Structural and functional characterization of eIF4E1 and eIF4E2 complexes involved in translational control

DISSERTATION

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Daniel Peter

aus Naila

Tübingen

2017

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:	12.12.2017
Dekan:	Professor Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Professor Dr. Elisa Izaurralde
2. Berichterstatter:	Professor Dr. Thilo Stehle
3. Berichterstatter:	Professor Dr. Markus Landthaler

The work described in this thesis was conducted in Prof. Dr. Elisa Izaurralde's laboratory in the Department of Biochemistry at the Max Planck Institute for Developmental Biology, Tübingen, Germany from November 2012 to October 2017. It was further supervised by Prof. Dr. Thilo Stehle at the Eberhard Karls University Tübingen, Germany, and was supported by the Max Planck Society. I declare that this thesis is the product of my own work. The parts that have been published or where other sources had been used were cited accordingly. Work that was conducted by my colleagues was also indicated accordingly.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

This thesis represents a summary of my scientific work and progress during the last years. However, it does not reflect the life experience that comes along with it and that changed many aspects of me as a person. Therefore, I want to thank all of my colleagues and friends who shared with me this memorable period of my life and without whom all of this would not have been possible.

First of all, I want thank my supervisor Prof. Dr. Elisa Izaurralde not only for offering me the opportunity to do my PhD in her lab, but also and most importantly for the constant support and guidance throughout this time. She always had an open ear for everything, and especially her help in the beginning of my PhD was the key to a very successful and unforgettable time that followed. Thank you for everything.

I am very grateful to Prof. Dr. Thilo Stehle at the Eberhard Karls Universität Tübingen for supervising my PhD thesis. Likewise I would also like to thank Prof. Dr. Ralf Jansen at the Eberhard Karls Universität Tübingen and Prof. Dr. Birte Höcker at the Universität Bayreuth for being members of my PhD defense committee.

All of this work would not have been possible without some guidance in the lab and in the projects. I deeply thank Cátia Igreja for her enormous support and the great collaboration. I enjoyed all of the projects we were working on together and all of the spontaneous discussions we had about interesting results. Thank you for showing me the excitement of science.

I also want to thank Tobias Raisch, Dr. Stefanie Jonas and Dr. Eugene Valkov for sharing their knowledge about x-ray crystallography with me and also for introducing me to other techniques. In this context I would also like to thank Dr. Oliver Weichenrieder for his help at the SLS.

My PhD project was supported by the Max Planck International PhD program. Therefore I want to thank the PhD program coordinator Dr. Dagmar Sigurdardottir for her support and especially for organizing many helpful courses that advanced my professional skills also outside pure research.

I could not really see myself as a teacher. But the gifted and at that time undergraduate students I worked together with made this job very easy for me. I would like to thank Carolin Köne, Ramona Weber, Lukas Langer and Felix Sandmeir

for their excellent work and motivation. I hope I could share some of my experience with you and I wish you all the best for your future.

I highly appreciate the great technical support I had from our technical and scientific staff. Therefore I want to thank Catrin Weiler, Min-Yi Chung, Lara Wohlbold, Regina Büttner, Gabriele Wagner, Dr. Heike Budde, Sigrun Helms and Maria Fauser for all your contributions you made to my scientific work.

All the past and current members of our Department II I would like to thank for nice and encouraging working atmosphere in the lab. Whomever I asked, I always obtained help and assistance. I want to especially thank the people with whom I shared my daily bench work: Dr. Anna Schneider, Dr. Valentina Ahl, Dr. Andreas Boland, Dr. Alain-Pierre Petit, Dr. Stefanie Jonas, Regina Büttner, Dr. Ying Chen, Dr. Mary Christie, Stefan Grüner, Tobias Raisch, Dr. Elena Khazina, Sowndarya Muthukumar, Eugene Valkov, Gabriele Wagner and Catrin Weiler. To Dr. Steffen Schmidt a special thanks for his quick and uncomplicated help with software-related problems. We all had good times together and I will miss you.

My special thanks again to Dr. Mary Christie. You were a great mentor and friend during some of the more difficult times.

Also I want thank my colleagues and friends Stefan Grüner, Ramona Weber and Dr. Moritz Thran for the many non-scientific discussions during lunch and coffee breaks. You guys provided many times a necessary distraction.

My gratitude goes also to my friends outside the lab. Thank you for trying to understand the crazy life of a scientist.

Abschließend möchte ich meiner Familie für ihren bedingungslosen Rückhalt und ihr Verständnis danken. Der Pfad eines Wissenschaftlers ist ein langer und er erfordert so manches Opfer. Meinem Vater Michael, meiner Mutter Antonie und meinem Bruder Martin, vielen Dank das ihr an mich glaubt und mich bestärkt.

Ebenso danke ich von ganzem Herzen meiner Frau Linda Ebertsch. Die viele Zeit im Labor war nicht immer einfach für uns, aber ohne dich hätte ich so manche Situation nicht so einfach bewältigt. Vielen Dank das du immer für mich da bist und mir den Rücken frei hältst.

1A. SUMMARY

Protein synthesis is one of the costliest processes in the cell. Therefore, the initiation of translation is a tightly regulated process. One major control mechanism targets the activity or formation of the so-called eIF4F (eukaryotic initiation factor 4F) complex bound to the 5' cap structure of an mRNA. This heterotrimeric complex, consisting of the RNA helicase eIF4A, the cap-binding protein eIF4E and the scaffold subunit eIF4G, is ultimately required for the recruitment of the 43S PIC (pre-initiation complex) to the mRNA, leading to subsequent scanning and initiation. The formation of the eIF4F complex is under the control of a group of inhibitory proteins known as eIF4E-binding proteins (4E-BPs), which bind to eIF4E and prevent its interaction with eIF4G. 4E-BPs comprise a group of functionally distinct proteins and include global translational repressors such as the three human proteins 4E-BP1-3, or large, multidomain proteins that likely act on an mRNA-specific level. Alternatively, the assembly of the eIF4F complex can be prevented by the eIF4E-homologous protein (4EHP or eIF4E2), which competes with eIF4E in binding to the 5'cap structure of an mRNA. Compared to the global repression by 4E-BPs, the later mechanism only acts on a message specific level.

Comprehensive molecular insight into eIF4E- and 4EHP-complexes involved in the regulation of translation initiation was lacking. My doctoral work provides a fundamental structural and mechanistic understanding of the formation of these regulatory complexes. In my initial studies, I characterized the binding of various 4E-BPs to eIF4E and provided the first structural insights into an extended eIF4E-binding mode of different 4E-BPs. The structures revealed a conserved mode of interaction with eIF4E, despite the lack of sequence conservation. Additionally, in a collaborative project, I observed that the eIF4E-binding mode characteristic of 4E-BP complexes is

also present in eIF4E-eIF4G complexes, expanding the knowledge on the mechanism of translation initiation and its regulation.

Another part of my doctoral studies focused on 4E-BPs very specific functions and architecture. Specifically, I investigated the binding mode of an invertebrate-specific 4E-BP called Mextli. My studies unveiled an unexpected variation and evolutionary plasticity in the eIF4E-binding mode of Mextli homologs across species, which confer distinct functional properties to the respective eIF4E-complexes.

I also studied 4EHP, the second member of the eIF4E protein family, and its specific interaction partners, the Grb10-interacting GYF domain-containing (GIGYF) proteins 1 and 2, and obtained the first crystal structures of theses 4EHP-specific binding partners bound to 4EHP. The molecular details of the 4EHP-GIGYF translational repressor complex explain why GIGYF proteins bind to 4EHP and not to eIF4E. Overall, my doctoral studies revealed new insights on eIF4E-related complexes and their diverse roles in posttranscriptional gene regulation.

1B. ZUSAMMENFASSUNG

Die Synthese von Proteinen zählt zu den kostspieligsten Prozessen der Zelle. Daher unterliegt die Translation von mRNAs einer strikten Kontrolle, insbesondere deren Initiation. An dessen Regulierung ist der sogenannte eIF4F- (eukaryotischer Initiationsfaktor 4F) Komplex beteiligt. Dieser heterotrimere Komplex besteht aus eIF4E, welches direkt an die 5'-Kappe der mRNA bindet, der RNA-Helikase eIF4A sowie eIF4G, welches das Gerüst des Komplexes bildet. Letztlich rekrutiert der eIF4F-Komplex den 43S PIC (pre-initiation complex; zu deutsch *Präinitiationskomplex*) an die mRNA und ermöglicht somit die Suche nach einem Startcodon, was zum Beginn der eigentlichen Proteinsynthese führt. Die Bildung des eIF4F-Komplexes wird durch eine Gruppe inhibitorischer Proteine reguliert, welche als eIF4E-bindende Proteine (4E-BPs) zusammengefasst werden, da sie an eIF4E binden und dadurch die Interaktion zwischen eIF4E und eIF4G verhindern. 4E-BPs umfassen eine funktionell diverse Gruppe von Proteinen, zu welcher sowohl generelle Translationsregulatoren wie z.B. die menschlichen Proteine 4E-BP1, 2 und 3 zählen, als auch komplexere Proteine, welche sehr spezifisch agieren und nur bestimmte mRNAs kontrollieren. Auch das eIF4E-homologe Protein (4EHP oder eIF4E2) kann die Formierung des eIF4F-Komplexes an der mRNA unterbinden, da es mit eIF4E um die Bindung an der 5'-Kappe der mRNA konkurriert. Im Gegensatz zu 4E-BPs, welche hauptsächlich an der generellen Inhibierung der Translation in der Zelle beteiligt sind, reguliert 4EHP ausschließlich die Expression spezifischer mRNAs.

eIF4E- und 4EHP-Komplexe spielen in der Kontrolle der Translations-initiation eine wichtige Rolle, allerdings fehlte bisher ein umfassender molekularer Einblick in diese Proteinkomplexe. Meine Doktorarbeit erbrachte daher ein grundlegendes strukturelles und mechanistisches Verständnis über diese

regulatorischen Komplexe. Zu Beginn meiner Studien charakterisierte ich die Interaktion verschiedener 4E-BPs mit eIF4E und konnte Kristallstrukturen dieser 4E-BPs in Komplex mit eIF4E bestimmen, welche einen erweiterten Bindungsmodus aufzeigten. Die Kristallstrukturen zeigten, dass die verschiedenen 4E-BPs trotz deutlich unterschiedlicher Proteinsequenzen dennoch auf sehr ähnliche Weise an eIF4E binden. Zudem konnte ich in einem weiteren, kollaborativen Projekt aufzeigen, dass dieser Bindungsmodus für eIF4E auch von eIF4G genutzt wird, was unser mechanistisches Verständnis über die Initiation und Regulation der Translation erweitert.

Ein weiterer Teil meiner Doktorarbeit beschäftigte sich mit 4E-BPs und eIF4E-Proteinen, welche sehr spezifische Funktionen erfüllen. Im Einzelnen untersuchte ich ein 4E-BP namens Mextli, das ausschließlich in Invertebraten vorkommt. Meine Arbeit zeigte eine evolutionäre Plastizität zwischen homologen Mextli Proteinen verschiedener Spezies bezüglich ihrer Bindung mit eIF4E auf. Diese Unterschiede im Bindungsmodus wirken sich zudem auf die Funktionsweise des jeweiligen eIF4E-Komplexes aus.

Schließlich untersuchte ich mit 4EHP ein weiteres eIF4E-Protein und dessen spezifische Interaktionspartner, die GIGYF-Proteine (Grb10-interagierende GYF Proteine) 1 und 2. Dabei konnte ich die erste Kristallstruktur dieser 4EHP-spezifischen Bindungspartner in Komplex mit 4EHP bestimmen. Die molekularen Details des 4EHP-GIGYF Komplexes erklären, warum GIGYF-Proteine an 4EHP und nicht an eIF4E binden. Insgesamt erbrachte meine Doktorarbeit neue Einblicke in eIF4E-umfassende Proteinkomplexe und deren Funktionen in der posttranskriptionellen Regulation der Genexpression.

2. INTRODUCTION

The flow of information from the genome in the nucleus to a protein in the cytoplasm requires a messenger molecule. This messenger RNA (mRNA) is the main player of a series of complex processes, starting with its biogenesis, export to and localization in the cytoplasm, translation and regulation, until its final decay (Moore, 2005).

mRNAs are transcribed from the genomic DNA in the nucleus. To produce a mature mRNA, the RNA transcript is processed in a series of reactions involving 5'-end capping, splicing, 3'-end cleavage and polyadenylation. The mature mRNAs assemble with proteins into messenger ribonucleoprotein (mRNP) complexes and are exported into the cytoplasm where their encoded information is used by ribosomes for protein synthesis before they are ultimately degraded (Moore, 2005).

However, this description is only a simplistic outline of the life cycle of an mRNA, which can be subjected to a variety of elaborate, posttranscriptional control mechanisms. One major node of regulation concerns the translation of mRNAs, as protein production entails high costs for the cell (Buttgereit and Brand, 1995; Kafri et al., 2016). mRNA translation can be divided into four phases: initiation, elongation, termination and ribosome recycling. Regulation of this process is mainly accomplished at the initiation phase and provides the quickest response in the gene regulation pathway, allowing for rapid changes of cellular factors (Sonenberg and Hinnebusch, 2009). Moreover, the tight control of translation is critical for developmental processes, cell differentiation and synaptic plasticity (Buffington et al., 2014; Sonenberg and Hinnebusch, 2009). Accordingly, the dysregulation of almost all initiation factors and regulators towards higher translational output is linked to a vast number of human cancers (Ruggero, 2013). Therefore, a lot of effort is being made to

comprehensively study translation initiation from a structural, functional, genetic and biological perspective to efficiently target the translation machinery in malignancies caused by an abnormally elevated protein synthesis (Bhat et al., 2015; Buffington et al., 2014; Pelletier et al., 2015).

2.1 Eukaryotic translation of mRNAs

The translation of mRNAs describes the process of protein synthesis by decoding the mRNA molecule. Translation takes place in all living cells and is carried out on ribosomes. However, the process is not fully conserved, and eukaryotic mRNA translation differs in some fundamental aspects from the mechanism in prokaryotes. Those differences concern for example the ribosomal composition, the factors that are involved in each of the phases of translation, but also the genomic organization and nature of the mRNA molecule itself.

In contrast to eukaryotic cells, prokaryotes do not have a nucleus and their genome is located in the cytoplasm. As such, prokaryotic transcription and translation occur simultaneously and can take place as a linked process. This so-called transcription-translation coupling is mediated by the pioneering ribosome proceeding on the nascent transcript synthesized by the RNA-polymerase. Concomitant translation promotes the rate of transcription by preventing pausing, backtracking and early termination (Kohler et al., 2017; Landick et al., 1985; Proshkin et al., 2010).

Prokaryotic mRNAs also differ from its eukaryotic counterparts also in structure and organization. In bacteria, mRNAs are mostly polycistronic, meaning that one transcript is coding for two or more proteins (Kozak, 1983). Additionally, bacterial mRNAs do not contain a cap structure at the 5’ end, and unlike eukaryotes, 3’ end polyadenylation leads to short poly(A)-tails that promote mRNA degradation

(Bandyra and Luisi, 2013; Regnier and Hajnsdorf, 2013; Sarkar, 1997; Steege, 2000). The distinct 5' ends of bacterial and eukaryotic mRNAs imply that the initiation step of translation differs significantly in the two systems.

Prokaryotic start codon recognition is mediated by a direct recruitment of the 30S small ribosomal subunit to the initiation region of the mRNA. This is achieved by base-pairing between an upstream Shine-Dalgarno (SD) sequence of the mRNA and an anti-SD sequence in the 16S rRNA. In eukaryotes, recruitment of the small ribosomal subunit usually requires the 5'cap structure and multiple translation initiation factors (eIFs) and involves an elaborated series of events. Therefore, also the mechanisms that control translation are different between prokaryotes and eukaryotes (Duval et al., 2015; Gualerzi and Pon, 2015). Since the work described in this thesis focuses on eukaryotic translation factors, only the eukaryotic process and mechanism will be explained in the following sections.

2.1.1 Translation initiation

The initiation phase of translation describes the positioning of a translation-competent ribosome on the start codon of the mRNA. In eukaryotes, the mechanism of canonical translation initiation requires the assembly of a multitude of eukaryotic initiation factors (eIFs) at the 5'cap structure (a methylated guanine nucleotide with a 5'-5' triphosphate linkage to the following nucleotide; m^7GpppN, where N is any nucleotide) of an mRNA (Figure 1). This multifactorial assembly is initiated by the heterotrimeric eIF4F complex, which comprises the cap-binding protein eIF4E, the DEAD-box RNA helicase eIF4A and the scaffolding protein eIF4G (Jackson et al., 2010). eIF4G not only mediates the interactions to eIF4A and eIF4E, but also causes the circularization of the mRNA molecule by additionally binding to the poly(A)-

binding protein (PABP) associated with the 3' poly(A)-tail of the mRNA. Further, eIF4G also recruits the 43S preinitiation complex (43S PIC) to the mRNA via interactions with subunits of the eIF3 complex (Jackson et al., 2010; Wells et al., 1998).

The 43S PIC comprises the initiator methionyl-tRNA (Met-tRNA$^{Met}_i$)-containing ternary complex (eIF2-GTP/ Met-tRNA$^{Met}_i$; TC), the 40S small ribosomal subunit and additional eIFs (eIF1, eIF1A, eIF3 and eIF5; Figure 1). Recruitment of the 43S PIC by the eIF4F complex results in the formation of the 48S particle, which scans the 5' untranslated region (5'UTR) of the mRNA in search for an AUG start codon. Scanning of highly structured 5'UTRs requires the ATP-dependent helicase activity of eIF4A to unwind secondary structures (Svitkin et al., 2001). The AUG triplet, usually located in an optimal context called the Kozak sequence (a purine base at -3 and +4 positions relative to the A in the AUG codon), is recognized by the 48S PIC as a result of base-pairing with the Met-tRNA$^{Met}_i$ [Figure 1; (Jackson et al., 2010; Kozak, 1991)]. Start codon recognition triggers conformational rearrangements of the 48S PIC, including the displacement of eIF1 and eIF1A and the release of eIF2-GDP. The GTPase-activating protein eIF5 plays a crucial role in dissociating the eIF2 heterotrimer (consisting of three subunits named α, β and γ), as it induces the GTP hydrolysis of the eIF2γ subunit (Maag et al., 2005; Unbehaun et al., 2004). Subsequently, eIF5B helps to dissociate the remaining eIFs and mediates the joining of the 60S large ribosomal subunit, before it catalyzes its own release from the assembled 80S ribosome by hydrolysis of bound GTP (Acker et al., 2006; Acker et al., 2009; Pestova et al., 2000).

