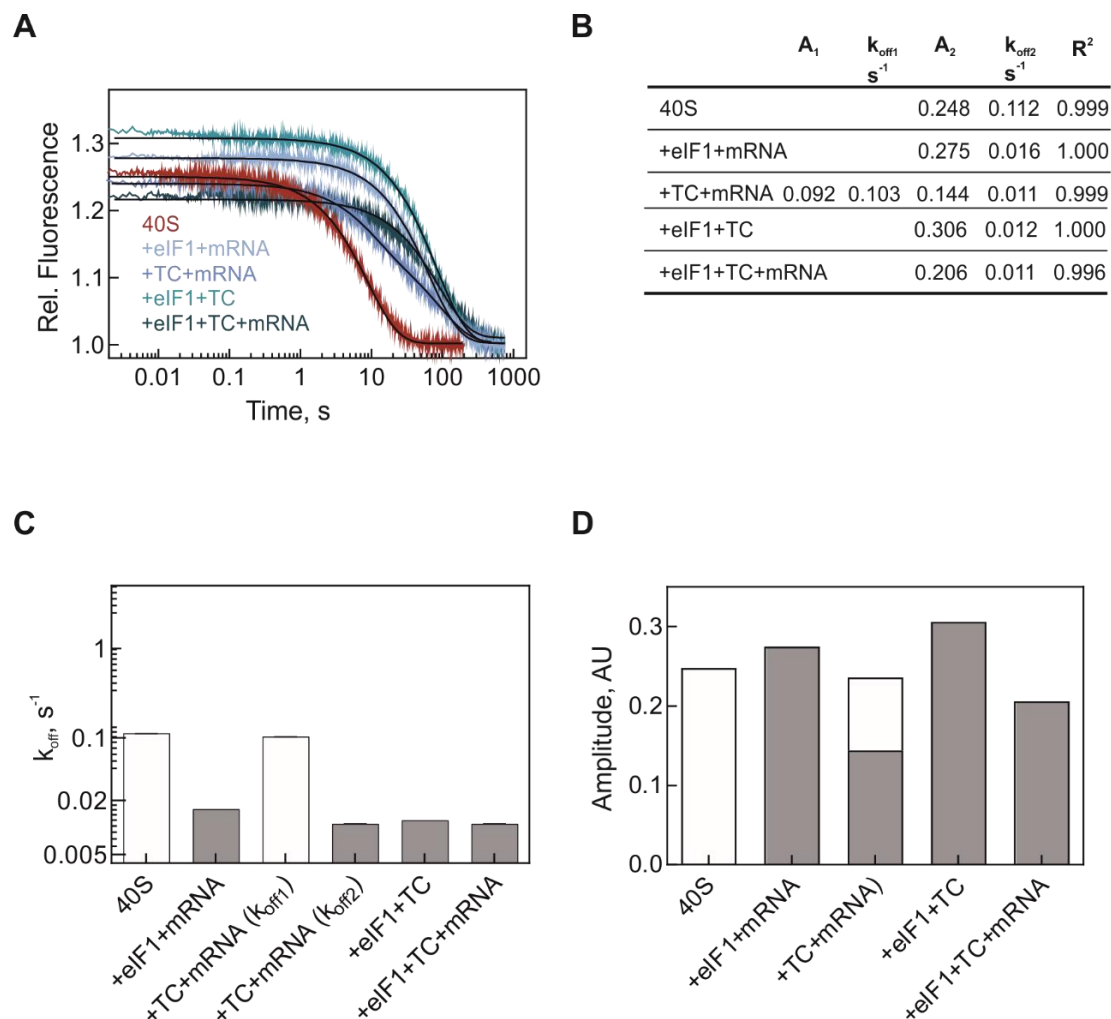


Figure 34. Dissociation of eIF1A from the 40S subunit with multiple initiation factors

(A) Kinetics of dissociation of eIF1A from the 40S subunit in the presence of different combinations of initiation factors. The time courses were analyzed by single-exponential fitting except for 40S +TC+mRNA time course, which is biphasic. **(B)** Summary of the k_{off1} and k_{off2} values and the fluorescence amplitudes A_1 and A_2 , respectively, from exponential fittings. R^2 is the coefficient of determination of the corresponding exponential fitting. **(C, D)** Bar charts of the k_{off} values and the respective fluorescence amplitudes. The states that stabilize eIF1A on the complexes are marked in grey, while the states with faster dissociation rates are marked in white. All SEMs of the k_{off} s are below 0.001. AU, arbitrary unit.

Next, I asked the question, which factor exerts control over 40S conformation in the 43S PIC. The 43S PIC is assembled with eIF1, eIF1A, TC, eIF3, eIF4A, eIF4B and the 40S subunit in the absence of mRNA. I included other combinations of initiation factors that contributes to the 43S PIC formation and measured the dissociation rates of eIF1A from these partial initiation complexes. These combinations of initiation factors are equivalent to complexes, in which one of the components of the 43S PIC is missing. The chase experiment shows that 43S PIC is a single-population complex with a low k_{off} , 0.034 s^{-1} (**Figure 35**) (or 0.012 s^{-1} in another independent experiment with different biological preparations of the native protein components (**Figure 37**)). In this respect, 43S PIC is similar to the 40S·eIF1·TC·eIF1A complex, which lacks eIF3, eIF4A, eIF4B and mRNA. However, these two complexes have different fluorescence intensity, suggesting that the environments of the fluorophore are different in those complexes. This might be due to conformational changes of the complexes or different orientations of eIF1A on the ribosomes. When eIF3 is omitted from the 43S PIC, the complex remains in a single population with a low k_{off} , 0.030 s^{-1} , but there is a large increase in the fluorescence amplitude, indicating that binding of eIF3 to the 43S PIC alters dramatically the microenvironment of eIF1A, potentially through direct interactions between its subunits and eIF1A. The 43S PIC without either eIF1, TC or eIF4A/4B shows two populations with fast and slow k_{off} s, but only small increases in the fluorescence intensities are observed, suggesting that eIF1, TC and eIF4A/4B might induce conformation changes of the 40S subunit or stabilize eIF1A on the 40S subunit in the context of the 43S PIC. eIF3 is the common factor in these three complexes. Since eIF3 is the scaffold protein that interacts with all initiation factors on the 40S subunit, this suggests that eIF3 senses the completeness of the 43S PIC and drive the equilibrium back towards the kinetically unfavored complex (state C). The facts that of 43S PIC without eIF3 has a low k_{off} and 43S PIC dynamics are induced by eIF3 in the absence of eIF1, TC or eIF4A/4B imply that eIF3 has a higher chance to join the 43 PIC after other factors have already bound to the 40S subunit. Without TC, the 43S PIC not only have two states but also the k_{off} values increase even when eIF1 is bound. In the presence of eIF3 in the 43S PIC, eIF1 is unable to stabilize a single ribosome conformation the way it does in the eIF1A·40S·eIF1·TC·mRNA complex. The 43S PIC without eIF1 also partitions into two states,

one with a low k_{off} and another with a dissociation rate similar to that of the 40S value. This suggests that the 43S PIC possesses the intrinsic ability to detect the vacancy of individual initiation factors and, when doing so, maintains access to the kinetically unstable state (C state). In summary, *en route* to the 43S PIC assembly, eIF1A has fast association rates to two different states of the 40S subunit and stabilizes one of them upon binding (**Figure 36**). The dissociation of eIF1A is further stabilized mainly by the presences of eIF1 and TC. The formation of 43S PIC leads the ribosome to another stable state that is ready to recruit the mRNA for scanning.

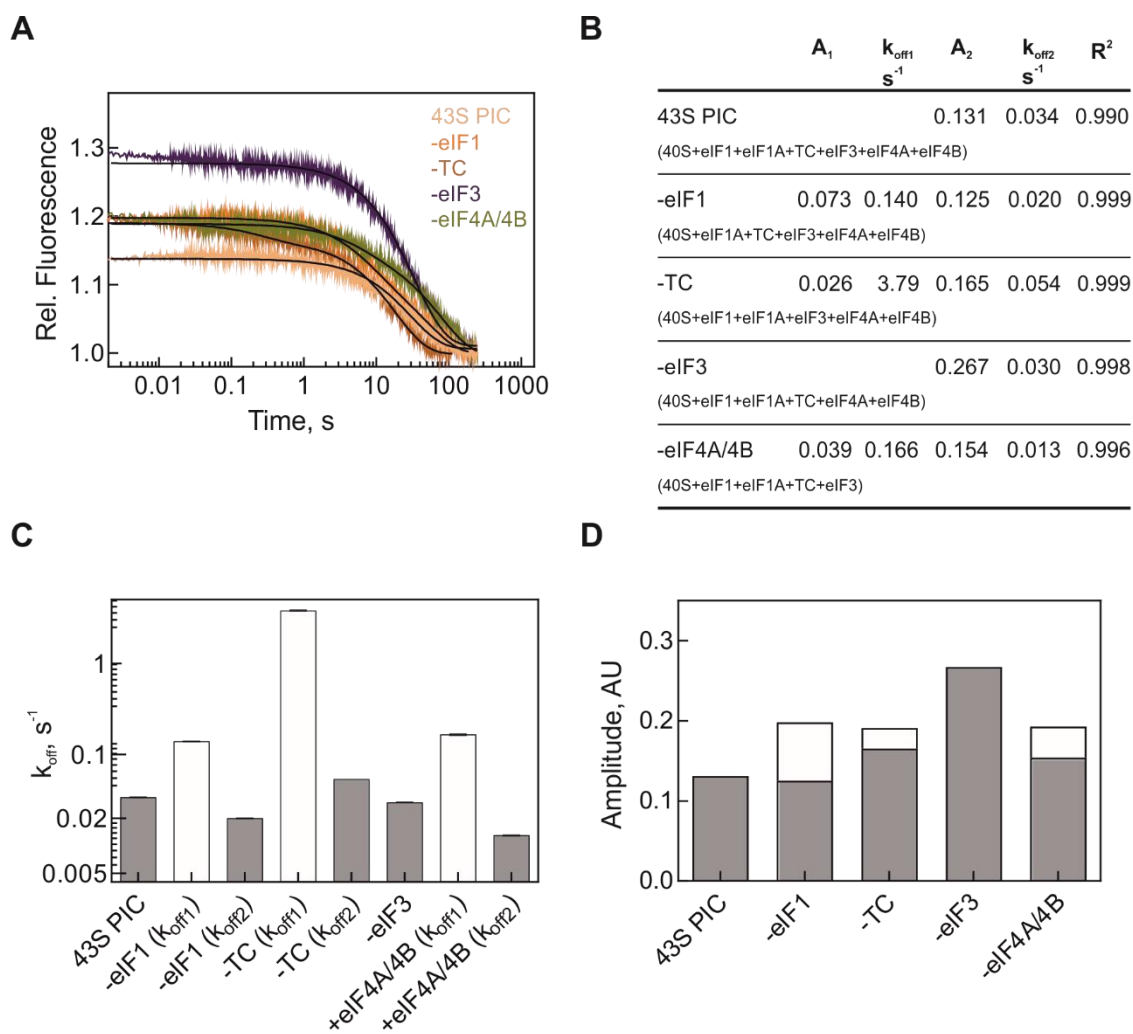


Figure 35. Dissociation of eIF1A from the 43S pre-initiation complex

(A) Kinetics of dissociation of eIF1A from the 40S subunit in the presence of different combinations of initiation factors. The time courses were analyzed by either one- or two-exponential functions. **(B)** Summary of the k_{off1} and k_{off2} values and the fluorescence amplitudes A_1 and A_2 , respectively, from exponential fittings. R^2 is the coefficient of determination of the corresponding exponential fitting. The composition of each complex is indicated in brackets. **(C, D)** Bar charts of the k_{off} values and the respective fluorescence amplitudes. The states that stabilize eIF1A on the complexes are marked in grey, while the states with faster dissociation rates are marked in white. All SEMs of the k_{off} s are below 0.06. AU, arbitrary unit.

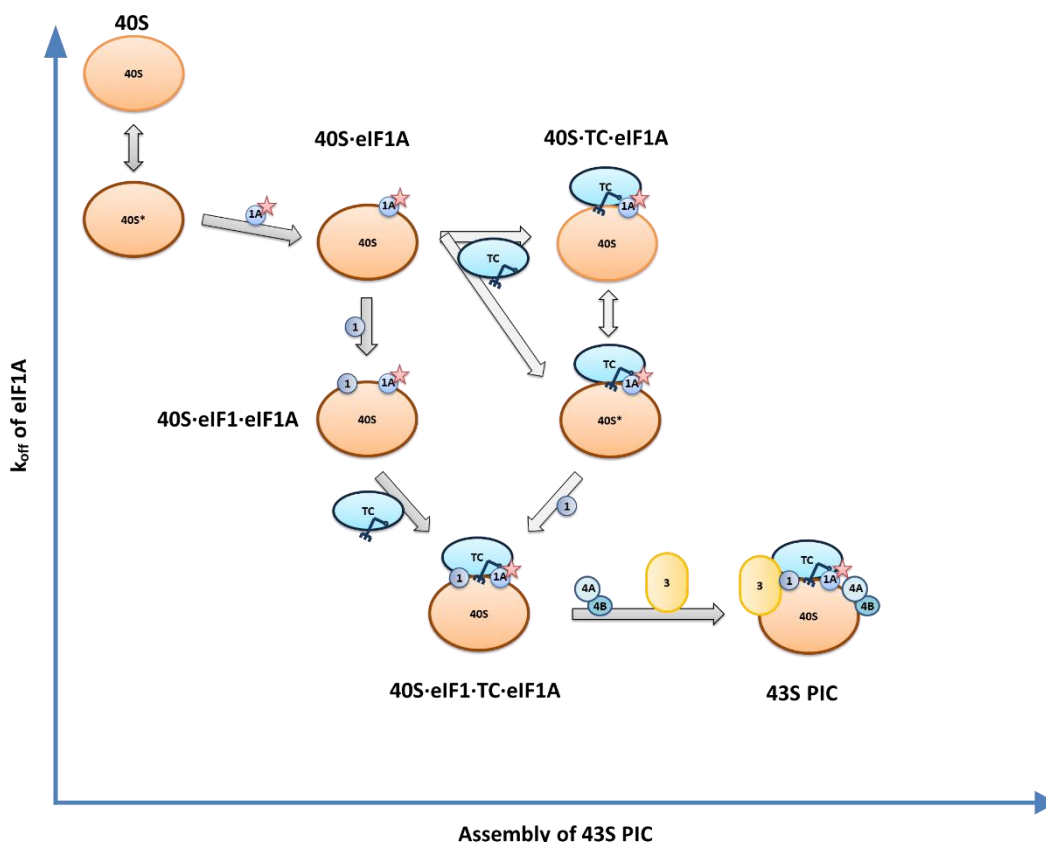


Figure 36. Hierarchy of initiation complexes *en route* to the 43S PIC formation probed by their stability of eIF1A binding

The potential pathway of 43S PIC assembly from the perspective of eIF1A. The initiation complexes with different bound-factors are presented as their states corresponding to the k_{off} s of eIF1A. 40S and 40S* indicates the two states of the 40S subunits with fast and slow k_{off} s, respectively.

Next, I added mRNA with cognate AUG start codon to the 43S IC, resulting in formation of the 48S (AUG) IC. The dissociation of eIF1A from 48S (AUG) IC is clearly biphasic, with two distinct k_{off} values, 0.110 s^{-1} and 0.005 s^{-1} (**Figure 37**), indicating two distinct sub-populations of the complexes, consistent with the previous results in yeast system (**Table 2**) (Llacer et al., 2018; Maag et al., 2006). Also, similarly to the results obtained with yeast 48S IC, the contribution of the rapid dissociation step to the overall fluorescence change was small, whereas the slow step contributed the majority of the fluorescence signal. To further probe the states of these two populations, I measured the k_{off} of eIF1A from the 48S pre-IC (PIC) with CUC model mRNA, which represents the mRNA-bound 48S PIC but without start codon

recognition. The 48S (CUC) PIC also biochemically imitates the scanning state of the initiation complexes. In the toeprinting assay, the 48S (CUC) PIC does not generate a characteristic toeprint either (**Figure 15**). Surprisingly, the 48S (CUC) PIC has a very slow k_{off} , 0.007 s^{-1} and comprises a single population (**Figure 37**). According to the scanning mechanism, once the full initiation complex attaches to the 5' end of the mRNA, the ribosome scans base-by-base towards the 3' end of the mRNA. As long as the ribosome does not encounter a start codon, it continues scanning and probably will stall or fall off from the mRNA at the 3' end. mRNA binding allows the 43S PIC to adopt a more kinetically favored state that can continuously scan on the mRNA without start codon recognition. The start codon recognition event not only retains some initiation complexes at the kinetically favored conformation but also induces some of ribosomes shift to a fast-dissociation state. Since I am using eIF1A as the observable, this could also imply that after the start codon recognition, eIF1A has different binding orientations or conformations on the ribosome. eIF1A might partially relocate or dispatch from its original position in the decoding center to prepare itself for the subsequent task, in assistance with the eIF5B for the large subunit joining. Therefore, I also included eIF5B into the reconstituted 48S (AUG) IC and measured the k_{off} of eIF1A from 48S+eIF5B IC complex. The result of 48S+eIF5B IC is similar to the 48S IC with two distinct but slightly slower k_{off} values, 0.089 s^{-1} and 0.004 s^{-1} (**Figure 37**). The addition of eIF5B also reduces the total fluorescence intensity, suggesting eIF5B either interacts directly with eIF1A and changes its microenvironment or induces a global conformational change of the initiation complex. In order to verify if the 2nd and 3rd AUG in the 3' UTR of the model mRNA affects the eIF1A dissociation from the 48S IC, I replaced these two AUG codons in the 3' UTR with CUC in mRNA with an AUG start codon and one without (i.e. CUC replacing AUG in the ORF). Chase experiments with 48S complexes assembled on these mRNAs, 48S (CUC) PIC and 48S (AUG) IC are identical to the original observation (**Figure 38**), where 48S (CUC) PIC comprises a stable single population and 48S (AUG) IC has two distinct states. The 2nd and 3rd AUGs are not selected by the 48S IC, because in the presence of both eIF1 and eIF1A, only the first AUG in the optimal Kozak sequence is recognized by the 48S IC. The 2nd and 3rd AUG codons are flanking in the poor Kozak context; therefore, they are mostly excluded by the ribosomes in

the presence of eIF1 and eIF1A, as it has also been shown by toeprinting (**Figure 15**). In conclusion, the ribosomes equilibrate from one intermediate complex, 43S PIC, to a further stable state, 48S (CUC) PIC (**Figure 39**). The kinetic checkpoints are driven by the recruitment of a complete set of initiation factors on the 43S PIC and binding of mRNA to form scanning 48S PIC. After start codon recognition, 48S (AUG) IC resumes dynamic interconversion between two states and some of the ribosomes convert to a conformation that allows eIF1A to dissociate faster from the 48S IC.

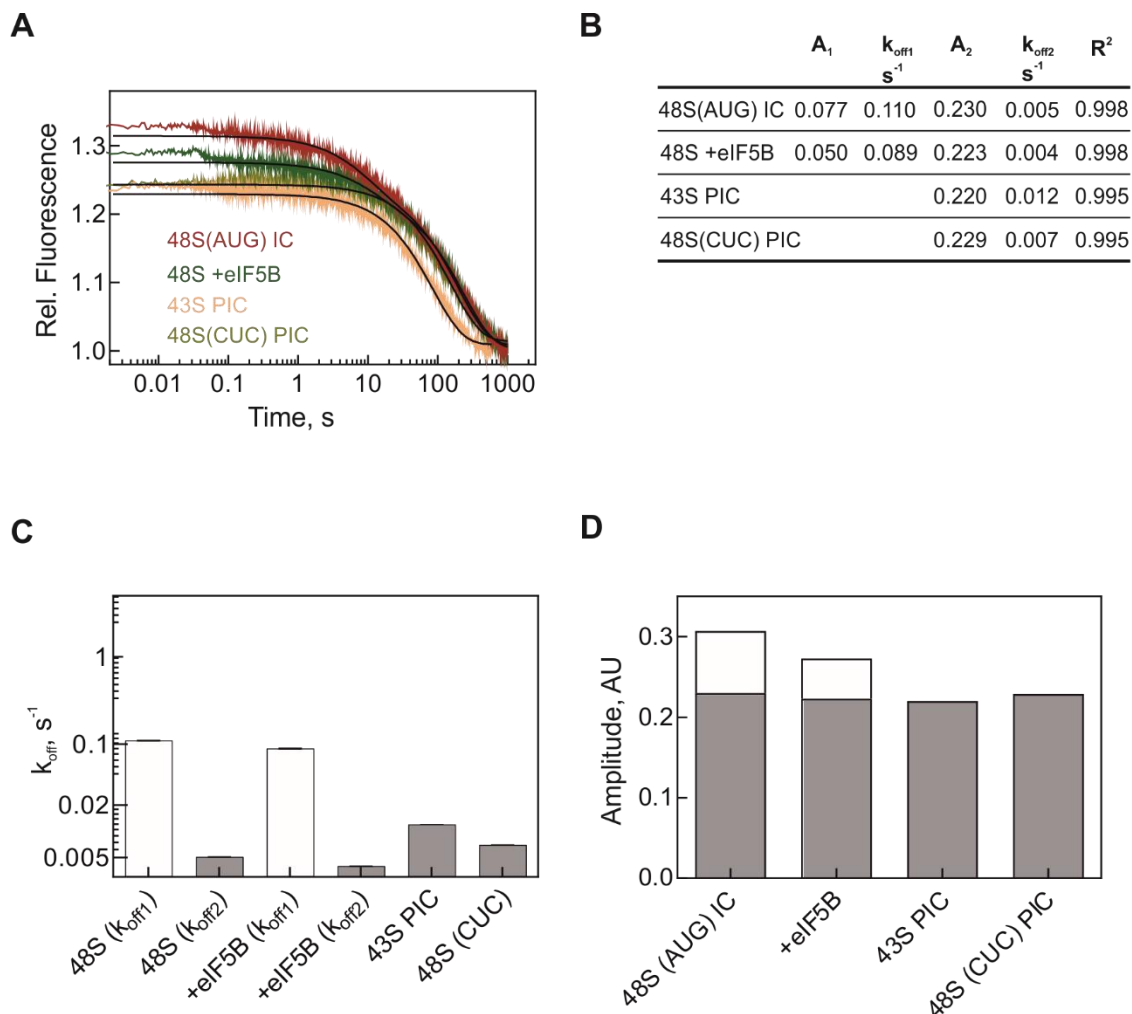


Figure 37. Codon recognition regulates the stability of eIF1A binding to the 48S IC

(A) Kinetics of dissociation of eIF1A from the 40S subunit in the presence of different combinations of initiation factors. The time courses were analyzed by one- or two-exponential functions. 48S (CUC) PIC is 48S complex formed in the presence of the mRNA with a CUC codon replacing the AUG start codon. **(B)** Summary of the k_{off1} and k_{off2} values and the fluorescence amplitudes A_1 and A_2 , respectively, from exponential fittings. R^2 is the coefficient of determination of the corresponding exponential fitting. **(C, D)** Bar charts of the k_{off} values and the respective fluorescence amplitudes. The states that stabilize eIF1A on the complexes are marked in grey, while the states with faster dissociation rates are marked in white. All SEMs of the k_{off} s are below 0.001. AU, arbitrary unit.

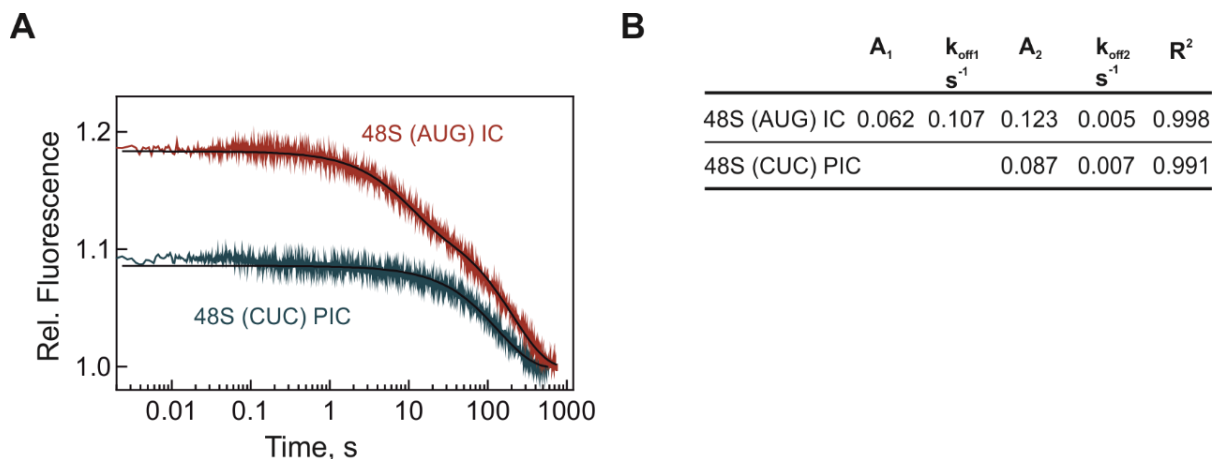


Figure 38. eIF1A dissociation from the 48S (CUC) PIC and 48S (AUG) IC is independent of AUG codons in the 3' UTR

(A) Time courses of eIF1A dissociation from 48S (CUC) PIC and 48S (AUG) IC. The 48S ICs are assembled with the AUG or CUC mRNAs in which the 2nd and 3rd AUGs in the 3' UTR are replaced with CUC. (B) Summary of the k_{off} values and amplitudes. The trace of the 48S (CUC) PIC is analyzed by single-exponential fitting, while the trace of the 48S (AUG) IC is biphasic as described previously. The 2nd and 3rd AUGs in the 3' UTR have little effects on the dissociation rates of eIF1A from the 48S (CUC) PIC and 48S (CUC) IC.

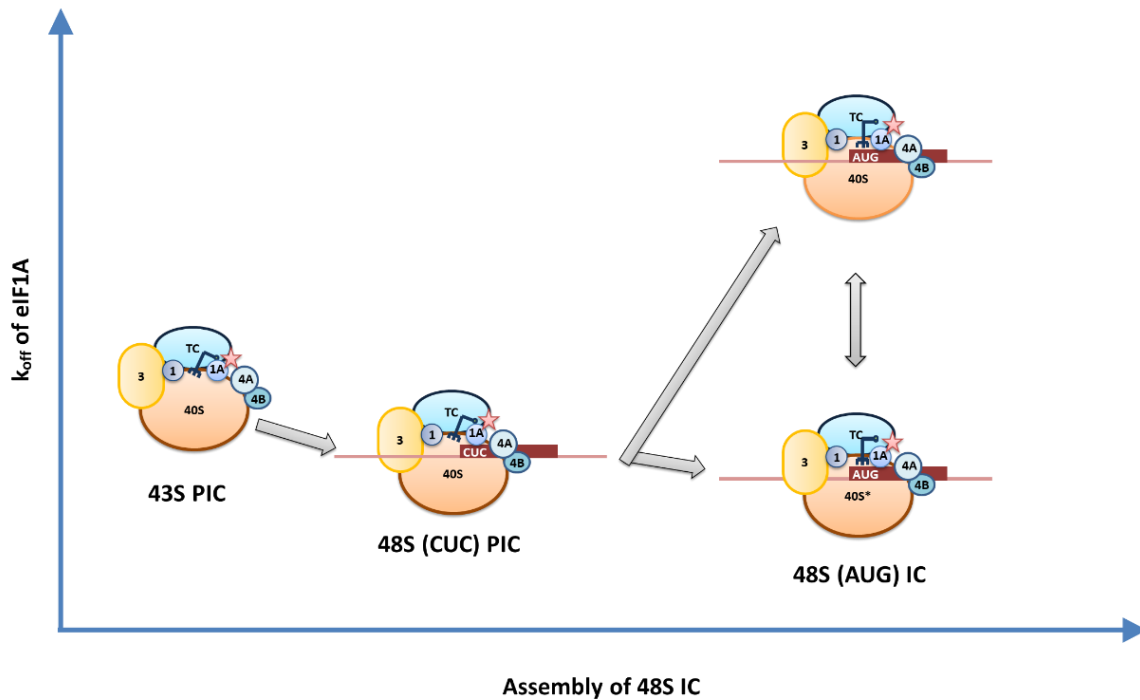


Figure 39. Hierarchy of initiation complexes *en route* to the 48S IC formation

The potential pathway of 48S IC assembly from the perspective of eIF1A. The initiation complexes of 43S PIC, 48S (CUC) PIC and 48S (AUG) IC are presented corresponding to the k_{off} of eIF1A. 40S and 40S* indicates the two states of the 40S subunits with fast and slow k_{off} s, respectively.

Exclusion of individual initiation factors from the 48S (AUG) IC has a dramatic effect on the stability of eIF1A binding. In the absence of eIF1 or eIF4A/4B, the 48S IC remains in two populations (**Figure 40**). From the toeprinting assay, excluding eIF1 or eIF4A/4B in the reconstituted 48S IC still results in a band corresponding to the 48S IC at the AUG position, albeit with a lower efficiency. Therefore, the two populations observed from the 48S-eIF1 and 48S-eIF4A/4B complexes could be functionally similar to the intact 48S IC. However, without eIF1, eIF1A dissociates significantly faster from both populations. In the previous chase experiment, eIF1 alone has been shown to stabilize eIF1A on the 40S subunit by reducing the k_{off} at least 5-fold. Since these two small proteins bind closely to each other in the decoding center, eIF1 could directly interact with eIF1A and tighten the binding of eIF1A on the 40S subunit. Thereby, the absence of eIF1 leads to faster dissociations of eIF1A from the initiation complexes. In the omissions of eIF4A and eIF4B, the two k_{off} s observed are the same as those

of the complete 48S (AUG) IC, yet, with much smaller fluorescence intensities, A_1 and A_2 . The functions of eIF4A and eIF4B are promoting the binding of mRNA to the ribosomes (Sokabe and Fraser, 2017). Since mRNA binding was shown to increase fluorescence in this complex (**Figure 26**), excluding these factors would significantly reduce the amount of 48S IC formation due to the lower efficiency of mRNA recruitment, which decreases the total fluorescence intensity (**Figure 40D**). Excluding either TC or eIF3, which are essential for 48S IC formation, diminishes the toeprints of the 48S (AUG) IC and leads to a single population of 48S IC with fast k_{off} s, 0.044 s^{-1} and 0.019 s^{-1} , for the 48S-TC and the 48S-eIF3 complexes, respectively. When the essential factors are not recruited to the ribosomes, the resulting 48S IC is not formed efficiently and eIF1A is not released rapidly from the ribosome. Since eIF1A is a pivotal factor that engages in the whole initiation process, it is important for it to stay tightly bound in the decoding center to regulate start codon recognition until the 48S IC is formed. The lifetime of eIF1A on the ribosomes is controlled largely by the dissociation rates. The slower the dissociate rate constant is, the tighter and longer the eIF1A binds to the ribosomes. Furthermore, ribosomes seem to possess the intrinsic ability to detect the absence of any essential initiation factor, initiator tRNA and mRNA and, thereby, influence the stability of bound-eIF1A. In conclusion, the fully-assembled 48S IC equilibrates between two populations that differ largely in the rates of eIF1A dissociation. One population binds eIF1A tightly, whereas the other population has a conformation that allows faster eIF1A dissociation after start codon recognition. The eIF1A-dissociation analysis of all partial complexes reveals sequential kinetic checkpoints that comprehensively regulate the formation of 48S IC (**Figure 41** and **Table 6**).