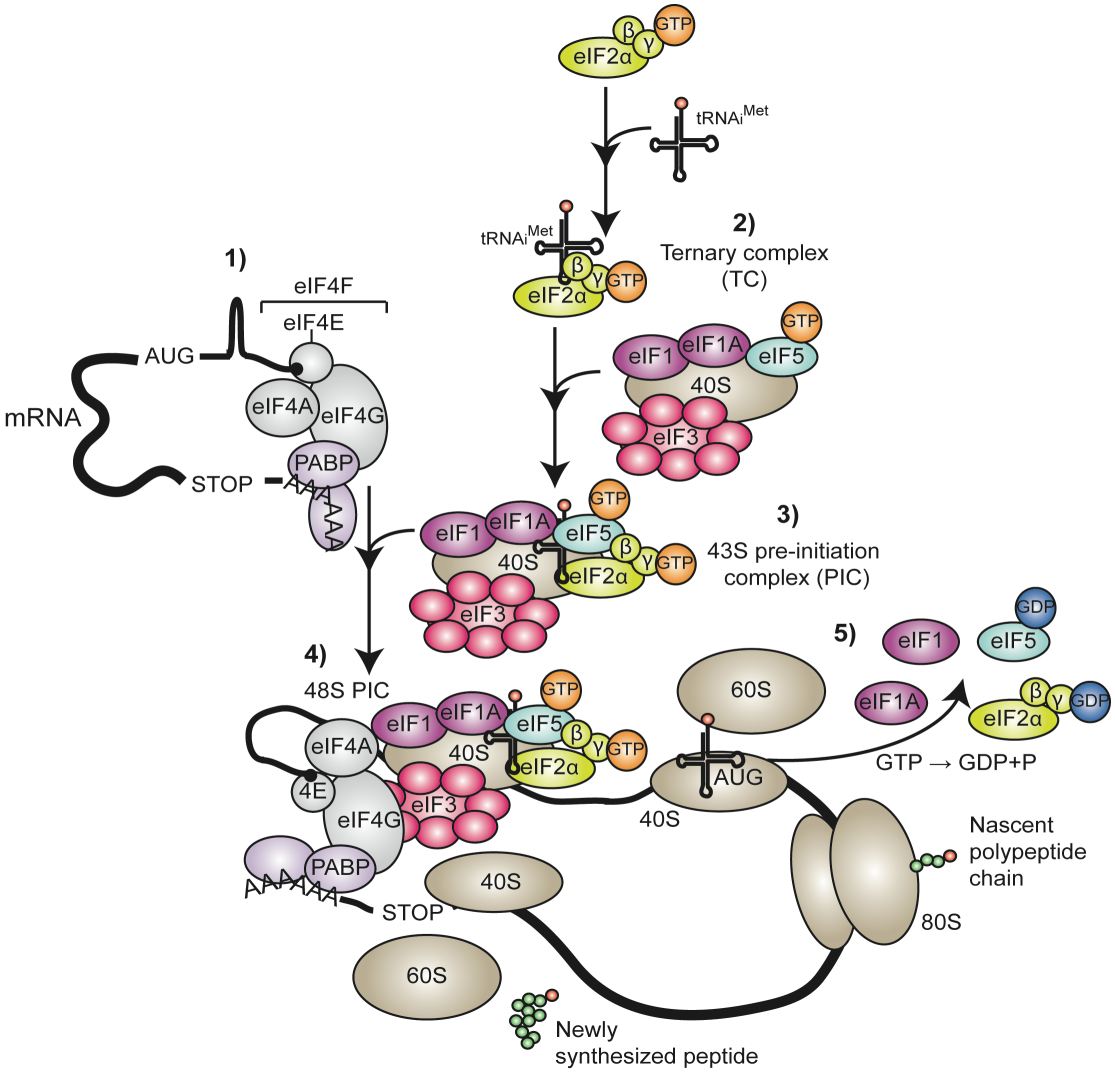


Figure 1: Canonical cap-dependent translation initiation. 1) Cap-dependent translation is initiated by the formation of the eIF4F complex (consisting of the initiation factors 4A, 4E and 4G) at the 5'cap of an mRNA. 2) The ternary complex (TC) is another subcomplex important in the initiation of translation and contains the initiation factor 2 proteins (eIF2; consisting of α, β and γ subunits), the initiator methionyl-tRNA (Met-tRNA$^{Met}_{i}$) and guanosine triphosphate (GTP). 3) Formation of the 43S preinitiation complex (43S PIC) containing the TC and the 40S small ribosomal subunit with bound initiation factors eIF1, eIF1A, eIF3 and eIF5. eIF3 itself is also a large complex consisting of 13 subunits (named a-m). 4) Recruitment of the 43S PIC by interactions between eIF4G of the eIF4F complex and eIF3 leads to the formation of the 48S PIC. This complex scans the 5' untranslated region (5'UTR) of the mRNA in search of an AUG start codon. 5) Start codon recognition upon base-pairing of the Met-tRNA$^{Met}_{i}$ with the AUG triplet triggers conformational changes in the 48S PIC, which results in a GTP-dependent release of several initiation factors and joining of the 60S large ribosomal subunit, forming the translationally competent 80S ribosome. The figure is adjusted from Bhat et al., 2015.

2.1.2 Elongation, Termination and Ribosome recycling

Compared to the initiation step of translation, all subsequent steps are more akin between prokaryotes and eukaryotes. Specifically, the elongation phase, meaning the incorporation of amino acids into a growing polypeptide, is most similar, since the core architecture as well as the *modus operandi* of prokaryotic and eukaryotic ribosomes are well conserved (Buskirk and Green, 2017; Frank and Spahn, 2006).

During the elongation cycle of translation, the tRNA with anticodon complementarity to the mRNA codon translocates between the A, P and E tRNA binding sites of the 80S ribosome (Steitz, 2008). In this process, aminoacyl-tRNAs bound in the A site are translocated into the P site after the peptidyl transferase reaction. This reaction links incoming amino acids to the growing peptide chain and occurs in the peptidyl transferase center of the large subunit of the ribosome [Figure 2; (Buskirk and Green, 2017; Steitz, 2008)]. Deacylated tRNAs from the P site move into the E site of the ribosome before dissociation. The two elongation factors eEF1A and eEF2 facilitate tRNA selection and translocation, respectively [Figure 2; (Buskirk and Green, 2017; Steitz, 2008)].

The elongation cycle is repeated until a stop codon (UAA, UAG and UGA) is positioned in the ribosomal A site (Figure 2). The translation termination factor eRF1 recognizes all of the three stop codons and is loaded into the A site. Together with eRF3, eRF1 promotes the deacylation of the peptidyl-tRNA on the P site in a GTP-dependent manner, resulting in the release of the nascent peptide [Figure 2; (Eyler et al., 2013; Frolova et al., 1996)].

Moreover, eRF1 is also involved in the recycling of the ribosome after termination. In this case, eRF1 recruits the ABC-family ATPase ABCE1 to the A site of the ribosome. ABCE1 induces the dissociation of the ribosome into the 60S subunit, the tRNA- and the mRNA bound 40S subunit [Figure 2; (Pisarev et al., 2010; Pisareva et al., 2011; Shoemaker and Green, 2011)]. The free ribosomal subunits and tRNA can subsequently be used in a new translation cycle.

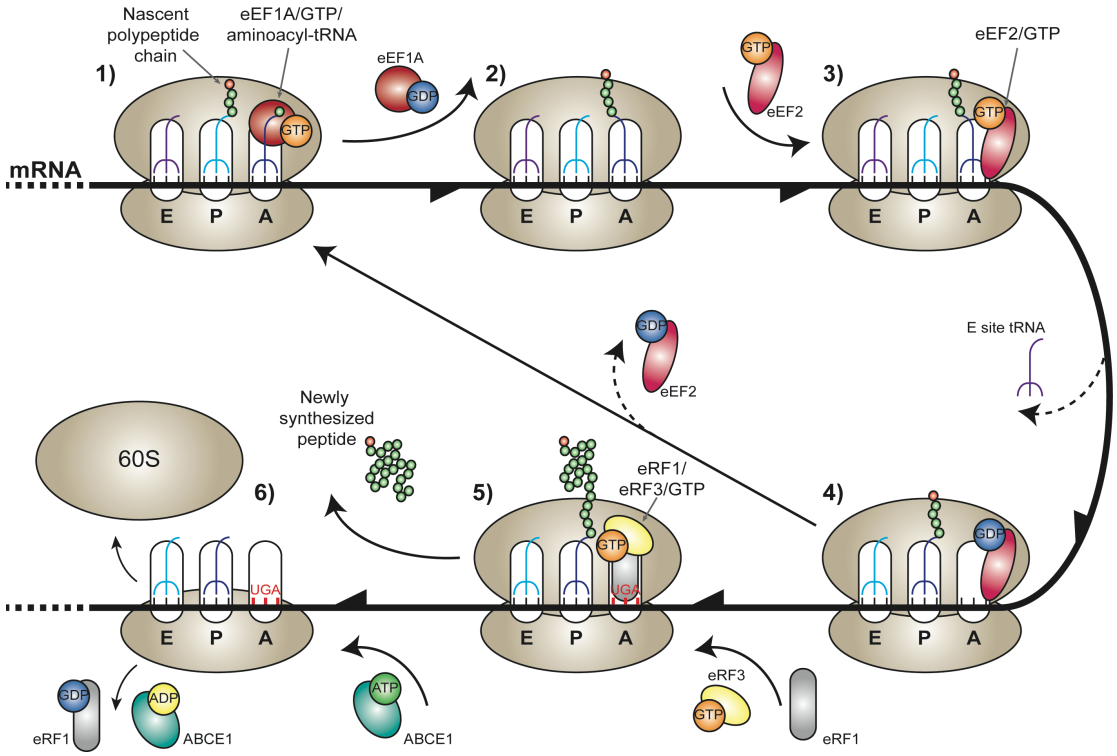


Figure 2: Elongation, termination and recycling steps of translation. 1) Translating ribosome with eEF1A/GTP/aminoacyl-tRNA bound to the A site, the peptidyl-tRNA in the P site and the deacylated tRNA in the E site. 2) Peptidyl transferase reaction results in the translocation of the peptide chain onto the A site tRNA. 3,4) GTP-dependent translocation of tRNAs from the P site to the E site and from the A site to the P site by eEF2. Upon dissociation of eEF2/GDP the ribosome is ready for the next round of elongation (step 1). 5) A stop codon (e.g. UGA) in the A site is recognized by eRF1 and triggers termination. Together with eRF3, the nascent peptide is released in a GTP-dependent manner. 6) eRF1 remains bound to the A site (not shown) and is recognized by ABCE1, which mediates the dissociation of the ribosome upon ATP hydrolysis. The figure is modified from Steitz et al., 2008 and from Pisarev et al., 2010.

2.2 Regulation of translation initiation

The control of translation has a central role in the regulation of gene expression at a post-transcriptional level. This type of regulation not only provides the possibility for the cell to quickly adapt to environmental changes, but it is also of crucial relevance in the cellular decision about the translational programs. As such, the regulation of translation has a major impact on cell growth, cell progression as well as on distinct developmental processes, as it ultimately controls the cellular proteome (Bhat et al., 2015; Hershey et al., 2012). The initiation process is the rate limiting and most regulated step of translation and can be modulated by many factors (Hershey et al., 2012; Morisaki et al., 2016; Wu et al., 2016). The regulatory mechanisms either impact on the activity of specific initiation factors, such as eIF2 and the eIF4F complex, or on the mRNA itself through specific RNA-binding proteins (RBPs) or microRNAs [miRNAs; (Hershey et al., 2012; Jackson et al., 2010)].

2.2.1 Regulation of translation by miRNAs

miRNAs are small non-coding RNAs of ~22 nucleotides (nt) in length derived from introns or exons of non-coding or coding transcripts (Kim et al., 2009). In essence, miRNAs are generated from precursor transcripts in multiple processing steps involving the RNAse III-type endonucleases Drosha and Dicer. The mature miRNAs form an RNA-induced silencing complex (RISC) together with Argonaute proteins (Ha and Kim, 2014).

RISC complexes are recruited to specific target mRNAs via base-pairing between the ‘seed’ region of the miRNA, which are nt 2-7 at the 5’end of the miRNA, and a complementary sequence in the 3’UTR of the mRNA (Bartel, 2009; Huntzinger

and Izaurralde, 2011). Therefore translational regulation by miRNAs is sequence-specific, but a single miRNA might control several mRNAs (Hershey et al., 2012).

In animals, miRNA-mediated control of gene expression occurs through mRNA degradation and translational repression. The GW182 proteins (TNRC6A/GW182, TNRC6B and TNRC6C in humans) play an important role in this mechanism. They associate with Argonaute proteins (AGO1-4 in humans) and are essential components of miRISC complexes since they recruit downstream effectors to induce translational repression and mRNA degradation (Behm-Ansmant et al., 2006a; Behm-Ansmant et al., 2006b; Chekulaeva et al., 2009; Eulalio et al., 2008). Degradation of miRNA-targets is accomplished through the direct recruitment of the CCR4-NOT complex by GW182 proteins, the major cytoplasmic deadenylase complex (reviewed in Jonas and Izaurralde, 2015). In contrast, much less is known about the mechanism of miRNA-mediated translational repression. Recent models are controversial but indicate that miRNAs affect the initiation step of the canonical cap-dependent translation (Jonas and Izaurralde, 2015; Kuzuoğlu-Öztürk et al., 2016; Ricci et al., 2013).

2.2.2 Phosphorylation of eIF2

The initiation factor eIF2 consists of three subunits, named eIF2α, -β and –γ, and forms the TC with GTP and Met-tRNA$^{Met}_{i}$. The TC is an integral part of the 48S PIC required for the recognition of the start codon. Its GTPase activity is also necessary to trigger the structural rearrangement of the 40S ribosomal subunit that allows translation initiation [Figure 1; (Maag et al., 2005; Unbehaun et al., 2004)].

eIF2B is a guanine nucleotide exchange factor that replaces hydrolyzed GDP of eIF2 with GTP for another initiation cycle. The GTP-bound state of eIF2 has an increased affinity for the Met-tRNA$^{Met}_{i}$ and can therefore be efficiently loaded to form a productive TC (Kapp and Lorsch, 2004).

Phosphorylation of eIF2 restricts the availability of productive eIF2 during the initiation of translation. As a result of various stress stimuli like nutrient deprivation or oxidative stress, different eIF2α kinases (e.g. GCN2, PERK, PKR and HRI) phosphorylate a specific serine residue of the eIF2α subunit (Dever et al., 2007; Donnelly et al., 2013). In the phosphorylated state, the TC can still form, and the affinity for the guanine nucleotide exchange factor eIF2B is even increased (Krishnamoorthy et al., 2001; Rowlands et al., 1988). However, the tight interaction between eIF2B and phosphorylated eIF2α inhibits the nucleotide exchange on the eIF2γ subunit, whereby Met-tRNA$^{Met}_{i}$ is also sequestered by this complex and removed from the translational pool (Krishnamoorthy et al., 2001; Rowlands et al., 1988). Therefore, phosphorylation of eIF2α constitutes an auto-inhibitory mechanism that affects the translation of most mRNAs (Hinnebusch and Lorsch, 2012; Jackson et al., 2010).

2.2.3 Regulation of translation by eIF4E-binding proteins

The eIF4E-binding proteins (4E-BPs) inhibit translation by interfering with the formation of the eIF4F complex on the mRNA 5'cap structure. 4E-BPs bind to eIF4E and prevent the interaction with eIF4G, therefore blocking the subsequent assembly steps of the initiation complex [Figure 3; (Jackson et al., 2010)].

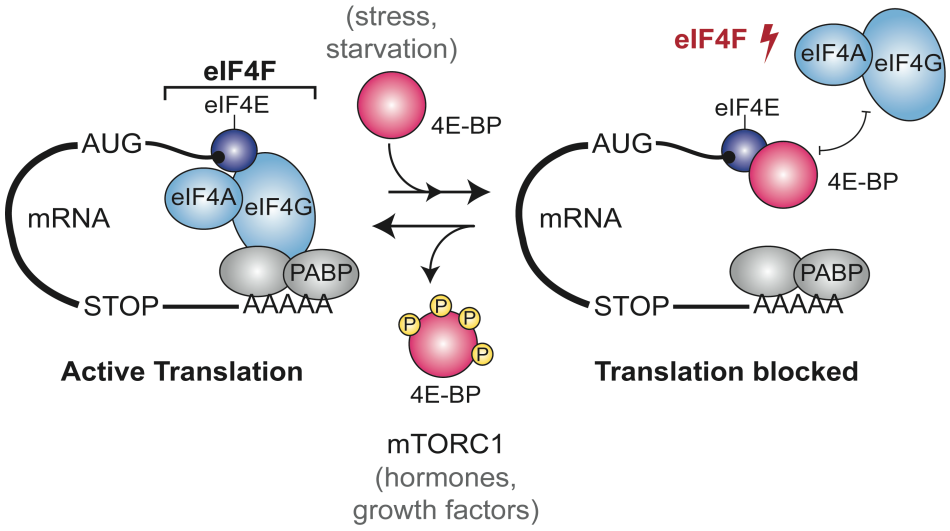


Figure 3: Regulation of translation by 4E-BPs. The assembly of the eIF4F complex (eIF4A, eIF4E and eIF4G, colored in blue) on the 5'cap structure of an mRNA is the initial step of the translation initiation process. Under stress conditions, 4E-BPs (magenta) interfere with the formation of the eIF4F complex by binding to eIF4E, an interaction that directly competes with eIF4G. Upon cellular stimuli that activate translation, the mTORC1 kinase complex phosphorylates 4E-BPs. As a result, hyperphosphorylated 4E-BPs cannot bind to eIF4E anymore and translation is activated. PABP binding to the poly(A)-tail is shown in gray and leads to the circularization of the mRNA through interactions with eIF4G.

The competition between 4E-BPs and eIF4G for eIF4E-binding is based on a similar binding motif of the sequence $YX_4L\Phi$ (where Y is tyrosine, X is any amino acid, L is leucine and Φ is a hydrophobic amino acid), which is termed the canonical eIF4E-binding motif [4E-BM; (Mader et al., 1995; Marcotrigiano et al., 1999)]. Since the canonical 4E-BM of 4E-BPs and eIF4G occupies the same binding site on eIF4E, described as the dorsal surface, the regulation of translation through 4E-BPs is based on a molecular mimicry to eIF4G, and the binding partner of eIF4E ultimately determines if cap-dependent translation is active or repressed [Figure 4; (Marcotrigiano et al., 1999)].

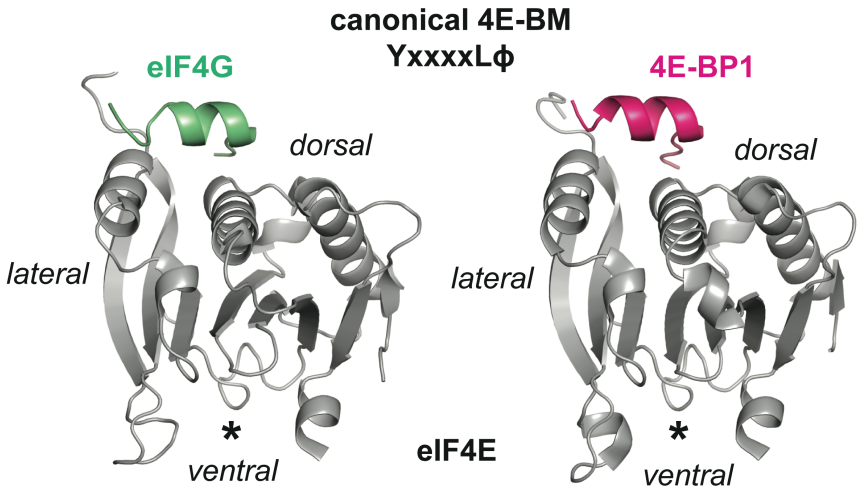


Figure 4: Structures of the canonical eIF4E-binding motifs of eIF4G and 4E-BP1 bound to eIF4E. The canonical 4E-BMs of eIF4G (green) and 4E-BP1 (magenta) bind to the dorsal surface of eIF4E (gray). The binding cavity for the 5' mRNA cap is located at the ventral site of eIF4E and indicated by an asterisk. The lateral side of eIF4E shown in this figure becomes relevant in the 'Results' sections of my thesis. The structures were published by Marcotrigiano et al., 1999, and the atomic coordinates can be found under the pdb accession codes '1ejh' (eIF4G complex) and '1ej4' (4E-BP1 complex), respectively (http://www.rcsb.org).

4E-BPs were initially identified as a group of small translational repressors that contain 3 members in mammals [4E-BP1, -2 and -3; (Bhat et al., 2015; Pause et al., 1994)]. Phosphorylation of 4E-BPs controls their repressive function by modulating the affinity for eIF4E. Hypophosphorylated 4E-BPs associate with eIF4E, but hyperphosphorylated 4E-BPs do not [Figure 3; (Gingras et al., 1999a; Gingras et al., 2001)]. 4E-BPs contain several phosphorylation sites that follow an hierarchical order of phosphorylation, and full phosphorylation is required to block the association with eIF4E (Gingras et al., 1999a; Gingras et al., 2001). A recent NMR study provided a structural explanation for the effect of the initial phosphorylation sites. 4E-BPs are generally unfolded when they are not bound to eIF4E, yet this study shows that the intrinsically disordered 4E-BP2 undergoes a partial disorder-to-order transition upon phosphorylation at specific positions [Figure 5; (Bah et al., 2015; Lukhele et al., 2013)].

The fold of hypophosphorylated 4E-BP2 partially encloses its canonical eIF4E-binding motif and consequently reduces the affinity for eIF4E, however, the molecular consequences of the subsequent phosphorylation sites still remained unclear [Figure 5; (Bah et al., 2015)].

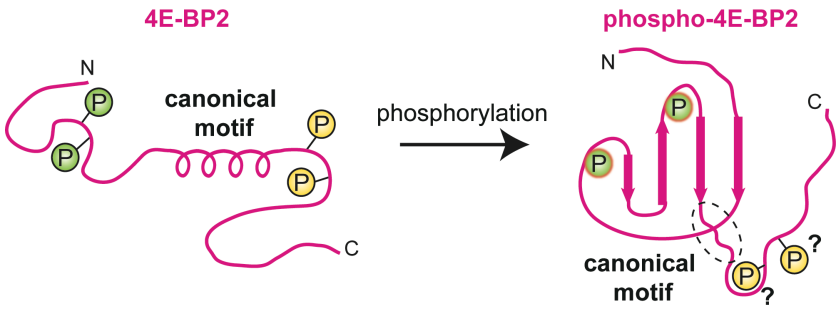


Figure 5: Schematic representation of the phosphorylation sites of 4E-BPs. The small translational repressors 4E-BP1, 2 and 3 contain multiple phosphorylation sites dispersed along the protein. 4E-BP2 is shown in magenta. It is unfolded in the unbound and unphosphorylated state. On the left, the canonical 4E-BM of 4E-BP2 is indicated as a helix, but it only forms upon binding to eIF4E. Phosphorylation of 4E-BPs follows a hierarchical order, whereby the phosphorylation of the later sites (yellow circles) requires the phosphorylation of the initial sites (green circles). Based on Bah et al., 2015, 4E-BP2 undergoes a folding transition upon phosphorylation of the initial sites (green circles), which partially occludes the canonical 4E-BM. However, the role of the later sites remains unclear. Note that 4E-BPs contain additional phosphorylation sites that are not indicated in the scheme for clarity reasons.