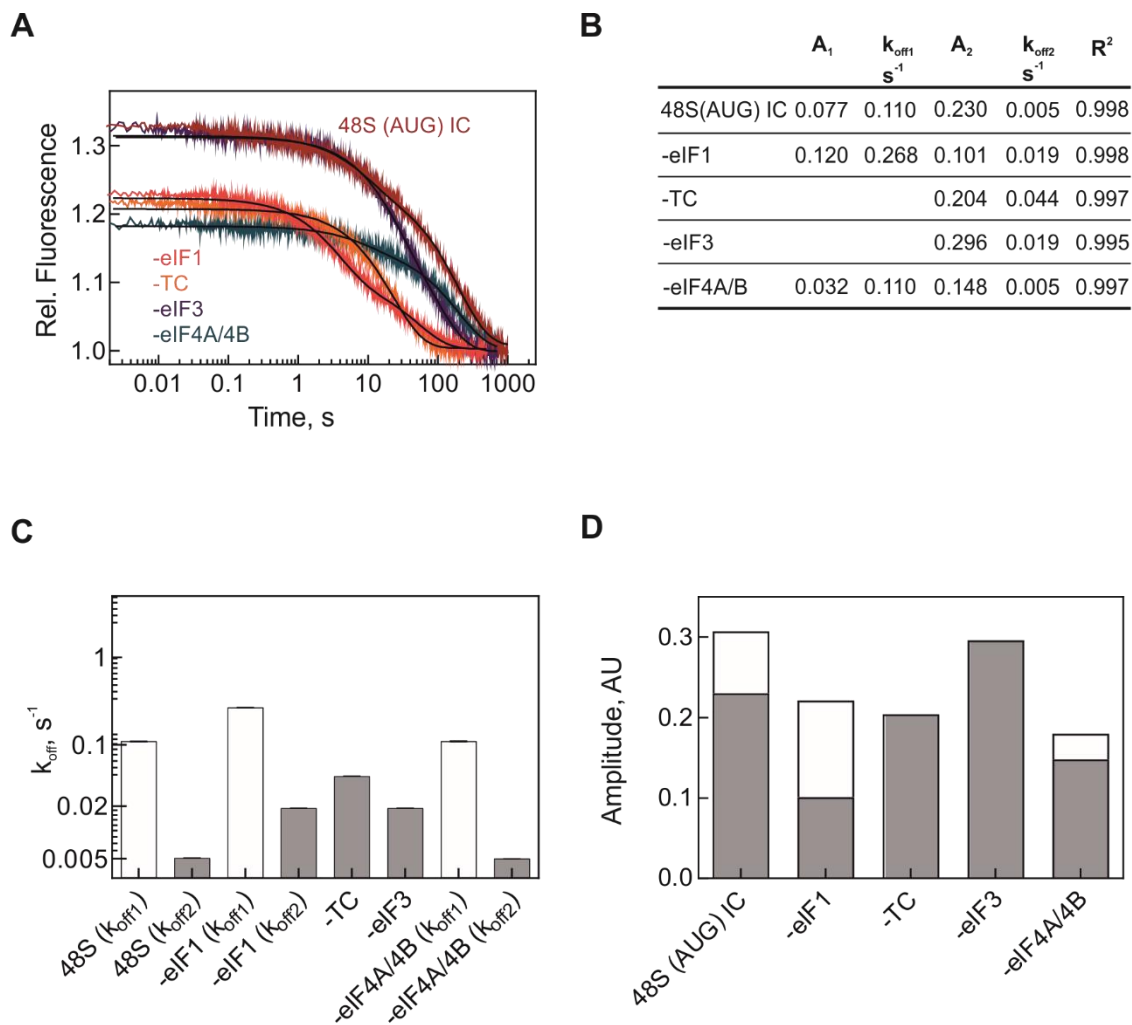


Figure 40. Role of individual eIFs in regulating eIF1A dissociation from the 48S IC

(A) Kinetics of eIF1A dissociation from partial 48S IC complexes in the presence of different combinations of eIFs. The time courses were analyzed by one- or two-exponential functions. **(B)** Summary of the k_{off1} and k_{off2} values and the fluorescence amplitudes A_1 and A_2 , respectively, from exponential fittings. R^2 is the coefficient of determination of the corresponding exponential fitting. **(C, D)** Bar charts of the k_{off} values and the respective fluorescence amplitudes. The states that stabilize eIF1A on the complexes are marked in grey, while the states with faster dissociation rates are marked in white. All SEMs of the k_{off} s are below 0.002. AU, arbitrary unit.

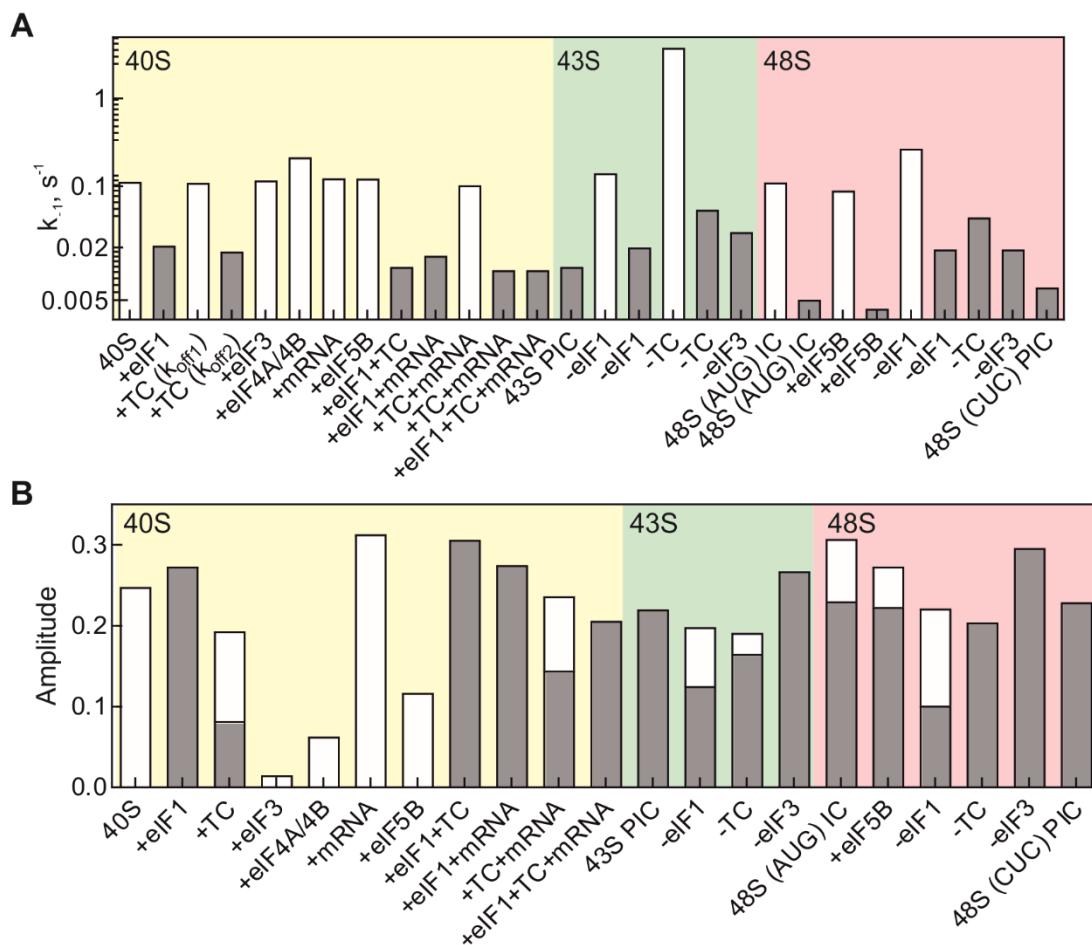


Figure 41. Summary of eIF1A dissociation rate constants from 40S subunit and various partial 43S and 48S ICs

(A) Summary of the k_{off} values of eIF1A-N4C-Alexa555 from initiation complexes. **(B)** Summary of the fluorescence amplitudes. The states with faster k_{off} are marked in white, while the states with slower k_{off} are marked in grey.

Table 6. Summary of dissociation rate constants of eIF1A-N4C-Alexa555 from initiation complexes

Complex	A ₁	k _{off1} (s ⁻¹)	A ₂	k _{off2} (s ⁻¹)	R ²	
40S	0.248	0.112			0.999	
40S+eIF1			0.273	0.021	0.999	
40S+TC	0.111	0.109	0.082	0.018	0.999	
40S+eIF3	0.015	0.116			0.963	
40S+eIF4A/4B	0.063	0.213			0.995	
40S+mRNA	0.313	0.123			0.999	
40S+eIF5B	0.117	0.122			0.999	
40S+eIF1+TC			0.306	0.012	1.000	
40S+eIF1+mRNA			0.275	0.016	1.000	
40S+TC+mRNA	0.092	0.103	0.144	0.011	0.999	
40S+eIF1+TC+mRNA			0.206	0.011	0.996	
43S PIC			0.131	0.034	0.990	*
43S-eIF1	0.073	0.140	0.125	0.020	0.999	
43S-TC	0.026	3.79	0.165	0.054	0.999	
43S-eIF3			0.267	0.030	0.998	
43S-eIF4A/4B	0.039	0.166	0.154	0.013	0.998	
43S+mRNA (48S IC)	0.068	0.112	0.127	0.006	0.996	*
48S (AUG) IC	0.077	0.110	0.230	0.005	0.998	
48S (CUC) PIC			0.229	0.007	0.995	
48S IC +eIF5B	0.050	0.089	0.223	0.004	0.998	
48S-eIF1	0.120	0.268	0.101	0.019	0.998	
48S-TC			0.204	0.044	0.997	
48S-tRNA _i			0.100	0.091	0.990	
48S-eIF3			0.296	0.019	0.995	
48S-eIF4A/4B	0.032	0.110	0.148	0.005	0.997	
48S-mRNA (43S PIC)			0.220	0.012	0.995	

The dissociation rate constants are determined from the time courses in the chase experiments. If the traces are analyzed by two-exponential fitting, the fast k_{off} is designated as k_{off1} and the slow one is designated as k_{off2}. A₁ and A₂ are the fluorescence amplitudes of k_{off1} and k_{off2}, respectively, from the fitting. R² is the coefficient of determination to the fitting in the chase experiments. The SEM of the k_{off1} of 43S -TC is 0.06, while all other SEMs of k_{offs} are below 0.002.

* The 43S PIC and 43S+mRNA (48S IC) are identical to the complexes of 48S-mRNA (43S PIC) and 48S (AUG) IC, respectively, but the experiments were conducted with different biological preparations, which explains the small differences in the rate constants.

The binding of eIF1A-N4C-Alexa555 to the 48S IC is biphasic

To further investigate the conformational states of the 48S IC, I tested whether the binding of eIF1A to the 48S IC is biphasic. 48S (AUG) IC prepared without eIF1A was rapidly mixed with eIF1A-N4C-Alexa555 in a stopped-flow apparatus. The amount of 48S IC is 6-fold higher than that of labeled eIF1A to ensure that all eIF1A-N4C-Alexa555 binds to the 48S IC at the given concentrations and permits maximum fluorescence changes. Notably, due to the low concentrations of the purified native initiation factors and the 40S subunit, it is not feasible to perform pseudo-first order titration experiments over a large range of 48S IC concentrations. The time course of eIF1A-N4C-Alexa555 binding to the 48S IC is two-exponential, in agreement with the two-population binding-equilibrium model (**Figure 42**). The biphasic curve also suggests that the 48S IC without eIF1A equilibrates between two conformations to which eIF1A binds with two distinct association rate constants (k_{on}). Since the 48S IC can still reach the AUG in the absence of eIF1A, these two conformations might represent the complexes that have recognized the start codon as the dissociations of eIF1A from the 48S IC. In conclusion, the binding of eIF1A-N4C-Alexa555 to the 48S IC further supports the two-population binding-equilibrium model in which ribosomes pre-exist in two conformational states that behave kinetically different.

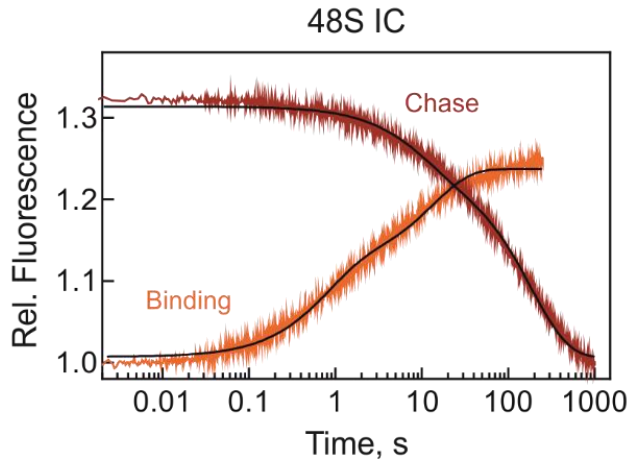


Figure 42. Binding of eIF1A to the 48S IC

Kinetics of eIF1A binding to the 48S IC in real time. The time courses are analyzed by two-exponential fitting.

2.4.3 Verification of kinetic checkpoints during initiation by fluorescent eIF1A variants S74C and T120C

In order to verify that the observed two-population binding-equilibrium model is not a result specific to the eIF1A label position, I tested two other Alexa555-labeled eIF1As, one in the OB domain at S74C and the other in the C-terminal tail at T120C. These labeled variants were reconstituted into 48S ICs for dissociation experiments from the partial initiation complexes. The Alexa555-labeled eIF1A variants at position S74C and T120C also report significant fluorescence changes during the 48S IC formation and both stimulate 48S IC formation in the toeprinting assay (**Figure 17**). The partial complexes that have been identified as the kinetic checkpoints from the eIF1A-N4C observation are selected for the dissociation experiments, which are the 40S subunit, 40S+eIF1, 40S+TC, 40S+mRNA, 43S PIC and 48S IC. In summary, the trends of the kinetic experiments are similar with eIF1A variants labeled in different positions (**Figure 43** and **Figure 44**). The total fluorescence intensities vary slightly due to the different label positions. The difference in the fluorescent amplitudes of each partial complexes also reflects characteristic microenvironments of the fluorophores at different stages during initiation. eIF1A-S74C-Alexa555 dissociates from the 40S subunit with a rate of

0.301 s⁻¹, which is reduced 10-fold in the presence of eIF1, while eIF1A-T120C-Alexa555 dissociates with a k_{off} of 0.565 s⁻¹ from the 40S subunit, and with a 4-fold slower rate from the 40S+eIF1 complex. Similarly, the presence of TC on the 40S subunit leads to two populations with distinct k_{off} values with all eIF1A variants, whereas the presence of mRNA on the 40S subunit does not alter the population distribution, albeit the k_{off} values are slightly higher with eIF1A S74C and T120C variants. The 43S PICs comprise a single population with k_{off} values 0.008 s⁻¹ and 0.0019 s⁻¹, for S74C and T120C variants, respectively. Additionally, the 48S (AUG) ICs with eIF1A-S74C-Alexa555 and eIF1A-T120C-Alexa555 all have biphasic dissociation time courses with k_{off} values similar to those of the 48S (AUG) IC with eIF1A-N4C-Alexa55. The $k_{\text{off}1}$ of the 48S IC with eIF1A variants N4C, S74C and T120C are 0.110 s⁻¹, 0.059 s⁻¹ and 0.111 s⁻¹, respectively, while the $k_{\text{off}2}$ values are 0.005 s⁻¹, 0.006 s⁻¹ and 0.008 s⁻¹. The population characterized by slow dissociation also contributes to the major changes in fluorescence intensity in all the variants tested. The total fluorescence intensity increases from 43S PIC to 48S IC upon binding to AUG mRNA with eIF1A variants N4C (**Figure 37**) and S74C (**Figure 43**) but decreases when the complexes are formed with eIF1A variants at position T120C (**Figure 44**) and D142C (**Figure 27**). The differences in the fluorescence changes in the presence of mRNA from different labeled positions support the evidences that the N- and C-terminal tails of eIF1A have different roles during start codon recognition. In conclusion, the two-population binding-equilibrium model of initiation complex formation holds true regardless of the observables on eIF1As. The observables from the NTT, the OB domain, and the CTT of eIF1A report the same population distributions of the partial initiation complexes and identify the same kinetic checkpoints along the pathway to 48S IC formation.

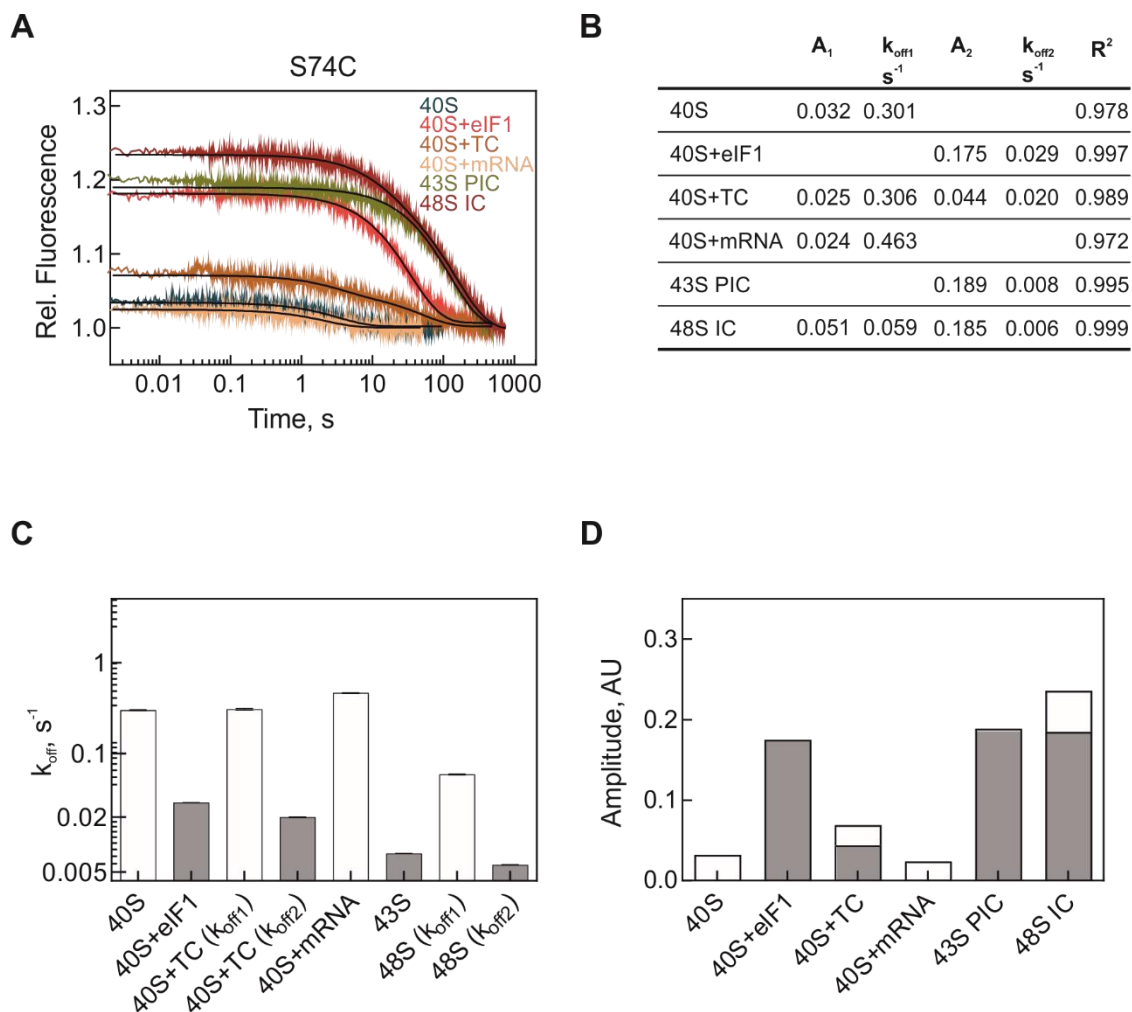


Figure 43. Dissociation of eIF1A-S74C-Alexa555 from initiation complexes

(A) Kinetics of dissociations of eIF1A-S74C-Alexa555 from the 40S subunit, 40S +eIF1, 40S +TC, 40S +mRNA, 43S PIC and 48S IC. The time courses are analyzed by single- or two-exponential functions, as appropriate. **(B)** Summary of the k_{off1} and k_{off2} are and the amplitude A_1 and A_2 , respectively, from exponential fittings. R^2 is the coefficient of determination of the corresponding exponential fitting. **(C, D)** Bar charts of the k_{off} values and the respective fluorescence amplitudes. The states that stabilize eIF1A on the complexes are marked in grey, while the states with faster dissociation rates are marked in white. All SEMs of the k_{off} s are below 0.007 in this figure. AU, arbitrary unit.

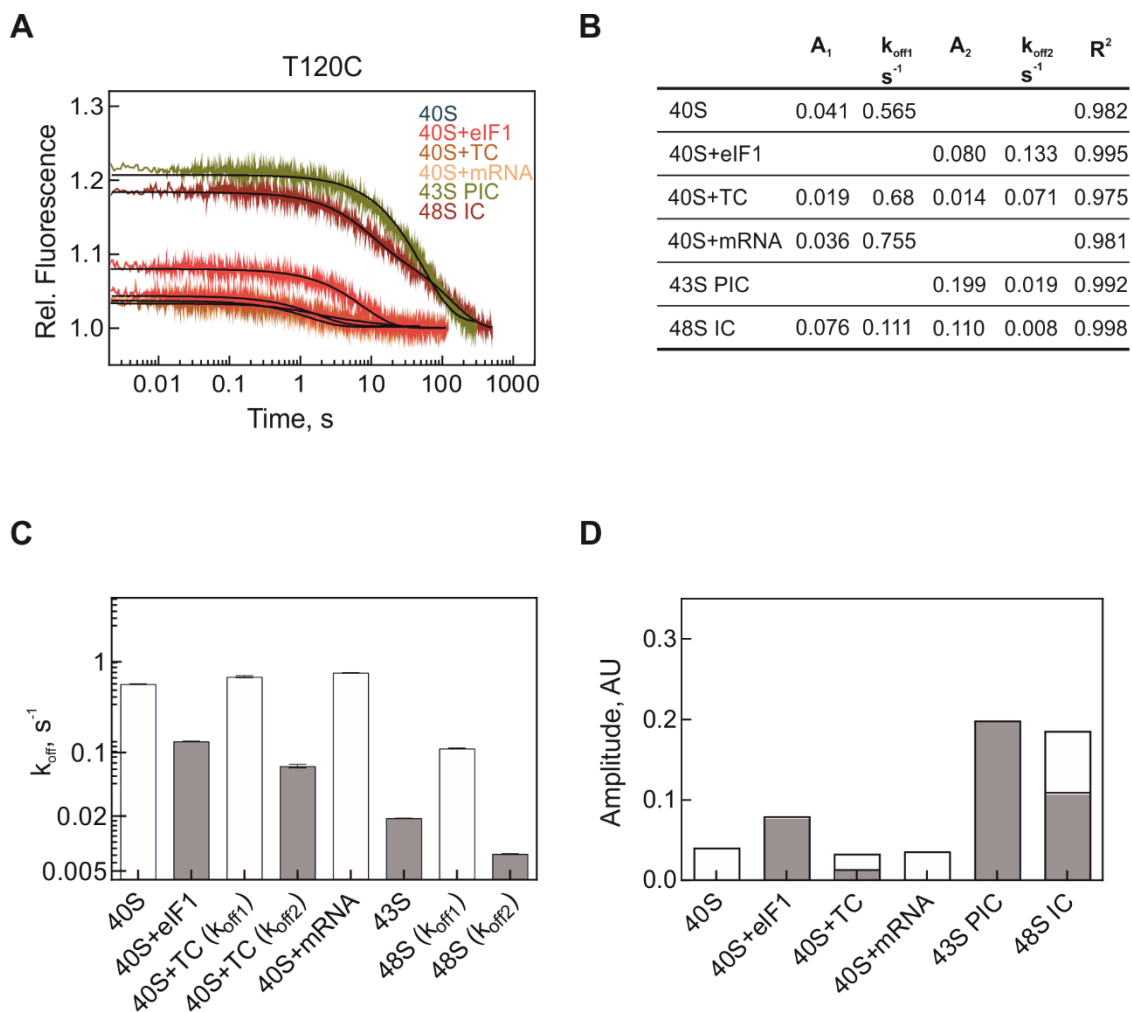


Figure 44. Dissociation of eIF1A-T120C-Alexa555 from initiation complexes

(A) Kinetics of dissociations of eIF1A-S74C-Alexa555 from the 40S subunit, 40S +eIF1, 40S +TC, 40S +mRNA, 43S PIC and 48S IC. The time courses are analyzed by one- or two-exponential functions, as appropriate. **(B)** Summary of the k_{off1} and k_{off2} are and the amplitude A_1 and A_2 , respectively, from exponential fittings. R^2 is the coefficient of determination of the corresponding exponential fitting. **(C, D)** Bar charts of the k_{off} values and the respective fluorescence amplitudes. The states that stabilize eIF1A on the complexes are marked in grey, while the states with faster dissociation rates are marked in white. All SEMs of the k_{off} s are below 0.0 in this figure. AU, arbitrary unit.

2.5 Cryo-EM structure of human 48S IC

To understand the structural arrangement of the 48S IC, I provided the complexes for a cryo-EM study performed by Dr. Erik Schliep, Dr. Niels Fischer and Valentyn Petrychenko from the research group of Prof. Holger Stark in the Department of Structural Dynamics at Max-Planck Institute of Biophysical Chemistry. Dr. Erik Schliep has solved the initial 48S IC structure at 3.7 Å resolution and Dr. Niels Fischer together with Valentyn Petrychenko have further improved the modeling of codon-anticodon base pairing. The 48S IC is reconstituted with eIF1, eIF1A, TC, eIF3, eIF4A, eIF4B, 40S subunit and the model AUG mRNA in the standard reaction buffer containing ATP and GTP. The 48S IC particles bearing occupancies of eIF3 octamer core on the solvent side of the ribosomes are sorted for the higher resolution analysis. The results showed that 10% of the complexes have an open conformation of the 40S subunit, while the other 90% of the complexes have a closed conformation of the 40S subunit (**Figure 45**). The open and closed form of the 48S IC is similar to the finding from yeast structures (**Figure 11**) (Llacer et al., 2015). However, the open and closed structures of yeast 48S IC were reconstituted independently from two 48S complexes with AUC or AUG mRNA, respectively. The observation of two conformations of mammalian 48S IC in one biochemically reconstituted complex is reported for the first time.

In the open form of the complex, the head of the 40S subunit is in an orientation that opens the mRNA binding channel. The electron density of eIF1 is found at the P site, while eIF1A locates at the A site with eIF2 β binding adjacent to it (**Figure 46**). The assignment of the electron density is further confirmed by cross-linking MS analysis in collaboration with Dr. Jasmin Corso and Dr. Andreas Linden from the Bioanalytical Mass Spectrometry research group of Prof. Henning Urlaub. In the closed form of the complex, the head of the 40S subunit moves largely towards the body of the 40S subunit, while TC accommodates into the P site, roughly 12 Å movement of the Met-tRNA_i^{Met}. The accommodation of TC results in base pairing between the codon of mRNA and anticodon of the initiator tRNA, indicating the start codon recognition in the closed conformation of 48S IC. Moreover, the density of eIF1 at the P site disappears in the closed state of the 48S IC, while some electron density is found along the TC near eIF2Y, which could be the density of eIF1 after its displacement from the P site. On the

other hand, eIF1A is still bound to the A site in the decoding center, but eIF2 β density diminishes in the closed state. Further classifications of eIF1A densities shows ribosome populations with different orientations of eIF1A at the A sites, implying that eIF1A binds to the A site in several possible conformations in the closed state (**Figure 47**). The average movement of the eIF1A density is around 10 Å between the open and closed states. Unfortunately, due to the high flexibility of eIF1A orientations, it is difficult to assign a solid binding model at this stage. The dynamic orientations of eIF1A could possibly explain the finding of two populations of 48S IC in the chase experiment that eIF1A with different orientations could have differences in k_{off} values. In the closed state, we also found that eIF1A has a stacking interaction to the Kozak sequence at +4 G position by its W70, which has also been seen in a recent structural study (Simonetti et al., 2020). In addition, the octamer core of eIF3 are clearly visible on the solvent side of the 40S subunit in both states, whereas the peripheral subunits of eIF3 are still too flexible to be assigned in a rigid model. The crosslinking MS result shows that eIF3 subunits contact the ribosome at both entry and exit of the mRNA channel as suggested from the literature (Brito Querido et al., 2020; Eliseev et al., 2018; Simonetti et al., 2020). Finally, the densities for eIF4A and eIF4B are not clearly seen on the ribosomes, but the crosslinking MS analysis suggests that they are in close vicinities with A, C, D, H, I, and L subunits of eIF3 and also contacts ribosomal RACK domain at the mRNA exit channel. In conclusion, the cryo-EM structural study also supports the structural dynamics of the ribosomes after start codon recognition. With cryo-EM technology, the steady-state conformations of the ribosomes are captured, whereas the real-time dynamics of the 48S IC are analyzed by rapid kinetic assays in this study. Both methods monitor the 48S IC in different time dimension but together support the notion of the dynamic behaviors of 48S IC upon start codon recognition.

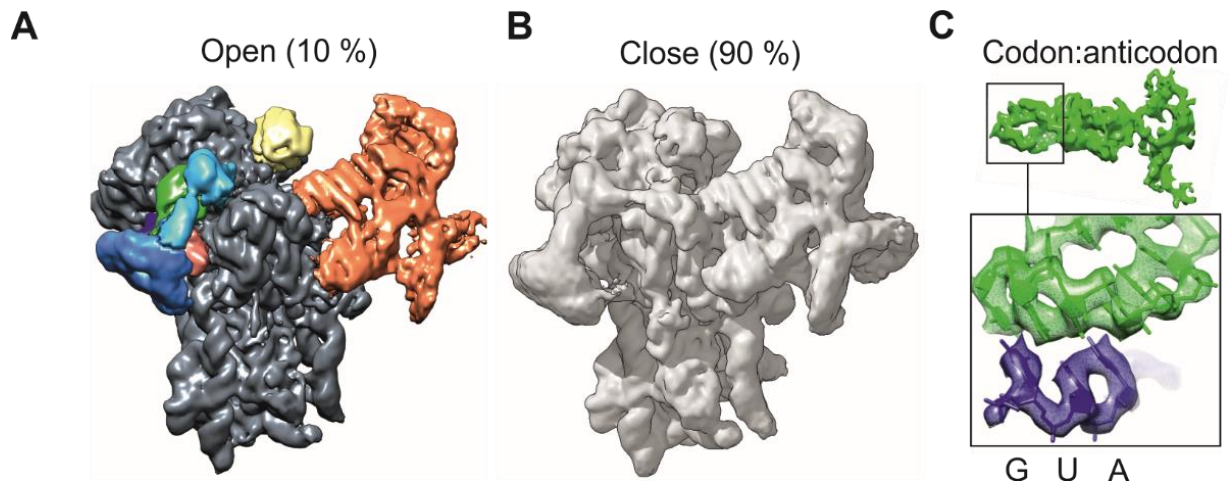


Figure 45. Overview of the Cryo-EM structure of the human 48S IC

(A) Open form of the 48S IC. **(B)** Closed form of the 48S IC. **(C)** Codon-anticodon basepairing in the closed form of the 48S IC. The populations of the open and closed form of the 48S IC are 10% and 90%, respectively. The figure of 48S IC structure is kindly provided by Dr. Erik Schliep and the figure of the codon-anticodon basepairing is from Dr. Niels Fischer.

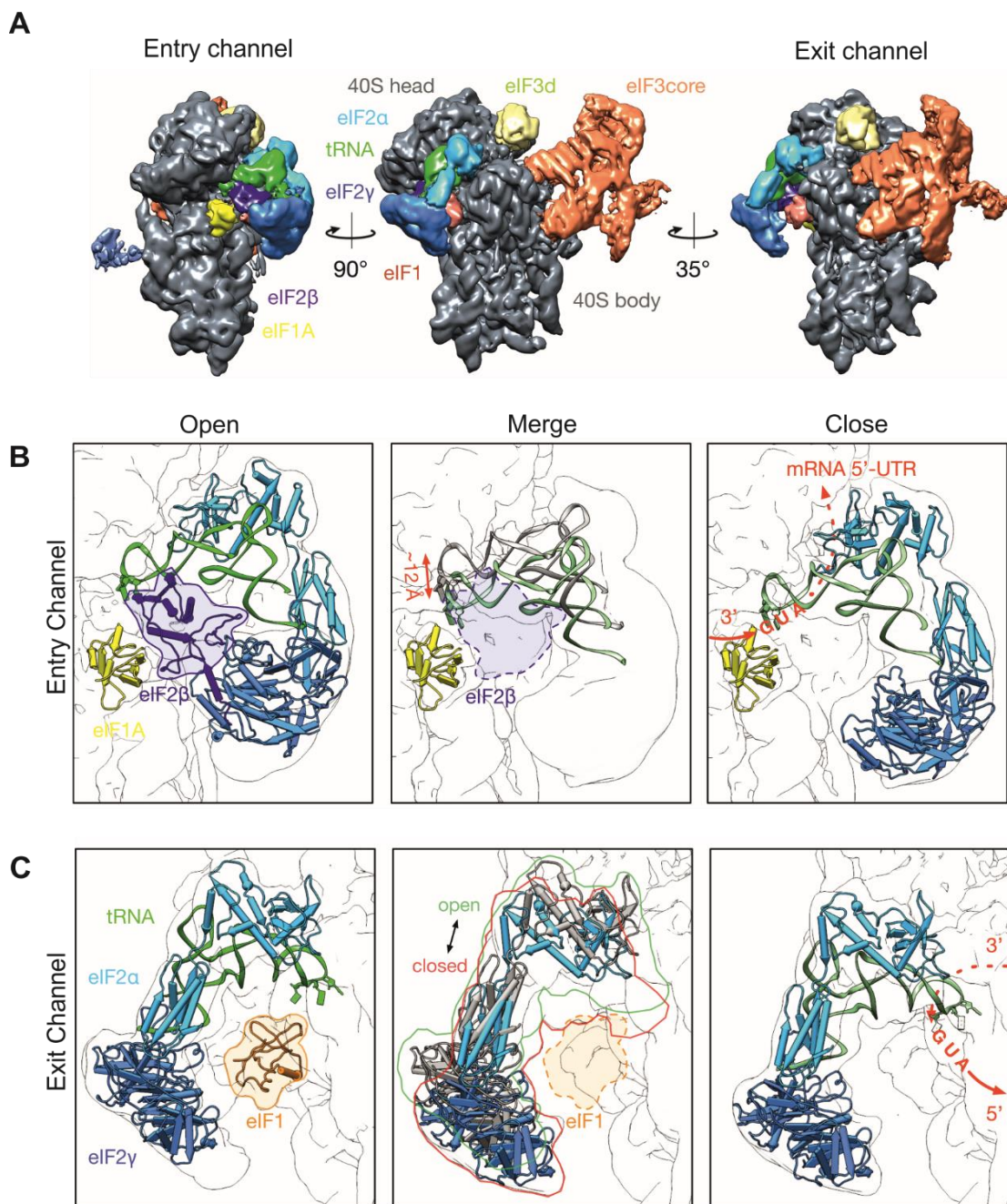


Figure 46. Structural dynamics in the decoding center

(A) Assignments of factors on the open form of the 48S IC from the entry and exit of the mRNA channel. **(B)** Comparison of the decoding center dynamic at the entry channel between open and closed form of the 48S IC. **(C)** Comparison of the decoding center dynamic at the exit channel between open and closed form of the 48S IC. The electron densities of eIF2 β and eIF1 disappear in the closed form, while TC and initiator tRNA have around 12 Å movement. For explanations, see text. The figure is kindly provided by Dr. Erik Schliep.

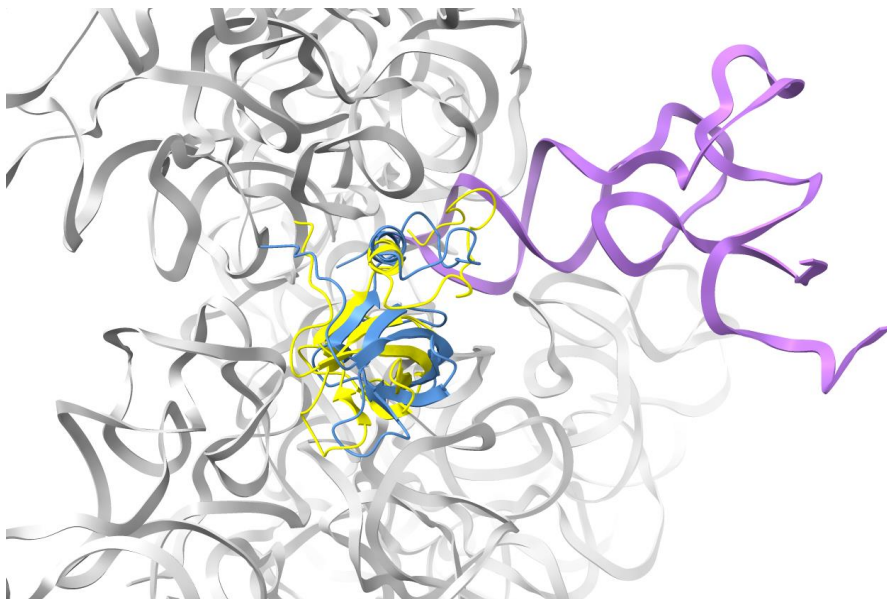


Figure 47. Binding sites of eIF1A in the open and closed form of 48S IC.

The eIF1A from the open and closed form of 48S IC is marked in yellow and blue, respectively. The OB domain has around 10 Å shift. The NTT and CTT are too flexible to be traced and were assigned arbitrary. The tRNA and the 40S subunit from the closed state are presented and labeled in purple and grey, respectively. The figure is kindly provided by Valentyn Petrychenko.

3 DISCUSSION

3.1 Reconstitution of *in vitro* mammalian translation system

3.1.1 Purification of the human initiation factors

In this study, I reconstituted the mammalian translation system *in vitro* using purified factors to study the kinetic mechanism of translation initiation. The native factors, eIF2, eIF3, eIF5B, and the ribosomes, 40S and 60S subunits, were purified from the cytoplasmic HeLa cell lysate, while single subunit factors, eIF1, eIF1A, eIF4A and eIF4B, were recombinantly expressed and purified from *E. coli*. The model mRNA and human initiator tRNA were prepared via T7-RNA polymerase *in vitro* transcription, followed by aminoacylation of tRNA_i^{Met} with [³H]-labeled Met.

The native initiation factors were purified from the polysome-bound fraction in the HeLa cell lysate instead of the free unbound factors from the cytosolic fraction. The amount of the cytosolic factors that can be purified from the same amount of cell lysate is higher than that from the polysome-bound fraction. However, due to unknown reasons, the cytosolic factors are unable to promote reconstitution of initiation complexes *in vitro*. For example, the cytosolic eIF3 is dysfunctional to promote 48S IC formation. This could be a consequence of inhibitory regulation to the cytosolic eIF3 by the post-translational modifications or binding of eIF3 inhibitors. Because ribosomes are in excess over the initiation factors in the cytoplasm (**Table 1**), it is rational for cells to adjust the overall translation efficiency by modulating the activity of the limiting initiation factors, e.g. phosphorylation of eIF2 or 4E-BP. Extensive phosphorylation has been also demonstrated in human eIF2, eIF3 and eIF4G (Andaya et al., 2014). Additionally, eIF3 has been shown to activate or suppress translation on specific growth-related mRNAs (Lee et al., 2015). It is very likely that some unknown regulatory proteins binds to the cytosolic eIF3 and changes its availability and activity for the 48S IC formation.

3.1.2 Assembly of the mammalian initiation complex

Mammalian initiation complex is reconstituted at optimized conditions in Hepes buffer pH 7.5 supplemented with 95 mM KOAc, 3.75 mM MgOAc, 1 mM ATP, 0.5 mM GTP, 0.25 mM Spermidine and 2 mM DTT at 37 °C. The toeprinting assay reveals that the 48S and 80S IC assemble on the AUG codon of the model mRNA (**Figure 13B**). Upon addition of eEF1A, eEF2 and aa-tRNAs, ribosomes form peptide bonds and translocate towards the 3' end of the mRNA. Further toeprinting assays of the stepwise elongations with addition of single tRNA at the defined step confirms that the movement of the ribosomes is due to the translocations of the ribosomes after each round of the peptide bond formation, indicating a full functionality of the reconstituted translation system (**Figure 13C**).

By omitting the individual initiation factors from the reconstitution system, I found that the minimum set of factors essential for the reconstituted 48S IC consists of eIF1, eIF1A, TC, and eIF3; for the 80S IC formation, eIF5B and 60S subunit are needed in addition (**Figure 14**). Because the model mRNA used in this study has an unstructured 5' UTR with CAA triplet repeats, the cap-binding complexes, eIF4F, and other eIF4-binding factors are not required for the recruitment of the 40S subunit onto the mRNA. However, the presence of eIF4A and eIF4B still increases the efficiency of initiation complexes assembly. Therefore, I always included them in the reconstituted reactions. On the other hand, eIF5 is considered non-essential in the reconstituted system because GTP hydrolysis by eIF2 is not required for the 48S IC formation (**Figure 18**), consistent with the early study in the mammalian reconstitution system (Pestova and Kolupaeva, 2002). Notably, preparations of native eIF5B have trace amount of carry-over eIF5 analyzed by mass-spectrometry. Since eIF5B is not an essential component for 48S IC reconstitution, the reconstituted 48S IC in this study is considered as an eIF5-free condition.

In conclusion, I reconstituted mammalian 48S IC with eIF1, eIF1A, eIF2, eIF3, eIF4A, eIF4B, 40S subunit, the model mRNA and Met-tRNA_i^{Met}. The 80S IC is formed with additional eIF5B and 60S subunit. To achieve higher reconstitution efficiency, eIFs are added in excess of 40S subunit. Typically, 150 – 500 nM of the 40S subunits are mixed with a 3-fold excess of recombinant eIF1, eIF1A, eIF4A, eIF4B, mRNA and tRNAs as well as a 1.5-fold excess of native

factors, eIF2, eIF3, eIF5B and the 60S subunits. Higher factor concentrations were not feasible in my experiments due to low stock concentration of the preparations. Since the binding affinities of the individual initiation factors to the ribosomes are in the nanomolar range, the concentration of the 40S subunit is probably higher than K_d s of most of the factors. Furthermore, initiation factors cooperatively increase their affinities to the 40S subunit in the presence of other factors (Sokabe and Fraser, 2014); therefore, a 1.5-fold excess of native factors is sufficient to promote a formation of biochemically homogenous initiation complex formation. The fluorescent intensities of the initiation complexes do not further change upon additions of native factors higher than 1.5-fold excess, indicating the completion of the initiation complex formation. Last, the concentrations of Met-tRNA^{Met}_i and model mRNA are 3 times higher than the 40S subunit to ensure homogenous TC formation and to saturate mRNA binding to the 40S subunits. Compared to yeast (Llacer et al., 2018) and other mammalian reconstitution systems (Kumar et al., 2016; Sokabe and Fraser, 2017), where the 48S IC is reconstituted with 1-fold of the 40S subunit (in 120 – 650 nM) in the presence of 2 - 3-folds of recombinant proteins and 1 - 2-folds excess of native protein, the condition used in this study ensures high efficiency and homogeneity of 48S IC formation.

3.1.3 Start-codon selection

The selection of the start codon is regulated in a sophisticated way by the *cis*-elements of the Kozak sequence flanking AUG (or a near-cognate codon) and the *trans*-factors, most important of which are eIF1 and eIF1A. The Kozak sequence is a strong determinant to the selection of the translation start site. In the presence of all essential initiation factors, the reconstituted 48S IC is able to distinguish AUG from CUC codons within optimal Kozak sequence context (**Figure 15**). The first AUG in an optimal Kozak sequence is the preferred translation start site. However, in the absence of either eIF1 or eIF1A, the selection of the start codons on the model mRNA depends on the strength of the near-cognate codons, Kozak sequence and the orders of the codons on the mRNA. On the CUC model mRNA, with the first AUG codon replaced by CUC, the 2nd and 3rd AUG codons in the sub-optimal contexts are then

selected by the fully assembled 48S IC, with a higher efficiency on the 2nd AUG due to its position and stronger Kozak sequence. In the absence of eIF1A, the ribosomes are less prone to scanning; therefore, most of the 48S IC are detected at the 2nd AUG only. In the absence of eIF1, start codon selection is less stringent. If the ribosome encounters a near-cognate codon, e.g. AUC, it can start translation at that codon before reaching the downstream AUG codon. This finding agrees with the functions of eIF1 and eIF1A in the fidelity of start codon selection and scanning, respectively.

The strength of the near-cognate codons also determines the usages of the start codon in the optimal Kozak sequence. After replacing the near-cognate codons with the AUG in the reporter ORF flanked by the optimal Kozak sequence, the strength of the near-cognate codons can be divided into two groups based on their selections by the 48S IC in the presence of eIF1 (**Figure 16**). The strong near-cognate codons AUC, UUG and GUG are selected by the 48S IC even in the presence of eIF1, whereas medium strength near-cognate codons AUU, AUA and CUG are used in the absence of eIF1 only. This finding is at variance with previous *in vivo* studies where CUG was the most efficient near-cognate codon followed by GUG, ACG and AUU (**Table 3**) (Kearse and Wilusz, 2017). In our reconstituted system, the *in vitro* transcribed Met-tRNA^{Met} is the only tRNA available for initiation. In contrast, in the previous studies, the assays were conducted in lysates which contain other tRNAs that could act to initiate translation and a mixture of regulatory cytoplasmic proteins. Interaction of unconventional initiation factors with these aa-tRNAs, e.g. eIF2A with Leu-tRNA, which recognizes CUG codons with ~40% higher efficiency than Met-tRNA^{Met} (Komar et al., 2005; Starck et al., 2016), might have contributed to the selection of the start codon *in vivo*.

The Kozak sequence affects the selection of near-cognate codons more than the position of the start codon. In the model AUC mRNA, the AUG in the reporter ORF is replaced by AUC. In the same mRNA, there is another AUC in the 5' UTR with -3 A and +4 C, which is a sub-optimal Kozak sequence, whereas the AUC in the reporter ORF is flanked by the optimal Kozak sequence, -3 A and +4 G. In the presence of eIF1, the 1st AUC in the sub-optimal context is not selected by the ribosomes, whereas the AUC in the optimal Kozak sequence is selected by a small portion of the ribosomes, but still less efficiently than the downstream AUG. The

AUC is mainly selected by the ribosomes in the absence of eIF1, which is consistent with the fidelity control function of eIF1 (**Figure 16**). Furthermore, in the absence of eIF1, the position of AUC becomes the main degenerating factor in the selection of the start codon by the ribosome in the presence of eIF1A. Because eIF1A promotes scanning, in the presence of eIF1A the ribosomes still have the ability to scan to the 2nd AUC in the optimal Kozak sequence. In contrast, the 1st AUC is selected in the absence of eIF1A.

This finding is generally consistent with the previous results from the mammalian reconstituted system (Lomakin et al., 2006; Yu et al., 2009). In the absence of eIF1, the toeprints of the 48S IC are found at the both upstream near-cognate AUU/GUG codon and downstream AUG codons, while eIF1 also increases the selection of AUG codon in a good Kozak sequence by the 48S IC (Lomakin et al., 2006). Additionally, eIF1A enhances the 48S IC to reach the downstream AUG codon and bypass the upstream GUG near-cognate codon in a weak Kozak sequence (Yu et al., 2009). In conclusion, eIF1 is the key player controlling the fidelity of AUG recognition, while eIF1A mostly promotes scanning. In the absence of eIF1, the selection of the translation start sites largely depends on the strength of the Kozak sequence in cooperation with the scanning-promoting function of eIF1A, which stimulates the ribosomes to scan through the mRNA until it reaches an AUG or near-cognate codon in the optimal Kozak sequence. With full 48S IC machinery, the determinants of the start codon selections begin with the strength of the codon (AUG is better than others), the Kozak sequence and then the order of the codons.

3.2 Fluorescence-labeled eIF1A as a reporter

To monitor the initiation complex assembly in real time, I labeled eIF1A with Alexa555 fluorescent dye at position N4C, S74C, T120C and D142C, which are in the N-terminal tail, OB domain and C-terminal tail (both T120C and D142C), respectively. eIF1A-D142C variant (labeled with fluorescein) is also used in the previous equilibrium fluorescence analysis (Fraser et al., 2007; Sokabe and Fraser, 2014), where the K_d value of eIF1A binding to the 40S subunit are 19 and 62 nM in two independent studies. Notably, yeast eIF1A was usually labeled at the

C-terminal end with additional dipeptide Cys-Lys to either a rhodamine dye (Maag and Lorsch, 2003) or a fluorescein dye (Saini et al., 2014). Since the N-terminal and C-terminal tails of eIF1A have distinct functions during initiation, addition of a dipeptide at the C-terminal end might interfere with its function. Therefore, in this study, I only used the eIF1A variant with single cysteine mutation for the fluorescence labeling.

All four labeled eIF1A variants report fluorescence changes during the stepwise formation of the 48S IC and 80S IC (**Figure 27**). The fluorescence intensity of all labeled eIF1A variants increases upon binding to the 40S subunit, with fluorescence intensity increasing by 15-70%. Because all labeled eIF1As have high affinity to the 40S subunit (**Figure 25**), the difference in the fluorescent intensities reflects directly the changes of the microenvironment in different parts of eIF1A upon binding to the 40S subunit at the given concentrations. Similarly, binding of eIF1 to the 40S subunit leads to the increase of the fluorescence intensity at position N4C, S74C and T120C, and decrease of the fluorescence intensity at position D142C. Fluorophores at N4C and T120C also reports dramatic fluorescence changes upon binding of eIF3 and TC, respectively. Furthermore, the fluorophore at position N4C induces significant fluorescence increase upon addition of the model AUG mRNA, which agrees with the function of N-terminal tail in stabilizing the tRNA_i in the P_{in} state upon start codon recognition. Additions of eIF5B and 60S subunits not only cause changes in fluorescence, but also decrease the anisotropy, indicating the dissociation of eIF1A upon the 60S subunit joining. In conclusion, the fluorescence-labeled eIF1A variants reflect the assembly events and conformational changes till late stages of the initiation. This makes them suitable reporters to monitor the initiation complex assembly in real time.

3.3 Kinetic analysis of eIF1A binding to the 40S subunit

3.3.1 Two-population binding-equilibrium model of eIF1A binding to the 40S subunit

The kinetics of eIF1A binding to the 40S subunit follows a biphasic time course, where both k_{app} values are linearly dependent to the concentrations of the 40S subunits (**Figure 28A**).

In the chase experiment, the time course of the dissociation of eIF1A from the 40S-eIF1A complexes is a single-exponential curve. This implies that eIF1A dissociates from a single kinetically predominant population of 40S-eIF1A complexes. If the two phases of the binding time course represent a fast binding reaction followed by a slow conformation change, the first apparent rate constant (k_{app}) will follow linear concentration dependence with a larger slope (k_{on}), while the second phase would follow a hyperbolic function with a slow conversion rate constant. However, the result does not satisfy the criteria of such a model. Additionally, the amplitude change in the chase experiments is almost identical to the change in the binding curves, suggesting that the bound eIF1A dissociates completely from the predominant complex. Therefore, a two-population binding model would better describe the interaction between eIF1A and the 40S subunits, where the two phases of the binding signal reflect two bi-molecular association reactions, i.e. eIF1A binds to two populations of the 40S subunits. Because a single-exponential curve is observed in the chase experiment, this suggests an equilibrium of the 40S-eIF1A complexes toward a predominant population after eIF1A binding. Therefore, I proposed a two-population binding-equilibrium model, where the 40S subunits equilibrate between two states before and after eIF1A binding. I have also found a similar result when eIF1A binds to the 40S subunit in the presence of eIF1 (**Figure 28B**), indicating a universal observation of the dynamics of the 40S subunit.

The two-population binding-equilibrium model assumes an equilibrium between two conformations of the 40S subunits before and after binding to the eIF1A (**Figure 29**). Binding of eIF1A stabilizes one of the conformations and shifts the equilibrium of the 40S-eIF1A complexes to a more thermodynamically-favored conformation (**Figure 30**). Therefore, the dissociation of the eIF1A in the chase experiment occurs from a single state, resulting in a monophasic dissociation time course. The presence of eIF1 on the 40S subunit alters neither the equilibrium of eIF1-40S complexes, nor the association rate constants of eIF1A to the 40S subunit. However, the k_{off} of eIF1A from the 40S subunit is 5-fold slower in the presence than that in the absence of eIF1 (**Figure 31**). This suggests that binding of eIF1A stabilizes the 40S subunit into a thermodynamic-favored state and eIF1 further enhances this state by slowing the dissociation of eIF1A from the 40S-eIF1-eIF1A complex. The dynamic structures of the 40S

subunit have also been shown in structural studies (Passmore 2007). Binding of eIF1 and eIF1A stabilized a conformation of the yeast 40S subunit from a closed state to an open state with an open mRNA binding channel. Recent crystal structures of 40S-eIF1-eIF1A and 40S-eIF1 complexes from *Tetrahymena thermophila* also suggest that binding of eIF1 on the 40S subunit induces a conformational change of the 40S subunit (Rabl et al., 2011; Weisser et al., 2013), providing a preformed binding site to the eIF1A and thereby, increasing the affinity of eIF1A to the 40S subunit. The structure also revealed that the eIF1A bridges the head and body of the 40S subunit at the A site (Lomakin and Steitz, 2013; Weisser et al., 2013), which might further promote the scanning and the start codon recognition. In line with the two-population binding-equilibrium model, the 40S subunit may equilibrate between two dynamic conformations and binding of eIF1A stabilizes one of the states.

3.3.2 Binding rates of eIF1A to the 40S subunit in cells

To estimate the actual rate of a bi-molecular reaction, we need not only the respective rate constant, but also the concentration of the components in the cell. If one of the two reactants is present in a large excess of the other, the reaction rate can be calculated by multiplying the concentrations of the excess reactant by the bi-molecular rate constants. In HeLa cells, the concentrations of the ribosomes and eIF1A are around 10 μM and 1 μM , respectively (**Table 1**). According to the estimations based on the two-population binding-equilibrium model, 40S subunits exist in two distinct populations at a ratio 32% to 68%, i.e. the concentrations of the excess reactant are 3.2 μM for the 40S* state and 6.8 μM for the 40S state (**Figure 30**). Therefore, the maximum binding rates of eIF1A to both 40S populations are 96 s^{-1} for the 40S* state and 34 s^{-1} for the 40S state. Considering the initiation rates of individual transcripts in yeast cells are in the range of 0.06 s^{-1} to 0.24 s^{-1} with a highest probable rate at 0.1 s^{-1} (Sharma et al., 2019), these rates are sufficiently high to be compatible with the rate of protein synthesis in cells, but are not close to the diffusion limit (100 – 1000 $\mu\text{M}^{-1}\cdot\text{s}^{-1}$). Fast-binding rates of eIF1A to the 40S subunit would ensure that the recycled 40S subunit from previous rounds of translation are saturated by eIF1A quickly after the 40S subunits are

released from the 80S ribosomes upon recycling. Binding of eIF1A to the 40S subunit would also facilitate rapid binding of other factors due to the cooperative interactions, e.g. eIF1, which would also further decrease the dissociation rates of eIF1A from the ribosomes. This may drive the ribosomes towards the formation of complete initiation complexes.

3.3.3 Binding affinity of eIF1A to the 40S subunit

The binding affinity can be determined by the steady-state binding assay in the fluorometer or calculated from the elemental rate constants. The binding affinity of eIF1A to the 40S subunit is 8.2 ± 0.5 nM, as determined from the equilibrium binding assay (**Figure 25A**), and 7 ± 2 nM, as calculated from the elemental rate constants with equation ($K_d = (k_{-3}/k_{+3}) * (k_{-2}/k_{+2})$) (**Figure 29** and **Table 5**). The values corroborate well with each other, providing support to the validity of the proposed kinetic model. In the presence of eIF1, the affinity of eIF1A increases to 5 ± 0.3 nM in the equilibrium measurement (**Figure 25A**), which is likely the upper limit for the rate constant defined by the sensitivity of the fluorescence titration assays. The calculated K_d value of 2.6 ± 0.2 nM from the kinetic model is probably more reliable in this affinity range (**Figure 30** and **Table 5**). This also provides a good example that the pre-steady state kinetic analysis is an accurate method to describe the reaction without the concentration limitation from the equilibrium measurements. Both methods show cooperative interactions of eIF1 and eIF1A on the 40S subunit with around 2- to 3-fold higher affinities compared to the single-factor binding. Previous work has also shown that the affinity of human eIF1A to the 40S subunit increases 3-fold, from $K_d = 62$ nM to 21 nM in the presence of eIF1 (Sokabe and Fraser, 2014). The 3-fold increase in the affinity agrees with the calculated affinity from the kinetic results. Differences in the absolute K_d values between our and previous experiments might be due to the position of the fluorescence label (N4 in this study and D142 in the previous literature) or the difference in the fluorescence dye (Alexa555 in this study and fluorescein in the previous literature), or the activity of the purified factors and ribosomal subunits. Notably, in the same publication, the authors also found that the affinity of eIF1A to the 40S subunit is reduced to 260 nM in the presence of eIF3 and remains similar at 64 nM in

the presence of TC (Sokabe and Fraser, 2014). However, in my experiment, the dissociation rates of eIF1A are slower in the presence of TC and remain similar in the presence of eIF3. Assuming the binding rate is unaffected, the affinity of eIF1A to the 40S·TC complexes is, therefore, higher and the affinity to 40S·eIF3 complexes is unchanged compared to free 40S subunit. The precise association rates of eIF1A to these two complexes have to be measured in the future to find out the actual binding affinity of eIF1A to these complexes. Regardless of these differences, under physiological condition where the concentration of initiation factors and ribosomes are in the 1 μ M range (Duncan and Hershey, 1983), the 40S subunit would be saturated by factors and achieve high efficiency of 43S PIC formation.

3.4 Kinetic analysis of 48S IC assembly

The kinetic analysis of discrete eIF1A binding states at different stages during initiation reveals the dynamic structures of the ribosomes along the pathway. Since the association rates of eIF1A to the 40S subunit remain similar in the absence and presence of eIF1, the stability of eIF1A binding to the 40S subunit and the lifetime of the 40S·eIF1A complex are mainly determined by the dissociation rates. With this in mind, I systematically measured the dissociation rate constants of eIF1A from all the potential partial initiation complexes. Since eIF1A is one of the essential factors throughout the initiation process, it is also expected that the longer the eIF1A binds to the 40S subunit, the more stable the partial initiation complexes are. In summary, eIF1 and TC alone have the strongest effect on the k_{off} values of eIF1A and stabilize eIF1A on the 40S subunit for longer lifetime. However, TC alters the distributions between two eIF1A-bound 40S subunits, while eIF1 stabilizes dominantly one of the states. The partial initiation complexes that assemble *en route* to the 43S PIC and 48S IC vary between one or two eIF1A-dissociation states depending on the combinations of initiation factors bound on the 40S subunit. The dissociations of eIF1A from the partial initiation complexes are generally slower than the 40S·eIF1A complex as more factors are bound to the 40S subunit. The 43S PIC and 48S (CUC) PIC are both stable complexes and comprise a single thermodynamically favored state before the start codon recognition. Once the AUG mRNA is

recruited to the initiation complexes, the start codon recognition event induces a rearrangement of the ribosomes, resulting in two kinetically distinct populations.