4E-BPs are one of the targets of the mechanistic target of rapamycin complex 1 [mTORC1, Figure 3; (Burnett et al., 1998; von Manteuffel et al., 1996)]. mTORC1 is a serine/threonine-specific protein kinase that acts as the downstream factor of the phosphoinositide 3-kinase (PI3K)/Akt signaling pathway, which links nutrient availability and growth factors to protein synthesis (Dowling et al., 2010; Hsieh et al., 2012). Importantly, the dysregulation of this signaling pathway is strongly related to some types of cancer and results in aberrant changes in the activity or amount of translation initiation factors or regulatory factors (Sonenberg and Hinnebusch, 2009). Overexpression of eIF4E is related to the tumorigenic transformation of cells, whereas the reduction of cellular eIF4E levels suppresses the oncogenic phenotype or provides resistance to tumorigenesis beforehand (Lazaris-Karatzas et al., 1990;

Rinker-Schaeffer et al., 1993; Ruggero et al., 2004; Truitt et al., 2015). Similarly, the total amount of 4E-BP1 and its phosphorylation status have implications in ovarian and breast cancer (Armengol et al., 2007). 4E-BP2 is an important factor for the regulation of protein synthesis in the brain tissue and controls the translation of mRNAs involved in learning, memory and behavior. As such, 4E-BP2 knock-out mice display a reduction in long-term memory and autistic phenotypes (Banko et al., 2005). Misregulation of translation has also been related to some neuropsychiatric disorders, such as the autism spectrum disorders (ASDs) or specifically to the fragile X mental retardation (FMR) syndrome (Gkogkas et al., 2013; Napoli et al., 2008). Therefore, the activity of translation initiation factors needs to be tightly regulated to maintain cellular homeostasis. Correspondingly, much effort is being made to target the translation machinery in cancer cells to dampen their translational level and reduce tumor growth (Bhat et al., 2015; Pelletier et al., 2015).

2.4 The broad class of eIF4E-binding proteins

Eukaryotic cells express a plethora of low and high molecular weight proteins with different complexities that interact with eIF4E. The small, low molecular weight 4E-BPs act as global translational repressors, i.e. the human 4E-BPs 1, 2 and 3, which have only one homolog in *Drosophila melanogaster* (*Dm*) called Thor. In contrast, the more complex, high molecular weight 4E-BPs link eIF4E to additional functions as they interact with different proteins to fulfill message-specific tasks [Figure 6; (Miron et al., 2003; Miron et al., 2001).

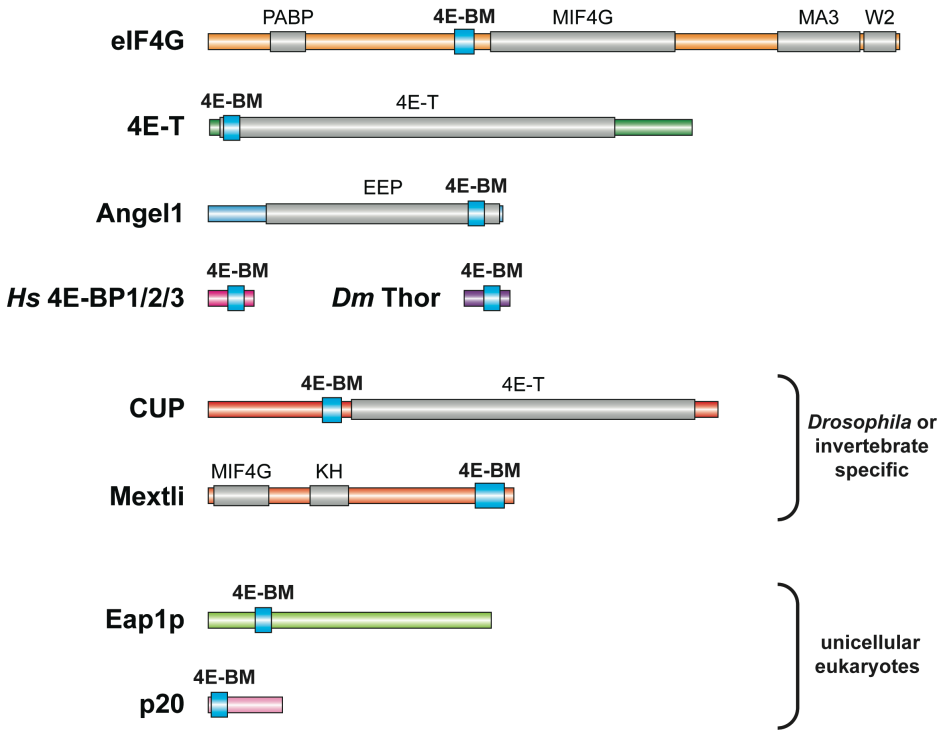


Figure 6: The variety of eIF4E-interacting proteins. Schematic representation of selected eIF4E-binding proteins. The eIF4E-binding motif (4E-BM) is highlighted with a blue box in all proteins. Additional domains are indicated with a gray box in each protein and are labeled above the box. For 4E-BPs that do not exist in humans the species-specific occurrence is indicated.

The Drosophila 4E-BP CUP controls translation during fly development

One of the high molecular weight 4E-BPs is the *Drosophila* protein CUP. CUP is a germline-specific protein that represses the translation of the maternal mRNAs *oskar* and *nanos* during oogenesis and embryogenesis while the mRNAs are transported to and localized at the posterior pole of the oocyte. The specificity for those mRNAs is achieved by binding of CUP to the RBPs Bruno and Smaug, which recognize specific elements in the 3'UTR of *oskar* and *nanos*, respectively (Nakamura et al., 2004; Nelson et al., 2004; Wilhelm et al., 2003; Zappavigna et al., 2004). Binding of CUP to eIF4E and to Bruno or Smaug is then required to repress translation. Additionally, CUP recruits the CCR4-NOT deadenylase complex to promote deadenylation of the target mRNA and to sustain translational repression (Igreja and Izaurralde, 2011).

The 4E-transporter protein is a paralog of Drosophila CUP

In terms of protein organization, CUP is mainly defined by its 4E-T domain, which is based on the similarity to the 4E-transporter protein (4E-T). 4E-T is another 4E-BP as it contains a canonical 4E-BM and, in contrast to CUP, is also present in vertebrates (Kamenska et al., 2014b). It was identified as a nucleocytoplasmic shuttling protein that transports eIF4E into the nucleus (Dostie et al., 2000). Although it has been shown that eIF4E is involved in the nuclear export of some mRNAs, no detailed nuclear function of eIF4E in complex with 4E-T is known (Culjkovic et al., 2005; Dostie et al., 2000). In the cytoplasm, 4E-T localizes to P-bodies, distinct foci in which the mRNA is stored together with proteins that are involved in translational repression and mRNA decay (Andrei et al., 2005; Decker and Parker, 2012; Eulalio et al., 2007; Ferraiuolo et al., 2005). Due to its canonical 4E-BM 4E-T is able to inhibit cap-dependent translation and, moreover, it is known to interact with several decay factors (Ferraiuolo et al., 2005; Kamenska et al., 2014a; Nishimura et al., 2015; Ozgur et al., 2015). However, the functional implications of the multiple 4E-T interaction partners or a detailed cellular function of 4E-T in terms of specific mRNA targets is not fully understood.

The invertebrate-specific 4E-BP Mextli promotes translation

In sharp contrast to other 4E-BPs, *Drosophila* Mextli was found to promote translation instead of repressing it. Mextli proteins also bind to eIF4E via a canonical 4E-BM and additionally exhibit a modular structure that allows further protein and RNA interactions [e.g. with subunits of the eIF3 complex and the DEAH-box DNA/RNA-helicase CG3225; (Hernandez et al., 2013)]. Therefore, parallels were drawn between Mextli and eIF4G in terms of function, and it was proposed that

Mextli forms alternative translation initiation complexes in specific cellular stages. Indeed, phenotypic analysis of mutant flies that show a strongly reduced expression of the Mextli protein indicated that Mextli plays a role in germ line stem cell maintenance and in early embryogenesis (Hernandez et al., 2013). Mextli orthologues can be found in several invertebrate species, e.g. in *Caenorhabditis elegans* (*Ce*), and more data from other organism could shed some light on the role of Mextli in translation.

The CCR4-like protein Angel1 recruits eIF4E to specific subcellular compartments

Several other factors have been reported to interact with eIF4E, many of which have not been shown to have a role in translation (Rhoads, 2009). The human protein Angel1 was identified *in silico* as a 4E-BP due to its canonical 4E-BM, and its interaction with eIF4E was validated in binding assays (Gosselin et al., 2013). Moreover, Angel1 provides an example that 4E-BPs might also control translation in specific cellular compartments as it localizes to the cytoplasmic perinuclear area, to the endoplasmic reticulum and to the Golgi (Gosselin et al., 2013). In addition to its canonical 4E-BM, Angel1 also harbors a CCR4-like endo/exonuclease phosphatase (EEP) domain that might fulfill a regulatory purpose. Due to the similarity between the Angel1 EEP domain and CNOT6, one of the two poly(A)-nucleases of the CCR4-NOT complex, Angel1 might also act directly on the mRNA, even so no deadenylation activity could be shown so far (Gosselin et al., 2013; Tucker et al., 2002). Therefore Angel1 might control the expression of a specific subset of mRNAs, but its molecular function is not yet understood.

Specific 4E-BPs in neuronal tissue

In neurons, CYFIP1 and Neuroguidin bind eIF4E and control translation during synaptic maturation and neurogenesis, respectively, in an mRNA-specific manner and as part of a bigger complex (Jung et al., 2006; Napoli et al., 2008).

4E-BPs in lower eukaryotes

In unicellular eukaryotes, like *Saccharomyces cerevisiae* (*Sc*), the two 4E-BPs Caf20 (also called p20) and Eap1, have been identified. Both of them have been shown to compete with eIF4G for binding to eIF4E and inhibit translation (Altmann et al., 1997; Cosentino et al., 2000). In this context, the *Sc* proteins Caf20 and Eap1 modulate the translation of different subsets of mRNAs (Cridge et al., 2010; Ibrahimo et al., 2006; Rendl et al., 2012).

Viral proteins can capture host eIF4E for their own translational program

Some viruses use specific proteins, called the viral protein genome-linked (VPg) factors, to hijack the host eIF4E protein in order to promote the translation of the viral proteome. However, the mode of binding to eIF4E is distinct for the VPgs compared to other 4E-BPs. The VPgs of plant potyviruses and human caliciviruses act as proteinaceous cap-substitutes that are covalently bound to the 5'end of the viral RNA (Murphy et al., 1996; Schaffer et al., 1980). VPgs bind to eIF4E without competing with eIF4G to recruit the host translation machinery to the viral RNA (Goodfellow et al., 2005; Leonard et al., 2000). Therefore, structural studies of the VPg-eIF4E interaction promise to reveal interesting and novel insights about the recruitment of eIF4E.

Overall, the diversity of proteins with the ability to bind to eIF4E allows the control of translation initiation in a diverse range of cellular and developmental processes. Furthermore, since many 4E-BPs are rather large proteins with additional domains and binding partners, they are able to couple the control of translation initiation with target specificity and other mechanisms that regulate gene expression post-transcriptionally (e.g. deadenylation, decapping, …).

2.5 The eIF4E-binding mode of 4E-BPs

eIF4G and 4E-BPs only have in common the presence of canonical 4E-BM. Apart from this motif, the eIF4E-interacting proteins including eIF4G and the various 4E-BPs do not share any similarity. These proteins generally differ in size, domain organization and even on the sequence level in the regions flanking the canonical 4E-BMs (see Figure 6). Remarkably, these flanking sequences diverge among orthologous proteins across the animal kingdom. Despite the sequence diversity, several studies indicated that additional binding elements in 4E-BPs and eIF4G also contribute to binding to eIF4E. These binding elements do not show an obvious similarity among 4E-BP or eIF4G proteins and are therefore termed as non-canonical eIF4E-binding motifs. Non-canonical 4E-BMs have been identified in the human 4E-BP2, where a bipartite binding region is required for the interaction with eIF4E (Lukhele et al., 2013). A NMR study shows that the eIF4E-4E-BP2 complex is a dynamic assembly, in which 4E-BP2 exists as a heterogeneous ensemble within the bound state (Lukhele et al., 2013). Other biophysical studies on the interaction between eIF4E and 4E-BPs or eIF4G also suggested the presence of a second element on these proteins (their non-canonical 4E-BM), which contributes to the overall affinity of 4E-BPs and eIF4G for eIF4E (Paku et al., 2012; Umenaga et al., 2011).

Moreover, a detailed molecular insight into the interaction of a non-canonical 4E-BM with eIF4E was provided by the crystal structure of the *D. melanogaster* protein CUP bound to eIF4E. CUP contains a conserved canonical 4E-BM, which binds to the dorsal surface of eIF4E in a similar way as eIF4G or 4E-BP1, forming a short α-helix (Kinkelin et al., 2012; Marcotrigiano et al., 1999). In the crystal structure, however, a second α-helix of the CUP peptide, the non-canonical 4E-BM, addresses a lateral, hydrophobic surface of eIF4E [Figure 7; (Kinkelin et al., 2012)].

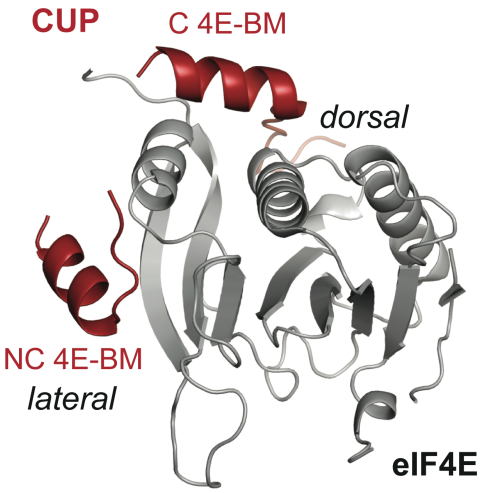


Figure 7: Structure of the eIF4E-binding region of CUP in complex with eIF4E. *Dm* CUP (red) binds to the dorsal surface of eIF4E (gray) with its canonical 4E-BM and to a lateral surface of eIF4E with its non-canonical 4E-BM. The structure was published by Kinkelin et al., 2012, and can be found under the pdb accession code '4axg' (http://www.rcsb.org).

Interestingly, the non-canonical 4E-BM of CUP is relevant for translational repression as it prevents decapping and complete degradation of CUP-regulated mRNAs (Igreja and Izaurralde, 2011). This data suggests that non-canonical 4E-BMs also play important roles for the function of 4E-BPs as translational repressors.

At the beginning of my doctoral work our structural understanding of those additional binding motifs was only limited to CUP, which is a germline-specific protein in *Drosophila* and it does not exist in mammals. Because outside of the canonical 4E-BM the sequence of different eIF4E-binding partners is highly diverse,

the information from the structural model of the eIF4E-CUP complex could not be transferred to eIF4G or other 4E-BPs. Especially the molecular details about the interaction of 4E-BP1, 2 and 3 or eIF4G with eIF4E remained unclear and limited our understanding on the binding mode and of the functional relevance of non-canonical contacts in the control of translation. Nevertheless, theses studies pointed towards the existence of an extended and more complex binding mode of different 4E-BPs and eIF4G for eIF4E and were the starting point of my doctoral work.

2.6 The 4E-homologous protein and its interaction partners

Eukaryotic cells express two or more eIF4E proteins that share a conserved protein sequence, which constitutes the folded core of eIF4E (Joshi et al., 2005). The different eIF4E proteins are classified into three classes (class I, II and III) based on minor variations of the conserved protein sequence (Joshi et al., 2004; Joshi et al., 2005). The binding capabilities, interaction partners, and the type of mRNA cap-structure recognized by eIF4E proteins differ between the three classes (Rhoads, 2009). According to theses properties, it is suggested that only one eIF4E (eIF4E1) in each organism is responsible for global translation, whereas the others (eIF4E2/3) perform more specific functions related to translation (Hernandez and Vazquez-Pianzola, 2005).

Caenorhabditis elegans (*Ce*) contains five eIF4E proteins termed initiation factor of *elegans* 1-5 (IFE1-5), which exhibit differences in binding specificity between monomethylguanosine (MMG; m^7GTP) and trimethylguanosine (TMG; $m_3^{2,2,7}GTP$) mRNA 5'caps. Whereas IFE-1, -2 and -5 can bind to MMG and TMG caps, IFE-3 and IFE-4 can only bind MMG caps (Keiper et al., 2000; Miyoshi et al., 2002). The TMG cap is acquired by a process called trans-splicing in which the

5'UTR including the MMG cap (called 'outron') is replaced by a short spliced leader (SL) sequence that carries a TMG cap (Lasda and Blumenthal, 2011). The TMG cap can stimulate translation in nematodes, and since it is only recognized by some of the IFE proteins, it might be related to specific translation programs (Keiper et al., 2000; Lall et al., 2004; Wallace et al., 2010).

Drosophila melanogaster contains eight eIF4E-related proteins (*Dm* eIF4E1-8) that are differently expressed throughout the development of the fly (Hernandez et al., 2005; Hernandez et al., 1997). All of the *Dm* eIF4E proteins are able to bind to the 5' cap of mRNAs (usually MMG caps), and besides *Dm* eIF4E-6, which is C-terminally truncated, and *Dm* eIF4E-8, all of them are able to bind eIF4G or 4E-BPs (Hernandez et al., 2005). *Dm* eIF4E-8, also called the 4E-homologous protein (4EHP), is classified as a class II eIF4E-family member, and it is the most diverse member among the *Dm* eIF4E proteins (Joshi et al., 2005). In *Drosophila*, 4EHP is expressed during early embryogenesis and is involved in the translational repression of the maternal mRNAs *caudal* and *hunchback*, to which it is recruited via interactions with the RBPs Bicoid and Brain tumor, respectively (Cho et al., 2006; Cho et al., 2005; Hernandez et al., 2005).

Likewise, human cells also contain three eIF4E proteins (*Hs* eIF4E1-3), and each one is part of a different class (Joshi et al., 2005). Among them, *Hs* eIF4E1 is the major variant driving the canonical cap-dependent translation, whereas *Hs* eIF4E2 is the human homolog to *Dm* 4EHP and does not bind to eIF4G. Conversely, *Hs* eIF4E3 does bind to eIF4G but not to 4E-BPs (Gingras et al., 1999b; Joshi et al., 2004; Joshi et al., 2005; Rom et al., 1998).

Similar to the *Drosophila* 4EHP, human 4EHP plays a role in the translational repression of specific mRNAs. Together with the Grb10-interacting protein 2 (GIGYF2), 4EHP is recruited to mRNAs containing AU-rich elements (AREs) in their 3'UTR by the specific RNA-binding protein Tristetraprolin [TTP, Figure 8; (Fu et al., 2016; Morita et al., 2012; Tao and Gao, 2015)]. As TTP triggers repression and degradation of cytokine mRNAs, which usually contain AREs, the 4EHP-GIGYF translational repressor complex has been implicated in the regulation of cytokine expression during immune responses (Barreau et al., 2005; Fu et al., 2016). Because 4EHP does not bind to eIF4G, it is proposed to interfere with translation initiation by competing with eIF4E1 for the mRNA cap-structure, ultimately preventing the assembly of the eIF4F complex [Figure 8; (Hernandez et al., 2005)]. A similar mechanism has been proposed in mice, in which the Zn-finger protein ZNF598 recruits the 4EHP-GIGYF complex to yet unknown targets, but it plays an important role in the mouse embryonic development (Giovannone et al., 2009; Morita et al., 2012).

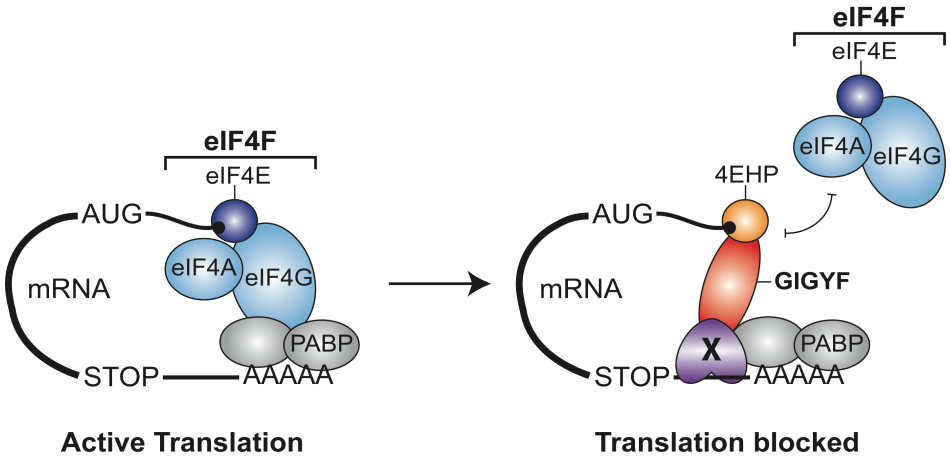


Figure 8: Translational repression by 4EHP. 4EHP (orange) competes with eIF4E (dark blue) for binding to the 5'cap structure of an mRNA. Because 4EHP does not bind to eIF4G (light blue), it cannot recruit the 43S PIC to initiate translation. Together with GIGYF (red), 4EHP forms a translational repressor complex that is recruited to specific mRNAs by a RNA-binding protein X (purple). The protein X can be TTP, ZNF598 or a yet unknown factor. eIF4A as a part of the eIF4F complex is colored in light blue, PABP is colored in gray.

In conclusion, the diverse functional roles of the eIF4E-family members are in part related to their binding specificity for mRNA 5'cap structures, but also by their distinct interacting partners. However, the high similarity between the eIF4E proteins in sequence and structure is in stark contrast to their binding specificity (Joshi et al., 2005; Rosettani et al., 2007). For example, GIGYF proteins specifically bind to 4EHP but not to eIF4E1 *in vivo*, even though their characterized 4EHP-binding motif resembles the canonical 4E-BM of eIF4G or 4E-BPs [$YXYX_4L\Phi$ vs. $YX_4L\Phi$ for GIGYF and eIF4G/4E-BPs, respectively; (Morita et al., 2012)]. For that reason it is mechanistically not clear how eIF4G and GIGYF can discriminate between eIF4E and 4EHP. Furthermore, even though *Hs* 4E-BP1 does not bind to 4EHP *in vivo*, it does bind *in vitro* but with a reduced affinity compared to eIF4E1 (Rom et al., 1998; Rosettani et al., 2007). This provides an indication that distinct molecular features might be involved in determining the binding specificity for eIF4E or 4EHP, a question that was addressed in my doctoral work and is discussed below.

3. AIMS & OBJECTIVES

My doctoral work focused on two main aspects of translational regulation. In the first part of my work, I investigated different eIF4E-4E-BP complexes and the impact of their binding mode on the competition with eIF4G and the control of translational. The second part of my study aimed at the structural characterization of 4EHP-related complexes involved in an alternative mechanism of translational repression.

3.1 The contribution of non-conserved regions of 4E-BPs to eIF4E-binding and their impact on the function of 4E-BPs as translational repressors

The class of 4E-BPs comprises functionally diverse proteins that share a conserved canonical eIF4E-binding motif of the sequence $YX_4L\Phi$ (where Y is tyrosine, X is any amino acid, L is leucine and Φ is a hydrophobic amino acid) with eIF4G. Therefore, these proteins compete with eIF4G for binding to eIF4E and ultimately shut down translation.

4E-BPs fulfill a critical function in the control of translation, either in a global manner or in a mRNA-specific context, however a detailed molecular understanding of the mechanism on how 4E-BPs compete with eIF4G and displace it from eIF4E remained elusive. Several studies focusing on the biophysical characterization of the eIF4E-4E-BP and eIF4E-eIF4G interactions indicated the presence of additional eIF4E-binding sequences, known as non-canonical eIF4E-binding motifs, in 4E-BPs and eIF4G (Lukhele et al., 2013; Paku et al., 2012; Umenaga et al., 2011). Additionally, the crystal structure of the *Drosophila* specific 4E-BP CUP bound to eIF4E revealed a second, lateral binding site on eIF4E (Kinkelin et al., 2012). Nonetheless, due to the sequence diversity among 4E-BPs, no common features could

be derived from the eIF4E-CUP complex and no detailed structural insight into other 4E-BPs was available.

For this reason, in the first part of my doctoral work I focused on obtaining a general understanding of the contribution and the molecular details of the interaction between non-canonical eIF4E-binding motifs from various 4E-BPs and eIF4E. My aim was to generate structural insight into different eIF4E-4E-BP complexes in order to deduce commonalities in their binding mode as well as a mechanistic explanation regarding the competition mechanism with eIF4G.

Additionally, I also initiated and participated in a collaborative project that focused on the eIF4E-eIF4G interaction. This work aimed at obtaining molecular details of eIF4E-eIF4G complexes, which could help to understand the mechanism of translational regulation by 4E-BPs.

Furthermore, I also investigated a highly specialized, invertebrate-specific 4E-BP called Mextli because it was reported to initiate translation instead of repressing it (Hernandez et al., 2013). Therefore, the Mextli proteins promised to provide additional structural insights and functional understanding regarding eIF4E-complexes that initiate translation.

3.2 The binding selectivity of eIF4E- and 4EHP-specific protein partners

The final part of my doctoral work focused on an alternative mechanism of translational repression. In this mechanism, the key player is another eIF4E-protein, called 4EHP, rather than a 4E-BP. 4EHP does not bind to eIF4G, which is the reason why it acts as translational repressor, but it binds other proteins like GIGYF1/2. GIGYF1/2 bind specifically to 4EHP but not to eIF4E.

The major objective of this study was to examine the binding specificity of GIGYF1/2, especially because molecular details about their interaction with 4EHP were limited. To address this question, I performed structural studies on the 4EHP-GIGYF1/2 translational repressor complexes. This work provided novel insights on the selective binding of GIGYF proteins to 4EHP over eIF4E, and on how specific translational regulatory complexes might be maintained *in vivo*, in the presence of a plethora of 4E-BPs and eIF4E-like proteins.

4. RESULTS

4.1 Non-canonical eIF4E-binding motifs are a common feature of different 4E-BPs and are required for their function as translational repressors

The work summarized in this section was published by Igreja et al., 2014 and Peter et al., 2015a, and the papers including detailed experimental data and methods are included in the Appendix.

Binding studies on 4E-BP/eIF4E interactions

To understand and compare the binding of different 4E-BPs to eIF4E, I focused on *in vitro* approaches that enabled working in minimal systems with purified components. For this purpose, I cloned the extended eIF4E-binding regions containing a canonical and a putative non-canonical 4E-BM (C+NC) of the *Dm* proteins CUP, Thor (homolog to human 4E-BP1-3 proteins), 4E-T and eIF4G for expression in *E. coli* BL21(DE3) cells. To define these regions, I made use of the crystal structure of the CUP-eIF4E complex, or published biophysical data regarding the human 4E-BP1, 4E-T and eIF4G proteins. This information could be transferred to the *Dm* proteins based on sequence comparisons [Figure 9; (Kubacka et al., 2013; Lukhele et al., 2013; Paku et al., 2012; Umenaga et al., 2011)]. To ensure the solubility and stability of the eIF4E-binding regions, which are predicted to be unstructured, I introduced the B1 domain of protein G (GB1) as a non-cleavable small globular domain at the C-termini of the constructs and purified the eIF4E-binding regions (C+NC), or fragments (C and ΔC) to homogeneity (Figure 9).

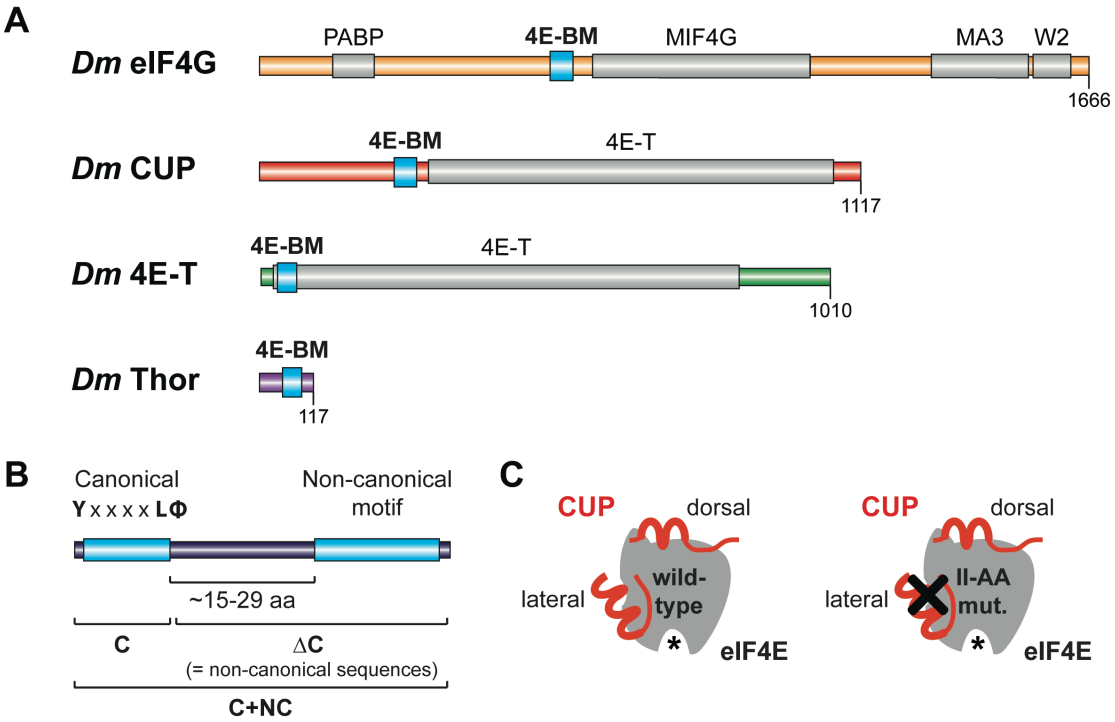


Figure 9: Schematic overview of different 4E-BPs from *Drosophila melanogaster*. (A) Schematic representation of *Dm* eIF4G, CUP, 4E-T and Thor. In gray are shown the PABP-binding region and the MIF4G, MA3 and W2 domains of eIF4G, and the 4E-T domain of CUP and 4E-T, which is based on the similarity to human 4E-T. In all proteins, the eIF4E-binding motif (4E-BM) is shown as a blue box. The numbers at the end of the protein schemes indicate their size in number of amino acids. (B) General construct design for the different 4E-BPs. The constructs contain either the canonical motif and the non-canonical sequences (C+NC) or derived fragments (C or ΔC). A linker of approximately 15 to 29 amino acids (aa) connects the canonical and non-canonical 4E-BMs. (C) Schematic representation of *Dm* CUP (red) bound to eIF4E (gray) based on the crystal structure (Kinkelin et al., 2012). The dorsal and lateral binding sites of eIF4E are indicated. Structure-based mutations disrupt the interaction at the lateral side of eIF4E (II-AA mut.).

The purified peptides of the different 4E-BPs were used in isothermal titration calorimetry (ITC) measurements to obtain the binding constants and thermodynamic parameters of the interaction with eIF4E, as well as to derive the binding contributions of each binding motif. The ITC data showed that the C+NC peptides of CUP, Thor and 4E-T strongly bind to eIF4E and exhibit dissociation constants (K_Ds) in the low nanomolar range comparable to the affinity of a similar peptide of eIF4G. Furthermore, the canonical motifs of CUP, Thor and 4E-T showed a reduced binding affinity for eIF4E compared to their C+NC peptides. This indicates that similar to CUP, also Thor and 4E-T do contain non-canonical sequences, which contribute to the binding of eIF4E.

Moreover, fragments of the eIF4E-binding region, which either consisted of the canonical motif or the non-canonical sequences, also revealed variations in the binding to eIF4E between the different 4E-BPs. First, the affinity of the canonical motif for eIF4E differs by at least one order of magnitude among the proteins, suggesting different contributions of the non-canonical sequences for eIF4E-binding (affinity of the canonical motifs: 4E-T > CUP > Thor). Second, only the CUP fragment lacking the canonical motif (ΔC) could be measured by ITC, as no changes in enthalpy could be detected when titrating corresponding Thor and 4E-T ΔC fragments into an eIF4E-containing solution. This finding is consistent with the higher entropic penalty for the interaction between eIF4E and CUP compared to Thor and 4E-T. The entropic component of the binding thermodynamics is a measure for the number of states in a closed system and can be seen as the disorder of the overall system. In the case of CUP, the entropy change of the reaction indicates a lower degree of conformational freedom in the bound state, which is in agreement with the non-canonical sequences of CUP forming an α-helix in the crystal structure of the complex with eIF4E (Kinkelin et al., 2012). Therefore, it can be speculated that the contribution and the mode of binding of the non-canonical sequences varies among different 4E-BPs, especially since the sequence composition is very diverse between the proteins.

Due to this diversity, I next tested if the non-canonical sequences of Thor and 4E-T bind to the same lateral, hydrophobic binding pocket of eIF4E as CUP. To address this question, a structure-based mutant of eIF4E, in which two hydrophobic residues of the lateral binding pocket (two isoleucines) were mutated to alanine (II-AA mutant), was used in ITC measurements to determine the binding affinities for the eIF4E-binding regions of eIF4G, CUP, Thor and 4E-T that comprise the canonical

motif and the non-canonical sequences (Figure 9). In agreement with the different binding contributions of the canonical motifs, binding of all of the proteins was affected by these mutations but to a different extent. In summary, we can conclude from the ITC measurements that all of the different 4E-BPs that were tested contain non-canonical sequences, and that a common lateral, hydrophobic pocket of eIF4E is involved in the interaction with these motifs. Additionally, the high affinity of the eIF4G fragment already suggests additional binding sequences outside its canonical motif.

To further investigate the relevance of the non-canonical interactions for the function of 4E-BPs as translational repressors, we performed competition assays *in vitro* using preassembled eIF4E-eIF4G complexes, in which eIF4E was either wild-type or mutated at the lateral binding pocket (II-AA mutant). Those complexes were immobilized on Ni-NTA affinity resin via a hexahistidine-tag on eIF4E. The immobilized complexes were then challenged with constructs of different 4E-BPs, which compete with eIF4G for eIF4E-binding. These assays were carried out together with my colleague Cátia Igreja. We could show that 4E-BP peptides (CUP, Thor and 4E-T) containing the canonical motif and the non-canonical sequences efficiently displace eIF4G from eIF4E. However, disrupting the non-canonical contacts either by deleting the non-canonical sequences of the competitor 4E-BP, or by mutating the lateral hydrophobic pocket of eIF4E, strongly reduced or abolished the ability of 4E-BPs to displace eIF4G that is in complex with eIF4E.

In agreement with the findings from the *in vitro* assays, *in vivo* studies confirmed the importance of the non-canonical sequences for the function of 4E-BPs as translational repressors, since mutations in the non-canonical motif abolished Thor- and CUP-mediated repression of the expression of a firefly luciferase reporter in

Drosophila cells (4E-T was excluded from the in vivo study because it causes mRNA degradation; the *in vivo* experiments were carried out by Cátia Igreja).

Structural studies on 4E-BP-eIF4E interactions

The functional importance of non-canonical sequences of different 4E-BPs is in stark contrast to the low sequence conservation outside their canonical motif. Due to this diversity among 4E-BPs, the structural insights from the *Dm* CUP-eIF4E complex could not be used to derive general principles for the non-canonical interactions. Therefore, in a continuation of the work with the *Dm* 4E-BPs Thor and 4E-T, I used X-ray crystallography to obtain detailed molecular insights into the mode of binding of those 4E-BPs to eIF4E. To this end, I solved the crystal structures of the eIF4E-binding region of *Dm* Thor, its human homolog 4E-BP1 and *Dm* 4E-T bound to the respective eIF4E (Figure 10).

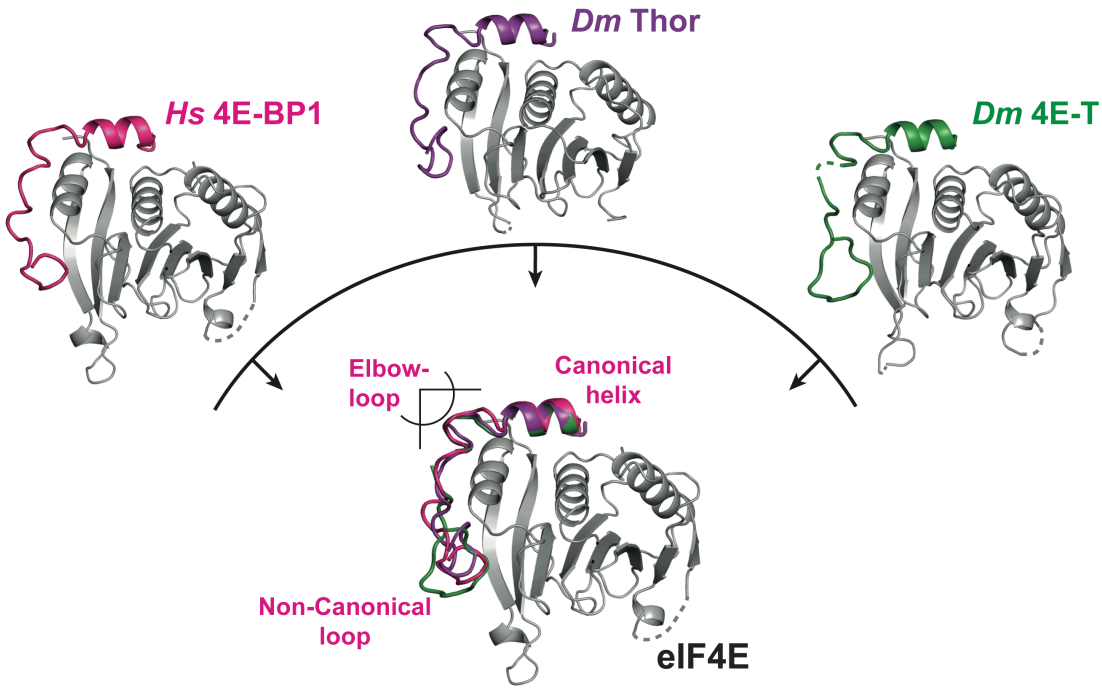


Figure 10: Comparison of the structures of different 4E-BPs bound to eIF4E. The upper part of the figure shows the structural overview of *Hs* 4E-BP1 (magenta), *Dm* Thor (purple) and *Dm* 4E-T (green) bound to eIF4E (gray) in a cartoon representation. The lower part shows an alignment of the eIF4E molecules of the three structures, whereby only the human eIF4E is shown for clarity. The structural elements that are common between the three 4E-BPs are indicated in magenta.

This comparative structural information revealed a common eIF4E-binding mode among distinct 4E-BPs. In the bound form, the eIF4E-binding region of *Hs* 4E-BP1, *Dm* Thor and *Dm* 4E-T arrange similarly and share three common structural elements: (1) the canonical motif forming an α-helix at the conserved dorsal surface of eIF4E, (2) an elbow-loop C-terminal to the canonical α-helix that mediates a sharp turn for directing the (3) unfolded non-canonical sequences (NC loops) to engage the lateral surface of eIF4E (Figure 10).

Besides the general arrangement, each of these three elements has specific molecular features. A comparative analysis of the canonical helices of *Hs* 4E-BP1, *Dm* Thor and *Dm* 4E-T revealed a conserved and important role of residues that flank the motif with long side chains containing a large aliphatic portion (mostly arginine or lysine residues). These residues are at the same position within the canonical motifs (position +2 and +9, whereby the tyrosine of $YX_4L\Phi$ is 0), they extend the interaction surface with eIF4E by shielding hydrophobic patches with the aliphatic portion of their side chain and mediate charge-charge interactions with conserved residues on eIF4E. Accordingly, substitutions of those residues to alanine in *Dm* Thor impaired its ability to compete with eIF4G for binding to eIF4E in competition assays. Due to the consistent role and conservation of those residues among 4E-BPs, we suggested an extended canonical motif of the sequence $YX[R/K]X_2L\Phi X_2[R/K]$.