3.4.1 Scanning and the dissociation rate of eIF1A

While the dissociation rate of eIF1 from the 48S IC controls the fidelity of start codon selection, the dissociation rate of eIF1A probably determines the scanning capability of the ribosomes. Both *in vivo* and *in vitro* reports suggest that fast-dissociating eIF1 mutants increase the selection on the near-cognate codons, while slow-dissociating eIF1 variants enhance the selection of AUG (Algire et al., 2005; Cheung et al., 2007; Martin-Marcos et al., 2014; Nanda et al., 2013). Genetic and thermodynamic experiments show that eIF1A promotes ribosome scanning (Fekete et al., 2007; Maag et al., 2006; Saini et al., 2010). Consistent to this, our data show that when two AUGs are both in suboptimal Kozak sequence, the absence of eIF1A from the 48S IC preferentially favors the selection of the upstream AUG codon (**Figure 16**), while the downstream AUG is only selected by the ribosomes in the presence of eIF1A. Therefore, I hypothesized that the dissociation rate of eIF1A is linked to the scanning competency of the ribosomes before the start codon recognition. Scanning is the critical process for searching the translation start site; hence, eIF1A must stay tightly bound in the decoding center to promote the scanning. The 43S PIC comprise a single population of complexes with a stably bound eIF1A (**Figure 37**); this complex is ready for mRNA binding and scanning. Upon addition of a CUC mRNA to the 43S PIC, the 48S (CUC) PIC forms and may constantly scan along the mRNA towards the 3' end. Therefore, the 48S (CUC) PIC resembles the scanning ribosome. The start codon recognition on an AUG mRNA leads to the formation of the 48S (AUG) IC. The 48S (AUG) IC has two distinct dissociation states of eIF1A. The slow k_{off} value remains similar before and after codon recognition (with 48S (AUG) IC (0.005 s^{-1}) and 48S (CUC) PIC (0.007 s^{-1})) (**Figure 37**). The fast-dissociation state (0.110 s^{-1}) is 22-fold faster than the slow state (**Figure 37**) and the value is similar to the rate of eIF1A from the vacant 40S subunit (0.112 s^{-1}) (**Figure 32**). Unfortunately, the distribution of the fast- and slow-dissociation eIF1A populations remains unknown. The slow-dissociation population possibly

resembles the ribosomes that reach the Kozak sequence and stay in a state prior the actual base-pairing between the codon and anticodon, while the fast-dissociation state may comprise ribosomes that have AUG base-paired with the anticodon in the P site. Since eIF1A still participates in the later stages of the initiation, e.g. facilitates eIF5B binding and the large subunit joining, rapid dissociation of eIF1A after start codon recognition does not suggest its actual release from the initiation complex. In fact, the anisotropy of the dye attached to eIF1A remains high, indicating eIF1A stays bound on the 48S IC (**Figure 26A**). Addition of eIF5B to the 48S IC does not change the kinetics of eIF1A dissociation (**Figure 37**), implying that binding of eIF5B does not trigger the actual dissociation of eIF1A either. eIF1A anisotropy decreases significantly only after the 60S subunit joining, which more likely indicates the release of eIF1A from the IC (**Figure 26A**). Therefore, the fast dissociation state of eIF1A on the 48S IC might reflect a different orientation or conformation of eIF1A on the ribosomes, which is not as tightly bound to the decoding center as its state during scanning. To further verify the potential reasons for changes in eIF1A dissociation rates after start codon recognition, single molecular fluorescence resonance energy transfer (smFRET) experiments should be employed to directly monitor the positions of eIF1A in respective to other components in the decoding center before and after start codon recognition.

3.4.2 Comparison of the eIF1A-dissociation states of the initiation complexes in lower and higher eukaryotes

Our finding is similar with earlier studies on the dissociation of eIF1A from yeast 48S IC. They also report two distinct states of 48S IC that differ in the rates of eIF1A dissociation and are interconvertible (**Table 2**) (Llacer et al., 2018; Maag et al., 2006). However, the authors assigned the fast dissociation state as open form of the ribosome and the slow dissociation state as closed form (Llacer et al., 2018; Maag et al., 2006). They also showed that eIF1A dissociation is generally faster from yeast 48S (UUG) IC (the rates of dissociation from the two populations are 0.022 s^{-1} and 0.0021 s^{-1}) than from the 48S (AUG) IC (0.006 s^{-1} and 0.0004 s^{-1} , respectively) (Llacer et al., 2018). The ratio of the two states increases in favor of the slow

dissociation state (closed state) in the 48S (AUG) IC upon start codon recognition and in the presence of eIF5. The authors interpreted that the codon-anticodon interaction enhances the closed structure of the ribosome and eIF5 shifts the equilibrium towards closed state. The yeast scanning ribosome, 48S (CUC) PIC, also has two populations with higher ratio of the fast state, indicating the major population is in the open structure during scanning (Maag et al., 2006). Notably, the dissociation rate constants of eIF1A reported from the yeast 48S (AUG) IC (0.006 s^{-1} and 0.0004 s^{-1} from (Llacer et al., 2018)) are also one magnitude slower than those we report for the mammalian 48S (AUG) IC (0.110 s^{-1} and 0.005 s^{-1}) (**Figure 37** or **Table 6**), probably due to the difference in the composition of reconstitution system used in the yeast 48S IC and in our mammalian 48S IC. The yeast 48S IC is reconstituted with GDPNP and eIF5 on a model mRNA bearing short 5' UTR in the absence of eIF3, eIF4A and eIF4B (Llacer et al., 2018). The mammalian 48S IC in this study is reconstituted with GTP on a model mRNA bearing long unstructured 5' UTR in the absence of eIF5. Because the yeast 48S IC is reconstituted with non-hydrolysable GDPNP, even in the presence of eIF5, the yeast 48S IC might be stalled at the state right before the GTP-hydrolysis event (but with start codon recognition). In contrast, the mammalian 48S IC in this study is reconstituted with GTP, where the GTP hydrolysis of TC might happen spontaneously after start codon recognition in the absence of eIF5. Notably, in the absence of eIF5, dissociation of eIF1A from the yeast 48S IC also shows two k_{off} values (0.015 s^{-1} and 0.0038 s^{-1}) in the same magnitude as that from the mammalian 48S IC (**Table 2** and **Table 6**) (Maag et al., 2006). This suggests that the presence of eIF5 in the yeast 48S IC stabilizes eIF1A in both states and has large impact on the k_{off} values of eIF1A. Furthermore, eIF3 is essential for the formation of mammalian 48S IC, which is dispensable for the reconstituted yeast 48S IC. Dissociation of eIF1A from yeast 48S IC in the presence of eIF3 (and eIF5) shows slightly higher k_{off} values (0.0068 s^{-1} and 0.00081 s^{-1}) than that of in the absence of eIF3 (**Table 2**) (Maag et al., 2006). Additionally, mammalian eIF4A and eIF4B not only have RNA helicase activity but also promote mRNA recruitment to the human ribosome (Sokabe and Fraser, 2017). Therefore, both mammalian eIF3, which has multi-functions in 48S IC assembly, and eIF4A/4B might further contribute to the difference in the k_{off} values of eIF1A between the yeast and mammalian system.

The number of states in the partial initiation complexes also differs in yeast and mammalian system. Yeast 43S PIC, 48S (CUC) PIC and 48S (AUG) IC all have two distinct states of eIF1A dissociation in the presence and absence of eIF5 (**Table 2**), whereas mammalian 43S PIC and 48S (CUC) PIC are identified as a single eIF1A-dissociation state (**Figure 37**). Notably, yeast 43S PIC only consists of eIF1, eIF1A and TC, which is actually a single-population complex in the mammalian system (**Figure 34**). In yeast, the majority of eIF1A dissociates from the fast state (0.110 s^{-1}) of the 48S (CUC) PIC (**Table 2**) (Maag et al., 2006), but in mammals, the 48S (CUC) PIC has a very slow eIF1A k_{off} (0.007 s^{-1}) that is almost similar to the slow state (0.005 s^{-1}) of mammalian 48S (AUG) IC (**Figure 37**). This implies that during scanning, the ribosome only has a single and stable eIF1A-binding state in the mammalian 48S (CUC) PIC that might differ from the open structure assigned for the yeast 48S (CUC) PIC. The dynamics of the mammalian ribosome resumes upon start codon recognition from a single state of 48S (CUC) PIC to two states of 48S (AUG) IC. Therefore, the assignment of the fast and slow states of the mammalian 48S (AUG) IC to the structural states of the ribosomes might be different from the yeast studies (see below).

In conclusion, yeast system suggests dynamic changes in the ratio of two kinetic states in 43S PIC, 48S (CUC) PIC and 48S (AUG) IC. The fast- and slow-dissociation states of these complexes refer to the open and closed structure of the ribosome, respectively. The ratio of these two states is mainly affected by the codon-anticodon interaction and the presence of eIF5. However, our results indicate that in mammalian system, the populations of the initiation complexes vary between one and two eIF1A-dissociation states and are determined by the combinations of initiation factors and the start codon recognition event. Although the exact rates of eIF1A dissociation and the assignments of the kinetic states to the structures may differ in yeast and mammalian systems, the dynamics of the 48S IC upon start codon recognition is consistent in lower and higher eukaryotes.

3.4.3 Roles of individual initiation factors during initiation complex assembly

eIF1 increases the binding affinity of eIF1A to 43S PIC and 48S IC by lowering the k_{off} of eIF1A (**Figure 35** and **Figure 40**). Previous studies using yeast components suggested that eIF1

is displaced or dissociates from the P site after start codon recognition (Nanda 2013; Ll  cer 2018). The chase experiments of eIF1A from 48S IC in the presence or absence of eIF1 suggest that even after the start codon recognition in the 48S IC, eIF1A is still stabilized by eIF1 in the complex (**Figure 40**). The timing of eIF1 dissociation remains unclear in the mammalian initiation pathway.

From our structural study, we found two states of 48S IC, the open and closed forms (**Figure 46**). The open form of the 48S IC has eIF1 density at the P site, but the codon-anticodon interaction is not observed. In the closed state, which has codon-anticodon base pairing in the decoding center, the density of eIF1 disappears from its original binding position at P site. Instead, some diffuse density is found in close proximity of the TC, which could possibly be the displaced eIF1. Since eIF1 stabilizes eIF1A in the decoding center, it makes considerable sense to assign the fast dissociation state of 48S IC to the closed form of the 48S IC structure. Because eIF1 is displaced from the P site and is relocated further away from eIF1A in the closed form upon start codon recognition (**Figure 46**), the interaction between eIF1A and eIF1 might be, therefore, weaker. The slow dissociation state would then correspond to the open form of the 48S IC, where eIF1 still binds at the P sites and stabilizes eIF1A prior to the start codon recognition. The recent structure of a scanning human ribosome, i.e. 48S (CUC) PIC, also indirectly supports this assignment (**Figure 5**) (Brito Querido et al., 2020). In this structure, the 40S subunit and Met-tRNA^{Met}_i are in a conformation that is neither the open nor the closed state of the yeast 48S IC, while eIF1A and eIF1 bind tightly at the A site and P site, respectively. Our result suggests a single population of 48S (CUC) PIC with a very slow k_{off} value, indicating that the slow eIF1A-dissociation state exists already during scanning (before start codon recognition). Therefore, the closed structure of the ribosome with codon-anticodon interaction might be associated with the fast eIF1A-dissociation state of 48S (AUG) IC that occurs upon start codon recognition. This results in opposite assignment of the two eIF1A-dissociation states of yeast 48S IC to the open and closed structures. This suggests that the dynamic of the ribosome during scanning and upon start codon recognition might differ in lower and higher eukaryotes.

Exclusion of TC from the 43S PIC and 48S IC have the most dramatic effect on the stability of eIF1A on these complexes (**Figure 35** and **Figure 40**), indicating a cooperative interaction between eIF1A and TC. When eIF3 is present, the lifetime of eIF1A on the ribosome lacking the TC decreases. TC binding and codon-anticodon interaction are key steps during initiation. Ribosomes lacking either mRNA or Met-tRNA fail to initiate translation. Notably, eIF3 plays very different roles in lower and higher eukaryotes. In contrast to mammalian 48S IC, yeast 48S IC can be reconstituted in the absence of eIF3. eIF3 not only acts as a scaffold protein that interacts with almost all the initiation factors during initiation, but also alters the dynamics to the ribosome when other factors are missing in the complexes (**Figure 35** and **Figure 40**). Even though the core of the eIF3 binds to the solvent side of the 40S subunit, some of the subunits contact the decoding center and are in close proximity to eIF1A. Binding of the eIF3 to the initiation complexes might not significantly alter the dissociation rate constants of the eIF1A, but the fluorescence intensity of the complexes decreases dramatically after eIF3 binding, indicating a different microenvironment of the eIF1A in the decoding center. Additionally, fluorophores at different parts of eIF1A also reflects different fluorescence changes upon binding of eIF3 (**Figure 27**). The fluorescence decreases when the fluorophore is at N4C, remains similar when the fluorophore is at S74C and T120C, and increases when the fluorophore is at D142C. This suggests that different parts of eIF1A have distinct interactions with eIF3 and thereby, different microenvironments around the fluorophore. From the cross-linking analysis, eIF3 also forms an extension outside the mRNA exit channel that stabilizes the binding of mRNA around the Kozak sequence region (Pisarev 2008). Studies using yeast translation system also show that mRNA with short 5' UTR is less stably bound to the ribosomes in the presence than in the absence of eIF3 (Aitken 2016). Therefore, even though the exact functions of eIF3 during initiation remains obscure, this factor has a major impact on the conformation of the ribosomes and seems to act as surveillance point to ensure the completeness of the initiation components.

3.5 The biochemical pathway of translation initiation in cells

In this study, I used the fluorescence-labeled eIF1A as a reporter to systematically analyze the lifetimes of eIF1A on the partial initiation complexes, as well as the structural dynamics during initiation complexes assembly. However, in cells, translation initiation does not start with the binding of eIF1A and other factors to the 40S subunit in a stepwise fashion, but rather from a mixture pool of initiation factors and ribosomes in the cytoplasm. Nevertheless, the dissociation analysis of eIF1A from the initiation complexes suggests a likely kinetically favored route during the initiation complex assembly. As expected, thermodynamically stable complexes would have longer lifetime in cells and, thus, would likely be the checkpoints to proceed forward in the initiation pathway. Comparing the dissociation rates of eIF1A from 40S·eIF1·eIF1A and 40S·TC·eIF1A complex, 40S·eIF1·eIF1A complex exists in a relative stable single population, while 40S·TC·eIF1A complex equilibrates between fast- and slow- dissociation states. The 40S·eIF1·eIF1A complex is, therefore, more stable than the 40S·TC·eIF1A complex. This implies that in cells, the 40S·eIF1·eIF1A complex would have a higher chance to reach the 40S·eIF1·TC·eIF1A complex in the pathway. Furthermore, *in vitro* studies in yeast system have shown that TC association to the 40S subunit is an mRNA-dependent reaction (Kolitz et al., 2009; Passmore et al., 2007). TC, more specifically the Met-tRNA_i^{Met} in the TC, accommodates to the 40S·eIF1·eIF1A complex in a slow binding and slow dissociations manner (**Table 7**). The association rate of TC to 40S·eIF1·eIF1A complex is increased over 1300-fold in the presence of mRNA, while the dissociation of TC is largely stabilized by eIF1A from 40S·eIF1A·mRNA complex. This suggests that TC would be recruited more likely to the 40S·eIF1·eIF1A complex or other partial complexes than to the 40S subunit alone.

Table 7. Kinetics of yeast TC binding to the 40S subunit in the presence of eIF1, eIF1A and mRNA

	k_{on} ($\mu\text{M}^{-1}\cdot\text{s}^{-1}$)	k_{off} (s^{-1}) (10^{-6})	k_{off}/k_{on} (nM)	Measured K_d (nM)
40S·eIF1·eIF1A	0.0042; 0.0005	25.6	60	67
40S·eIF1·mRNA	0.0028	17.0	6	18
40S·eIF1A·mRNA	0.155	<0.5	<0.003	<1
40S·eIF1·eIF1A·mRNA	5.6	5.6	0.01	<1

The binding and dissociation rates of TC from different 40S complexes. The TC is quantified from the radioactivity of the ^{35}S -Met-tRNA_i in a native gel-based assay. The time course of the binding of TC to the 40S·eIF1·eIF1A complexes is biphasic with two slow linear association rate constants. The results were taken from (Kolitz et al., 2009; Passmore et al., 2007)

Additionally, ribosomes lacking essential factors, e.g. TC, would also tend to lose other initiation factors and return to the preceding state, which prevents translation initiation with truncated complexes. 43S PIC lacking eIF1 or TC have two eIF1A dissociation states, while 43S PIC lacking eIF3 has similar eIF1A k_{off} as intact 43S PIC. Therefore, in cells, eIF3 might bind to the 43S PIC in the later stage or right before the mRNA recruitment. eIF3 alone can recognize mRNA features in the 5' UTR, e.g. m⁶A, and regulate the recruitment of the initiation machinery (Meyer et al., 2015; Zhou et al., 2015), suggesting that 43S PIC without eIF3 is recruited to the eIF3-bound mRNA in cells. The 6 peripheral subunits and octamer core of the mammalian eIF3 might contribute to the translational regulatory function of eIF3 in the mammalian 43S PIC, which is lacking in the yeast octameric eIF3.

In conclusion, there are several possible routes to assemble 43S PIC and 48S IC in cells. However, in the mixture of all the partial initiation complexes, the checkpoint complexes are more stable than others and therefore, have more chance to proceed along the initiation pathway. Using a reconstituted mammalian translation system, the k_{off} values of eIF1A from the partial initiation complexes were systematically analyzed for the first time. This study, thus, sheds light on the thermodynamic and kinetic landscape of the initiation complexes and

identifies the key components that stabilize the complexes into the checkpoints and drive the initiation reactions towards the start codon recognition.

3.5.1 Summary of the dynamic of the initiation complexes

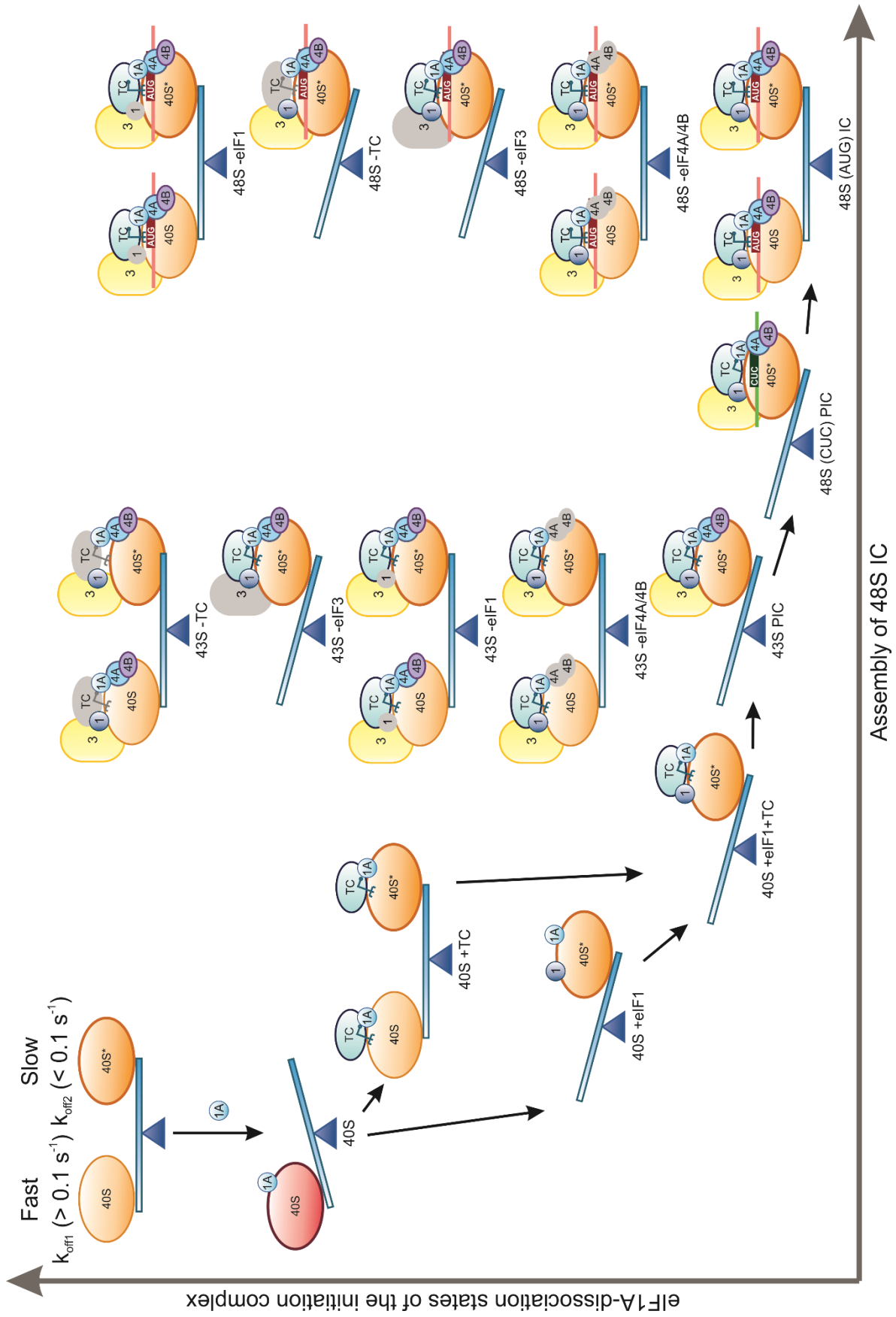
Taking all our data, we propose a model for translation initiation pathway in view of eIF1A. (**Figure 48**). Initially, free 40S subunits are dynamic and possess two kinetically distinct populations. Once the eIF1A binds to the 40S subunits, one of the populations becomes thermodynamically favored. The joining of eIF1 to the ribosomes further stabilizes eIF1A on the 40S·eIF1·eIF1A complex by decreasing the k_{off} 5-fold; hence, enhancing the thermodynamically favored state of the ribosomes and also the binding affinity of eIF1A to the 40S subunit. Joining of TC to the 40S subunit also partially stabilizes eIF1A on the 40S subunit. However, unlike eIF1, the 40S·TC·eIF1A complex equilibrate between two states. The ribosomes at this stage are not fully committed to receiving the TC because the accommodation of TC is facilitated by eIF1A together with eIF1 and other initiation factors. Therefore, eIF1A exists in two different states inside the decoding center, probably interacting with the TC in an equilibrium reaction. The recruitments of eIF1 and TC together on the 40S·eIF1·TC·eIF1A complex stabilize the ribosomes into a more thermodynamically favored state. The eIF1A has a relative long lifetime on the 40S·eIF1·TC·eIF1A complex. The completion of 43S PIC assembly is achieved upon binding of eIF3, eIF4A and eIF4B to the 40S subunit. For 43S PIC to be able to scan the mRNA, eIF1A should remain tightly bound. Even though the dissociation rate constants of eIF1A from the 40S·eIF1·TC·eIF1A complex and 43S PIC are similar in value, the fluorescence intensities of these two complexes differ remarkably from each other, implying a significant difference in the microenvironment inside the decoding center around eIF1A.

After binding of the CUC mRNA, a single population of the 48S (CUC) PIC forms and eIF1A remains tightly bound to the ribosome. The 48S (CUC) PIC resembles the scanning complex as all the essential components are recruited to the ribosome, but start codon recognition does not take place. Therefore, eIF1A is locked in the decoding center to promote

scanning. Binding of the AUG mRNA to the 43S PIC triggers the start codon recognition and leads to the formation of 48S (AUG) IC. The 48S (AUG) IC resumes its dynamic behavior and have two kinetically distinct states. Of the two k_{off} values, one is similar to the 48 (CUC) PIC and the other k_{off} is relatively fast. At this stage, eIF1A has fulfilled its function in promoting scanning, but is still required to facilitate the eIF5B binding and 60S subunits joining in the later stage. To achieve these tasks, eIF1A possibly have a new orientation/conformation or even a different binding site on the ribosomes that leads to the fast dissociation state of eIF1A from the 48S (AUG) IC. Translation initiation concludes with the eIF5B-mediated 60S subunit joining and the formation of the 80S IC.

Figure 48. Dynamic of the eIF1A-dissociation states of the initiation complex (see below)

Overview of the eIF1A-dissociation states of the initiation complex. The fast- and slow-dissociation states of the ribosomes are indicated in light and dark circle and presented on the left and right side of the seesaw, respectively. Here, the fast- and slow- states attribute to the k_{off1} and k_{off2} in **Table 6** that are over or below 0.1 s^{-1} , respectively. Therefore, the 40S subunit ($k_{\text{off1}} = 0.112 \text{ s}^{-1}$, in red circle) are indicated differently than the one in **Figure 30**. Number of the complexes on the seesaw is corresponding to the number of eIF1A-dissociation states of the complex. The leaning seesaw indicates the dominant state, while flat seesaw represents an equilibrium between two states. The initiation factors are designated with their respective number. The omitted initiation factors are colored in grey, while the ORF is presented as maroon and green box with codons AUG or CUC, respectively. The size and binding site of each factor are arbitrary.



3.6 Perspectives

In this study, I focused on the kinetic analysis of the conventional scanning mechanism of the initiation complex assembly from the perspective of eIF1A. In cells, protein biosynthesis also starts with alternative unconventional initiation pathways, e.g. near-cognate start codons, which are often employed under stress conditions. To further explore the regulation of initiation in these circumstances, different mRNAs can be designed and included into the reconstitution system. For example, to understand the control of start codon fidelity, near-cognate codon mRNAs can be used for the reconstitution of the complexes and the stability of these complexes can be measured with the dissociation rates of eIF1A. Furthermore, the regulation of IRES-mediated translation can also be elucidated through the analysis of the stability of the partial complexes. This could help potential drug design to prevent the hostages of the virus to the human translation system.

To illustrate the panoramic assembly landscape of mammalian initiation complex, more observables on the ribosomes have to be established in the future. In this study, the stability of the partial initiation complexes is measured through the dissociations of eIF1A. Similar procedures can be also applied with fluorescently labeled eIF1 since both of them are pivotal players in the decoding center and are crucial to stabilize the ribosomes during scanning and start codon recognition. Furthermore, eIF1 is suggested to be the gate keeper of the fidelity and is displaced after start codon recognition by eIF5. Predictably, eIF1 would dissociate faster on the 48S IC after start codon recognition. The strengths of each near-cognate codons can also be systematically determined with the dissociation rates of eIF1 from the initiation complexes assembled on the corresponding codons. Furthermore, even though native proteins, eIF2, eIF3, eIF4F, eIF5B, and the ribosomes are still not an easy target for fluorescently labeling, recombinant proteins, eIF1, eIF4A and eIF4B, and RNA components have relevant potentials to generate further fluorescent observables for the initiation complexes. The dynamic changes of the ribosomes during initiation can be illustrated thoroughly from different perspectives with these reporters. Moreover, with real-time GTP hydrolysis assay, the checkpoints of start codon regulation and GTP hydrolysis can be

investigated simultaneously on 48S IC with different components and near-cognate codon mRNAs.

The kinetic regulation of late stage initiation in the mammalian translation system, which occurs after start codon recognition, also remains obscure. With reconstituted system, the timing of 60S subunit joining can be measured in the stopped-flow apparatus with time courses of light scattering. The requirement of GTP by the eIF5B during the large subunit joining can also be studied with GTP analogs to identify exact timing of GTP hydrolysis in the late stage of initiation. Furthermore, the regulation of reinitiating efficiency with uORF and main ORF can be measured with adjusting the ratio of the key components, e.g. eIF3, to the ribosomes. The reconstituted mammalian translation system provides an elegant platform for studying of all biochemical events during protein biosynthesis in both equilibrium and pre-steady state conditions.

4 MATERIALS AND METHODS

4.1 Reconstitution of mammalian translation system

4.1.1 Purification of the native human factors

4.1.1.1 HeLa Cell Processing

In this study, the cytoplasmic HeLa cell lysate is the main material for the preparation of native human ribosomes and initiation factors, eIF2, eIF3, eIF4F and eIF5B, except for eIF5, which is purified from rabbit reticulocyte lysate (RRL). Cytoplasmic HeLa cell lysate is kindly provided by the HeLa cell facility from Prof. Dr. Holger Stark's research group. To generate cytoplasmic lysate, cells are lysed and the lysate is cleared from debris and the nuclei fraction. As soon as I received the freshly prepared HeLa cell lysates, 40 mL of the lysate is distributed into 50 mL Falcon tube and snap-frozen in liquid nitrogen. The lysate is stored in the -80 °C freezer. To begin the purification, polysomes are first isolated from the lysates, which are further separated into fractions of ribosomes and native factors. The fractions containing native factors are then subject to the fast protein liquid chromatography (FPLC) for the purification of individual factors, while the native ribosomes are purified from the ribosomal pellets. Importantly, instead of the cytosolic initiation factors, only the polysome-bound initiation factors are purified for the reconstitution system, because they are the highly active fractions in the cytoplasm. Detailed purification conditions are described below.