In contrast to the canonical motif, the sequence composition of the subsequent C-terminal regions deviates between the different 4E-BPs. Nonetheless, I could derive a common structural principle regarding their arrangement, which involves the formation of similar elbow-loop structures that bend the 4E-BPs backbone towards the lateral surface of eIF4E. In general, the elbow-loops consist of a helical half-turn starting with a proline or glycine residue. Further the elbow loops are stabilized by a

network of intramolecular interactions and are tethered to the surface of eIF4E close to the lateral binding pocket by interactions with conserved residues of eIF4E. In addition, the structures of *Dm* Thor and *Hs* 4E-BP1 uncovered the location of two of the conserved phosphorylation sites (a serine and a threonine in both cases) placed within and after the elbow loop (Gingras et al., 1999a; Gingras et al., 2001; Miron et al., 2003). These two phosphorylation sites in *Hs* 4E-BP1-3 are the last to be modified in a phosphorylation cascade inhibiting the interaction with eIF4E (Gingras et al., 1999a; Gingras et al., 2001). However, phosphorylation of these serine and threonine residues did not measurably impact the affinity of the eIF4E-binding region of *Dm* Thor for *Dm* eIF4E, as determined in ITC measurements with synthetic peptides of *Dm* Thor containing or not the phospho-modifications (wild-type peptide vs. pSpT peptide). Nevertheless, phosphorylation of these two residues had significant effects on a functional level. In competition assays *in vitro*, the pSpT peptide of *Dm* Thor was unable to replace eIF4G in preassembled eIF4E-eIF4G complexes compared to the wild-type peptide. This finding is supported by the effect of a phosphomimetic mutant of *Dm* Thor (the respective serine and threonine replaced by aspartates) *in vivo*, which, in contrast to wild-type Thor, does not repress the translation of a firefly luciferase reporter mRNA (the *in vivo* experiments were carried out by Cátia Igreja). These results indicate that phosphorylation at these two specific positions does not affect the interaction with eIF4E but destabilizes or hampers the formation of the elbow loop of *Dm* Thor or *Hs* 4E-BP1-3 upon binding to eIF4E, since the serine residue is involved in the intramolecular stabilization of the elbow loop. The potential hindrance introduced by the two strongly negative charged phosphates might be detrimental for 4E-BPs when another competing molecule like eIF4G is already bound to eIF4E. Moreover, the synergy of multiple phosphorylation sites on full

length *Hs* 4E-BP1-3 or *Dm* Thor might have additional regulatory effects, since the stepwise phosphorylation of *Hs* 4E-BP2 induces folding and increases its stability in the unbound state, but also strongly reduces its binding affinity for eIF4E (Bah et al., 2015; Yanagiya et al., 2012). *Dm* 4E-T does not contain similar phosphorylation sites and it is not known how its interaction with eIF4E is regulated.

Following the elbow-loop, *Hs* 4E-BP1, *Dm* Thor and *Dm* 4E-T approach the lateral side of eIF4E with their non-canonical motif, but all diverge in the path of their polypeptide chain and do not form a defined secondary structure like *Dm* CUP (Figure 10). Comparative analysis of the three eIF4E-complexes, however, shows that *Hs* 4E-BP1, *Dm* Thor and *Dm* 4E-T align again around the lateral hydrophobic pocket of eIF4E. This lateral binding site of eIF4E is conserved across species and lined by a phenylalanine (which is a tyrosine in *Dm* eIF4E) and two isoleucines in *Hs* eIF4E. In contrast to the lateral surface of eIF4E, the non-canonical motifs of 4E-BPs are highly variable at the sequence level. Nonetheless, the structural insight uncovered similar binding principles between *Hs* 4E-BP1, *Dm* Thor and *Dm* 4E-T. All of these proteins engage the lateral hydrophobic pocket of eIF4E with similar hydrophobic contacts, which involve either valine, leucine or isoleucine residues for *Hs* 4E-BP1 and *Dm* Thor, and bulky aromatic residues (phenylalanine and tryptophan) in the case of *Dm* 4E-T. These interactions are associated with analogous main-chain contacts between eIF4E and the 4E-BP polypeptides, leading to an alignment of the 4E-BPs backbone along the lateral side of eIF4E. In conclusion, the sequence variability of the non-canonical motifs of 4E-BPs is overcome on the structural level by similar interactions with the lateral hydrophobic pocket of eIF4E. Therefore it can be assumed that all 4E-BPs, even those not yet studied at the structural level, generally bind to eIF4E in a bipartite manner and involve similar dorsal and lateral surfaces of eIF4E.

Targeting the eIF4E-eIF4G interaction in disease

The overexpression of eIF4E is a hallmark of a variety of tumor cells and is linked to oncogenic transformation (Avdulov et al., 2004; De Benedetti and Graff, 2004; Lazaris-Karatzas et al., 1990; Mamane et al., 2004). Since 4E-BPs are able to revert cellular transformation, the presented structural insights provide valuable information for alternative strategies to target eIF4E in malignancies (De Benedetti and Graff, 2004; Grzmil and Hemmings, 2012; Mamane et al., 2004). One possible strategy is the use of small molecule compounds that target protein-protein interactions. There are some inhibitors described to target the eIF4E-eIF4G interaction, e.g. a synthetic molecule called 4EGI-1 (eIF4E/eIF4G Interaction Inhibitor-1) that mimics the function of 4E-BP1, or the small molecules called 4E1RCat/4E2RCat that are a dual inhibitor for the eIF4E-eIF4G and eIF4E-4E-BP interactions (Cencic et al., 2011; Moerke et al., 2007).

Since such compounds have several drawbacks regarding efficiency, affinity or specificity, we made use of our structural knowledge to design a more potent, peptide-based 4E-BP mimic. For this purpose, we combined the canonical motif of *Dm* 4E-T, the elbow-loop of *Dm* Thor and the non-canonical motif of *Dm* CUP, which forms an α-helix (NC helix), to generate a 4E-BP chimera (Figure 11). The selection of the elements was based on the biophysical and structural studies described in the previous section, and only the connection between the Thor elbow-loop and the CUP NC helix required the addition of a glycine residue as a spacer.

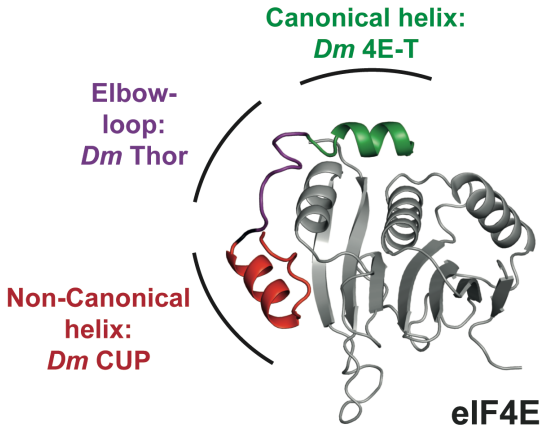


Figure 11: Design of the 4E-BP chimera. Crystal structure of the 4E-BP chimera bound to eIF4E (gray). The elements of the 4E-BP chimera are shown in different colors for the parental proteins with *Dm* 4E-T in green, *Dm* Thor in purple and *Dm* CUP in red, whereby the glycine spacer is colored in black.

The designed 4E-BP chimera was validated by ITC measurements, which revealed a low nanomolar affinity for *Dm* eIF4E that is in the same range as other 4E-BPs. I also determined the crystal structure of the 4E-BP chimera in complex with *Dm* eIF4E, which showed that there were no deviations from the expected arrangement. Interestingly, the 4E-BP chimera displays significant functional improvements. In comparison to its parental proteins, the 4E-BP chimera was a more efficient competitor of the eIF4E-eIF4G interaction *in vitro*. Furthermore, *in vivo* it displayed a stronger translational repressor activity compared to *Dm* Thor (the competition assays *in vitro* were performed together with Ramona Weber, the *in vivo* experiments were carried out by Cátia Igreja). The improved function might result from advantages in the binding mode and also from the fact that, when compared to *Dm* Thor or *Hs* 4E-BP1, it lacks the regulatory recognition sequences, so it cannot be regulated by phosphorylation. Moreover, even though the 4E-BP chimera is based on *Dm* proteins it efficiently disrupts the eIF4E-eIF4G interaction in human cells (the *in vivo* experiments were carried out by Lara Wohlbold). Therefore it presents a promising starting point for the future development of structure-based drug approaches.

Overall, these studies on 4E-BPs uncovered intrinsic molecular properties regarding the recognition of eIF4E by functionally distinct 4E-BPs. Moreover, the structural insights generated in theses studies can be used in therapeutic approaches targeting eIF4E that aim at reducing translation in cancer cells. The potential use of our research for therapeutic approaches is a topic that will be discussed below.

4.2 eIF4G and 4E-BPs bind similarly to eIF4E

The work summarized in this section was published by Grüner et al., 2016, and the paper including detailed experimental data and methods is included in the Appendix.

The bipartite binding mechanism to two orthogonal surfaces of eIF4E is a prerequisite for 4E-BPs to act as competitors of eIF4G and consequently as translational repressors. However, a comparable structural insight into the mode of binding between eIF4G and eIF4E was lacking, even though previous studies also indicated the presence of a second eIF4E-binding site on the human eIF4G and auxiliary binding regions in yeast eIF4G (Gross et al., 2003; Umenaga et al., 2011). In a collaborative approach, we elucidated the eIF4E-binding mechanism of metazoan eIF4G. In this work, I initiated the project and further supported it with functional experiments and analysis, whereby my colleague Stefan Grüner did the major work and was responsible for the structure determination and the biophysical analysis.

The findings of our work revealed that metazoan eIF4G (human and *Dm*) also addresses the lateral hydrophobic pocket of eIF4E in a similar way as 4E-BPs. This result was unexpected, since a NMR structure of the *Sc* eIF4E-eIF4G complex suggested that eIF4G does not bind to the lateral surface of eIF4E (Gross et al., 2003). In essence, the NMR structure of the *Sc* eIF4E-eIF4G complex showed that auxiliary sequences N- and C-terminal to the canonical motif of *Sc* eIF4G contribute to complex formation by forming a bracelet-like structure around the N-terminus of eIF4E (Gross et al., 2003). Such an arrangement would leave the lateral side of eIF4E open for non-canonical interactions with 4E-BPs, thus providing an anchoring point for the competition with eIF4G. However, our structural data about metazoan eIF4E-

eIF4G complexes revises this idea of an exclusive 4E-BP binding site on eIF4E outside of the yeast system and extents the model of the 4E-BP competition mechanism to a complete molecular mimicry of 4E-BPs to eIF4G (Figure 12).

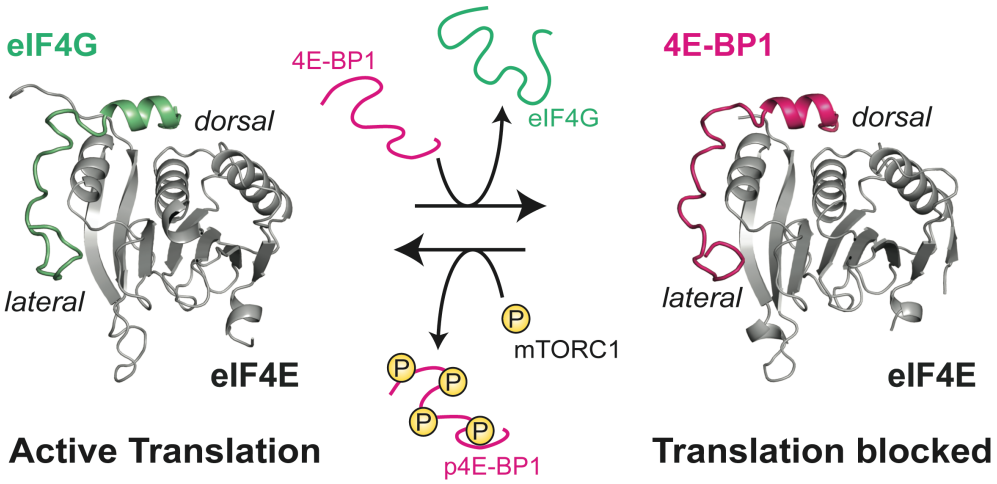


Figure 12: Extended molecular mimicry between eIF4G and 4E-BPs for binding to eIF4E. Both eIF4G (green) and 4E-BP1 (magenta) contact the dorsal and lateral surfaces of eIF4E (gray). The mode of binding between both complexes is highly similar but decides upon activation or inhibition of mRNA translation. However, molecular differences in the binding mode enable 4E-BPs to efficiently displace eIF4G from eIF4E-complexes. Phosphorylation of 4E-BPs by mTORC1 is required to regulate 4E-BPs.

Despite the structural similarity with the eIF4E-binding region of eIF4G, 4E-BPs are efficient competitors of the eIF4E-eIF4G interaction. In order to delineate the molecular details of their competitive advantage, we performed competition assays *in vitro* in which we tested several chimeric competitors for their ability to displace eIF4G from preassembled eIF4E-eIF4G complexes. Those chimeric competitors consisted of interchanged, defined elements between *Dm* Thor and *Dm* eIF4G, namely their canonical and non-canonical motifs, and the linker connecting both. These assays revealed that the linker and the non-canonical motif of *Dm* Thor confer a competitive advantage over eIF4G although the affinity of the different chimeric competitors was not significantly altered compared to *Dm* Thor or *Dm* eIF4G (the competition assays were done together with Ramona Weber, the ITC measurements

were performed by Stefan Grüner). Moreover, the competitive advantage of *Dm* Thor could partially be attributed to the rigidity of its linker, since an exchange of a double proline sequence in *Dm* Thor to the structural equivalents of *Dm* eIF4G (lysine and glutamine) impaired its ability to displace eIF4G from eIF4E compared to wild-type *Dm* Thor. In agreement with the important role of the linker and NC motifs of 4E-BPs for competition with eIF4G, further ITC analysis of *Dm* Thor revealed that 4E-BPs have a stronger requirement for the lateral side of eIF4E. Taken together, these results suggest a competition mechanism where 4E-BPs strongly rely on the lateral side of eIF4E to dissociate bound eIF4G and ultimately repress translation initiation (discussed below, see Figure 19).

4.3 Diverse eIF4E-binding modes among Mextli proteins confer distinct functional properties

The work summarized in this section was published by Peter et al., 2015b, and the paper including detailed experimental data and methods is included in the Appendix.

Mextli proteins are invertebrate-specific 4E-BPs that, in contrast to most other 4E-BPs, exhibit a modular structure consisting of a MIF4G-like (middle domain of eukaryotic initiation factor 4G) domain and a KH- (hnRNP K homology-) domain (Figure 13). Besides the interaction with eIF4E mediated by a canonical motif, Mextli binds RNA and other protein factors involved in translation. The specific role of Mextli is thought to be the control of translation initiation in germline stem cells and during early embryogenesis in *Drosophila* (Hernandez et al., 2013). Indeed, Mextli is the only eIF4E-interacting protein currently known to activate translation besides eIF4G. In analogy to eIF4G, Mextli has been reported to interact with an uncharacterized DEAH-box DNA/RNA-helicase (encoded by CG3225) as well as with eIF3 subunits, leading to the recruitment of the 43S PIC to the bound mRNA [Figure 13; (Hernandez et al., 2013)]. Therefore, Mextli proteins might be involved in an alternative mechanism to initiate cap-dependent translation.

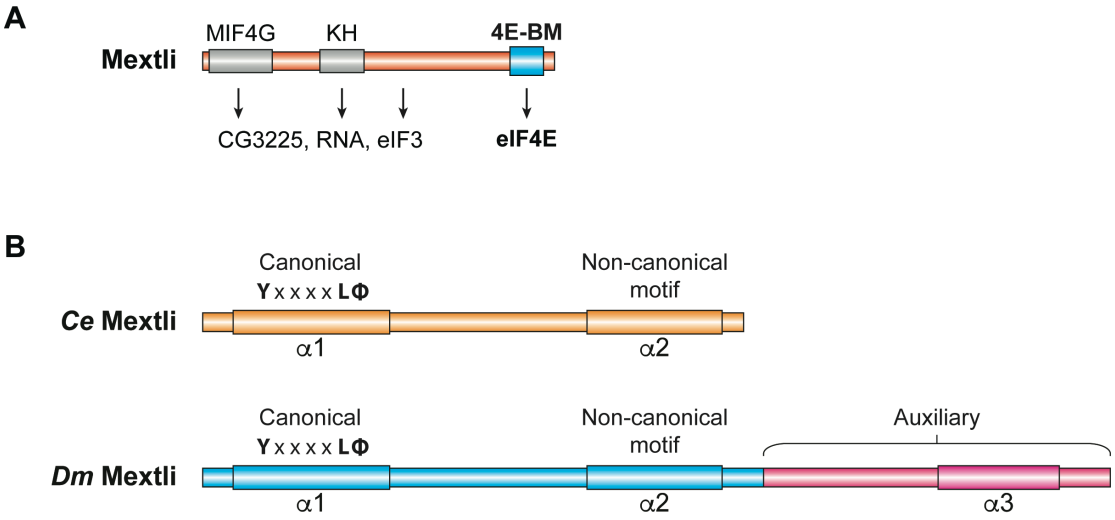


Figure 13: The modular architecture of Mextli proteins. (A) Schematic representation of Mextli. The MIF4G and KH domains are shown in gray, the 4E-binding motif (4E-BM) in blue. Reported interactions are indicated below the scheme. (B) Schematic comparison of the 4E-BMs of *Ce* and *Dm* Mextli. Besides the canonical and non-canonical motifs present in *Ce* (orange) and *Dm* (blue) Mextli, *Dm* Mextli contains C-terminal auxiliary regions (magenta). Secondary structure elements are indicated below the cartoons.

To shed some light into this unusual pathway of translation, we analyzed the interaction of *Dm* and *Caenorhabditis elegans* (*Ce*) Mextli proteins with eIF4E from the respective species. Interestingly, *Dm* and *Ce* Mextli display differences in their eIF4E-binding region already at the sequence level. In both species, the eIF4E-binding site of Mextli is located at the very C-terminus of the protein and consists of a canonical motif followed by a region of helical propensity and conserved hydrophobic residues, which could likely act as a non-canonical motif (α2). However, in fly Mextli the C-terminus extends even further and comprises another auxiliary helical region (α3) that is conserved in flies and some insects, but absent in worms like *Ce* (Figure 13). This peculiarity between Mextli proteins on the sequence level also affects the binding to eIF4E *in vivo* and *in vitro*. Both *Dm* and *Ce* Mextli proteins respond to a lateral side mutant of their respective eIF4E (II-AA mutant in *Dm*, VI-AA mutant in *Ce*), which strongly reduces or abolishes the interaction in pull-down assays *in vitro*. However, fragments of *Dm* Mextli that consist of the canonical motif (C) and the non-canonical sequences (NC), but contain or not the fly-specific auxiliary sequences

(including the helical region α3; fragments therefore described as C+NC+α3 or C+NC; Figure 13) were pulled down by eIF4E from cell lysates to a different extend. In particular, the lack of the auxiliary region of *Dm* Mextli resulted in a significantly reduced binding to eIF4E (the *in vivo* experiments were conducted by Cátia Igreja). This finding is in agreement with *in vitro* pull-down assays, in which a C+NC+α3 fragment of *Dm* Mextli could partially bind to the II-AA mutant of eIF4E, whereas the C+NC fragment could not.

For a quantitative assessment of the contribution of the auxiliary region of *Dm* Mextli for binding to eIF4E, I performed ITC measurements comparing the same fragments described above that contain or not the auxiliary sequences (C+NC+α3 vs. C+NC; Figure 13). The addition of the auxiliary region of *Dm* Mextli (C+NC+α3) increases the affinity for eIF4E around 10-fold and is associated with a significantly higher entropic penalty of the interaction compared to the C+NC fragment. This result indicates that the auxiliary region of *Dm* Mextli contributes to the interaction with eIF4E and is highly restrained in its flexibility in the bound state, likely undergoing folding upon binding. To also obtain insight into the unbound state of the eIF4E-binding region of *Dm* Mextli, we performed nuclear magnetic resonance (NMR) measurements (NMR experiments and evaluation were performed together with Vincent Truffault). Therefore, I expressed the C+NC+α3 eIF4E-binding region of *Dm* Mextli in minimal medium supplemented with ammonium-^{15}N chloride as the only nitrogen source. The purified ^{15}N-labelled peptide was used for ^{1}H-^{15}N heteronuclear single quantum coherence (HSQC) experiments, which rely on the transfer of magnetization from proton to ^{15}N nuclei and back. Thus, every bonded nitrogen-proton pair of the protein (i.e. backbone amides and side chains with nitrogen-bound protons) yields a peak in the two-dimensional HSQC spectrum with a specific

assigned resonance. A broad dispersion of the peaks in the HSQC-spectrum indicates a well-folded protein since the residues of the protein are found in different chemical environments (solvent-exposed vs. globular fold). In contrast, a central cluster of peaks is a strong indicator of an unfolded state of a polypeptide, because most of the residues are similarly exposed to the solvent. This is the case for the C+NC+α3 peptide of *Dm* Mextli in isolation, which confirms that, similar to other 4E-BPs, the eIF4E-binding region of Mextli is mainly unfolded in the unbound state (Lukhele et al., 2013; Tait et al., 2010). Upon addition of an excess of unlabeled eIF4E, the resonance dispersion of the Mextli peptide vastly increases and many peaks disappeared from the disordered region of the spectrum, reflecting a disorder-to-order transition upon binding. In conclusion, the NMR data of the unbound and bound state of the eIF4E-binding region of *Dm* Mextli supports the idea that the different entropic contribution of the two Mextli fragments observed in the ITC measurements is caused by additional folding of the auxiliary region of *Dm* Mextli.