The following protocol is adapted and revised from the published method (Pisarev et al., 2007b) and the lab notes from Dr. Akanksha Goyal. First, 160 mL of HeLa cell lysate is thawed in the 25 °C water bath for one batch of the process. The thawed lysate is supplemented with 1 tablet of cOmplete EDTA-free protease inhibitor (Roche) and stirred in a beaker on ice until the tablet dissolved completely. 12 mL of lysate is then layered on top of 12 mL sucrose cushion (20 mM Tris-HCl (pH 7.5), 2 M sucrose, 20 mM KCl, 4 mM MgCl₂, 2 mM DTT) and centrifuged for 18 hr at 45000 rpm in the Beckman Ti 50.2 rotor at 4 °C to isolate the polysomes from the cytosol and remove the cell debris at the interface of the cushion. The supernatant is discarded and each polysomal pellet is gently resuspended in 2 mL Buffer A (20

mM Tris-HCl (pH 7.5), 0.25 M sucrose, 50 mM KCl, 4 mM MgCl₂, 2 mM DTT). The polysomes can be snap-frozen in liquid nitrogen for the storage. The absorbance at 260 nm (A_{260}) of the polysome solution is measured in the spectrophotometer and diluted to 100 A_{260} U/mL with buffer A. To separate the bound factors from the polysomes, 4 M KCl is dropwise added into the diluted polysomal resuspension with stirring on ice until the KCl concentration is 0.5 M. The suspension is then continuously stirred on ice for another 30 min to allow complete dissociations of the factors from the polysomes. The factors are then separated from the polysomes by centrifugation for 4.5 hr at 45000 RPM in the Beckman Ti 50.2 rotor at 4 °C. After centrifugation, the supernatant, which is named ribosomal salt wash (RSW), contains all the factors and is the material for the native factor purification, whereas the ribosomal pellet is used for the native human ribosomes preparation. The ribosomal pellet is removed from the centrifugation tube and snap-frozen in the 15 mL falcon tube for storage in the -80 °C freezer. The RSW is collected for the subsequent ammonium sulfate precipitation (ASP) and can also be snap-frozen at this step for storage. The native factors, eIF3, eIF4F, eIF2 and eIF5B, are first separated from the RSW by the ammonium sulfate precipitation. The bulky factors, eIF3 and eIF4F, precipitate in the fraction of the 0 – 40% (NH₄)₂SO₄, whereas eIF2 and eIF5B precipitate in the 40 – 50% (NH₄)₂SO₄. Importantly, (NH₄)₂SO₄ is first grinded into fine powder by a blender before adding into the RSW to avoid high local concentration of AS. The AS is then slowly added into the RSW with gentle stirring on ice to reach 40% AS, following by another extra 30 min stirring on ice. The 40% AS containing RSW is centrifuged at 15000 RPM in JA 25.50 rotor for 20 min at 4 °C to isolate the precipitations. The pellet containing eIF3, eIF4F and other larger proteins and is marked as 0 – 40% ASP. The supernatant is supplemented with extra AS to reach 50% AS, followed by same procedure for the stirring and centrifugation. The 40 – 50% ASP contains eIF2, eIF5B and other medium size proteins, while the supernatant contains eIF1, eIF1A, eIF4A, eIF4B and other small factors which can be further purified but will not be used in this study. The supernatant is marked as > 50% ASP fraction and snap-frozen for the storage. Each 0-40% ASP is resuspended in 5 mL of buffer B (20 mM Tris-HCl (pH 7.5), 0.1 mM EDTA, 100 mM KCl, 10% glycerol, 2 mM DTT) and dialyzed against the same buffer in the 3.5 KDa snake skin dialysis membrane for 2 hr and overnight. Likewise, each 40 – 50% ASP is

resuspended in 3 mL buffer B and, thereafter, dialyzed against buffer B with the same procedure. After dialysis, the 0 – 40% AS resuspension (AS SN) is usually milky and cloudy while 40 – 50% AS SN is relative transparent. The resuspensions of both fractions are collected and centrifuged at 10000 RPM for 10 min at 4 °C in a bench top centrifuge to remove the precipitates. The 0-40% fractions usually result in relatively big pellets. The supernatants are collected and marked as 0-40% AS SN and 40-50% AS SN, which are the starting materials for the native factor purification. To increase the yield of the native factors, it is recommended to proceed the 0-40% AS SN for eIF3 and eIF4F purification directly. However, both resuspensions can also be snap-freezed at this step and used in the near future.

4.1.1.2 The 40S and 60S subunits

Human 40S and 60S subunits are purified from factor-free ribosomal pellets. Two ribosomal pellets are thawed in 10 mL buffer C (20 mM Tris-HCl (pH 7.5), 4 mM MgCl₂, 50 mM KCl, 2 mM DTT) with gentle shaking in the cold room for roughly 2 - 3 hr. The pellets are gently mixed by pipetting up and down with a cut 1 ml tip during the rehydration. The dissolved ribosome resuspensions are distributed into 1.5 mL Eppendorf tubes and centrifuged at 14000 RPM at 4 °C for 10 min in a bench top centrifuge. The transparent upper layers are collected for the subsequent purification, whereas the bottom cloudy fractions are discarded. The clear ribosomal fraction is diluted with buffer C to 300 A₂₆₀ U/mL and distributed into 1.4 mL per 2-mL Eppendorf. 14 µL of 106 mM puromycin is slowly added to each Eppendorf and mixed thoroughly. The puromycin-treated ribosomes are incubated first for 10 min on ice to allow the puromycin bind to the ribosomes and then in 37 °C for another 10 min to hydrolyze the preexisting peptide in the ribosomes. The small subunits and large subunits are further separated by slowly adding 200 µL 4M KCl to each Eppendorf to raise the final KCl concentration to 0.5 M. 1.5 mL of the ribosome solution is layered slowly on top of a 10-30% sucrose gradients. The 10-30% sucrose gradient is prepared by loading 19.2 mL of 30% sucrose gradient buffer (20 mM Tris-HCl (pH 7.5), 30% sucrose, 4 mM MgCl₂, 500 mM KCl, 2 mM DTT) in the bottom of the Beckman coulter disposable tubes (25x89 mm) and layering 19.2 mL 10%

sucrose gradient buffer (20 mM Tris-HCl (pH 7.5), 10% sucrose, 4 mM MgCl₂, 500 mM KCl, 2 mM DTT) on top, following by mixing at the gradient master with 10-30% SW28 program. The ribosomes are centrifuged at 22700 RPM in the swinging bucket rotor SW 32 Ti for 16.5 hours at 4 °C. The gradient is fractionized for 750 µL from the top to the bottom of the tube. A₂₆₀ of each fraction is measured and the profile (**Figure 49**) is generated in the GraphPad Prism software. The first peak contains the 40S subunit, usually fraction 22 - 27, whereas second peak contains 60S subunit, fraction 38 - 43. 5 - 6 fractions of the 40S and 60S subunits on the tip of the corresponding peaks are collected, avoiding cross-contamination from the 40S and 60S in the valley between 40S peak and 60S peak. The 40S subunit is diluted with 1 volume of buffer D (20 mM Tris-HCl (pH 7.5), 4 mM MgCl₂, 50 mM KCl, 2 mM DTT) and the 60S subunit is diluted with 2 volumes of buffer D to reduce the sucrose concentration from the gradient. The diluted ribosomes are concentrated in the 100 kDa cut-off 50 ml centricon at 4500 g at 4 °C until the volume is less than 1 mL. 8 mL of buffer D with 0.25 M sucrose is added to the ribosomes, following by concentration until the volume reaches 1 mL. The A₂₆₀ of the ribosomes are measured to determine the concentration. 1 A₂₆₀ U/mL for the 40S subunit is 50 pmole/mL, while 1 A₂₆₀ U/mL for the 60S subunit is 25 pmole/mL. The ribosomes are divided into aliquots with adequate size and snap-frozen in liquid nitrogen. The ribosomes are stored in the -80 °C freezer.

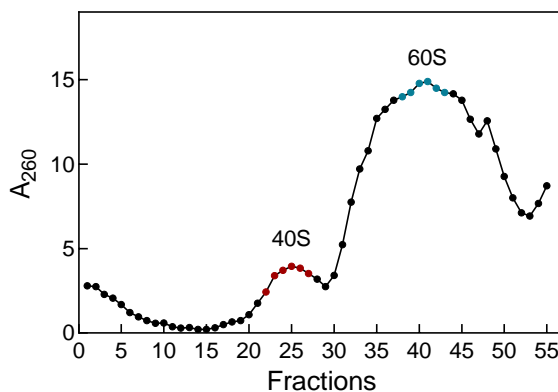


Figure 49. Purification of the 40S and 60S subunits on a 0–30% sucrose gradient

The fractions of the 40S subunit are marked in red, while the fractions of the 60S subunits are marked in blue.

4.1.1.3 eIF2 and eIF5B

For native eIF2 and eIF5B purification, 40-50% AS SN is used as starting material. The 40-50% AS SN is centrifuged at 10000 RPM for 10 min at 4 °C in a bench top centrifuge to remove the precipitates before loading to the fast protein liquid chromatography (FPLC) column. The clear supernatant is applied to the 5 mL Hitrap Q HP column (GE healthcare), which has been equilibrated with buffer E (20 mM Tris-HCl (pH 7.5), 0.1 mM EDTA, 5% glycerol, 2 mM DTT) + 100 mM KCl, with flow rate 1 mL/mL. eIF2 and eIF5B are eluted with a 100 – 1000 mM KCl gradient in buffer E in 25 column volume (CV) with flow rate 2 mL/mL and 1 mL per fraction. The fractions are analyzed by 12% SDS-PAGE and those containing eIF2 and eIF5B are pooled together. Usually eIF2 elutes at 300 mM KCl and eIF5B elutes later at 450 mM KCl. The pooled fractions are dialyzed in buffer E + 250 mM KCl for 2 hr and then in buffer E + 100 mM KCl overnight. Next, the dialyzed eIF2 + eIF5B eluate is applied on two 1 mL Hitrap SP HP columns (GE Healthcare), which were equilibrated with buffer E + 100 mM KCl. eIF2 and eIF5B are eluted with a 100 – 1000 mM KCl gradient in buffer E in 25 CV with flow rate 1 mL/mL and collected into 0.5 mL fractions. eIF2 and eIF5B usually elute together at 350 mM KCl. Fractions containing eIF2 and eIF5B are pooled and dialyzed in 250 mM KCl and 100 mM KCl in buffer E for 2 hr and overnight. Last, the dialyzed eIF2 + eIF5B eluate is applied to 1 mL MonoQ 5/50 GL column equilibrated with buffer E + 100 mM KCl. eIF2 and eIF5B are eluted with a 100 – 600 mM KCl gradient in buffer E in 25 CV with flow rate 1 mL/mL and collected into 0.4 mL fractions. Peaks containing eIF2 or eIF5B are collected separately this time and dialyzed against 250 mM KCl and 100 mM KCl in buffer E with 10% glycerol for 2 hr and overnight. The eIF2 elutes at 300 mM KCl while eIF5B elutes at 450 mM KCl (**Figure 50**). The dialyzed eIF2 and eIF5B fractions are snap-frozen in liquid nitrogen and stored at -80 °C freezer. Notably, eIF5 is purified by the same procedure but from the RRL.

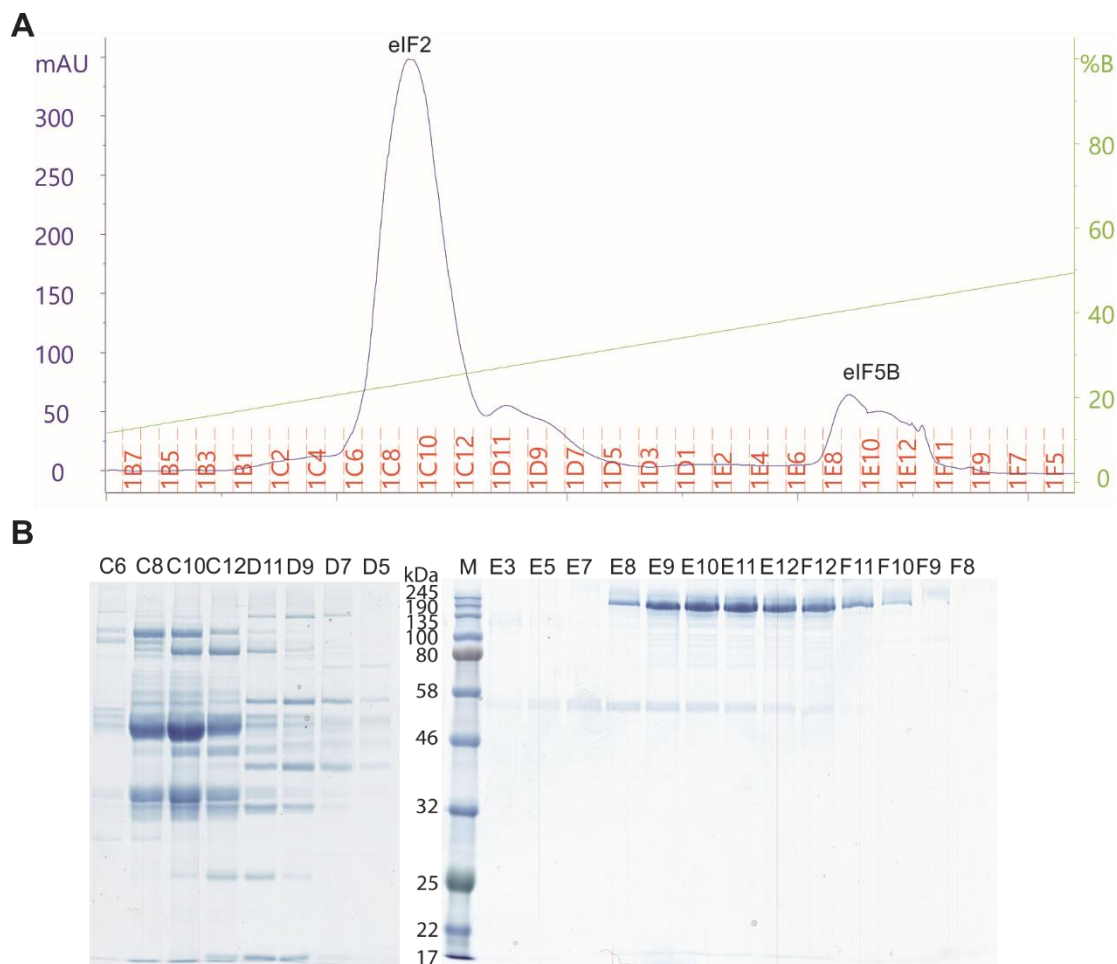


Figure 50. Elution profile of eIF2 and eIF5B from MonoQ 5/50 GL column

(A) Elution profile of eIF2 and eIF5B from MonoQ 5/50 GL column. The absorbance of 280 nm is the blue line and the percentage of buffer E + 600 mM KCl is the green line. **(B)** 12% SDS-PAGE analysis of the fractions with Coomassie Brilliant blue staining. eIF2 is collected from fraction C7 - C12 and eIF5B is collected from fraction E8 – F11 in this profile.

4.1.1.4 eIF3 and eIF4F

For eIF3 and eIF4F purification, similar procedures are applied as described above for the eIF2 and eIF5B. 0 – 40% AS SN is used to purify the eIF3 and eIF4F. Precipitates in the 0 – 40% AS SN are removed by centrifugation before applying to the FPLC column. The clear supernatant is loaded onto 5 mL Hitrap Q HP column equilibrated with buffer E + 100 mM KCl.

The eIF3 and eIF4F are eluted with a 100 – 1000 mM KCl gradient in buffer E in 25 CV with flow rate 2 mL/mL and collected into 1 mL fractions. eIF3 and eIF4F are eluted together usually between 330 mM KCl and 460 mM KCl. Fractions are pooled and dialyzed against 250 mM KCl and 100 mM KCl in buffer E for 2 hr and overnight. The dialyzed eIF3 and eIF4F eluates are then applied to two 1 mL Hitrap SP HP columns equilibrated with buffer E + 100 mM KCl. The eIF3 and eIF4F are eluted directly at 400 mM in buffer E and subsequently dialyzed back to 100 mM KCl in buffer E. Finally, the dialyzed eluate is applied to 1 mL MonoQ 5/50 GL column equilibrated with buffer E + 100 mM KCl for polishing. Notably, it is important to avoid overloading the MonoQ column to have better separation of eIF3 and eIF4F, less than 20 mg of protein should be applied to the column at once. eIF3 and eIF4F are eluted with a 100 – 600 mM KCl gradient in buffer E in 25 CV with flow rate 1 mL/mL and collected into 0.4 mL fractions. The concentrated eIF3 and eIF4F are collected separately this time and dialyzed against 250 mM KCl and 100 mM KCl in buffer E with 10% glycerol for 2 hr and overnight. eIF4F usually elutes from 320 mM KCl, while eIF3 elutes from 410 mM KCl (**Figure 51**). The eIF3 and eIF4F can be further purified by sucrose gradient purification. Higher purities are gained at the cost of large losses in yield. The dialyzed factors are snap-frozen in liquid nitrogen and stored at -80 °C.

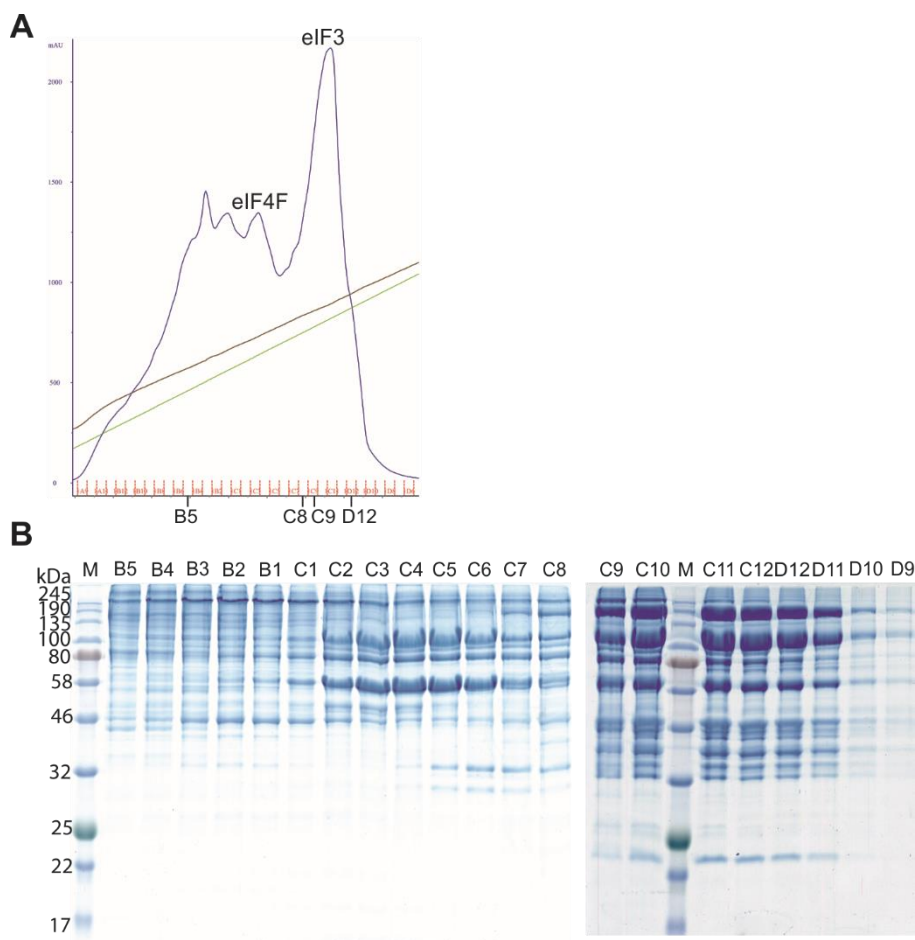


Figure 51. Elution profile of eIF3 and eIF4F from MonoQ 5/50 GL column

(A) Elution profile of eIF3 and eIF4F from MonoQ 5/50 GL column. The absorbance of 280 nm is the blue line and the percentage of buffer E + 600 mM KCl is the green line. **(B)** 12% SDS-PAGE analysis of the fractions with Coomassie Brilliant blue staining. eIF3 is collected from fraction C9-D12 and eIF4F is collected from fraction B5 – C8 in this profile.

4.1.2 Purification of the recombinant proteins

4.1.2.1 Expression of the recombinant proteins in *E. coli*

Human cDNAs of eIF1, eIF1A, eIF4A and eIF4B are constructed into pET24a plasmid for non-tag recombinant protein purification and these constructs are kindly provided by Dr Akanksha Goyal. pET24a-eIF1 and pET24a-eIF1A are transformed into *E. coli* strain BL21 codon-

plus RILP (Agilent Technologies), while pET24a-eIF4A and pET24a-eIF4B are expressed in *E. coli* strain Rosetta. For large scale expression, the cells are grown in 12 L lysogeny broth (LB) medium containing kanamycin, spectinomycin and chloramphenicol from 0.1 OD₆₀₀ till 0.8 OD₆₀₀ at 37 °C with 180 RPM shaking. Isopropyl β- d-1-thiogalactopyranoside (IPTG) is added into the cell culture with final 1 mM concentration to induce the protein expression for another 3 hr at 37 °C. The cells are harvested by centrifugation at 7000 RPM in the Beckman Coulter JLA8.100 rotor for 20 min at 4 °C. The cell pellets are collected and snap-frozen in the liquid nitrogen. The cell pellets are analyzed on the 12% SDS-PAGE to confirm the expression of the recombinant proteins. To lyse the cells, the cell pellets are resuspended in 3 mL lysis buffer per gram of cells (20 mM Hepes-KOAc (pH 7.5), 10 mM MgCl₂, 50 mM KCl, 1 mM EDTA (pH 8), 2 mM DTT) with 1 mg of lysozyme per gram of cells and 100 μL of DNase. The resuspended cells are lysed through the emulsification and centrifuged at 30000 RPM Beckman Coulter JA 30.50 rotor for 45 min at 4 °C. The supernatant is ready for FPLC column purification.

4.1.2.2 eIF1

The supernatant containing eIF1 is applied to a 5 mL Hitrap SP HP column equilibrated with buffer F (20 mM Hepes-KOAc (pH 7.5), 0.1 mM EDTA (pH 8), 2 mM DTT) + 50 mM KCl. The eIF1 is eluted with a 50 – 1000 mM KCl gradient in buffer F in 20 CV with flow rate 2 mL/min and collected into 2 mL fractions. The fractions are analyzed by 15% SDS-PAGE and those containing eIF1 are collected, usually from 280 mM KCl. eIF1 eluate is dialyzed against 200 mM and 50 mM KCl in buffer F for 2 hr and overnight. The dialyzed eIF1 is then applied to a 6 mL Resource Q column equilibrated with buffer F + 50 mM KCl. The eIF1 is eluted with a 50 – 250 mM KCl in 18 CV with flow rate 2 mL/min and collected into 1.5 mL fractions. The eIF1 is collected from 100 mM KCl and dialyzed against 100 mM KCl in buffer F with 10% glycerol for 2 hr and overnight. The final eIF1 preparation is snap-frozen in liquid nitrogen and stored in the -80 °C freezer.

4.1.2.3 eIF1A

The supernatant containing eIF1A is applied to a 5 mL Hitrap SP HP column equilibrated with buffer F + 50 mM KCl. eIF1A is eluted with a 50 – 2000 mM KCl gradient in buffer F in 20 CV with flow rate 2 mL/mL and collected into 2 mL fractions. The fractions are analyzed by 15% SDS-PAGE and those containing eIF1A are collected, usually from 540 mM KCl. The eIF1A eluate is dialyzed against 250 mM and 100 mM KCl in buffer F for 2 hr and overnight. The dialyzed eIF1A is then applied to two 1 mL Hitrap Q HP columns equilibrated with buffer F + 100 mM KCl. The eIF1A is eluted with a 100 – 1000 mM KCl in 20 CV with flow rate 1 mL/mL and collected into 0.5 mL fractions. The eIF1A is collected from 400 mM and dialyzed against 250 mM KCl and 100 mM KCl in buffer F with 10% glycerol for 2 hr and overnight. The final eIF1A preparation is snap-frozen in liquid nitrogen and stored in the -80 °C freezer.

4.1.2.4 eIF4A

The supernatant containing eIF4A is applied to a 5 mL Hitrap Q HP column equilibrated with buffer F + 50 mM KCl. eIF4A is eluted with a 100 – 500 mM KCl gradient in buffer F in 20 CV with flow rate 2 mL/mL and collected into 2 mL fractions. The fractions are analyzed by 15% SDS-PAGE and the those containing eIF4A are collected. The eIF4A eluate is dialyzed twice against 100 mM KCl in buffer F for 2 hr and overnight. The dialyzed eIF4A is then applied to a 5 mL Blue Sepharose column equilibrated with buffer F + 100 mM KCl and eluted with a 100 – 1000 mM KCl in 20 CV with flow rate 1 mL/mL and collected into 0.5 mL fractions. The fractions containing eIF4A are collected and dialyzed against 250 mM KCl and 100 mM KCl in buffer F with 10% glycerol for 2 hr and overnight. The final eIF4A preparation is snap-frozen in liquid nitrogen and stored in the -80 °C freezer. The eIF4A used in this study is kindly provided by Dr. Akanksha Goyal.

4.1.2.5 eIF4B

The supernatant containing eIF4B is applied to a 5 mL Hitrap SP HP column equilibrated with buffer G (20 mM Hepes-KOAc (pH 7.5), 2 mM MgCl₂, 2 mM DTT) + 100 mM KCl. The eIF4B is eluted with a 100 – 1000 mM KCl gradient in buffer F in 20 CV with flow rate 2 mL/min and collected into 2 mL fractions. The fractions are analyzed by 15% SDS-PAGE and the those containing eIF4B are collected. The eIF4B eluate is dialyzed against 250 mM and 100 mM KCl in buffer G for 2 hr and overnight. The dialyzed eIF4B is then applied to two 1 mL Hitrap Q HP columns equilibrated with buffer G + 100 mM KCl. The eIF4B is eluted with a 100 – 1000 mM KCl in 20 CV with flow rate 1 mL/min and collected into 0.5 mL fractions. The fractions containing eIF4B are collected and dialyzed against 250 mM KCl and 100 mM KCl in buffer G with 10% glycerol for 2 hr and overnight. The final eIF4B preparation is snap-frozen in liquid nitrogen and stored in the -80 °C freezer.

4.1.3 *in vitro* Transcription

4.1.3.1 Preparation of Met-tRNA_i^{Met}

Human initiator tRNA sequence (AGCAGAGTGGCGCAGCGGAAGCGTGCTGGGCCCATAA CCCAGAGGTCGATGGATCGAAACCATCCTCTGCTACCA) is constructed under the expression of the T7 promoter (TAATACGACTCACTATA) on the pIDTSMART-AMP (Integrated DNA Technologies, IDT). The PCR product for the *in vitro* T7 polymerase transcription is amplified with the forward primer (TAATACGACTCACTATAAGCAGAGTGG) and the reverse primer (TGGTAGCAGAGGATGGTTTCGATC) from 35 thermocycles of 10 sec at 98 °C, 15 sec at 50 °C and 5 sec at 72 °C. Importantly, 2 mL of total PCR mixture is divided into PCR tube with 50 µL aliquots to achieve homogenous temperature throughout the PCR reaction. For the *in vitro* T7 transcription reaction, PCR product is mixed with T7 reaction buffer (40 mM Tris-HCl (pH7.5), 15 mM MgCl₂, 2 mM Spermidine, 10 mM NaCl, 3 mM ATP, 3 mM UTP, 3 mM GTP, 3 mM CTP, 10 mM DTT, 0.4 U/µL RiboLock RNase-Inhibitor (Thermo Scientific), 2 U/µL inorganic pyrophosphatase (Sigma), 1 U/µL T7 RNA polymerase (in-house made)). The *in vitro*

transcribed RNA product is purified through Hitrap Q HP column (GE Healthcare) and precipitated with 100% ethanol. The final transcript is resuspended in the water. The initiator tRNA is aminoacylated with in-house made S100 extract prepared from *E. coli* and [³H]-labeled Methionine. The Met-tRNA^{Met} is purified by the reverse-phase chromatography using published process (Milon 2007). The fractions containing high radioactivity of [³H] and A₂₆₀ are collected, precipitated in ethanol and dissolved in water.

4.1.3.2 Preparation of the model mRNA

Model mRNA (255 nt, AUG - underlined; Reporter ORF - bold; Toeprinting primer binding site – italics underlined)

GGG CAA CAA CAA CAA GCU AGC CACAA CAA CAA CAA CAA CAA GUC GAC CAA CAA
CAA CAA CAA CAA CAA CAA CUC GAG CAA CAA CAA CAA CAA CAA CAA CAA GGA UCC
AAA ACA GAC CACC AUG **GUA CGU UUC AAG GCU UGA** GCC AUG GUC ACU GCC CUG UGG
GGC AAG GUG AAU GUG GAA GAA GUU GGU GGU GAG GCC CUG GGC AGG CUG CAG AGU
GUG AGG GAA GGU CUG GUU GUC UAC CAA

13XCAA model mRNA (170 nt, AUG - underline; Reporter ORF - bold; Toeprinting primer binding site – italics underlined)

GGG CAA CAA CAA CAA CAA CAA CAA CAA CAA CAA CAA CAA CC AAA ACA GAC AGA AUG
GUA UUC UGA GCC AUG GUC ACU GCC CUG UGG GGC AAG GUG AAU GUG GAA GAA GUU
GGU GGU GAG GCC CUG GGC AGG CUG CAG AGU GUG AGG GAA GGU CUG GUU GUC UAC
CAA

The model mRNA sequence and the 13XCAA model mRNA sequence are constructed under the expression of the T7 promotor (TAATACGACTCACTATA) on pUC19 and pEX-A2 (Eurofins), respectively. The PCR products are amplified with the forward primer (ATCTAGAATTCTAATACGACTCACTATAGG) and reverse primer (TTGGTAGACAACCAGACCTTCC) from 35 thermocycles of 10 sec at 98 °C, 15 sec at 54 °C and 15 sec at 72 °C. For the *in vitro* T7 transcription reaction, PCR product is mixed with T7 reaction buffer (40 mM Tris-HCl (pH7.5),

15 mM MgCl₂, 2 mM Spermidine, 10 mM NaCl, 3 mM ATP, 3 mM UTP, 3 mM GTP, 3 mM CTP, 10 mM DTT, 0.4 U/μL RiboLock RNase-Inhibitor (Thermo Scientific), 2 U/μL inorganic pyrophosphatase (Sigma), 1 U/μL T7 RNA polymerase (in-house made)). The *in vitro* transcribed RNA product is purified through Hitrap Q HP column (GE Healthcare), precipitated with 100% ethanol and subsequently dissolved in water.