To obtain a detailed molecular insight into the differences in eIF4E-binding between the Mextli homologs from *Dm* and *Ce*, I determined the crystal structures of their eIF4E-binding regions bound to the respective eIF4E (Figure 14). Both, *Dm* and *Ce* Mextli contact eIF4E with canonical and non-canonical motifs, similarly to the 4E-BPs described above. This classical bipartite mode of binding involves two orthogonal surfaces (dorsal and lateral) of eIF4E. At the dorsal surface of eIF4E, conserved interactions are established by the $YX_4L\Phi$ consensus motif ubiquitously present in 4E-BPs and also in Mextli proteins. Interactions at the lateral surface of eIF4E are mainly mediated through hydrophobic contacts of Mextli and are structurally equivalent to the interactions of other 4E-BPs at this hydrophobic pocket

of eIF4E. Again this finding highlights a common structural principle for non-canonical sequences of 4E-BPs despite the lack of conservation on the sequence level.

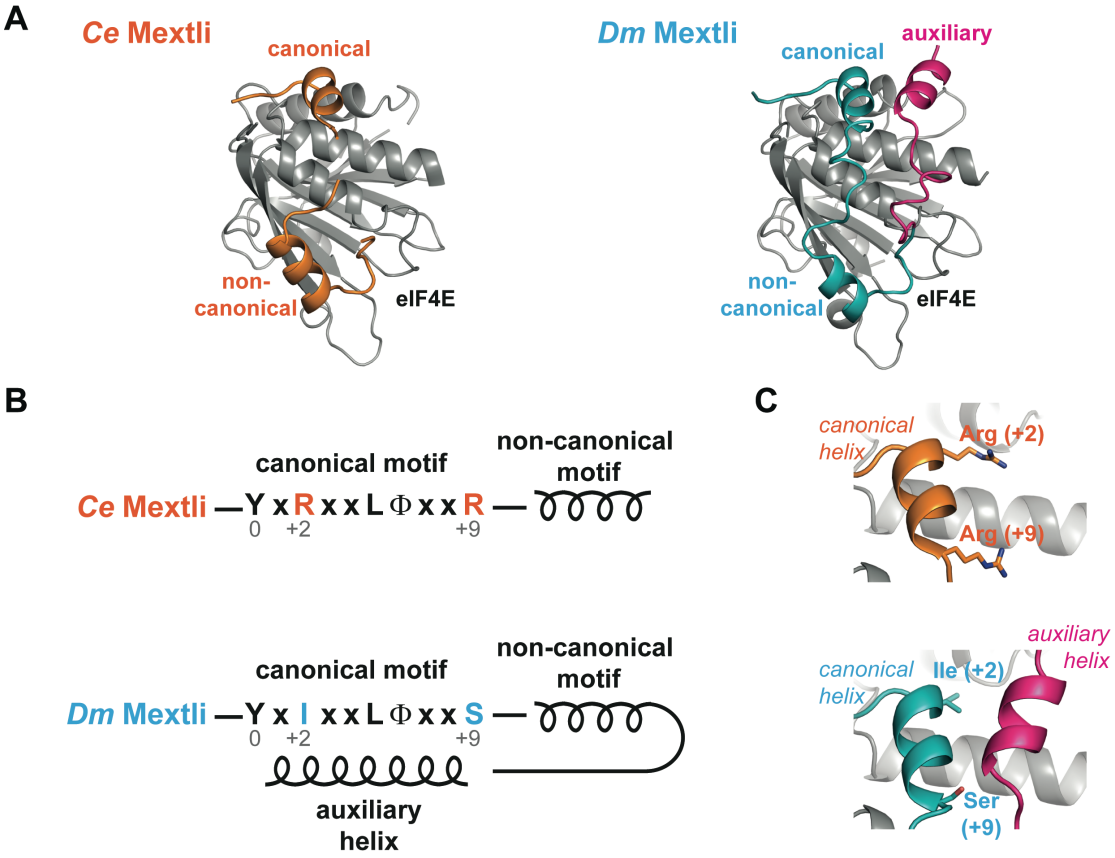


Figure 14: Structures of eIF4E-Mextli complexes. (A) Structural overviews of *Ce* Mextli (orange) and *Dm* Mextli (cyan and magenta) bound to the species-related eIF4E (gray). (B) Schematic comparison of the *Ce* and *Dm* Mextli 4E-BM. Residues of the extended canonical motif are indicated in the consensus sequence and the positions +2 and +9 are highlighted, with the initial tyrosine (Y) is counted as 0. (C) Close-up view on the arrangement of *Ce* (top) and *Dm* (bottom) Mextli on the dorsal surface of eIF4E. The canonical helices of Mextli are colored in orange (*Ce*) or cyan (*Dm*) and the *Dm* Mextli auxiliary helix is colored in magenta. Residues at positions +2 and +9 of the extended canonical motif are shown as sticks. The figure is adapted from the book chapter of Hernandez et al., 2016 to which I contributed to and to which I provided the original figure.

The Mextli non-canonical motifs adopt an α-helical conformation (α2), a feature so far only described for CUP (Kinkelin et al., 2012). In the complex of *Dm* Mextli with eIF4E, the non-canonical helix is connected to the canonical motif by a well-ordered linker that forms the characteristic elbow loop structure of 4E-BPs. In contrast, the linker bridging both helices in *Ce* Mextli is rather flexible and adopts different conformations between the copies from the two independent crystal structures of the complex. The lacking coordination of the linker region of *Ce* Mextli might be compensated by the extension of the non-canonical helix compared to *Dm*, which contributes to the hydrophobic interaction with the lateral surface of eIF4E.

Interestingly, *Dm* Mextli extends this classical bipartite binding mode in an exceptional manner. Following the non-canonical motif, the auxiliary sequences extend the interaction back to the dorsal surface of eIF4E, where *Dm* Mextli folds into a third, auxiliary α-helix (α3). The auxiliary helix is placed in close proximity and antiparallel to the canonical helix, and both form intramolecular contacts. This leads to an overall U-shaped arrangement of *Dm* Mextli on dorsal and lateral surfaces of eIF4E. Besides the three helical elements, also the linker regions of *Dm* Mextli contribute to the tight interaction with eIF4E and are well organized. For the coordination of the *Dm* Mextli linker regions, two highly conserved residues (an asparagine and a histidine) of eIF4E fulfill an important role. These residues are located at the edge between the dorsal and the lateral side of eIF4E and dually coordinate the linker region between the canonical and non-canonical helices (non-canonical linker), and between the non-canonical and auxiliary helices (auxiliary linker). Mutations of those two residues to glutamate in *Dm* eIF4E (NH-EE mutation) abolished the interaction with *Dm* Mextli, but did not affect the interaction to other eIF4E-binding partners like eIF4G, CUP, Thor and 4E-T in immunoprecipitation

assays (IPs) from *Dm* S2 cells (the IPs were carried out by Cátia Igreja). This result confirms the importance of these two residues of eIF4E for the binding of *Dm* Mextli and illustrates the diversity of this binding mode in comparison to bipartite 4E-BPs.

The plasticity in the eIF4E-binding mode among the Mextli orthologs can be attributed to a distinct structural feature in their canonical motifs. As suggested from the structural studies of bipartite 4E-BPs described above, residues with long aliphatic side chains usually flank the canonical helix and strongly contribute to the interaction with eIF4E. Due to the conserved position of those residues, we proposed an extended canonical 4E-BM of the sequence YX[R/K]X_2LΦX_2[R/K]. This extended canonical 4E-BM is also present in *Ce* Mextli, which contains arginine residues on both positions (+2 and +9, whereby the tyrosine is counted as 0). In *Dm* Mextli, however, both positions are substituted by residues with shorter side chains (isoleucine at position +2, serine at position +9; Figure 14). For that reason, the auxiliary helix of *Dm* Mextli can be accommodated on the dorsal surface of eIF4E in proximity to the canonical helix. The surface that is contacted on eIF4E by the auxiliary helix of *Dm* Mextli is primarily overlapping with that covered by the long aliphatic side chains present in the eIF4E-complexes with bipartite 4E-BPs like the *Ce* Mextli. Moreover, the conserved interactions generally mediated by the long side chains of the R/K residues in bipartite 4E-BPs are replaced by equivalent contacts mediated by residues in the auxiliary sequences of *Dm* Mextli. Particularly, two aromatic residues establish these contacts, namely a phenylalanine shields the hydrophobic surface of eIF4E, and a tyrosine mediates hydrophobic and polar interactions with eIF4E and the canonical helix of *Dm* Mextli. In conclusion, the structural analysis of the *Dm* eIF4E-Mextli complex indicated that the presence of long side chains on positions +2 and +9 of the canonical motif would interfere with the arrangement of the auxiliary helix of

Dm Mextli. As such, 4E-BPs containing the extended canonical 4E-BM cannot adopt a similar auxiliary arrangement, which seems to be unique to *Dm* Mextli. Furthermore, the presence of specific residues in the canonical motif can be used as an indication for a bipartite or non-bipartite 4E-BP.

We also investigated the consequences of the tripartite binding mode of *Dm* Mextli on a functional level. For this purpose, we analyzed the behavior of bipartite and tripartite 4E-BPs in competition assays *in vitro* (the competition assays were performed with the help of Ramona Weber). From these assays, we could derive valuable insights into specific structure-function relations:

First, a bipartite 4E-BP competes efficiently with eIF4G for eIF4E-binding. In contrast, the tripartite *Dm* Mextli C+NC+α3 fragment was a weak competitor of the eIF4E-eIF4G interaction.

Second, the *Dm* eIF4E-Mextli complex is resistant to the competition by bipartite 4E-BPs. The stability of the *Dm* eIF4E-Mextli complexes in the presence of competitor proteins was assessed in competition assays with preassembled eIF4E-Mextli complexes containing or lacking the auxiliary sequences. In the presence of different competitor proteins (eIF4G or the chimeric 4E-BP) only the *Dm* Mextli complex lacking the auxiliary sequences could be competed out. Hence, the auxiliary sequences of *Dm* Mextli confer stability and resistance to the eIF4E-Mextli complex against bipartite 4E-BP competitors.

Third, the long aliphatic side chains at positions +2 and +9 of the extended canonical 4E-BM ($YX[R/K]X_2L\Phi X_2[R/K]$; Y is counted as 0) are an important feature of a bipartite 4E-BP. As stated before, substitution of these residues impairs the ability of 4E-BPs to compete with eIF4G. The functional significance of these residues can also be observed in the Mextli proteins. In *Dm* Mextli, these two

positions are substituted by smaller residues (isoleucine and serine) to accommodate the auxiliary helix. A *Dm* Mextli peptide lacking the auxiliary sequences competes poorly with other 4E-BPs or eIF4G for binding to eIF4E and, *vice versa*, is quickly displaced from eIF4E by other 4E-BPs. However, mutating positions +2 and +9 of the extended canonical motif of the *Dm* Mextli peptide that lacks the auxiliary sequences to arginine (IS-RR mutant) resulted in a bipartite *Dm* Mextli mutant with properties similar to other bipartite 4E-BPs. This bipartite *Dm* Mextli peptide carrying the IS-RR substitutions displaced eIF4G from preassembled eIF4E-eIF4G complexes as efficiently as *Ce* Mextli, which is a bipartite 4E-BP and already contains arginines at both positions. Further, the IS-RR mutant of *Dm* Mextli displayed an increased resistance to competition by other 4E-BPs when it was bound to eIF4E, but not to the some extent as the *Dm* Mextli peptide containing the auxiliary region. Therefore, the behavior of the bipartite *Dm* Mextli IS-RR mutant is clearly distinct from that mediated by the auxiliary sequences. The arginine substitutions do not provide similar properties to the eIF4E-complex but are the key features to turn a functionally impaired peptide of *Dm* Mextli that lacks the auxiliary sequences into a normal bipartite 4E-BP. These results underline the functional importance of the extended canonical motif for a bipartite 4E-BP.

In summary, *Dm* Mextli is an interesting 4E-BP, not only because of its proposed function as a translational activator, but also in terms of its eIF4E-binding region. On a structural level, the novel tripartite eIF4E-binding mode of *Dm* Mextli underlines the relevance of the extended canonical motif. From a functional perspective, however, this binding mode implies a different regulatory mechanism to disassemble *Dm* eIF4E-Mextli complexes.

4.4 The eIF4E-homologous protein provides binding specificity to GIGYF through a surface that is not conserved in eIF4E

The work summarized in this section is described in the manuscript by Peter et al., 2017 and is accepted for publication in *Genes & Development*. The accepted manuscript including detailed experimental data and methods is included in the Appendix.

The eIF4E-homologous protein (4EHP, or also eIF4E-2) is one of the three eIF4E proteins in vertebrates. However, compared to the major eIF4E protein in cells, eIF4E-1, 4EHP does not bind to eIF4G and therefore acts as a translational repressor (Morita et al., 2012). Instead 4EHP interacts with additional proteins involved in diverse cellular and developmental functions. Among those partners are the human Grb10-interacting GYF domain-contain (GIGYF) proteins 1 and 2. The 4EHP-GIGYF repressive complex is recruited to specific mRNAs by additional interactions with certain RNA-binding proteins. Indeed, the Zn-finger protein ZNF598 has been shown to interact with GIGYF and is required for mouse embryonic development, whereas in the human context it has been shown that Tristetraprolin (TTP) recruits the 4EHP-GIGYF complex to mRNAs containing AU-rich elements in their 3'-UTRs, which play a role in the inflammatory cytokine response (Fu et al., 2016; Giovannone et al., 2009; Morita et al., 2012; Tao and Gao, 2015).

GIGYF proteins have been reported to specifically bind to 4EHP but not to eIF4E *in vivo*, despite the approximately 30% identity and 42% similarity between both human eIF4E proteins (Morita et al., 2012; Rosettani et al., 2007). To investigate whether GIGYF can also discriminate between eIF4E and 4EHP at a molecular level,

i.e. in the absence of a possible regulatory mechanism, we performed pull-down assays *in vitro* (the pull-down assays were performed with the support of Sigrun Helms). In these assays eIF4E or 4EHP were pulled down in the presence of a N-terminal fragment of GIGYF1/2 containing the 4EHP-binding region, or the eIF4E-binding region of 4E-BP1 (Figure 15). The GIGYF1/2 proteins were strongly pulled down by 4EHP, but only weakly by eIF4E, whereas 4E-BP1 bound similarly to both. These results indicate that GIGYF1/2 discriminates between eIF4E and 4EHP *in vitro*, whereas 4E-BP1 does not as previously reported (Rosettani et al., 2007).

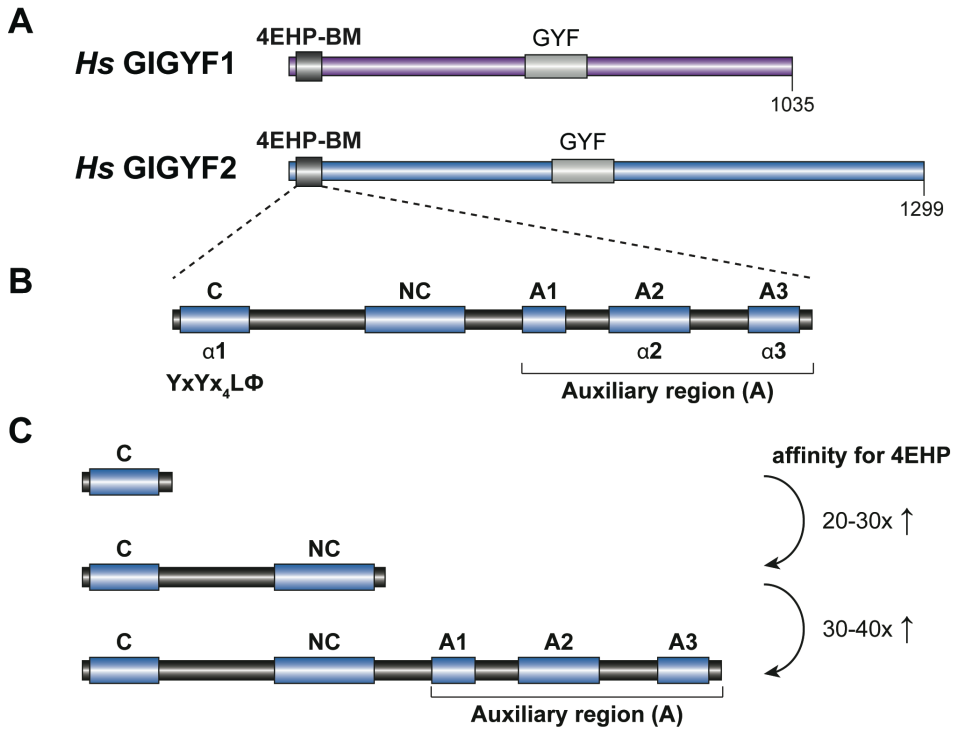


Figure 15: 4EHP-binding region of the human GIGYF proteins. (A) Schematic representation of the two human GIGYF proteins. GIGYF1 (purple) and GIGYF2 (blue) contain a GYF domain (gray) and a N-terminal 4EHP-binding motif (black). (B) Close-up view on the 4EHP-BMs of GIGYF. Identified binding elements are shown with a blue box and refer to the canonical (C), non-canonical (NC) and auxiliary (A) binding sites. The canonical 4EHP-BM as well as predicted secondary structure elements are indicated below the corresponding regions. (C) Fragments of GIGYF that were used for *in vitro* binding studies with 4EHP. The relative change in affinity for 4EHP according to ITC measurements is indicated on the right.

The sequence and structural similarity between eIF4E and 4EHP suggests that 4EHP also contains dorsal and lateral binding sites analogous to eIF4E. To address this question, 4EHP structure-based mutants were generated to disrupt interactions at the dorsal (W-A mutant) or lateral (IM-AA mutant) surfaces of 4EHP. Both mutations affected the interaction of 4EHP with endogenous GIGYF1/2 and abolished the interaction with overexpressed 4E-BP1 in pull-down assays from human cell lysates, showing that 4EHP indeed uses these two orthogonal surfaces to interact with its binding partners (the *in vivo* experiments were carried out by Lara Wohlbold). This result also shows that both the dorsal and the lateral binding surface are conserved within the eIF4E-protein family.

The N-terminal 4EHP-binding region of *Hs* GIGYF1/2 is characterized by a canonical 4EHP-binding motif ($YXYX_4L\Phi$; canonical 4EHP-BM), which is similar to the canonical 4E-BM [Figure 15; (Morita et al., 2012)]. Moreover, GIGYF1/2 proteins contain additional binding elements C-terminal to their canonical 4EHP-BM to address the lateral surface of 4EHP. Accordingly, a non-canonical 4EHP-BM, which displays a high content of hydrophobic residues is located within a characteristic distance to the canonical 4EHP-BM. Furthermore, an additional C-terminal extension, termed the auxiliary region (A), is also present in these proteins (Figure 15). The auxiliary region comprises several elements that either displayed high sequence conservation among GIGYF proteins or additionally helical propensity (Figure 15).

To analyze the contribution of the different binding elements of *Hs* GIGYF1/2 to the interaction with 4EHP, we compared the association of *Hs* GIGYF1/2 proteins that contain or not the auxiliary region (C+NC+A vs. C+NC; Figure 15) with 4EHP in binding assays *in vitro*. These GIGYF proteins bound differently to the 4EHP mutants

(W-A and IM-AA) compared to wild-type. In these binding assays, only *Hs* GIGYF1/2 proteins that lack the auxiliary region (C+NC) were strongly affected in their interaction with 4EHP by the mutations on 4EHP. The GIGYF1/2 proteins containing the auxiliary region (C+NC+A) bound to wild-type or mutant 4EHP to the same extent (the pull-down assays were done with the help of Sigrun Helms). The W-A and IM-AA mutations of 4EHP abolished the interaction with *Hs* 4E-BP1 *in vitro*. Therefore, the auxiliary regions of *Hs* GIGYF1/2 contribute to the assembly of the 4EHP-GIGYF complex, as they compensate for the mutations on the dorsal and lateral surfaces of 4EHP and confer properties to the complex that are distinct from the bipartite model of the eIF4E-4E-BP1 complex.

Moreover, ITC measurements of *Hs* GIGYF1/2 protein fragments confirmed the contribution of non-canonical and auxiliary sequences for binding to 4EHP. Compared to the canonical 4EHP-BM alone, a GIGYF protein containing the non-canonical sequences in addition to the canonical motif displayed a strong gain in affinity for 4EHP (approximately 20-30 fold; C vs. C+NC fragments; Figure 15). An additional increase in the binding affinity was observed when the GIGYF protein fragments also included the auxiliary region (approximately 30-40 fold; C+NC vs. C+NC+A fragments; Figure 15). The final binding affinity of GIGYF proteins for 4EHP is in the sub-nanomolar range (C+NC+A; Figure 15). Of further note is the fact that the presence of the auxiliary region of *Hs* GIGYF1/2 (C+NC+A) imposes a significant increase in the entropic penalty compared to the fragments lacking these sequences (C+NC). In agreement with the structural studies below, this increase in entropic contribution can be explained by additional disorder-to-order transitions upon binding of the auxiliary region to 4EHP.

In order to get a detailed molecular insight into the 4EHP-GIGYF complex, I solved the crystal structure of the 4EHP-binding region of *Hs* GIGYF1/2 including the canonical motif, non-canonical sequences and the auxiliary region bound to 4EHP (Figure 16). Both 4EHP-complexes (*Hs* GIGYF1 and *Hs* GIGYF2) are highly similar to each other, and reveal unexpected insights into this interaction.

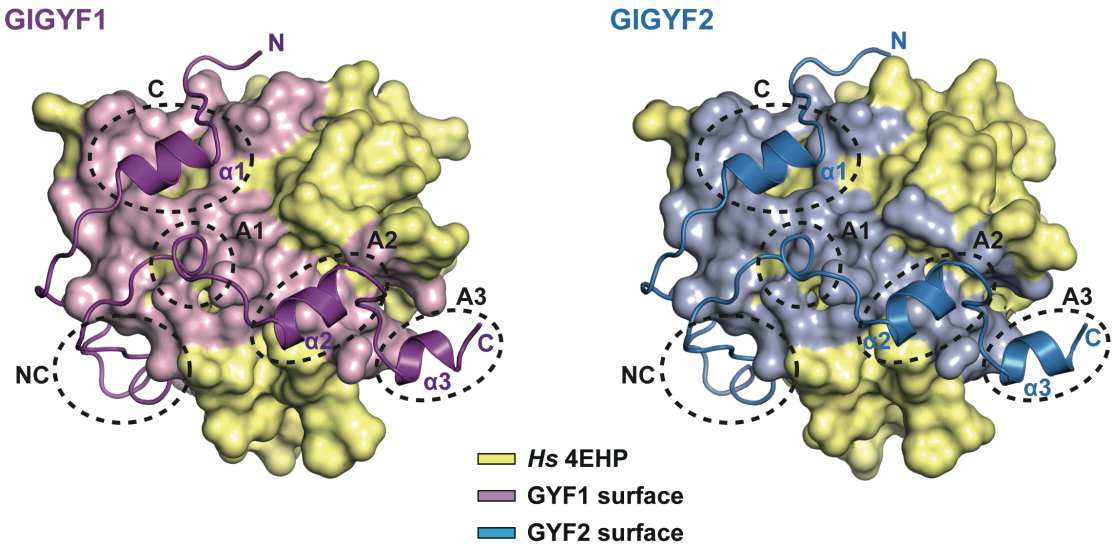


Figure 16: Structures of the human GIGYF1/2 proteins bound to 4EHP. The 4EHP-binding region of GIGYF1 (purple) and GIGYF2 (blue) is shown in a cartoon representation with the particular binding elements (canonical, C: non-canonical, NC; auxiliary1-3, A1-3) highlighted by a black dashed circle. The secondary structure elements of GIGYF are indicated. The surface of 4EHP is shown in yellow, whereas the interface with GIGYF1/2 is highlighted in light purple or light blue, respectively.

First, GIGYF proteins contact the dorsal and the lateral surface of 4EHP in a manner similar to how 4E-BPs bind to eIF4E. On the dorsal surface of 4EHP, GIGYF forms a canonical helix ($YXXYX_4L\Phi$) that mediates conserved interactions with 4EHP analogous to those observed for the canonical 4E-BM of other 4E-BPs or eIF4G ($YX_4L\Phi$), as the motif and the dorsal surface are conserved features (Marcotrigiano et al., 1999). Moreover, the initial tyrosine of the canonical 4EHP-BM (**Y**$XXYX_4L\Phi$) does not fulfill a function in discriminating between eIF4E and 4EHP as it was previously proposed (Cho et al., 2005; Villaescusa et al., 2009). This tyrosine residue

only contributes to the binding with weak van der Waals contacts and could potentially be accommodated as well in a complex with eIF4E. The most interesting feature about the arrangement of GIGYF on the dorsal surface of 4EHP concerns the interactions with the auxiliary sequences. The canonical helix of GIGYF1/2 together with 4EHP form a composite surface to which the A1 motif binds in a shape-complementary manner. A key factor involved in this interaction is an aromatic residue (phenylalanine in *Hs* GIGYF1, tyrosine in *Hs* GIGYF2) that is positioned at the prominent position +9 of the extended canonical 4E-BM, instead of the long aliphatic residue usually present in 4E-BPs (extended canonical 4E-BM of 4E-BPs: YX[R/K]X_2LΦX_2[**R/K**]; see above). Thus, besides the case of *Dm* Mextli described above, GIGYF proteins are another interesting example on the influence of the +9 position of the extended canonical 4E-BM on the binding mode to members of the eIF4E-family of proteins.

The coordination of the linker connecting canonical and non-canonical motifs of GIGYF by 4EHP is different from 4E-BPs bound to eIF4E, as GIGYF proteins lack the characteristic elbow-loop structure. A triad of a tyrosine, an isoleucine and a methionine mainly lines the lateral surface of 4EHP, forming a hydrophobic pocket with contributions from surrounding residues. Overall the lateral binding site of 4EHP is structurally very similar to eIF4E, and the mode of binding of GIGYF resembles other 4E-BPs as it involves placing of hydrophobic or aromatic residues at critical positions. Therefore it is surprising that a GIGYF1/2 peptide containing canonical motif and non-canonical sequences can still discriminate between eIF4E and 4EHP. In comparison with structures of the eIF4E-4E-BP complexes, the non-canonical interactions of GIGYF with 4EHP do not display any clear feature that would preclude binding to eIF4E. Indeed, GIGYF introduces a critical phenylalanine residue

into the hydrophobic pocket of 4EHP, which establishes an elaborate hydrophobic interaction network. Consequently, the similarity of interactions at the dorsal and lateral surfaces of 4EHP compared to eIF4E-complexes does not shed light on the binding specificity of GIGYF for 4EHP over eIF4E. Additional mutational studies on the 4EHP-binding region of GIGYF are therefore required to address this unsolved question.

Following the non-canonical motif, GIGYF1/2 contain auxiliary sequences that contact and cross the dorsal surface of 4EHP and finally bind in a groove of 4EHP almost opposite to the non-canonical interactions (Figure 16). The auxiliary region of GIGYF1/2 harbors a conserved PLAL motif (sequence stretch of proline-leucine-alanine-leucine) as well as two additional helical elements (α2 and α3; Figure 16). The PLAL motif binds to the composite surface formed by the canonical helix and 4EHP. The helices α2 and α3 are connected by a tightly coordinated linker sequence. Interestingly, the coordination of the auxiliary region of *Hs* GIGYF1/2 involves many residues of 4EHP that are diverse in eIF4E, indicating that similar interactions cannot be formed within eIF4E-complexes and suggesting why GIGYF prefers 4EHP over eIF4E *in vivo*.

To validate that the 4EHP-surface contacted by the auxiliary region of *Hs* GIGYF1/2 is specific to GIGYF, we analyzed the interaction of wild-type and mutant 4EHP (i.e. residues involved in the interaction with the auxiliary region of *Hs* GIGYF1/2) with GIGYF2 (endogenous) or overexpressed 4E-BP1 in human cells (the cell-based assays were carried out by Lara Wohlbold). These residues are 4EHP-specific and comprise an arginine and a glutamate residue, which coordinate the auxiliary region of *Hs* GIGYF1/2 and were substituted by leucine residues present at the corresponding position in eIF4E (RE-LL mutant). In this assay, the RE-LL mutant

of 4EHP did not interact with GIGYF2, although it bound to 4E-BP1. Therefore, this specific surface of 4EHP seems to fulfill an important role for the selection of GIGYF proteins over other 4E-BPs that might be present in cells. To further test the role of the auxiliary interactions for the selective binding of 4EHP to GIGYF, we performed competition assays *in vitro* (the competition assays were performed with the support of Sigrun Helms). In this assays, we challenged purified 4EHP-4E-BP1 complexes with *Hs* GIGYF2 peptides that comprised the canonical motif, non-canonical sequences and the auxiliary region and traced the amount of 4EHP-bound 4E-BP1 over time. We compared 4EHP-4E-BP1 complexes in which the 4EHP molecule is either wild-type or carries the RE-LL mutation. In agreement with the pull-down assays, the *Hs* GIGYF2 peptide displaced 4E-BP1 efficiently from wild-type 4EHP-complexes, but was strongly impaired in competing with 4E-BP1 in the context of the RE-LL mutant of 4EHP. Both assays show that 4EHP-GIGYF complexes gain affinity and specificity through a specific interface formed by the auxiliary region of GIGYF proteins and residues that are only present in 4EHP and not in eIF4E. Even if bipartite 4E-BPs, like the human 4E-BP1, can bind to both eIF4E and 4EHP *in vitro*, its binding affinity for 4EHP is lower compared to GIGYF proteins, as it does not contain similar auxiliary binding sites that contact additional surfaces of 4EHP. For that reason, we propose that the additional auxiliary interactions observed in the 4EHP-GIGYF complex provide stability and specificity to this translational repressor complex *in vivo*, where both proteins are less abundant than other competing 4E-BPs.

As observed in the pull-down assays, 4EHP and eIF4E bind to 4E-BP1. However, it is known that the interaction of 4E-BP1 with eIF4E is approximately tenfold stronger than with 4EHP (Rosettani et al., 2007). The molecular reason underlying this difference is not clear. To obtain a molecular insight into the

interactions of 4E-BP1 with 4EHP, we also determined the crystal structure of the 4E-BP1 peptide containing the canonical motif and non-canonical sequences bound to 4EHP (Figure 17).

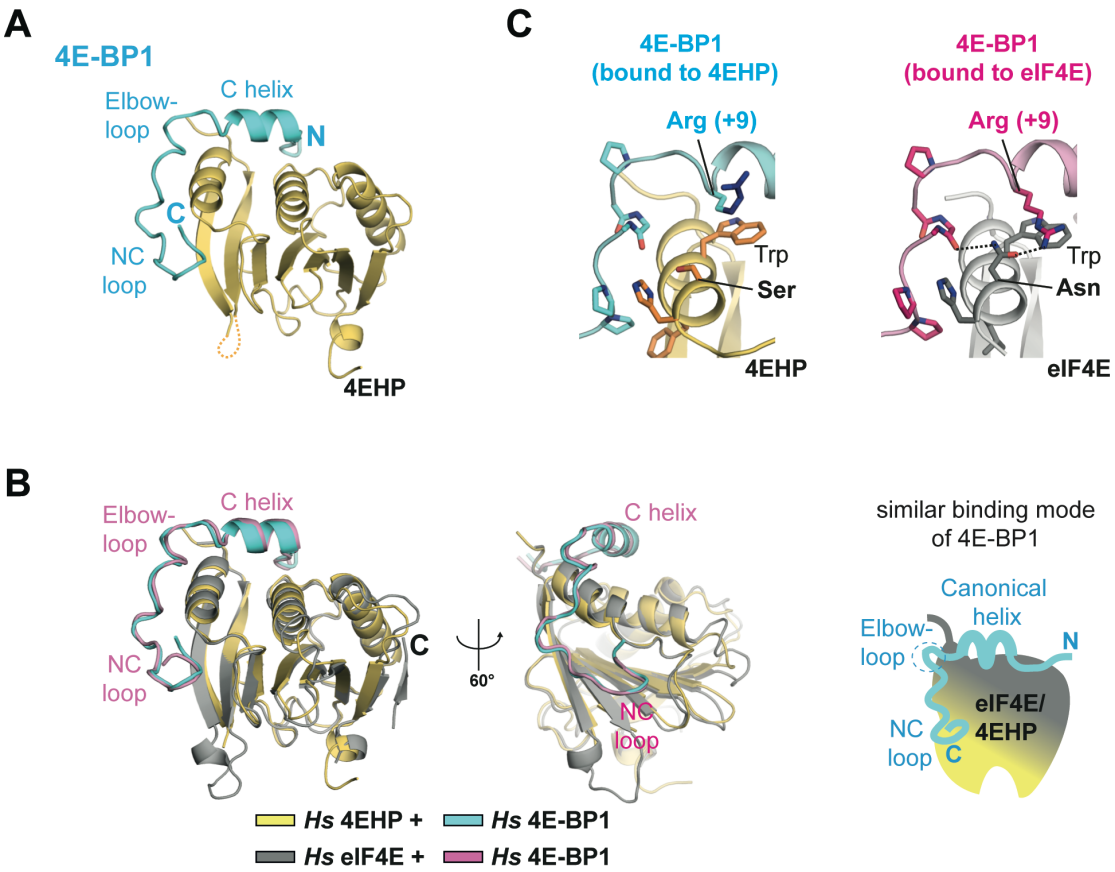


Figure 17: The structure of the human 4E-BP1 bound to 4EHP. (A) Structural overview of the human 4E-BP1 (cyan) bound to 4EHP (yellow). The structural elements of 4E-BP1 (canonical helix, C helix; elbow loop; non-canonical loop, NC loop) are indicated. (B) Superposition of the 4E-BP1 complexes with eIF4E and 4EHP in two views. The 4E-BP1 peptide bound to eIF4E (gray) is shown in pink and its structural elements are indicated. The cartoon on the right depicts the similar binding mode of 4E-BP1 to either eIF4E or 4EHP. (C) Detailed view on the coordination of the linker region of 4E-BP1 between 4EHP and eIF4E. The arginine at position +9 is colored in dark blue from its Cγ atom on because it appeared disordered in the crystal structure of the 4EHP complex. In contrast, in the eIF4E-complex an asparagine residue coordinates the same arginine.

The overall complex is highly similar to the human eIF4E-4E-BP1 complex described above (Figure 17). Nevertheless, a single amino acid substitution between eIF4E and 4EHP within the binding site of 4E-BP1 affects the strength of the interaction. This substitution concerns a conserved asparagine residue in eIF4E, which forms hydrogen bonds bridging and coordinating the canonical 4E-BM and the

linker region of 4E-BP1. Since this residue is replaced by a serine in 4EHP, this coordination is absent and results in missing electron density of an otherwise strictly fixed arginine side chain of 4E-BP1 in the crystal structure. To analyze the effects of this specific position in binding, we generated a mutant of 4EHP in which the serine residue was substituted by asparagine as observed in eIF4E (S-N 4EHP mutant). Interestingly, this single amino acid substitution is sufficient to explain the reduced affinity of 4E-BP1 for 4EHP compared to eIF4E as seen in ITC measurements, since the 4EHP S-N mutant displays a dissociation constant (K_D) similar to eIF4E and tenfold lower than wild-type 4EHP. This result shows that also molecular features in the common dorsal and lateral binding surface of eIF4E proteins contribute to the specificity for some binding partners. Once again, the +9 position of the extended canonical 4E-BM has a prominent role in the interaction with eIF4E-proteins. In this way, 4EHP likely evolved towards a specific complex with GIGYF by generating new binding surfaces for GIGYF1/2 not present in eIF4E on the one hand, and on the other hand by modifying interactions that are otherwise used by 4E-BPs like 4E-BP1.

Another interesting aspect that might contribute to the selective assembly of the 4EHP-GIGYF complexes concerns their oligomeric state in solution. In the crystal environment, the 4EHP-GIGYF1/2 complexes form close contacts to a neighboring complex within the asymmetric unit (ASU), involving a large interface of around 1000 $Å^2$. The interface between the two complexes involves elements of the canonical 4EHP-BM as well as of the auxiliary region of GIGYF1/2. To verify if this close inter-complex interaction *in cristallo* also occurs in solution, we performed small-angle X-ray scattering (SAXS) measurements (Figure 18).

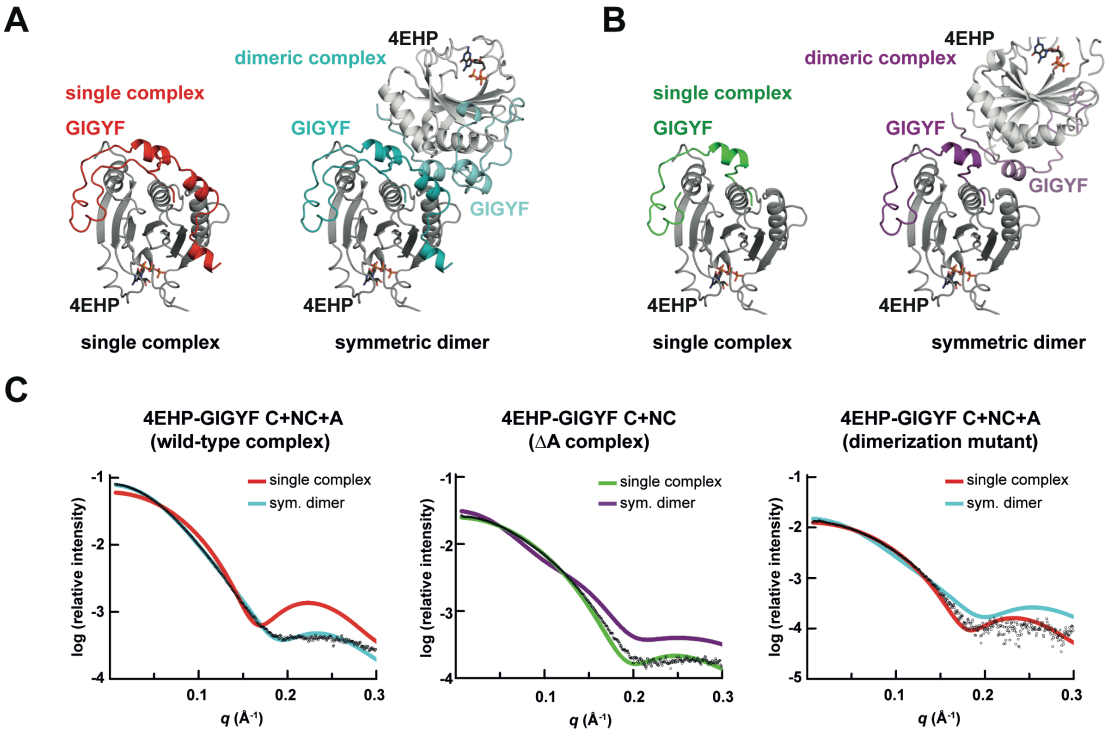


Figure 18: Small-angle X-ray scattering measurements on 4EHP-GIGYF complexes. (A) Crystallographic models of a single 4EHP-GIGYF complex (GIGYF in red) and the symmetric dimer of two 4EHP-GIGYF complexes (GIGYF in cyan) found in the asymmetric unit of the crystal. 4EHP is shown in gray. (B) Crystal structure of a 4EHP-GIGYF2 complex in which GIGYF is lacking the auxiliary region. For the single complex, GIGYF is colored in green and for the symmetric dimer in purple. Note that this complex also crystallized in a similar dimeric arrangement that is slightly distorted compared to the complex containing the auxiliary region. (C) SAXS profiles of different 4EHP-GIGYF complexes including the fits of the respective models. The data was collected from different complexes (wild-type, deletion of the auxiliary region of GIGYF and specific dimerization mutants) as indicated above each profile. The experimental data is shown as dots, the fits of the respective model as a colored line. The colors of the fits are indicated in each graph and correspond to the colors of the GIGYF peptides in the structures shown in panels A and B, which were used for the fitting of the data.

The SAXS data revealed that 4EHP-GIGYF1/2 complexes form indeed dimers of two complexes in solution, as the scattering profile of the 4EHP-GIGYF protein solution could be explained by the symmetric dimer of two complexes observed in the ASU of the crystal, but not by a single complex (Figure 18). The oligomeric state of the 4EHP-GIGYF complexes changes towards a single complex when i) the auxiliary region of GIGYF is absent (ΔA complex), or ii) specific residues of the dimer interface are mutated (Figure 18). The dimerization mutants of 4EHP and the lack of the auxiliary region did not affect the formation of the complex between GIGYF and 4EHP. The latter could also be confirmed by the crystal structure of the 4EHP-GIGYF2 complex that does not contain the auxiliary region of GIGYF2,

showing that the arrangement of the canonical motif and the non-canonical sequences of GIGYF2 are not affected by the lack of the auxiliary region (Figure 18). In summary, the SAXS measurements show that 4EHP-GIGYF complexes have the propensity to form specific, symmetric dimers in solution, which requires the auxiliary region of GIGYF as well as 4EHP-specific residues. However, the biological relevance of this oligomeric arrangement of 4EHP-GIGYF complexes is not clear. It can only be speculated that this dimerization increases the avidity of 4EHP close to the 5'cap structure of an mRNA, enhancing its efficiency to compete with eIF4E for binding and ultimately to repress translation.