4.1.4 Preparation of the fluorescence-labeled eIF1A variants

The major challenges of studying the fast kinetic in the higher eukaryotic are the high demanding of the materials for the stop-flow experiment and the lack of suitable observables to monitor the initiation process. An ideal observable should stay on the ribosome throughout the initiation process or at least within the major events for the observation. Genetically, it should also be easily manipulated to allow the introduction of site-specific labeled fluorophores. Therefore, eIF1A is an ideal candidate as a fluorescence reporter to monitor the formation of 48S and 80S IC.

To label eIF1A with the Alexa Fluor™ 555 C₂ maleimide dye (Thermo Scientific), single cysteine is introduced at the positions N4C, S74C, T120C or D142C of eIF1A. The mutant eIF1As are expressed and purified with the same procedure as the WT eIF1A. However, in the last dialysis step, the DTT is excluded in the dialysis buffer to prevent quenching of the maleimide dye. To label the protein, 1 mg of Alexa Fluor™ 555 C₂ maleimide (Alexa555, Thermo Scientific) is freshly dissolved in 80 μL DMSO as 10 mM stock solution. The eIF1A mutant is mixed with 4-fold excess of the Alexa555 dye and incubated in the dark at 25 °C with gentle shaking for 1 hr. The precipitant of the dye is first removed by the centrifugation. The supernatant containing labeled protein, unlabeled protein and excess dye is applied to a 5 mL Hitrap SP HP column and the excess free dye is eluted by washing the column with 20 CV of buffer F + 50 mM KCl. The eIF1A variants are eluted with the same procedure as the WT eIF1A. The fractions are analyzed by the 15% SDS-PAGE and fluorescently imaged by the fluorescence scanner (Typhoon™ FLA7000 (GE Healthcare)). The fractions containing labeled eIF1A are collected and

dialyzed in buffer F + 100 mM KCl with 10% glycerol for 2 hr and overnight and stored at -80 °C freezer.

4.2 Reconstitution of the Initiation complexes

4.2.1 48S IC formation

48S IC is reconstituted with purified factors, eIF1, eIF1A, eIF2, eIF3, eIF4A, eIF4B, the 40S subunit, Met-tRNA_i^{Met}, and the model mRNA. To achieve higher reconstitution efficiency, in most of the assays, the 40S subunit is saturated by excess amounts of the factors, except in the toeprinting assay, where the mRNA is the reporter and is the limiting component in the reconstitution. Since the binding affinities of the individual factor to the 40S subunit are in the 10 - 100 nM range, the ideal concentration of the 40S subunit is 0.1 – 0.5 μM. Higher concentrations would be better for some assays but are technically difficult due to the low concentration of the 40S subunit stock, which is usually 2 – 5 μM. For the reconstitution, 40S subunit is mixed with 3-fold excess of eIF1, eIF1A and eIF4A, 1.5-fold excess of eIF2 and eIF3, 1.5-fold excess of eIF4B, 3-fold excess of Met-tRNA_i^{Met}, and the model mRNA in the buffer R (20 mM Hepes (pH 7.5), 95 mM KOAc, 3.75 mM Mg(OAc)₂, 1 mM ATP, 0.5 mM GTP, 0.25 mM Spermidine, 2 mM DTT, 0.4 U/μL RiboLock RNase-Inhibitor (Thermo Scientific)) and incubated at 37 °C for 10 min. It is important to consider the potassium salt carried over from the factor preparation in order to maintain the final KOAc between 80 – 120 mM. Likewise, when the buffer is supplemented with ATP or GTP, extra Mg(OAc)₂ needs to be added to maintain the free Mg²⁺ concentration at 2 – 3 mM. The ideal order of the additions of the factors to the buffer is 40S, eIF3, eIF2, eIF1, eIF1A, eIF4A, eIF4B, Met-tRNA_i^{Met}, and the model mRNA to allow the binding of the factors to the ribosomes immediately after they are added to the mixture and prevent the factors stick to the test tube. The concentration of the 48S IC is, theoretically, the concentration of the 40S subunit or, in the case of toeprinting assay, the concentration of the model mRNA.

4.2.2 80S IC formation

80S IC is reconstituted by promoting the 60S subunit joining to the 48S IC. First, the 48S IC is formed with the procedure described above. Two-fold excess of eIF5B (comparing to the 48S IC concentration) is mixed with buffer R and incubated at 37 °C for 5 min to allow the exchange of eIF5B·GDP to eIF5B·GTP. Then, 2-fold excess of the 60S subunit (comparing to the 48S IC concentration) is added to the eIF5B solution. Subsequently, the 60S·eIF5B mixture is added to the 48S IC and incubated at 37 °C for 10 min to allow the large subunit joining. To obtain a higher concentration of the 80S ribosomes, equal or half volume of the 60S·eIF5B mixture is usually added to 1 volume of the 48S IC. Therefore, the final concentration of the 80S IC is 1/2 or 1/3 of the 48S IC concentration, respectively. The 80S ribosomes are relative stable and can resume elongation activity even after being snap-frozen in the liquid nitrogen.

4.3 Functional assays

4.3.1.1 Toeprinting assay

Toeprinting assay is used to detect the positions of 48S IC, 80S IC and 80S EC on mRNA. In this assay, mRNA is the reporter and the limiting component. Therefore, 1-fold of the model mRNA (usually 0.1 μ M in 20 μ L reaction) is mixed with 2 times of the 40S subunit and excess of other factors in the same procedure described above. The mRNA in the 48S IC and 80S IC reconstitution are annealed with fluorescently 5'-labeled DNA primer (Atto647N – GACCTTCCCTCACACTCTG (IDT)) in the 3' UTR and the primer is reverse-transcribed in the reverse-transcription reaction buffer (0.5 mM dNTP (New England Biolabs), 0.05 μ M toeprinting primer, 8 mM MgCl₂, 0.15 U/ μ L Reverse Transcriptase (Invitrogen SuperScript III)) at 37 °C for 45 min. The reserves transcription is quenched by adding 10 volumes of the stop solution (12.5 mM EDTA (pH 8.0), 0.25% SDS) to the reaction. The cDNA products are extracted with phenol : chloroform : isoamyl alcohol (Sigma-Aldrich) and precipitated in 100% Ethanol with 0.015 mg/mL GlycoBlue (Invitrogen). The cDNA is dissolved in the RNA loading dye (95% Deionized formamide, 0.025% Bromophenol blue, 5 mM EDTA (pH 8.0)) and separated on the

8% UREA-PAGE with TBE buffer (90 mM Tris-HCl, 90 mM Boric acid, 0.2 mM EDTA (pH 8.0)) at 70 V for 90 min. The fluorescent image of the UREA-PAGE gel is analyzed in the fluorescent scanner (Typhoon™ FLA7000 (GE Healthcare)) with Cy5 program at 635 nm excitation wavelength.

The revers transcribed cDNAs can be detected via radioactivity or fluorescence. The convenience of handling fluorescent over radioactive samples is one of the most important advantage of the fluorescent-based toeprinting assay. Since I utilized an optimal and relative short mRNA for assembly the initiation complexes, the nucleotide-resolution can be easily achieved with a small UREA-PAGE electrophoresis system, rather than a complex radio-based sequencing gel. Small UREA-PAGE is also easier to prepare and require less separation time during the electrophoresis, which is usually a time-consuming process in conventional sequencing gel.

The fluorescent reporter is very sensitive and, therefore, requires less amount of the materials for the detection. The high demand of the materials is one of the challenges when reconstituting mammalian translation system. In this toeprinting assay, I only required minimum 2 pmol of the model mRNA in each sample with only 3 pmol of the 40S subunit in the presence of other excess factors. This little amount of material requirement allows me to test various conditions within the same biological protein preparations. However, the high sensitivity of the fluorescent reporter also reveals the small difference between experiments from different handling batch. Therefore, to compare the relative intensity between the samples, it is very important to prepare the samples from the same mixture and purify the RT products all together to synchronize the systematic variants. The results from different experiments can only be compared qualitatively but not quantitatively.

4.3.2 Peptides formation assay

4.3.2.1 Oligopeptide formation

To promote the oligopeptide (Met-Val-Arg-Phe-Lys-Ala) formation, the 80S IC is assembled as described above and mixed with 3-fold amount of the corresponding [^{14}C]-labeled aminoacyl-tRNAs (Val, Arg, Phe, Lys, Ala), native human eEF1A and eEF2, which are kindly provided by Dr. Dmitry Burakovskiy, in the buffer R at 37 °C for 10 min. The oligopeptides are hydrolyzed in 0.5 M KOH at 37 °C for 30 min and neutralized with 10% acetate and 0.1% trifluoroacetic acid (TFA). The hydrolyzed oligopeptides are applied on to LiChrospher RP-8 column (Merck) on the HPLC system. The oligopeptides are eluted with a 0 – 65% acetonitrile gradient in 0.1% TFA in 30 min at 1 mL/min flow rate. 1 mL fractions are collected and mixed with 2 mL scintillation cocktail (PerkinElmer). The radioactivity of [^3H] from the methionine and [^{14}C] from the individual amino acids are calculated for each fraction and converted to the amount of the oligopeptides.

4.3.2.2 Dipeptide formation

The dipeptide (Met-Val) formation applied similar procedure as oligopeptide assay except the eEF2 is excluded and the elution gradient for the dipeptide is different. First, the 80S IC is assembled as described above and mixed with 3 times amount of the [^{14}C]-labeled Val-tRNA and native human eEF1A in the buffer R at 37 °C for 10 min. The Met-Val dipeptides are hydrolyzed in 0.5 M KOH at 37 °C for 30 min and neutralized with 10% acetate and 0.1% trifluoroacetic acid (TFA). The hydrolyzed dipeptides are applied on to LiChrospher RP-8 column (Merck) on the HPLC system. The dipeptides are eluted with a 0 – 3.25% acetonitrile gradient in 0.1% TFA in 20 min at 1 mL/min flow rate. 1 mL fractions are collected and mixed with 2 mL scintillation cocktail (PerkinElmer). The radioactivity of [^3H] from the methionine and [^{14}C] from the valine are calculated for each fraction. To have a better signal to noise ratio, 3 pmol of 80S ribosomes is usually used for each dipeptide assay, which leads to 1.5 – 2 pmol of the Met-Val dipeptides.

4.4 Steady-state fluorescence measurements

4.4.1 Gel shift assay

Since the 40S subunit and eIF3 are huge complexes, which migrate too slowly in the native PAGE, agarose gel system is an alternative assay to separate the huge complexes with their small binding ligand, e.g. mRNA. To increase the sensitivity of the binding assay, 3' fluorescein-labeled mRNA is used as reporter. However, unlabeled dark mRNA can also be used in this assay by staining with DNA Stain G dye (SERVA). 5 pmole of the 3' fluorescein-labeled model mRNA is mixed with increasing amount of eIF3 or the 40S subunit in the buffer R at 37 °C for 10 min. The protein/RNA complex is loaded into 0.7% agarose gel with equal volume of the loading dye (50% glycerol, 0.025% bromophenol blue) in a pre-cold TBE buffer (90 mM Tris base, 90 mM boric acid, 2 mM EDTA (pH 8.0)) + 100 mM KOAc. The electrophoresis is conducted on ice in the cold room at 50 V for 3 hr to prevent heating up the buffer and ensure the stability of the protein/RNA complexes during the process. The gel is imaged by a fluorescent scanner (Typhoon™ FLA7000 (GE Healthcare)) with Alexa488 program and stained with conventional Coomassie Brilliant Blue to compare the locations of the bound/free mRNA and the protein complexes.

4.4.2 Binding affinity of eIF1A to the 40S subunit

To determine the binding affinity of a fluorescent protein to its binding partner, the steady-state anisotropy of the fluorescent protein upon binding is monitored in the fluorometer (FluoroMax-4 (HORIBA Scientific)). 5 nM of eIF1A-N4C-Alexa555 is mixed in the observation cuvette in buffer R + 0.05% BSA to prevent eIF1A stick to the cuvette wall. Increasing amounts of the 40S subunits are added to the mixture and incubated at 25 °C for 2 min before measuring to allow the reaction reach equilibrium. 5 technical replicates of anisotropy are measured for each concentration of the 40S subunit. To obtain the binding affinity, the binding curve is analyzed using the GraphPad Prism software with the quadratic equation, $Y = r_f + ((r_b - r_f) / 2L) * [(L + X + K_d) - ((L + X + K_d)^2 - 4 * L * X)^{1/2}]$, where Y is the anisotropy

of the eIF1A-N4C-Alexa555 at X concentration of the 40S subunit, r_b is the bound anisotropy of the eIF1A-N4C-Alexa555, r_f is the free anisotropy of the eIF1A-N4C-Alexa555, L is the total concentration of the eIF1A-N4C-Alexa555 and K_d is the binding affinity. To have more accurate measurement of the binding affinity, it is important to keep the concentration of the fluorescent component as low as possible in the reaction, but the fluorescent signals should be still higher than 10,000 count per second (cps) during the measurement. 555 nm excitation wavelength and 568 nm emission wavelength with 3 nm slits are used for the fluorometer measurement.

4.4.3 Competition assay of unlabeled WT eIF1A to the 40S·eIF1A-N4C-Alexa555 complex

To determine the binding affinity of the unlabeled eIF1A to 40S subunit, 5 nM of eIF1A-N4C-Alexa555 is first bound with 30 nM of the 40S subunit in the buffer R + 0.05% BSA and then mixed with increasing amounts of unlabeled eIF1A in the same reaction mixture. 5 technical replicates of anisotropy are measured for each addition of the dark eIF1A at 25 °C. The half maximal inhibitory concentration (IC_{50}) of the competition reaction is calculated using the GraphPad prism software with the equation $Y = r_f + (r_b - r_f) / (1 + 10^{(X - \text{Log}IC_{50})})$, where Y is the anisotropy of the eIF1A-N4C-Alexa555 at X concentration of the unlabeled eIF1A, r_b is the bound anisotropy of the eIF1A-N4C-Alexa555, r_f is the free anisotropy of the eIF1A-N4C-Alexa555, and IC_{50} is the half maximal inhibitory concentration of the eIF1A to the 40S·eIF1A-N4C-Alexa555 complexes. The competition affinity (K_i) of eIF1A to the 40S subunit is converted from the IC_{50} according to the published equation (Burnett et al., 2007) and its online converter (https://bioinfo-abcc.ncifcrf.gov/IC50_Ki_Converter/index.php).

4.4.4 Fluorescence changes of eIF1A-Alexa555 during 48S and 80S IC formation

To monitor the formation of the 48S IC and 80S IC, 30 nM of the fluorescently labeled eIF1A mutants are mixed with the buffer R + 0.05% BAS in the cuvette of the fluorometer. 90

nM 40S subunit, 150 nM eIF1, 120 nM eIF2, 240 nM Met-tRNA^{iMet}, 120 nM eIF3, 200 nM eIF4A, 90 nM eIF4B and 240 nM model mRNA are subsequently added into the observation cuvette to stepwise form the 48S IC and incubated at 25 °C for two min to allow the binding reaching equilibrium. The fluorescence intensity and anisotropy are measured 5 times for each addition at 25 °C with excitation at 555 nm and emission at 568 nm + 3 nm slit. For the 80S IC assembly, 60 nM eIF5, 80 nM eIF5B and 100 nM 60S subunit are included subsequently after the 48S IC formation. 900 nM unlabeled WT eIF1A is added in the last step of the measurement for the chase experiment to confirm the specific binding of eIF1A to the 40S subunit. Fluorescence intensities are normalized to the initial fluorescence intensity of the unbound eIF1A mutants to compare the relative changes in each step.

4.5 Rapid kinetic assays

4.5.1 Association kinetic assay

The pre-steady state kinetic experiments in this study are carried out in a stopped-flow apparatus (SX-20MV (Applied Photophysics)). To elucidate the association kinetics of eIF1A to the 40S subunit, 15 nM of eIF1A-N4C-Alexa555 is rapidly mixed with 45, 60, 90, 120, 150, 225 nM of the 40S subunits in buffer R + 0.05% BSA at 25 °C. The fluorescence intensity is recorded for 4000 time points logarithmically in 200 sec with a 535 nm LED lamp source for the excitation and an OG570 emission filter coupling to a photomultiplier tube (PMT) detector at 350 voltage. At least 3 technical replicates of raw traces from each condition are taken for average as results. For the binding of eIF1A to the 40S-eIF1 complex, 1 μM of the eIF1 is incubated with the same titration concentrations of the 40S subunits as describe above at 37 °C for 10 min in buffer R + 0.05% BSA. The binding reactions of eIF1A to the 40S-eIF1 complex are recorded with the same parameters as with the 40S subunit. The binding curves are analytically fitted with two exponential equation to obtain the apparent rate constants. The analytical association rate constants are calculated from the linear dependence of the apparent rate constants to the concentrations of the 40S subunits by Graphpad prism. The

numerical model is analyzed via KinTek Explorer with parameters described in the result section.

4.5.2 Dissociation kinetic assay

The dissociation rate constants of eIF1A from the 40S subunit and initiation complexes are determined by the chase experiments. 90 nM of the 40S subunit is pre-bound with 15 nM of fluorescently labeled eIF1A in the presence of 300 nM individual eIF as indicated in each condition in buffer R + 0.05% BSA at 37°C for 10 min. The 40S·eIF1A·eIFs complexes are rapidly mixed with 1350 nM of unlabeled WT eIF1A in buffer R + 0.05% BSA at 25°C in the stopped-flow apparatus. The fluorescence intensity is recorded for 4000 time points logarithmically in 200 - 1000 sec until the trace reaches a stable baseline with a 535 nm LED lamp source for the excitation and an OG570 emission filter coupling to a photomultiplier tube (PMT) detector at 350 voltage. At least 3 technical replicates of raw traces from each condition are taken for average as results. For the dissociation of eIF1A from 43S PIC and 48S IC, the initiation complexes are assembled with 30 nM fluorescently labeled eIF1A, 180 nM 40S subunit, 550 nM of eIF1, eIF4A, the model mRNA (excluded for the 43S PIC) and Met-tRNA^{Met}, 270 nM eIF3 and eIF4B, and 360 nM eIF2 in buffer R + 0.05% BSA at 37°C for 10 min. The initiation complexes are rapidly mixed with 3600 nM unlabeled WT eIF1A in buffer R + 0.05% BSA at 25°C in the stopped-flow apparatus. The fluorescence intensity is recorded as described above. The dissociation curves are analytically fitted with either one or two exponential equation to obtain the dissociation rate constants.

4.6 Human 48S IC for the cryo-EM structure study

For the cryo-EM structure study, the 48S IC is assembled with 500 nM 40S subunit, 1000 nM of eIF1, eIF1A, eIF4A, Met-tRNA^{Met}, and the model mRNA, and, 700 nM eIF2, eIF3 and eIF4B in cryo buffer (20 mM Hepes (pH 7.5), 95 mM KOAc, 3.75 mM Mg (OAc)₂, 1 mM ATP, 0.5 mM GTP) at 37°C for 10 min. After the 48S IC formation, the complex is purified from the

Zeba spin desalting column (Thermo Scientific) to remove the trace amount of glycerol from the protein purification. The complex is then subject to the sucrose gradient containing glutathione to separate the intact 48S IC from other truncated complexes and cross-link the factors on the 40S subunit with published GraFix method (Stark, 2010). The purified 48S IC is ready for cryo-EM analysis.

5 REFERENCES

- Abramson, R.D., Dever, T.E., Lawson, T.G., Ray, B.K., Thach, R.E., and Merrick, W.C. (1987). The ATP-dependent interaction of eukaryotic initiation factors with mRNA. *J Biol Chem* 262, 3826-3832.
- Acker, M.G., Shin, B.S., Dever, T.E., and Lorsch, J.R. (2006). Interaction between eukaryotic initiation factors 1A and 5B is required for efficient ribosomal subunit joining. *Journal of Biological Chemistry* 281, 8469-8475.
- Acker, M.G., Shin, B.S., Nanda, J.S., Saini, A.K., Dever, T.E., and Lorsch, J.R. (2009). Kinetic Analysis of Late Steps of Eukaryotic Translation Initiation. *Journal of Molecular Biology* 385, 491-506.
- Adams, S.L., Safer, B., Anderson, W.F., and Merrick, W.C. (1975). Eukaryotic Initiation Complex-Formation - Evidence for 2 Distinct Pathways. *Journal of Biological Chemistry* 250, 9083-9089.
- Agalarov, S., Sakharov, P.A., Fattakhova, D., Sogorin, E.A., and Spirin, A.S. (2014). Internal translation initiation and eIF4F/ATP-independent scanning of mRNA by eukaryotic ribosomal particles. *Sci Rep* 4, 4438.
- Aitken, C.E., Beznoskova, P., Vlckova, V., Chiu, W.L., Zhou, F., Valasek, L.S., Hinnebusch, A.G., and Lorsch, J.R. (2016). Eukaryotic translation initiation factor 3 plays distinct roles at the mRNA entry and exit channels of the ribosomal preinitiation complex. *Elife* 5.
- Algire, M.A., Maag, D., and Lorsch, J.R. (2005). Pi release from eIF2, not GTP hydrolysis, is the step controlled by start-site selection during eukaryotic translation initiation. *Mol Cell* 20, 251-262.
- Andaya, A., Villa, N., Jia, W., Fraser, C.S., and Leary, J.A. (2014). Phosphorylation stoichiometries of human eukaryotic initiation factors. *Int J Mol Sci* 15, 11523-11538.
- Asano, K., Clayton, J., Shalev, A., and Hinnebusch, A.G. (2000). A multifactor complex of eukaryotic initiation factors, eIF1, eIF2, eIF3, eIF5, and initiator tRNA(Met) is an important translation initiation intermediate in vivo. *Genes Dev* 14, 2534-2546.
- Aylett, C.H., Boehringer, D., Erzberger, J.P., Schaefer, T., and Ban, N. (2015). Structure of a yeast 40S-eIF1-eIF1A-eIF3-eIF3j initiation complex. *Nat Struct Mol Biol* 22, 269-271.
- Babendure, J.R., Babendure, J.L., Ding, J.H., and Tsien, R.Y. (2006). Control of mammalian translation by mRNA structure near caps. *RNA* 12, 851-861.
- Barth-Baus, D., Bhasker, C.R., Zoll, W., and Merrick, W.C. (2013). Influence of translation factor activities on start site selection in six different mRNAs. *Translation (Austin)* 1, e24419.

- Battiste, J.L., Pestova, T.V., Hellen, C.U.T., and Wagner, G. (2000). The eIF1A solution structure reveals a large RNA-binding surface important for scanning function. *Molecular Cell* 5, 109-119.
- Beerman, R.W., and Jongens, T.A. (2011). A non-canonical start codon in the *Drosophila* fragile X gene yields two functional isoforms. *Neuroscience* 181, 48-66.
- Benne, R., and Hershey, J.W.B. (1978). Mechanism of Action of Protein-Synthesis Initiation-Factors from Rabbit Reticulocytes. *Journal of Biological Chemistry* 253, 3078-3087.
- Berthelot, K., Muldoon, M., Rajkowitsch, L., Hughes, J., and McCarthy, J.E. (2004). Dynamics and processivity of 40S ribosome scanning on mRNA in yeast. *Mol Microbiol* 51, 987-1001.
- Brito Querido, J., Sokabe, M., Kraatz, S., Gordiyenko, Y., Skehel, J.M., Fraser, C.S., and Ramakrishnan, V. (2020). Structure of a human 48S translational initiation complex. *Science* 369, 1220-1227.
- Browning, K.S., Gallie, D.R., Hershey, J.W., Hinnebusch, A.G., Maitra, U., Merrick, W.C., and Norbury, C. (2001). Unified nomenclature for the subunits of eukaryotic initiation factor 3. *Trends Biochem Sci* 26, 284.
- Burnett, J.C., Ruthel, G., Stegmann, C.M., Panchal, R.G., Nguyen, T.L., Hermone, A.R., Stafford, R.G., Lane, D.J., Kenny, T.A., McGrath, C.F., *et al.* (2007). Inhibition of metalloprotease botulinum serotype A from a pseudo-peptide binding mode to a small molecule that is active in primary neurons. *Journal of Biological Chemistry* 282, 5004-5014.
- Calvo, S.E., Pagliarini, D.J., and Mootha, V.K. (2009). Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. *Proc Natl Acad Sci U S A* 106, 7507-7512.
- Carberry, S.E., Friedland, D.E., Rhoads, R.E., and Goss, D.J. (1992). Binding of protein synthesis initiation factor 4E to oligoribonucleotides: effects of cap accessibility and secondary structure. *Biochemistry* 31, 1427-1432.
- Cate, J.H. (2017). Human eIF3: from 'blobology' to biological insight. *Philos Trans R Soc Lond B Biol Sci* 372.
- Ceci, M., Gaviraghi, C., Gorrini, C., Sala, L.A., Offenhauser, N., Marchisio, P.C., and Biffo, S. (2003). Release of eIF6 (p27(BBP)) from the 60S subunit allows 80S ribosome assembly. *Nature* 426, 579-584.
- Chang, K.J., and Wang, C.C. (2004). Translation initiation from a naturally occurring non-AUG codon in *Saccharomyces cerevisiae*. *J Biol Chem* 279, 13778-13785.

- Cheung, Y.N., Maag, D., Mitchell, S.F., Fekete, C.A., Algire, M.A., Takacs, J.E., Shirokikh, N., Pestova, T., Lorsch, J.R., and Hinnebusch, A.G. (2007). Dissociation of eIF1 from the 40S ribosomal subunit is a key step in start codon selection in vivo. *Genes Dev* 21, 1217-1230.
- Clements, J.M., Laz, T.M., and Sherman, F. (1988). Efficiency of translation initiation by non-AUG codons in *Saccharomyces cerevisiae*. *Mol Cell Biol* 8, 4533-4536.
- Coots, R.A., Liu, X.M., Mao, Y., Dong, L., Zhou, J., Wan, J., Zhang, X., and Qian, S.B. (2017). m(6)A Facilitates eIF4F-Independent mRNA Translation. *Mol Cell* 68, 504-514 e507.
- Cowling, V.H. (2009). Regulation of mRNA cap methylation. *Biochem J* 425, 295-302.
- Das, S., Ghosh, R., and Maitra, U. (2001). Eukaryotic translation initiation factor 5 functions as a GTPase-activating protein. *J Biol Chem* 276, 6720-6726.
- Das, S., Maiti, T., Das, K., and Maitra, U. (1997). Specific interaction of eukaryotic translation initiation factor 5 (eIF5) with the beta-subunit of eIF2. *J Biol Chem* 272, 31712-31718.
- Das, S., and Maitra, U. (2000). Mutational analysis of mammalian translation initiation factor 5 (eIF5): role of interaction between the beta subunit of eIF2 and eIF5 in eIF5 function in vitro and in vivo. *Mol Cell Biol* 20, 3942-3950.
- Das, S., and Maitra, U. (2001). Functional significance and mechanism of eIF5-promoted GTP hydrolysis in eukaryotic translation initiation. *Prog Nucleic Acid Res Mol Biol* 70, 207-231.
- Dennis, M.D., Person, M.D., and Browning, K.S. (2009). Phosphorylation of plant translation initiation factors by CK2 enhances the in vitro interaction of multifactor complex components. *J Biol Chem* 284, 20615-20628.
- Dmitriev, S.E., Terenin, I.M., Andreev, D.E., Ivanov, P.A., Dunaevsky, J.E., Merrick, W.C., and Shatsky, I.N. (2010). GTP-independent tRNA Delivery to the Ribosomal P-site by a Novel Eukaryotic Translation Factor. *Journal of Biological Chemistry* 285, 26779-26787.
- Duncan, R., and Hershey, J.W. (1983). Identification and quantitation of levels of protein synthesis initiation factors in crude HeLa cell lysates by two-dimensional polyacrylamide gel electrophoresis. *J Biol Chem* 258, 7228-7235.
- Elfakess, R., and Dikstein, R. (2008). A translation initiation element specific to mRNAs with very short 5'UTR that also regulates transcription. *PLoS One* 3, e3094.
- Elfakess, R., Sinvani, H., Haimov, O., Svitkin, Y., Sonenberg, N., and Dikstein, R. (2011). Unique translation initiation of mRNAs-containing TISU element. *Nucleic Acids Res* 39, 7598-7609.
- Eliseev, B., Yeramala, L., Leitner, A., Karuppasamy, M., Raimondeau, E., Huard, K., Alkalaeva, E., Aebersold, R., and Schaffitzel, C. (2018). Structure of a human cap-dependent 48S translation pre-initiation complex. *Nucleic Acids Res* 46, 2678-2689.

Fekete, C.A., Applefield, D.J., Blakely, S.A., Shirokikh, N., Pestova, T., Lorsch, J.R., and Hinnebusch, A.G. (2005). The eIF1A C-terminal domain promotes initiation complex assembly, scanning and AUG selection in vivo. *EMBO J* 24, 3588-3601.

Fekete, C.A., Mitchell, S.F., Cherkasova, V.A., Applefield, D., Algire, M.A., Maag, D., Saini, A.K., Lorsch, J.R., and Hinnebusch, A.G. (2007). N- and C-terminal residues of eIF1A have opposing effects on the fidelity of start codon selection. *EMBO J* 26, 1602-1614.

Feoktistova, K., Tuvshintogs, E., Do, A., and Fraser, C.S. (2013). Human eIF4E promotes mRNA restructuring by stimulating eIF4A helicase activity. *Proc Natl Acad Sci U S A* 110, 13339-13344.

Fernandez, I.S., Bai, X.C., Hussain, T., Kelley, A.C., Lorsch, J.R., Ramakrishnan, V., and Scheres, S.H.W. (2013). Molecular Architecture of a Eukaryotic Translational Initiation Complex. *Science* 342, 824-+.

Fraser, C.S., Berry, K.E., Hershey, J.W., and Doudna, J.A. (2007). eIF3j is located in the decoding center of the human 40S ribosomal subunit. *Mol Cell* 26, 811-819.

Fringer, J.M., Acker, M.G., Fekete, C.A., Lorsch, J.R., and Dever, T.E. (2007). Coupled release of eukaryotic translation initiation factors 5B and 1A from 80S ribosomes following subunit joining. *Mol Cell Biol* 27, 2384-2397.

Garcia-Garcia, C., Frieda, K.L., Feoktistova, K., Fraser, C.S., and Block, S.M. (2015). RNA BIOCHEMISTRY. Factor-dependent processivity in human eIF4A DEAD-box helicase. *Science* 348, 1486-1488.

Giess, A., Torres Cleuren, Y.N., Tjeldnes, H., Krause, M., Bizuayehu, T.T., Hiensch, S., Okon, A., Wagner, C.R., and Valen, E. (2020). Profiling of Small Ribosomal Subunits Reveals Modes and Regulation of Translation Initiation. *Cell Rep* 31, 107534.

Gingras, A.C., Gygi, S.P., Raught, B., Polakiewicz, R.D., Abraham, R.T., Hoekstra, M.F., Aebersold, R., and Sonenberg, N. (1999a). Regulation of 4E-BP1 phosphorylation: a novel two-step mechanism. *Genes Dev* 13, 1422-1437.

Gingras, A.C., Raught, B., Gygi, S.P., Niedzwiecka, A., Miron, M., Burley, S.K., Polakiewicz, R.D., Wyslouch-Cieszyńska, A., Aebersold, R., and Sonenberg, N. (2001). Hierarchical phosphorylation of the translation inhibitor 4E-BP1. *Genes Dev* 15, 2852-2864.

Gingras, A.C., Raught, B., and Sonenberg, N. (1999b). eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. *Annu Rev Biochem* 68, 913-963.

Goodfellow, I., Chaudhry, Y., Gioldasi, I., Gerondopoulos, A., Natoni, A., Labrie, L., Laliberte, J.F., and Roberts, L. (2005). Calicivirus translation initiation requires an interaction between VPg and eIF 4 E. *EMBO Rep* 6, 968-972.

Grifo, J.A., Abramson, R.D., Satler, C.A., and Merrick, W.C. (1984). RNA-stimulated ATPase activity of eukaryotic initiation factors. *J Biol Chem* 259, 8648-8654.

Grifo, J.A., Tahara, S.M., Morgan, M.A., Shatkin, A.J., and Merrick, W.C. (1983). New initiation factor activity required for globin mRNA translation. *J Biol Chem* 258, 5804-5810.

Gross, J.D., Moerke, N.J., von der Haar, T., Lugovskoy, A.A., Sachs, A.B., McCarthy, J.E., and Wagner, G. (2003a). Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E. *Cell* 115, 739-750.

Gross, J.D., Moerke, N.J., von der Haar, T., Lugovskoy, A.A., Sachs, A.B., McCarthy, J.E.G., and Wagner, G. (2003b). Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E. *Cell* 115, 739-750.

Gruner, S., Peter, D., Weber, R., Wohlbold, L., Chung, M.Y., Weichenrieder, O., Valkov, E., Igreja, C., and Izaurralde, E. (2016). The Structures of eIF4E-eIF4G Complexes Reveal an Extended Interface to Regulate Translation Initiation. *Mol Cell* 64, 467-479.

Harms, U., Andreou, A.Z., Gubaev, A., and Klostermeier, D. (2014). eIF4B, eIF4G and RNA regulate eIF4A activity in translation initiation by modulating the eIF4A conformational cycle. *Nucleic Acids Res* 42, 7911-7922.

Hashem, Y., des Georges, A., Dhote, V., Langlois, R., Liao, H.Y., Grassucci, R.A., Hellen, C.U., Pestova, T.V., and Frank, J. (2013a). Structure of the mammalian ribosomal 43S preinitiation complex bound to the scanning factor DHX29. *Cell* 153, 1108-1119.

Hashem, Y., des Georges, A., Dhote, V., Langlois, R., Liao, H.Y., Grassucci, R.A., Pestova, T.V., Hellen, C.U., and Frank, J. (2013b). Hepatitis-C-virus-like internal ribosome entry sites displace eIF3 to gain access to the 40S subunit. *Nature* 503, 539-543.

Hinnebusch, A.G. (2014). The scanning mechanism of eukaryotic translation initiation. *Annu Rev Biochem* 83, 779-812.

Hinnebusch, A.G. (2017). Structural Insights into the Mechanism of Scanning and Start Codon Recognition in Eukaryotic Translation Initiation. *Trends Biochem Sci* 42, 589-611.

Hinnebusch, A.G., and Lorsch, J.R. (2012). The mechanism of eukaryotic translation initiation: new insights and challenges. *Cold Spring Harb Perspect Biol* 4.

Hinton, T.M., Coldwell, M.J., Carpenter, G.A., Morley, S.J., and Pain, V.M. (2007). Functional analysis of individual binding activities of the scaffold protein eIF4G. *J Biol Chem* 282, 1695-1708.

Holz, M.K., Ballif, B.A., Gygi, S.P., and Blenis, J. (2005). mTOR and S6K1 mediate assembly of the translation preinitiation complex through dynamic protein interchange and ordered phosphorylation events. *Cell* 123, 569-580.

Homma, M.K., Wada, I., Suzuki, T., Yamaki, J., Krebs, E.G., and Homma, Y. (2005). CK2 phosphorylation of eukaryotic translation initiation factor 5 potentiates cell cycle progression. *Proc Natl Acad Sci USA* *102*, 15688-15693.

Huang, H.K., Yoon, H., Hannig, E.M., and Donahue, T.F. (1997). GTP hydrolysis controls stringent selection of the AUG start codon during translation initiation in *Saccharomyces cerevisiae*. *Gene Dev* *11*, 2396-2413.

Hussain, T., Llacer, J.L., Fernandez, I.S., Munoz, A., Martin-Marcos, P., Savva, C.G., Lorsch, J.R., Hinnebusch, A.G., and Ramakrishnan, V. (2014). Structural changes enable start codon recognition by the eukaryotic translation initiation complex. *Cell* *159*, 597-607.

Imataka, H., Olsen, H.S., and Sonenberg, N. (1997). A new translational regulator with homology to eukaryotic translation initiation factor 4G. *EMBO J* *16*, 817-825.

Imataka, H., and Sonenberg, N. (1997). Human eukaryotic translation initiation factor 4G (eIF4G) possesses two separate and independent binding sites for eIF4A. *Mol Cell Biol* *17*, 6940-6947.

Ingolia, N.T., Ghaemmaghami, S., Newman, J.R., and Weissman, J.S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* *324*, 218-223.

Ingolia, N.T., Lareau, L.F., and Weissman, J.S. (2011). Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell* *147*, 789-802.

Ivanov, I.P., Firth, A.E., Michel, A.M., Atkins, J.F., and Baranov, P.V. (2011). Identification of evolutionarily conserved non-AUG-initiated N-terminal extensions in human coding sequences. *Nucleic Acids Res* *39*, 4220-4234.

Ivanov, I.P., Loughran, G., and Atkins, J.F. (2008). uORFs with unusual translational start codons autoregulate expression of eukaryotic ornithine decarboxylase homologs. *Proc Natl Acad Sci U S A* *105*, 10079-10084.

Ivanov, I.P., Loughran, G., Sachs, M.S., and Atkins, J.F. (2010). Initiation context modulates autoregulation of eukaryotic translation initiation factor 1 (eIF1). *Proc Natl Acad Sci U S A* *107*, 18056-18060.

Jackson, R.J. (1991). The ATP requirement for initiation of eukaryotic translation varies according to the mRNA species. *Eur J Biochem* *200*, 285-294.

Jennings, M.D., Kershaw, C.J., Adomavicius, T., and Pavitt, G.D. (2017). Fail-safe control of translation initiation by dissociation of eIF2alpha phosphorylated ternary complexes. *Elife* *6*.

Jennings, M.D., Kershaw, C.J., White, C., Hoyle, D., Richardson, J.P., Costello, J.L., Donaldson, I.J., Zhou, Y., and Pavitt, G.D. (2016). eIF2beta is critical for eIF5-mediated GDP-dissociation inhibitor activity and translational control. *Nucleic Acids Res* 44, 9698-9709.

Jennings, M.D., and Pavitt, G.D. (2010). eIF5 has GDI activity necessary for translational control by eIF2 phosphorylation. *Nature* 465, 378-381.

Johnstone, T.G., Bazzini, A.A., and Giraldez, A.J. (2016). Upstream ORFs are prevalent translational repressors in vertebrates. *EMBO J* 35, 706-723.

Kapp, L.D., and Lorsch, J.R. (2004). GTP-dependent recognition of the methionine moiety on initiator tRNA by translation factor eIF2. *J Mol Biol* 335, 923-936.

Kearse, M.G., Green, K.M., Krans, A., Rodriguez, C.M., Linsalata, A.E., Goldstrohm, A.C., and Todd, P.K. (2016). CGG Repeat-Associated Non-AUG Translation Utilizes a Cap-Dependent Scanning Mechanism of Initiation to Produce Toxic Proteins. *Molecular Cell* 62, 314-322.

Kearse, M.G., and Wilusz, J.E. (2017). Non-AUG translation: a new start for protein synthesis in eukaryotes. *Genes Dev* 31, 1717-1731.

Kolitz, S.E., Takacs, J.E., and Lorsch, J.R. (2009). Kinetic and thermodynamic analysis of the role of start codon/anticodon base pairing during eukaryotic translation initiation. *RNA* 15, 138-152.

Komar, A.A., Gross, S.R., Barth-Baus, D., Strachan, R., Hensold, J.O., Kinzy, T.G., and Merrick, W.C. (2005). Novel characteristics of the biological properties of the yeast *Saccharomyces cerevisiae* eukaryotic initiation factor 2A. *Journal of Biological Chemistry* 280, 15601-15611.

Kozak, M. (1980a). Evaluation of the "scanning model" for initiation of protein synthesis in eucaryotes. *Cell* 22, 7-8.

Kozak, M. (1980b). Role of ATP in binding and migration of 40S ribosomal subunits. *Cell* 22, 459-467.

Kozak, M. (1981). Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. *Nucleic Acids Res* 9, 5233-5252.

Kozak, M. (1984). Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucleic Acids Res* 12, 857-872.

Kozak, M. (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* 44, 283-292.

Kozak, M. (1987a). An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res* 15, 8125-8148.

- Kozak, M. (1987b). At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. *J Mol Biol* *196*, 947-950.
- Kozak, M. (1990). Downstream secondary structure facilitates recognition of initiator codons by eukaryotic ribosomes. *Proc Natl Acad Sci U S A* *87*, 8301-8305.
- Kozak, M. (1991a). An analysis of vertebrate mRNA sequences: intimations of translational control. *J Cell Biol* *115*, 887-903.
- Kozak, M. (1991b). Effects of long 5' leader sequences on initiation by eukaryotic ribosomes in vitro. *Gene Expr* *1*, 117-125.
- Kozak, M. (1991c). A short leader sequence impairs the fidelity of initiation by eukaryotic ribosomes. *Gene Expr* *1*, 111-115.
- Kozak, M., and Shatkin, A.J. (1978). Migration of 40 S ribosomal subunits on messenger RNA in the presence of edeine. *J Biol Chem* *253*, 6568-6577.
- Kulak, N.A., Pichler, G., Paron, I., Nagaraj, N., and Mann, M. (2014). Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat Methods* *11*, 319-324.
- Kumar, P., Hellen, C.U., and Pestova, T.V. (2016). Toward the mechanism of eIF4F-mediated ribosomal attachment to mammalian capped mRNAs. *Genes Dev* *30*, 1573-1588.
- Kwan, T., and Thompson, S.R. (2019). Noncanonical Translation Initiation in Eukaryotes. *Cold Spring Harb Perspect Biol* *11*.
- Lawless, C., Pearson, R.D., Selley, J.N., Smirnova, J.B., Grant, C.M., Ashe, M.P., Pavitt, G.D., and Hubbard, S.J. (2009). Upstream sequence elements direct post-transcriptional regulation of gene expression under stress conditions in yeast. *BMC Genomics* *10*, 7.
- Lee, A.S., Kranzusch, P.J., and Cate, J.H. (2015). eIF3 targets cell-proliferation messenger RNAs for translational activation or repression. *Nature* *522*, 111-114.
- Lee, A.S., Kranzusch, P.J., Doudna, J.A., and Cate, J.H. (2016). eIF3d is an mRNA cap-binding protein that is required for specialized translation initiation. *Nature* *536*, 96-99.
- Lee, S., Liu, B., Lee, S., Huang, S.X., Shen, B., and Qian, S.B. (2012). Global mapping of translation initiation sites in mammalian cells at single-nucleotide resolution. *Proc Natl Acad Sci U S A* *109*, E2424-2432.
- Leppek, K., Das, R., and Barna, M. (2018). Functional 5' UTR mRNA structures in eukaryotic translation regulation and how to find them. *Nat Rev Mol Cell Biol* *19*, 158-174.

Liang, H., Chen, X., Yin, Q., Ruan, D., Zhao, X., Zhang, C., McNutt, M.A., and Yin, Y. (2017a). PTENbeta is an alternatively translated isoform of PTEN that regulates rDNA transcription. *Nat Commun* 8, 14771.

Liang, H., Chen, X., Yin, Q., Ruan, D.H., Zhao, X.Y., Zhang, C., McNutt, M.A., and Yin, Y.X. (2017b). PTEN beta is an alternatively translated isoform of PTEN that regulates rDNA transcription. *Nature Communications* 8.

Lieberman, N., Gandin, V., Svitkin, Y.V., David, M., Virgili, G., Jaramillo, M., Holcik, M., Nagar, B., Kimchi, A., and Sonenberg, N. (2015). DAP5 associates with eIF2beta and eIF4AI to promote Internal Ribosome Entry Site driven translation. *Nucleic Acids Res* 43, 3764-3775.

Lin, K.Y., Nag, N., Pestova, T.V., and Marintchev, A. (2018). Human eIF5 and eIF1A Compete for Binding to eIF5B. *Biochemistry* 57, 5910-5920.

Lin, Y., Li, F., Huang, L., Polte, C., Duan, H., Fang, J., Sun, L., Xing, X., Tian, G., Cheng, Y., *et al.* (2020). eIF3 Associates with 80S Ribosomes to Promote Translation Elongation, Mitochondrial Homeostasis, and Muscle Health. *Mol Cell* 79, 575-587 e577.

Linder, P., and Slonimski, P.P. (1989). An essential yeast protein, encoded by duplicated genes TIF1 and TIF2 and homologous to the mammalian translation initiation factor eIF-4A, can suppress a mitochondrial missense mutation. *Proc Natl Acad Sci U S A* 86, 2286-2290.

Llacer, J.L., Hussain, T., Marler, L., Aitken, C.E., Thakur, A., Lorsch, J.R., Hinnebusch, A.G., and Ramakrishnan, V. (2015). Conformational Differences between Open and Closed States of the Eukaryotic Translation Initiation Complex. *Mol Cell* 59, 399-412.

Llacer, J.L., Hussain, T., Saini, A.K., Nanda, J.S., Kaur, S., Gordiyenko, Y., Kumar, R., Hinnebusch, A.G., Lorsch, J.R., and Ramakrishnan, V. (2018). Translational initiation factor eIF5 replaces eIF1 on the 40S ribosomal subunit to promote start-codon recognition. *Elife* 7.

Lomakin, I.B., Shirokikh, N.E., Yusupov, M.M., Hellen, C.U., and Pestova, T.V. (2006). The fidelity of translation initiation: reciprocal activities of eIF1, IF3 and YciH. *EMBO J* 25, 196-210.

Lomakin, I.B., and Steitz, T.A. (2013). The initiation of mammalian protein synthesis and mRNA scanning mechanism. *Nature* 500, 307-311.

Loughran, G., Sachs, M.S., Atkins, J.F., and Ivanov, I.P. (2012). Stringency of start codon selection modulates autoregulation of translation initiation factor eIF5. *Nucleic Acids Res* 40, 2898-2906.

Lu, P.D., Harding, H.P., and Ron, D. (2004). Translation reinitiation at alternative open reading frames regulates gene expression in an integrated stress response. *J Cell Biol* 167, 27-33.

Lu, W.T., Wilczynska, A., Smith, E., and Bushell, M. (2014). The diverse roles of the eIF4A family: you are the company you keep. *Biochem Soc Trans* 42, 166-172.

Lukaszewicz, M., Feuermann, M., Jerouville, B., Stas, A., and Boutry, M. (2000). In vivo evaluation of the context sequence of the translation initiation codon in plants. *Plant Sci* 154, 89-98.

Luna, R.E., Arthanari, H., Hiraishi, H., Akabayov, B., Tang, L., Cox, C., Markus, M.A., Luna, L.E., Ikeda, Y., Watanabe, R., *et al.* (2013). The interaction between eukaryotic initiation factor 1A and eIF5 retains eIF1 within scanning preinitiation complexes. *Biochemistry* 52, 9510-9518.

Ma, J., Haldar, S., Khan, M.A., Sharma, S.D., Merrick, W.C., Theil, E.C., and Goss, D.J. (2012). Fe²⁺ binds iron responsive element-RNA, selectively changing protein-binding affinities and regulating mRNA repression and activation. *Proc Natl Acad Sci U S A* 109, 8417-8422.

Maag, D., Algire, M.A., and Lorsch, J.R. (2006). Communication between eukaryotic translation initiation factors 5 and 1A within the ribosomal pre-initiation complex plays a role in start site selection. *J Mol Biol* 356, 724-737.

Maag, D., Fekete, C.A., Gryczynski, Z., and Lorsch, J.R. (2005). A conformational change in the eukaryotic translation preinitiation complex and release of eIF1 signal recognition of the start codon. *Mol Cell* 17, 265-275.

Maag, D., and Lorsch, J.R. (2003). Communication Between Eukaryotic Translation Initiation Factors 1 and 1A on the Yeast Small Ribosomal Subunit. *Journal of Molecular Biology* 330, 917-924.

Mader, S., Lee, H., Pause, A., and Sonenberg, N. (1995). The translation initiation factor eIF-4E binds to a common motif shared by the translation factor eIF-4 gamma and the translational repressors 4E-binding proteins. *Mol Cell Biol* 15, 4990-4997.

Majumdar, R., Bandyopadhyay, A., Deng, H.T., and Maitra, U. (2002). Phosphorylation of mammalian translation initiation factor 5 (eIF5) in vitro and in vivo. *Nucleic Acids Research* 30, 1154-1162.

Majumdar, R., Bandyopadhyay, A., and Maitra, U. (2003). Mammalian translation initiation factor eIF1 functions with eIF1A and eIF3 in the formation of a stable 40 S preinitiation complex. *J Biol Chem* 278, 6580-6587.

Majumdar, R., and Maitra, U. (2005). Regulation of GTP hydrolysis prior to ribosomal AUG selection during eukaryotic translation initiation. *EMBO J* 24, 3737-3746.

Marcotrigiano, J., Gingras, A.C., Sonenberg, N., and Burley, S.K. (1999). Cap-dependent translation initiation in eukaryotes is regulated by a molecular mimic of eIF4G. *Mol Cell* 3, 707-716.

Marintchev, A., Kolupaeva, V.G., Pestova, T.V., and Wagner, G. (2003). Mapping the binding interface between human eukaryotic initiation factors 1A and 5B: A new interaction between old partners. *P Natl Acad Sci USA* 100, 1535-1540.

- Marintchev, A., and Wagner, G. (2004). Translation initiation: structures, mechanisms and evolution. *Q Rev Biophys* 37, 197-284.
- Martin-Marcos, P., Cheung, Y.N., and Hinnebusch, A.G. (2011). Functional elements in initiation factors 1, 1A, and 2beta discriminate against poor AUG context and non-AUG start codons. *Mol Cell Biol* 31, 4814-4831.
- Martin-Marcos, P., Nanda, J., Luna, R.E., Wagner, G., Lorsch, J.R., and Hinnebusch, A.G. (2013). beta-Hairpin loop of eukaryotic initiation factor 1 (eIF1) mediates 40 S ribosome binding to regulate initiator tRNA(Met) recruitment and accuracy of AUG selection in vivo. *J Biol Chem* 288, 27546-27562.
- Martin-Marcos, P., Nanda, J.S., Luna, R.E., Zhang, F., Saini, A.K., Cherkasova, V.A., Wagner, G., Lorsch, J.R., and Hinnebusch, A.G. (2014). Enhanced eIF1 binding to the 40S ribosome impedes conformational rearrangements of the preinitiation complex and elevates initiation accuracy. *RNA* 20, 150-167.
- Mayr, C. (2017). Regulation by 3'-Untranslated Regions. *Annu Rev Genet* 51, 171-194.
- Mazumder, B., Seshadri, V., and Fox, P.L. (2003). Translational control by the 3'-UTR: the ends specify the means. *Trends Biochem Sci* 28, 91-98.
- Merrick, W.C. (2015). eIF4F: a retrospective. *J Biol Chem* 290, 24091-24099.
- Merrick, W.C., and Pavitt, G.D. (2018). Protein Synthesis Initiation in Eukaryotic Cells. *Cold Spring Harb Perspect Biol* 10.
- Meyer, K.D., Patil, D.P., Zhou, J., Zinoviev, A., Skabkin, M.A., Elemento, O., Pestova, T.V., Qian, S.B., and Jaffrey, S.R. (2015). 5' UTR m(6)A Promotes Cap-Independent Translation. *Cell* 163, 999-1010.
- Mignone, F., Gissi, C., Liuni, S., and Pesole, G. (2002). Untranslated regions of mRNAs. *Genome Biol* 3, REVIEWS0004.
- Milon, P., Maracci, C., Filonava, L., Gualerzi, C.O., and Rodnina, M.V. (2012). Real-time assembly landscape of bacterial 30S translation initiation complex. *Nat Struct Mol Biol* 19, 609-615.
- Mitchell, S.F., Walker, S.E., Algire, M.A., Park, E.H., Hinnebusch, A.G., and Lorsch, J.R. (2010). The 5'-7-methylguanosine cap on eukaryotic mRNAs serves both to stimulate canonical translation initiation and to block an alternative pathway. *Mol Cell* 39, 950-962.
- Modelska, A., Turro, E., Russell, R., Beaton, J., Sbarrato, T., Spriggs, K., Miller, J., Graf, S., Provenzano, E., Blows, F., *et al.* (2015). The malignant phenotype in breast cancer is driven by eIF4A1-mediated changes in the translational landscape. *Cell Death Dis* 6, e1603.

Mohammad, M.P., Munzarova Pondelickova, V., Zeman, J., Gunisova, S., and Valasek, L.S. (2017). In vivo evidence that eIF3 stays bound to ribosomes elongating and terminating on short upstream ORFs to promote reinitiation. *Nucleic Acids Res* **45**, 2658-2674.

Moran, U., Phillips, R., and Milo, R. (2010). SnapShot: key numbers in biology. *Cell* **141**, 1262-1262 e1261.

Muckenthaler, M., Gray, N.K., and Hentze, M.W. (1998). IRP-1 binding to ferritin mRNA prevents the recruitment of the small ribosomal subunit by the cap-binding complex eIF4F. *Mol Cell* **2**, 383-388.

Nag, N., Lin, K.Y., Edmonds, K.A., Yu, J., Nadkarni, D., Marintcheva, B., and Marintchev, A. (2016). eIF1A/eIF5B interaction network and its functions in translation initiation complex assembly and remodeling. *Nucleic Acids Res* **44**, 7441-7456.

Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M., and Snyder, M. (2008). The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* **320**, 1344-1349.

Nanda, J.S., Cheung, Y.N., Takacs, J.E., Martin-Marcos, P., Saini, A.K., Hinnebusch, A.G., and Lorsch, J.R. (2009). eIF1 Controls Multiple Steps in Start Codon Recognition during Eukaryotic Translation Initiation. *Journal of Molecular Biology* **394**, 268-285.

Nanda, J.S., Saini, A.K., Munoz, A.M., Hinnebusch, A.G., and Lorsch, J.R. (2013). Coordinated movements of eukaryotic translation initiation factors eIF1, eIF1A, and eIF5 trigger phosphate release from eIF2 in response to start codon recognition by the ribosomal preinitiation complex. *J Biol Chem* **288**, 5316-5329.

Naveau, M., Lazennec-Schurdevin, C., Panvert, M., Dubiez, E., Mechulam, Y., and Schmitt, E. (2013). Roles of yeast eIF2alpha and eIF2beta subunits in the binding of the initiator methionyl-tRNA. *Nucleic Acids Res* **41**, 1047-1057.

Niedzwiecka, A., Marcotrigiano, J., Stepinski, J., Jankowska-Anyszka, M., Wyslouch-Cieszyńska, A., Dadlez, M., Gingras, A.C., Mak, P., Darzynkiewicz, E., Sonenberg, N., *et al.* (2002a). Biophysical studies of eIF4E cap-binding protein: recognition of mRNA 5' cap structure and synthetic fragments of eIF4G and 4E-BP1 proteins. *J Mol Biol* **319**, 615-635.

Niedzwiecka, A., Stepinski, J., Darzynkiewicz, E., Sonenberg, N., and Stolarski, R. (2002b). Positive heat capacity change upon specific binding of translation initiation factor eIF4E to mRNA 5' cap. *Biochemistry* **41**, 12140-12148.

Nielsen, K.H., Behrens, M.A., He, Y., Oliveira, C.L., Jensen, L.S., Hoffmann, S.V., Pedersen, J.S., and Andersen, G.R. (2011). Synergistic activation of eIF4A by eIF4B and eIF4G. *Nucleic Acids Res* **39**, 2678-2689.

Noderer, W.L., Flockhart, R.J., Bhaduri, A., Diaz de Arce, A.J., Zhang, J., Khavari, P.A., and Wang, C.L. (2014). Quantitative analysis of mammalian translation initiation sites by FACS-seq. *Mol Syst Biol* 10, 748.

O'Leary, S.E., Petrov, A., Chen, J., and Puglisi, J.D. (2013). Dynamic recognition of the mRNA cap by *Saccharomyces cerevisiae* eIF4E. *Structure* 21, 2197-2207.

Obayashi, E., Luna, R.E., Nagata, T., Martin-Marcos, P., Hiraishi, H., Singh, C.R., Erzberger, J.P., Zhang, F., Arthanari, H., Morris, J., *et al.* (2017). Molecular Landscape of the Ribosome Pre-initiation Complex during mRNA Scanning: Structural Role for eIF3c and Its Control by eIF5. *Cell Rep* 18, 2651-2663.

Olsen, D.S., Savner, E.M., Mathew, A., Zhang, F., Krishnamoorthy, T., Phan, L., and Hinnebusch, A.G. (2003). Domains of eIF1A that mediate binding to eIF2, eIF3 and eIF5B and promote ternary complex recruitment in vivo. *Embo Journal* 22, 193-204.

Ozes, A.R., Feoktistova, K., Avanzino, B.C., and Fraser, C.S. (2011). Duplex unwinding and ATPase activities of the DEAD-box helicase eIF4A are coupled by eIF4G and eIF4B. *J Mol Biol* 412, 674-687.

Park, H.S., Himmelbach, A., Browning, K.S., Hohn, T., and Ryabova, L.A. (2001). A plant viral "reinitiation" factor interacts with the host translational machinery. *Cell* 106, 723-733.

Parsyan, A., Shahbazian, D., Martineau, Y., Petroulakis, E., Alain, T., Larsson, O., Mathonnet, G., Tettweiler, G., Hellen, C.U., Pestova, T.V., *et al.* (2009). The helicase protein DHX29 promotes translation initiation, cell proliferation, and tumorigenesis. *Proc Natl Acad Sci U S A* 106, 22217-22222.

Passmore, L.A., Schmeing, T.M., Maag, D., Applefield, D.J., Acker, M.G., Algire, M.A., Lorsch, J.R., and Ramakrishnan, V. (2007). The eukaryotic translation initiation factors eIF1 and eIF1A induce an open conformation of the 40S ribosome. *Mol Cell* 26, 41-50.

Paulin, F.E., Campbell, L.E., O'Brien, K., Loughlin, J., and Proud, C.G. (2001). Eukaryotic translation initiation factor 5 (eIF5) acts as a classical GTPase-activator protein. *Curr Biol* 11, 55-59.

Pavitt, G.D., Ramaiah, K.V., Kimball, S.R., and Hinnebusch, A.G. (1998). eIF2 independently binds two distinct eIF2B subcomplexes that catalyze and regulate guanine-nucleotide exchange. *Genes Dev* 12, 514-526.

Paytubi, S., Wang, X., Lam, Y.W., Izquierdo, L., Hunter, M.J., Jan, E., Hundal, H.S., and Proud, C.G. (2009). ABC50 promotes translation initiation in mammalian cells. *J Biol Chem* 284, 24061-24073.

Peabody, D.S. (1989). Translation initiation at non-AUG triplets in mammalian cells. *J Biol Chem* 264, 5031-5035.

Pesole, G., Mignone, F., Gissi, C., Grillo, G., Licciulli, F., and Liuni, S. (2001). Structural and functional features of eukaryotic mRNA untranslated regions. *Gene* 276, 73-81.

Pestova, E.V., and Morrison, D.A. (1998). Isolation and characterization of three *Streptococcus pneumoniae* transformation-specific loci by use of a lacZ reporter insertion vector. *J Bacteriol* 180, 2701-2710.

Pestova, T.V., Borukhov, S.I., and Hellen, C.U.T. (1998a). Eukaryotic ribosomes require initiation factors 1 and 1A to locate initiation codons. *Nature* 394, 854-859.

Pestova, T.V., and Kolupaeva, V.G. (2002). The roles of individual eukaryotic translation initiation factors in ribosomal scanning and initiation codon selection. *Genes Dev* 16, 2906-2922.

Pestova, T.V., Lomakin, I.B., Lee, J.H., Choi, S.K., Dever, T.E., and Hellen, C.U.T. (2000). The joining of ribosomal subunits in eukaryotes requires eIF5B. *Nature* 403, 332-335.

Pestova, T.V., Shatsky, I.N., Fletcher, S.P., Jackson, R.J., and Hellen, C.U. (1998b). A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initiation of hepatitis C and classical swine fever virus RNAs. *Genes Dev* 12, 67-83.

Peterson, D.T., Merrick, W.C., and Safer, B. (1979). Binding and release of radiolabeled eukaryotic initiation factors 2 and 3 during 80 S initiation complex formation. *J Biol Chem* 254, 2509-2516.

Pisarev, A.V., Hellen, C.U., and Pestova, T.V. (2007a). Recycling of eukaryotic posttermination ribosomal complexes. *Cell* 131, 286-299.

Pisarev, A.V., Kolupaeva, V.G., Pisareva, V.P., Merrick, W.C., Hellen, C.U., and Pestova, T.V. (2006). Specific functional interactions of nucleotides at key -3 and +4 positions flanking the initiation codon with components of the mammalian 48S translation initiation complex. *Genes Dev* 20, 624-636.

Pisarev, A.V., Kolupaeva, V.G., Yusupov, M.M., Hellen, C.U., and Pestova, T.V. (2008). Ribosomal position and contacts of mRNA in eukaryotic translation initiation complexes. *EMBO J* 27, 1609-1621.

Pisarev, A.V., Unbehaun, A., Hellen, C.U., and Pestova, T.V. (2007b). Assembly and analysis of eukaryotic translation initiation complexes. *Methods Enzymol* 430, 147-177.

Pisareva, V.P., Hellen, C.U.T., and Pestova, T.V. (2007). Kinetic analysis of the interaction of guanine nucleotides with eukaryotic translation initiation factor eIF5B. *Biochemistry* 46, 2622-2629.

Pisareva, V.P., and Pisarev, A.V. (2014). eIF5 and eIF5B together stimulate 48S initiation complex formation during ribosomal scanning. *Nucleic Acids Research* 42, 12052-12069.

- Pisareva, V.P., and Pisarev, A.V. (2016). DHX29 and eIF3 cooperate in ribosomal scanning on structured mRNAs during translation initiation. *RNA* 22, 1859-1870.
- Pisareva, V.P., Pisarev, A.V., Komar, A.A., Hellen, C.U., and Pestova, T.V. (2008). Translation initiation on mammalian mRNAs with structured 5'UTRs requires DExH-box protein DHX29. *Cell* 135, 1237-1250.
- Poyry, T.A., Kaminski, A., and Jackson, R.J. (2004). What determines whether mammalian ribosomes resume scanning after translation of a short upstream open reading frame? *Genes Dev* 18, 62-75.
- Rabl, J., Leibundgut, M., Ataide, S.F., Haag, A., and Ban, N. (2011). Crystal structure of the eukaryotic 40S ribosomal subunit in complex with initiation factor 1. *Science* 331, 730-736.
- Resch, A.M., Ogurtsov, A.Y., Rogozin, I.B., Shabalina, S.A., and Koonin, E.V. (2009). Evolution of alternative and constitutive regions of mammalian 5'UTRs. *BMC Genomics* 10, 162.
- Richter, J.D., and Sonenberg, N. (2005). Regulation of cap-dependent translation by eIF4E inhibitory proteins. *Nature* 433, 477-480.
- Rogers, G.W., Jr., Lima, W.F., and Merrick, W.C. (2001a). Further characterization of the helicase activity of eIF4A. Substrate specificity. *J Biol Chem* 276, 12598-12608.
- Rogers, G.W., Jr., Richter, N.J., Lima, W.F., and Merrick, W.C. (2001b). Modulation of the helicase activity of eIF4A by eIF4B, eIF4H, and eIF4F. *J Biol Chem* 276, 30914-30922.
- Romani, A.M.P. (2011). Intracellular magnesium homeostasis. In *Magnesium in the Central Nervous System*, R. Vink, and M. Nechifor, eds. (Adelaide (AU)).
- Rowlands, A.G., Panniers, R., and Henshaw, E.C. (1988). The catalytic mechanism of guanine nucleotide exchange factor action and competitive inhibition by phosphorylated eukaryotic initiation factor 2. *J Biol Chem* 263, 5526-5533.
- Roy, B., Vaughn, J.N., Kim, B.H., Zhou, F., Gilchrist, M.A., and Von Arnim, A.G. (2010). The h subunit of eIF3 promotes reinitiation competence during translation of mRNAs harboring upstream open reading frames. *RNA* 16, 748-761.
- Rozen, F., Edery, I., Meerovitch, K., Dever, T.E., Merrick, W.C., and Sonenberg, N. (1990). Bidirectional RNA helicase activity of eucaryotic translation initiation factors 4A and 4F. *Mol Cell Biol* 10, 1134-1144.
- Rubio, C.A., Weisburd, B., Holderfield, M., Arias, C., Fang, E., DeRisi, J.L., and Fanidi, A. (2014). Transcriptome-wide characterization of the eIF4A signature highlights plasticity in translation regulation. *Genome Biol* 15, 476.

- Safer, B., Adams, S.L., Kemper, W.M., Berry, K.W., Lloyd, M., and Merrick, W.C. (1976). Purification and characterization of two initiation factors required for maximal activity of a highly fractionated globin mRNA translation system. *Proc Natl Acad Sci U S A* 73, 2584-2588.
- Saini, A.K., Nanda, J.S., Lorsch, J.R., and Hinnebusch, A.G. (2010). Regulatory elements in eIF1A control the fidelity of start codon selection by modulating tRNA(i)(Met) binding to the ribosome. *Genes Dev* 24, 97-110.
- Saini, A.K., Nanda, J.S., Martin-Marcos, P., Dong, J., Zhang, F., Bhardwaj, M., Lorsch, J.R., and Hinnebusch, A.G. (2014). Eukaryotic translation initiation factor eIF5 promotes the accuracy of start codon recognition by regulating Pi release and conformational transitions of the preinitiation complex. *Nucleic Acids Res* 42, 9623-9640.
- Saini, H.K., and Fischer, D. (2007). Structural and functional insights into Mimivirus ORFans. *BMC Genomics* 8, 115.
- Scheel, H., and Hofmann, K. (2005). Prediction of a common structural scaffold for proteasome lid, COP9-signalosome and eIF3 complexes. *BMC Bioinformatics* 6, 71.
- Schmitt, E., Panvert, M., Lazennec-Schurdevin, C., Coureux, P.D., Perez, J., Thompson, A., and Mechulam, Y. (2012). Structure of the ternary initiation complex aIF2-GDPNP-methionylated initiator tRNA. *Nat Struct Mol Biol* 19, 450-454.
- Schott, J., Reitter, S., Philipp, J., Haneke, K., Schafer, H., and Stoecklin, G. (2014). Translational regulation of specific mRNAs controls feedback inhibition and survival during macrophage activation. *PLoS Genet* 10, e1004368.
- Schreier, M.H., Erni, B., and Staehelin, T. (1977). Initiation of mammalian protein synthesis. I. Purification and characterization of seven initiation factors. *J Mol Biol* 116, 727-753.
- Schreier, M.H., and Staehelin, T. (1973). Translation of rabbit hemoglobin messenger RNA in vitro with purified and partially purified components from brain or liver of different species. *Proc Natl Acad Sci U S A* 70, 462-465.
- Sendoel, A., Dunn, J.G., Rodriguez, E.H., Naik, S., Gomez, N.C., Hurwitz, B., Levorse, J., Dill, B.D., Schramek, D., Molina, H., *et al.* (2017). Translation from unconventional 5' start sites drives tumor initiation. *Cancer Research* 77.
- Sette, M., vanTilborg, P., Spurio, R., Kaptein, R., Paci, M., Gualerzi, C.O., and Boelens, R. (1997). The structure of the translational initiation factor IF1 from E-coli contains an oligomer-binding motif. *Embo Journal* 16, 1436-1443.
- Shah, P., Ding, Y., Niemczyk, M., Kudla, G., and Plotkin, J.B. (2013). Rate-limiting steps in yeast protein translation. *Cell* 153, 1589-1601.

Sharifulin, D.E., Bartuli, Y.S., Meschaninova, M.I., Ven'yaminova, A.G., Graifer, D.M., and Karpova, G.G. (2016). Exploring accessibility of structural elements of the mammalian 40S ribosomal mRNA entry channel at various steps of translation initiation. *Biochim Biophys Acta* 1864, 1328-1338.

Sharma, A.K., Sormanni, P., Ahmed, N., Ciryam, P., Friedrich, U.A., Kramer, G., and O'Brien, E.P. (2019). A chemical kinetic basis for measuring translation initiation and elongation rates from ribosome profiling data. *PLoS Comput Biol* 15, e1007070.

Shatsky, I.N., Terenin, I.M., Smirnova, V.V., and Andreev, D.E. (2018). Cap-Independent Translation: What's in a Name? *Trends Biochem Sci* 43, 882-895.

Shenvi, C.L., Dong, K.C., Friedman, E.M., Hanson, J.A., and Cate, J.H. (2005). Accessibility of 18S rRNA in human 40S subunits and 80S ribosomes at physiological magnesium ion concentrations--implications for the study of ribosome dynamics. *RNA* 11, 1898-1908.

Shin, B.S., Kim, J.R., Walker, S.E., Dong, J., Lorsch, J.R., and Dever, T.E. (2011). Initiation factor eIF2gamma promotes eIF2-GTP-Met-tRNAⁱ(Met) ternary complex binding to the 40S ribosome. *Nat Struct Mol Biol* 18, 1227-1234.

Simonetti, A., Brito Querido, J., Myasnikov, A.G., Mancera-Martinez, E., Renaud, A., Kuhn, L., and Hashem, Y. (2016). eIF3 Peripheral Subunits Rearrangement after mRNA Binding and Start-Codon Recognition. *Mol Cell* 63, 206-217.

Simonetti, A., Guca, E., Bochler, A., Kuhn, L., and Hashem, Y. (2020). Structural Insights into the Mammalian Late-Stage Initiation Complexes. *Cell Rep* 31, 107497.

Singh, C.R., Watanabe, R., Chowdhury, W., Hiraishi, H., Murai, M.J., Yamamoto, Y., Miles, D., Ikeda, Y., Asano, M., and Asano, K. (2012). Sequential eukaryotic translation initiation factor 5 (eIF5) binding to the charged disordered segments of eIF4G and eIF2beta stabilizes the 48S preinitiation complex and promotes its shift to the initiation mode. *Mol Cell Biol* 32, 3978-3989.

Sinvani, H., Haimov, O., Svitkin, Y., Sonenberg, N., Tamarkin-Ben-Harush, A., Viollet, B., and Dikstein, R. (2015). Translational tolerance of mitochondrial genes to metabolic energy stress involves TISU and eIF1-eIF4G cooperation in start codon selection. *Cell Metab* 21, 479-492.

Skabkin, M.A., Skabkina, O.V., Dhote, V., Komar, A.A., Hellen, C.U.T., and Pestova, T.V. (2010). Activities of Ligatin and MCT-1/DENR in eukaryotic translation initiation and ribosomal recycling. *Gene Dev* 24, 1787-1801.

Skabkin, M.A., Skabkina, O.V., Hellen, C.U.T., and Pestova, T.V. (2013). Reinitiation and Other Unconventional Posttermination Events during Eukaryotic Translation. *Molecular Cell* 51, 249-264.

Slusher, L.B., Gillman, E.C., Martin, N.C., and Hopper, A.K. (1991). mRNA leader length and initiation codon context determine alternative AUG selection for the yeast gene MOD5. *Proc Natl Acad Sci U S A* **88**, 9789-9793.

Sokabe, M., and Fraser, C.S. (2014). Human eukaryotic initiation factor 2 (eIF2)-GTP-Met-tRNA_i ternary complex and eIF3 stabilize the 43 S preinitiation complex. *J Biol Chem* **289**, 31827-31836.

Sokabe, M., and Fraser, C.S. (2017). A helicase-independent activity of eIF4A in promoting mRNA recruitment to the human ribosome. *Proc Natl Acad Sci U S A* **114**, 6304-6309.

Sokabe, M., Fraser, C.S., and Hershey, J.W. (2012). The human translation initiation multi-factor complex promotes methionyl-tRNA_i binding to the 40S ribosomal subunit. *Nucleic Acids Res* **40**, 905-913.

Soto-Rifo, R., Rubilar, P.S., Limousin, T., de Breyne, S., Decimo, D., and Ohlmann, T. (2012). DEAD-box protein DDX3 associates with eIF4F to promote translation of selected mRNAs. *EMBO J* **31**, 3745-3756.

Spirin, A.S. (2009). How does a scanning ribosomal particle move along the 5'-untranslated region of eukaryotic mRNA? Brownian Ratchet model. *Biochemistry* **48**, 10688-10692.

Srivastava, S., Verschoor, A., and Frank, J. (1992). Eukaryotic initiation factor 3 does not prevent association through physical blockage of the ribosomal subunit-subunit interface. *J Mol Biol* **226**, 301-304.

Starck, S.R., Jiang, V., Pavon-Eternod, M., Prasad, S., McCarthy, B., Pan, T., and Shastri, N. (2012). Leucine-tRNA initiates at CUG start codons for protein synthesis and presentation by MHC class I. *Science* **336**, 1719-1723.

Starck, S.R., Tsai, J.C., Chen, K.L., Shodiya, M., Wang, L., Yahiro, K., Martins-Green, M., Shastri, N., and Walter, P. (2016). Translation from the 5' untranslated region shapes the integrated stress response. *Science* **351**.

Stark, H. (2010). GraFix: stabilization of fragile macromolecular complexes for single particle cryo-EM. *Methods Enzymol* **481**, 109-126.

Svitkin, Y.V., Pause, A., Haghighat, A., Pyronnet, S., Witherell, G., Belsham, G.J., and Sonenberg, N. (2001). The requirement for eukaryotic initiation factor 4A (eIF4A) in translation is in direct proportion to the degree of mRNA 5' secondary structure. *RNA* **7**, 382-394.

Tahara, S.M., Morgan, M.A., and Shatkin, A.J. (1981). Two forms of purified m7G-cap binding protein with different effects on capped mRNA translation in extracts of uninfected and poliovirus-infected HeLa cells. *J Biol Chem* **256**, 7691-7694.

Takahashi, K., Maruyama, M., Tokuzawa, Y., Murakami, M., Oda, Y., Yoshikane, N., Makabe, K.W., Ichisaka, T., and Yamanaka, S. (2005). Evolutionarily conserved non-AUG translation initiation in NAT1/p97/DAP5 (EIF4G2). *Genomics* 85, 360-371.

Trachsel, H., and Staehelin, T. (1979). Initiation of mammalian protein synthesis. The multiple functions of the initiation factor eIF-3. *Biochim Biophys Acta* 565, 305-314.

Unbehaun, A., Borukhov, S.I., Hellen, C.U., and Pestova, T.V. (2004). Release of initiation factors from 48S complexes during ribosomal subunit joining and the link between establishment of codon-anticodon base-pairing and hydrolysis of eIF2-bound GTP. *Genes Dev* 18, 3078-3093.

Unbehaun, A., Marintchev, A., Lomakin, I.B., Didenko, T., Wagner, G., Hellen, C.U., and Pestova, T.V. (2007). Position of eukaryotic initiation factor eIF5B on the 80S ribosome mapped by directed hydroxyl radical probing. *EMBO J* 26, 3109-3123.

Valasek, L.S. (2012). 'Ribozoomin'--translation initiation from the perspective of the ribosome-bound eukaryotic initiation factors (eIFs). *Curr Protein Pept Sci* 13, 305-330.

Vassilenko, K.S., Alekhina, O.M., Dmitriev, S.E., Shatsky, I.N., and Spirin, A.S. (2011). Unidirectional constant rate motion of the ribosomal scanning particle during eukaryotic translation initiation. *Nucleic Acids Res* 39, 5555-5567.

Vattem, K.M., and Wek, R.C. (2004). Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. *Proc Natl Acad Sci U S A* 101, 11269-11274.

Villa, N., Do, A., Hershey, J.W., and Fraser, C.S. (2013). 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.

von der Haar, T., and McCarthy, J.E. (2002). Intracellular translation initiation factor levels in *Saccharomyces cerevisiae* and their role in cap-complex function. *Mol Microbiol* 46, 531-544.

Wei, J., Zhang, Y., Ivanov, I.P., and Sachs, M.S. (2013). The stringency of start codon selection in the filamentous fungus *Neurospora crassa*. *J Biol Chem* 288, 9549-9562.

Weisser, M., Schafer, T., Leibundgut, M., Bohringer, D., Aylett, C.H.S., and Ban, N. (2017). Structural and Functional Insights into Human Re-initiation Complexes. *Molecular Cell* 67, 447-+.

Weisser, M., Voigts-Hoffmann, F., Rabl, J., Leibundgut, M., and Ban, N. (2013). The crystal structure of the eukaryotic 40S ribosomal subunit in complex with eIF1 and eIF1A. *Nat Struct Mol Biol* 20, 1015-1017.

Wek, R.C. (2018). Role of eIF2alpha Kinases in Translational Control and Adaptation to Cellular Stress. *Cold Spring Harb Perspect Biol* 10.

Wilson, J.E., Pestova, T.V., Hellen, C.U., and Sarnow, P. (2000a). Initiation of protein synthesis from the A site of the ribosome. *Cell* 102, 511-520.

Wilson, J.E., Powell, M.J., Hoover, S.E., and Sarnow, P. (2000b). Naturally occurring dicistronic cricket paralysis virus RNA is regulated by two internal ribosome entry sites. *Molecular and Cellular Biology* 20, 4990-4999.

Wolfe, A.L., Singh, K., Zhong, Y., Drewe, P., Rajasekhar, V.K., Sanghvi, V.R., Mavrakis, K.J., Jiang, M., Roderick, J.E., Van der Meulen, J., *et al.* (2014). RNA G-quadruplexes cause eIF4A-dependent oncogene translation in cancer. *Nature* 513, 65-70.

Yamamoto, H., Unbehaun, A., Loerke, J., Behrmann, E., Collier, M., Burger, J., Mielke, T., and Spahn, C.M.T. (2014). Structure of the mammalian 80S initiation complex with initiation factor 5B on HCV-IRES RNA. *Nature Structural & Molecular Biology* 21, 721-727.

Yamamoto, Y., Singh, C.R., Marintchev, A., Hall, N.S., Hannig, E.M., Wagner, G., and Asano, K. (2005). The eukaryotic initiation factor (eIF) 5 HEAT domain mediates multifactor assembly and scanning with distinct interfaces to eIF1, eIF2, eIF3, and eIF4G. *Proc Natl Acad Sci U S A* 102, 16164-16169.

Yanagiya, A., Svitkin, Y.V., Shibata, S., Mikami, S., Imataka, H., and Sonenberg, N. (2009). Requirement of RNA binding of mammalian eukaryotic translation initiation factor 4G1 (eIF4G1) for efficient interaction of eIF4E with the mRNA cap. *Mol Cell Biol* 29, 1661-1669.

Yoffe, Y., David, M., Kalaora, R., Povodovski, L., Friedlander, G., Feldmesser, E., Ainbinder, E., Saada, A., Bialik, S., and Kimchi, A. (2016). Cap-independent translation by DAP5 controls cell fate decisions in human embryonic stem cells. *Genes Dev* 30, 1991-2004.

Yu, Y.P., Marintchev, A., Kolupaeva, V.G., Unbehaun, A., Veryasova, T., Lai, S.C., Hong, P., Wagner, G., Hellen, C.U.T., and Pestova, T.V. (2009). Position of eukaryotic translation initiation factor eIF1A on the 40S ribosomal subunit mapped by directed hydroxyl radical probing. *Nucleic Acids Research* 37, 5167-5182.

Zhang, F., Saini, A.K., Shin, B.S., Nanda, J., and Hinnebusch, A.G. (2015). Conformational changes in the P site and mRNA entry channel evoked by AUG recognition in yeast translation preinitiation complexes. *Nucleic Acids Res* 43, 2293-2312.

Zhou, J., Wan, J., Gao, X., Zhang, X., Jaffrey, S.R., and Qian, S.B. (2015). Dynamic m(6)A mRNA methylation directs translational control of heat shock response. *Nature* 526, 591-594.

Zinoviev, A., Hellen, C.U.T., and Pestova, T.V. (2015). Multiple Mechanisms of Reinitiation on Bicistronic Calicivirus mRNAs. *Molecular Cell* 57, 1059-1073.

Zitomer, R.S., Walthall, D.A., Rymond, B.C., and Hollenberg, C.P. (1984). *Saccharomyces cerevisiae* ribosomes recognize non-AUG initiation codons. *Mol Cell Biol* 4, 1191-1197.

Zoll, W.L., Horton, L.E., Komar, A.A., Hensold, J.O., and Merrick, W.C. (2002). Characterization of mammalian eIF2A and identification of the yeast homolog. *Journal of Biological Chemistry* 277, 37079-37087.

6 APPENDIX

6.1 List of figures

Figure 1. The 5' m ⁷ G cap structure in eukaryotic mRNA.....	7
Figure 2. Overview of canonical cap-dependent translation initiation pathway.....	8
Figure 3. Structure of human 48S PIC at 3.1 Å	12
Figure 4. eIF1A structure	13
Figure 5. Structures of 40S subunit and 40S·eIF1·eIF1A complex.....	15
Figure 6. Interspecies variation in mRNA 5' UTR length	23
Figure 7. mRNA interactions with mammalian 48S IC	24
Figure 8. Non-canonical initiation pathway	30
Figure 9. Different types of IRES.....	34
Figure 10. Replacement of eIF1 by eIF5 at the P site in the closed state of yeast 48S PIC after start codon recognition	37
Figure 11. Open and closed conformations of yeast 48S PIC.....	41
Figure 12. Components for the reconstitution of the mammalian translation systems	51
Figure 13. Toeprinting assay of reconstituted the 48S IC, 80S IC and 80S EC.....	57
Figure 14. Functions of individual initiation factor in the reconstituted 48S IC and 80S IC.....	60
Figure 15. The 48S IC formation on the model AUG/CUC mRNA and short 13XCAA model mRNA	62
Figure 16. 48S IC formation on mRNAs bearing near-cognate start codons	65
Figure 17. Toeprinting assay using different eIF1A variants	66
Figure 18. Toeprinting assay using initiator tRNA and GTP analogs	67
Figure 19. Met-Val dipeptide formation by 80S IC.....	69
Figure 20. Oligopeptide formation by 80S IC	70
Figure 21. MV dipeptide formation at different concentrations of KOAc and MgOAc	71
Figure 22. mRNA binding to eIF3.....	73
Figure 23. mRNA binding to the 40S subunit	74
Figure 24. Binding kinetics of the 40S subunit to the Alexa488-labeled mRNA	76
Figure 25. Binding of eIF1A to the 40S subunit	79

Figure 26. Fluorescence profiles of eIF1A-N4C-Alexa555 during 80S IC formation.....	82
Figure 27. Fluorescence profiles of eIF1A-Alexa555 variants during 80S IC formation.....	84
Figure 28. Kinetic analysis of binding and dissociation of the eIF1A-N4C-Alexa555 to the 40S subunit.....	86
Figure 29. Two-population binding-equilibrium model of eIF1A binding to the 40S subunit .	90
Figure 30. Fitting the data to the binding-equilibration model by numerical integration	92
Figure 31. Kinetic model of eIF1A binding to the 40S-eIF1 complex.....	94
Figure 32. Dissociation of eIF1A from the 40S subunit in the presence of other individual initiation factor	97
Figure 33. Fluorescence changes during titration of the 40S-eIF1A complex with TC.....	98
Figure 34. Dissociation of eIF1A from the 40S subunit with multiple initiation factors	99
Figure 35. Dissociation of eIF1A from the 43S pre-initiation complex	102
Figure 36. Hierarchy of initiation complexes en route to the 43S PIC formation probed by their stability of eIF1A binding	103
Figure 37. Codon recognition regulates the stability of eIF1A binding to the 48S IC	106
Figure 38. eIF1A dissociation from the 48S (CUC) PIC and 48S (AUG) IC is independent of AUG codons in the 3' UTR.....	107
Figure 39. Hierarchy of initiation complexes en route to the 48S IC formation	108
Figure 40. Role of individual eIFs in regulating eIF1A dissociation from the 48S IC	110
Figure 41. Summary of eIF1A dissociation rate constants from 40S subunit and various partial 43S and 48S ICs.....	111
Figure 42. Binding of eIF1A to the 48S IC	114
Figure 43. Dissociation of eIF1A-S74C-Alexa555 from initiation complexes	116
Figure 44. Dissociation of eIF1A-T120C-Alexa555 from initiation complexes	117
Figure 45. Overview of the Cryo-EM structure of the human 48S IC.....	120
Figure 46. Structural dynamics in the decoding center	121
Figure 47. Binding sites of eIF1A in the open and closed form of 48S IC.....	122
Figure 48. Dynamic of the eIF1A-dissociation states of the initiation complex (see below) .	142
Figure 49. Purification of the 40S and 60S subunits on a 0–30% sucrose gradient.....	150
Figure 50. Elution profile of eIF2 and eIF5B from MonoQ 5/50 GL column	152
Figure 51. Elution profile of eIF3 and eIF4F from MonoQ 5/50 GL column.....	154

6.2 List of tables

Table 1. Number of human initiation factors and ribosomes per HeLa cell	9
Table 2. Summary of dissociation rates of eIF1A from yeast 43S PIC and 48S IC	39
Table 3. Efficiency of Non-AUG start codons in various assays	44
Table 4. Summary of analytical rate constants of eIF1A binding to 40S and 40S·eIF1 complex	88
Table 5. Summary of K_{ds} of eIF1A binding to 40S and 40S·eIF1 complex by kinetic and thermodynamic analysis.....	95
Table 6. Summary of dissociation rate constants of eIF1A-N4C-Alexa555 from initiation complexes.....	112
Table 7. Kinetics of yeast TC binding to the 40S subunit in the presence of eIF1, eIF1A and mRNA.....	140

ACKNOWLEDGEMENTS

I am extremely thankful to Prof. Dr. Marina V. Rodnina for giving me the opportunity to pursue my doctoral study with this ambitious project. I thank Marina for your scientific guidance and supporting me with full resources to accomplish the project. Every discussion with you is a grandiose lesson for me that improves my scientific thinking. I wish myself to be a scientist who understands the details of each scientific theory and experiment as you do. I would also like to thank my second supervisor Dr. Sarah Adio for having in-depth discussions with me and giving me the freedom to explore the project independently. Working together with you expands my understanding of the system from a different perspective. I enjoyed the excitement of the scientific debates when we uncovered the puzzling findings. Besides, I would also like to thank my thesis advisory committee members and the extended examination board, Prof. Dr. Markus T. Bohnsack, Prof. Dr. Kai Tittmann, Prof. Dr. Holger Stark and Prof. Dr. Henning Urlaub, for their individual scientific inputs and fruitful collaborations.

I am especially thankful to my former supervisor, Dr. Akanksha Goyal. This project was initially initiated by Dr. Akanksha Goyal during her Ph.D. time and I joined on board afterward as my Ph.D. project. I thank Aki for leading me into the field with extreme patient and showing me all the essential techniques. To establish a new system is not an easy task and from you, I learned that only when we understood the principles of the assay down to the ground, we would be able to troubleshoot the problems. The ability of troubleshooting is what I gained the most during my Ph.D. study. Thanks for being a model in the beginning of my Ph.D.!

To accomplish this project, I would not achieve without the fantastic technical supports from Theresia Steiger, Michael Zimmermann, Lena Preiser, Olaf Geintzer and Vanessa Herold. I thank Theo for processing the HeLa cell lysate and purifying native factor eIF2, eIF3 and eIF5B, Lena for purifying recombinant eIF1 and eIF4B and fluorescently labeling eIF1A variants, and Michael for purifying the native human ribosomes and recombinant eIF1A. I also appreciate the excellent preparation of tRNA from Olaf and IVT of mRNA from Vanessa. All of your contribution makes this challenging project possible.

I would also like to thank all the collaborators in human 48S IC project, the cryo-EM structure collaborators, Dr. Erik Schliep Dr. Niels Fischer, Valentyn Petrychenko, and the cross-linking MS collaborators, Dr. Jasmin Corso and Dr. Andreas Linden. A special thank goes to the HeLa cell facility, Dr. Ashwin Chari and Hossein Kohansal, for the kind support of the high-quality HeLa lysate. It is a pleasant for me to collaborate with you and is also a treasured opportunity for me to explore new fields.

I am thankful to our IMPRS Molbio program coordinators, Dr. Steffen Burkhardt and Kerstin Grüniger, for helping us starting our life in Göttingen, organizing various events, and funding me with travel grants to various international conferences. Molbio is a fruitful Ph.D. program and I am proud to be part of it.

I sincerely appreciate Kai, Gerald, Yifan, Dr. Luisa Hallmaier-Wacker, Dr. Evan Mercier and Dr. Kärt Denks for proofreading the manuscript of this thesis. I am thankful to Dr. Evan Mercier for the assistance on the kinetic experiments and discussion on the kinetic model. I also thank Evan for being a good friend who always gives me constructive advice whenever I needed. I deeply appreciate Dr. Namit Ranjan and Dr. Sandra Blanchet for sharing the experience in establishing the yeast translation system and helping me with various experiments, especially dipeptide assay. I thank Dr. Cristina Maracci and Dr. Prajwal Karki for being amazing supervisors during my rotation and inspiring me to pursue my Ph.D. here. I thank Dr. Riccardo Belardinelli for helping me the stopped-flow apparatus, Dr. Ingo Wohlgemuth for the MS analysis and Dr. Manisankar Maiti for the FCS measurement. I thank Dr. Bee-Zen Peng for showing me the HPLC system and Dr. Marija Liutkute for discussing the KinTek explorer software. I thank Ninadini Sharma and Sofia Ainatzi for their contribution to the project during their rotations here. I thank Dimitra Papastavrou and Mario Klein for handling tedious paper works and solving the random computer problems. I also want to mention all my lovely colleagues, Tessa, Panos, Sakshi, Lena, Jan, Cathie, Xiaolin and Kärt, for the joyful atmosphere in the lab and during the lunchtime.

I want to thank all my friends here in Göttingen and in Europe. I thank Kai for being a fantastic travel partner and exploring Europe with me. I thank all my teammates, especially

Hagen, Knut, Jan, Jeff, Yu, Josi, Yohi, Tessa, Leander, Philipp, Luisa, and Evan, from Göttingen Allstars Baseball club for the cheer moments in the baseball games and our two times champions in these 5 years. I thank Xenia, Pauline and Saskia for our Chinese-German hybrid conversations. I thank Flávia, Yin and Kathinka for diverse cultural adventures. I thank Mayr for the delightful time in the coronavirus pandemic. Finally, my special thank goes to Yen-Yun, Gerald and Luisa who have always been supportive friends and give me lifetime advice during the hard and joyful time in Göttingen.

I am grateful to my family in Taitung in Taiwan. I thank Fang for always cheering me with positive vibes. I thank Tzu-Ying and my brother-in-law for visiting me in Germany and greeting me with lots of home treats. I thank my parents for always being supportive and encouraging me to explore the world with a kind heart. Finally, I thank my grandmas, who will always be in my mind, for giving me a lovely family. I am a beloved son and I devoted my Ph.D. work to all of you.