Taken together, the structural studies on 4EHP-GIGYF complexes revealed a detailed molecular insight into this 4EHP-selective complex. These structures illustrate the tight and extensive interaction between both partners, provide insights into the specificity of their complex formation and further demonstrate an unprecedented oligomeric arrangement. Nevertheless, it is still unclear why GIGYF proteins do not bind to eIF4E and, *vice versa*, why eIF4G does not bind to 4EHP *in vitro*. Even with the high-resolution structural insight into all the respective complexes, this issue remains unsolved. More studies are necessary to address the kinetics of individual eIF4E- and 4EHP-complexes, also with respect to the individual binding elements of the interacting partners. Due to the similarity of the canonical 4E-BM and canonical 4EHP-BM, the answer to the binding specificity of GIGYF and eIF4G might lie in their non-canonical sequences.

5. DISCUSSION

5.1 A revised molecular model for the repression of translation by 4E-BPs

The conservation of the canonical motif among eIF4E-interacting partners enables 4E-BPs to compete with eIF4G for binding to eIF4E based on a mechanism of molecular mimicry (Marcotrigiano et al., 1999). Several studies indicate that additional, non-conserved contacts are established between eIF4G or 4E-BPs and eIF4E. Among different eIF4E-binding partners, these additional binding elements may exhibit distinct binding contributions (Kinkelin et al., 2012; Lukhele et al., 2013; Paku et al., 2012; Umenaga et al., 2011). However, a detailed structural insight into the eIF4E-complexes with either eIF4G or 4E-BPs was missing when I started my doctoral work. Therefore, the molecular model for the repression of translation by 4E-BPs relied on the binding discrepancies between translational activators and repressors. Solely based on the structural insights of *Dm* CUP bound to eIF4E, this model proposed that a second, lateral binding site on eIF4E is exclusively used by 4E-BPs for docking and displacing bound eIF4G (Kinkelin et al., 2012).

My structural and functional work on 4E-BPs revealed that this lateral surface of eIF4E is a common binding site for different 4E-BPs. Moreover, my data on the *Drosophila* proteins suggests that eIF4G depends less on the lateral side of eIF4E for binding, making the lateral side binding a hallmark for the competitive mechanism of 4E-BPs. Surprisingly however, the follow-up structural work on eIF4E-eIF4G complexes together with my colleague Stefan Grüner revealed that eIF4G proteins also contain non-canonical motifs that bind to the lateral side of eIF4E, similarly to 4E-BPs (Figure 19).

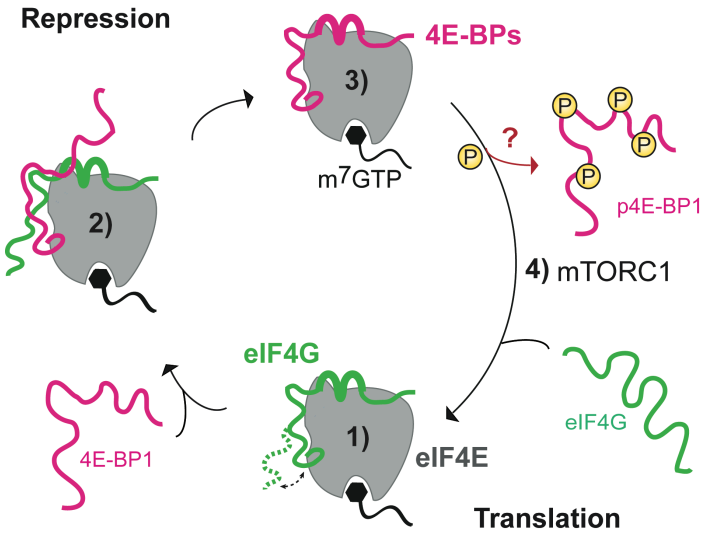


Figure 19: Competition model of 4E-BP-based translational regulation. 1) eIF4G (green) is bound to eIF4E (gray) which leads to the activation of cap-dependent translation. However, the lateral interaction site with eIF4E contributes less and might be more dynamic in the case of eIF4G. 2) Non- or partially phosphorylated 4E-BPs (magenta) might capture a free lateral side of eIF4E and start to dissociate eIF4G. 3) Fully bound 4E-BPs repress translation initiation by preventing the assembly of the eIF4F complex on the 5'cap of an mRNA. 4) Conversely, 4E-BPs get regulated by the mTOR signaling pathway and cannot bind to eIF4E when fully phosphorylated (p4E-BP1), leading to the activation of translation by eIF4G. Yet, it is not completely understood how eIF4E-4E-BP complexes are regulated to release eIF4E to participate in translation.

Despite the similarity in the binding mode between 4E-BPs and eIF4G, our binding studies and competition assays indicated that the sequence composition of the non-canonical motif and the N-terminal connecting linker region is the determinant for the lateral side dependence of eIF4G and 4E-BPs. The fact that the binding of eIF4G to the lateral side of eIF4E contributes much less to its overall affinity and that 4E-BPs strongly require the lateral side of eIF4E to compete with eIF4G supported the idea that the non-canonical sequences of eIF4G might be more dynamic in the complex with eIF4E compared to 4E-BPs. Regarding the competition model of 4E-BPs with eIF4E-eIF4G complexes, we could therefore speculate that 4E-BPs occupy a transiently free lateral side of eIF4E and start to dissociate eIF4G from the complex to achieve translational repression (Figure 19). However, more studies regarding the dynamics of the unbound as well as the eIF4E-bound state of 4E-BPs and eIF4G, or

even the capture of a putative intermediate trimeric eIF4E-4E-BP-eIF4G state are required to validate this.

Another aspect that indicates differences in the lateral side binding between 4E-BPs and eIF4G is provided by studies on the small molecule inhibitor 4EGI-1, which targets the eIF4E-eIF4G interaction (Moerke et al., 2007). A crystal structure of this compound bound to eIF4E revealed that it also binds to the lateral side of eIF4E (Papadopoulos et al., 2014). Moreover, a structural comparison shows that the binding of 4EGI-1 directly overlaps with the non-canonical motifs of eIF4G and 4E-BPs at the lateral side of eIF4E, and thus might be mutually exclusive for both. Strikingly, 4EGI-1 seems to specifically affect the eIF4E-eIF4G interaction and was reported to be compatible with 4E-BP1 binding (Moerke et al., 2007; Papadopoulos et al., 2014; Sekiyama et al., 2015). In the light of our results, the stronger lateral side binding of 4E-BPs might not allow the inhibitor to fully access its binding site and therefore 4E-BP1 remains bound to eIF4E. Accordingly, the weaker lateral side binding of eIF4G might be competed out by 4EGI-1, enhancing the dissociation of eIF4G from eIF4E or blocking its stable association. Thus, these results agree with our structural and binding studies, and indicate that 4EGI-1 also utilizes the differential contribution in the lateral side binding of eIF4G and 4E-BPs to block translation initiation. Moreover, our data provides an alternative and complementary explanation to the proposed allosteric effect of this small molecule compound. In our model, 4EGI-1 directly competes with eIF4G for binding to the lateral surface of eIF4E, and upon binding it causes allosteric changes that ultimately lead to the dissociation of the eIF4E-eIF4G complex (Papadopoulos et al., 2014; Salvi et al., 2016; Sekiyama et al., 2015).

5.2 Regulation of eIF4E-4E-BP complexes

Although we have expanded our understanding on the molecular mechanism underlying the repression of translation by 4E-BPs, there are still many open questions concerning the regulation of eIF4E-4E-BP complexes. Indeed, 4E-BPs are regulated by phosphorylation through the mTORC1 complex, and hyperphosphorylated 4E-BPs cannot associate with eIF4E (Bah et al., 2015; Gingras et al., 1999a; Gingras et al., 2001; Miron et al., 2003). In contrast, hypophosphorylated 4E-BPs bind to eIF4E and might occupy a part of the cellular eIF4E-pool as a mean to control the availability of the cap-binding protein and fine-tune global translational levels. In fact, a reduction of the cellular eIF4E levels does not impact on global protein synthesis in normal cellular conditions and even prevented the oncogenic transformation of cells (Truitt et al., 2015). Therefore, eIF4E is likely present in excess in cells and the pool of eIF4E available for translation is controlled by 4E-BPs depending on their phosphorylation status. The phosphorylation status of 4E-BPs also regulates their abundance, since hypophosphorylated 4E-BPs that are not bound to eIF4E are ubiquitinated and targeted for proteasomal degradation (Yanagiya et al., 2012). This regulatory mechanism adds another layer of translational control, since the activation of 4E-BPs must occur from their hyperphosphorylated or their eIF4E-bound pool.

However, it remains unclear how eIF4E is released from 4E-BPs to participate in translation. A piece of this puzzle is laid with the cryo-electron microscopy structure of the mTORC1 complex, which provided a comprehensive insight into the architecture of this machinery, including its specific adaptor protein Raptor that recognizes 4E-BPs via binding to their C-terminal TOR signaling (TOS) motif (Aylett et al., 2016; Schalm et al., 2003). The structural insight into this complex suggests

that 4E-BPs are likely phosphorylated in their extended conformation, and not when they are bound to eIF4E as already proposed in previous studies (Youtani et al., 2000). This might specifically concern the phosphorylation sites known to be modified only at the end of the phosphorylation cascade and which become inaccessible in the eIF4E-bound form [see Figure 5; (Gingras et al., 1999a; Gingras et al., 2001)]. Therefore, to increase the pool of eIF4E available for translation, it is necessary to dissociate the eIF4E-4E-BP complexes. This would enable the mTORC1 complex to fully phosphorylate 4E-BPs in order to prevent re-association with eIF4E. To assess the dissociation of eIF4E-4E-BP complexes, however, more quantitative kinetic studies are required that also consider the different phosphorylation states of 4E-BPs. One possibility is that the activation of the mTORC1 complex facilitates the dissociation of the eIF4E-complex by binding to 4E-BPs, which could affect their interaction with eIF4E. Another possibility would be that hypophosphorylated 4E-BPs (i.e. modified only at the initial phosphorylation sites; see Figure 5) display an increased dissociation rate from eIF4E compared to their non-phosphorylated form, which would increase mTORC1 accessibility and therefore the likelihood of being phosphorylated. Indeed, our data on partially phosphorylated 4E-BPs suggests that phosphorylation affects 4E-BPs in their ability to compete with eIF4G when it is already bound to eIF4E. Conversely, hypophosphorylated 4E-BPs might also be easier displaced from eIF4E by other factors like eIF4G. Therefore, understanding the molecular mechanism underlying the regulation of 4E-BPs and the fine-tuning of cellular eIF4E levels requires a deeper insight into the kinetics of these complexes and their turnover *in vivo*.

5.3 The eIF4E-eIF4G interaction as a target for therapeutic applications in cancer

A common perturbation in human cancers concerns malfunctions in the control of translation. These malfunctions ultimately increase overall protein synthesis, which in turn facilitates cancer progression and resistance to cancer treatments (Grzmil and Hemmings, 2012). Therefore, the translation machinery and its regulators are major therapeutic targets in cancer treatment, and a variety of inhibitors for almost every component involved in translation have been characterized (an overview can be found in Bhat et al., 2015).

Among the translational components, the eIF4F complex offers some opportunities for therapeutic treatment (Steinberger et al., 2016). For example, efforts have been made to target the eIF4E-eIF4G interaction with small molecule inhibitors, and the 4EGI-1 compound discussed above is only one example. In my doctoral thesis, I used a peptide-based approach to target the eIF4E-eIF4G interaction. Our structural and functional studies on 4E-BPs resulted in a fusion peptide with elements from different 4E-BPs (see Figure 11). We could show that this chimeric 4E-BP displayed improved functional properties as a competitor of eIF4G and as an efficient translational repressor. These results exemplify the feasibility of 4E-BP based peptides as tools for therapeutic applications. The structural details provided by our research can be used as a template for structure-guided drug design, a strategy to improve existing compounds or new lead compounds in terms of affinity and selectivity (Valkov et al., 2012). In fact, the direct use of small peptides as pharmacological drugs has been evaluated, as there is a significant interest in targeting the eIF4E-eIF4G interaction (Fosgerau and Hoffmann, 2015; Zhou et al., 2012). Alternatively, 4E-BP derived peptides have also been applied in cell based

assays as a tool to suppress cancer cell growth. For instance, fusion of a 4E-BP like peptide to an analog of the gonadotropin-releasing hormone (GnRH) reduced the tumor growth in ovarian tumor xenograft mouse models without toxic effects (Ko et al., 2009). Since the majority of epithelial ovarian cancer types express GnRH receptor (GnRH-RI), the 4E-BP like peptide was taken up into GnRH-RI expressing tumor cells (Ko et al., 2009). Alternative delivery strategies can also be used to target other types of tumors. An interesting approach utilizes the pH-induced transmembrane structure (pHLIP), a peptide that folds into transmembrane helix in low pH conditions. As the extracellular environment of solid tumors is usually acidic, the folding of the pHLIP peptide will be induced and will transport a fused macromolecule across the cellular barrier of such tissues (Cheng et al., 2015). Considering the various approaches, our research might provide a framework for targeting the eIF4E-eIF4G interaction in a multitude of biological and disease-relevant contexts.

5.4 Regulation of specific eIF4E-4E-BP complexes: the eIF4E-Mextli complex

Binding to eIF4E by eIF4E-binding partners other than eIF4G is generally thought to repress translation in a global or message specific manner, depending on the type of 4E-BP that is involved. *Drosophila* Mextli, however, seems to be an exception from this general assumption, since it is the only 4E-BP that was reported to promote translation besides eIF4G (Hernandez et al., 2013). Its modular domain architecture and interaction network are functionally mimicking eIF4G, however we could also show that the eIF4E-binding mode of *Dm* Mextli involving canonical, non-canonical and auxiliary helices is unique among 4E-BPs.

In the *Dm* eIF4E-Mextli complex, the accommodation of the auxiliary helix on the dorsal surface of eIF4E requires specific amino acid substitutions in the canonical motif. These are two arginine residues, which are present in the canonical motif of *Ce* Mextli, but which are replaced by residues with smaller side chains (i.e. a serine and a isoleucine) in *Dm* Mextli. These substitutions are required in the case of *Dm* Mextli to avoid clashes with the auxiliary helix, which in turn compensates for the missing interactions of the arginine side chains. Furthermore, the incompatibility between the long arginine side chains and the auxiliary helix of *Dm* Mextli implies that other previously studied 4E-BPs cannot form similar auxiliary arrangements.

Together with the functional studies on the tripartite and bipartite binding modes of the Mextli proteins, the structural insights again underpin the relevance of the long aliphatic side chains that take similar positions in our previously defined extended canonical eIF4E-binding motif (YX[**R/K**]X_2LΦX_2[**R/K**]). The extended canonical motif is present in all bipartite 4E-BPs known so far and in eIF4G (where +9 position can also be a glutamine in some species instead), and in all cases the role of these long aliphatic side chains is to cover a large part of the dorsal surface of

eIF4E. Therefore, a more detailed knowledge on key interactions of 4E-BPs might allow a predictive assumption on the binding mode of yet uncharacterized 4E-BPs, and might lead to the improvement of motif-based searches to identify novel 4E-BPs *in silico* by reducing false-positive hits.

Moreover, the unique tripartite eIF4E-binding mode influences the functional properties of *Dm* Mextli compared to bipartite 4E-BPs. Namely, the tripartite assembly of *Dm* Mextli with eIF4E guarantees the formation of a very stable complex that is resistant to the competition by bipartite 4E-BPs (Figure 20). This exceptional feature might allow *Dm* Mextli to fulfill its specific functions during embryogenesis and germline stem cell maintenance in *Drosophila* in the presence of other 4E-BPs. However, this raises the question of how the *Dm* eIF4E-Mextli complexes are regulated and why the binding of Mextli to eIF4E differs in worms like *Ce*.

One possible scenario concerns the population of 4E-BPs and eIF4E in worms. In fact, worms do not contain many 4E-BPs. Besides Mextli, only IFET-1 (a homologue to 4E-T) has been described as a 4E-BP in worms (Sengupta et al., 2013). In contrast to the low number of 4E-BPs, worms contain different eIF4E isoforms as well as mRNAs with different cap-structures as a result of trans-splicing events (Keiper et al., 2000; Van Doren and Hirsh, 1990). Since some eIF4E isoforms were found to be specific for a certain mRNA cap, they are related to specific translational programs (Dinkova et al., 2005; Keiper et al., 2000; Miyoshi et al., 2002). This could reduce the necessity to compete for eIF4E-binding in worms, since there are already specialized types of eIF4E rather than a common pool. Nevertheless, which 4E-BPs exist in worms like *Ce*, whether they exhibit specificity for a particular eIF4E isoform, or if there are other regulatory mechanisms involved is so far unknown.

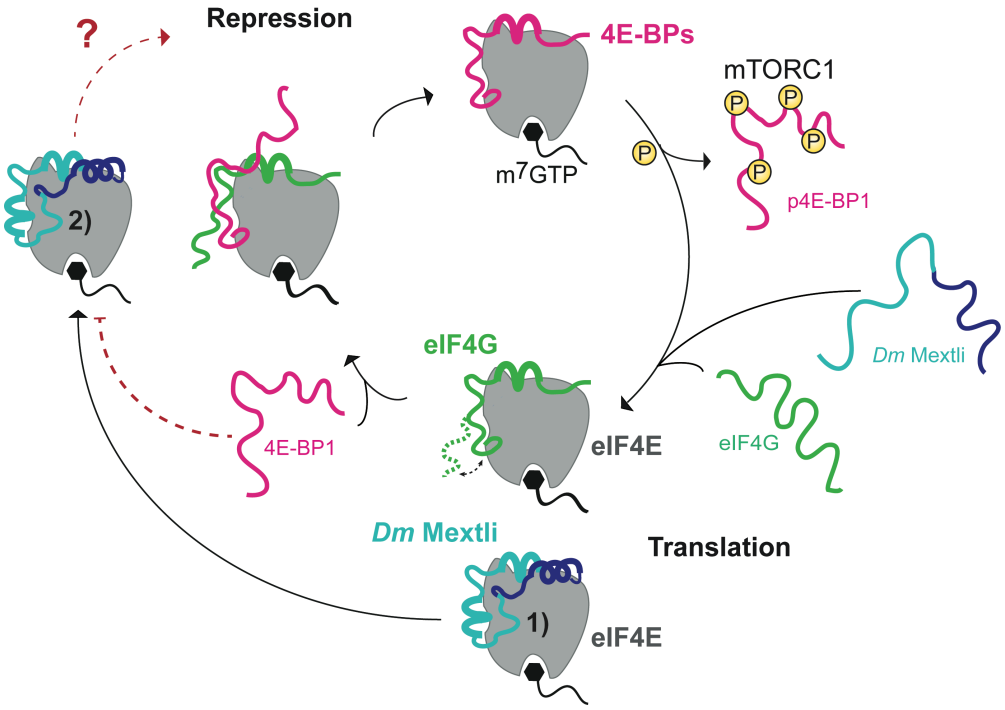


Figure 20: The regulatory mechanism controlling the binding of *Dm* Mextli to eIF4E remains unclear. The figure relates to the 4E-BP based competition model shown in **Figure 19**. *Dm* Mextli (cyan and dark blue) is added to this model based on the insights generated in this work. 1) Under certain cellular conditions or developmental processes in *Drosophila*, Mextli is able to bind to eIF4E and activates translation by functionally mimicking eIF4G. 2) However, the regulation of the *Dm* eIF4E-Mextli complex is unclear, since bipartite 4E-BPs are not able to displace Mextli.

Several eIF4E proteins also exist in *Drosophila*, and *Dm* Mextli does not bind to all of the eight eIF4E paralogs present in flies (Hernandez et al., 1997; Hernandez et al., 2013). It is not known whether *Dm* Mextli uses a specific eIF4E paralog to initiate translation, and the cellular conditions in which *Dm* Mextli promotes translation are unclear. *Dm* Mextli might be mRNA-specific and only promote specific translational programs in addition to the eIF4G-mediated global translation in the cell. Additionally, the unique tripartite eIF4E-binding mode of *Dm* Mextli also allows it to sustain protein synthesis under conditions where 4E-BPs are hyperactive and shut down eIF4G-dependent translation. However, the resistance against the regulation by 4E-BPs implies that an alternative mechanism controls *Dm* eIF4E-Mextli complexes (Figure 20). One possibility could be the regulation by phosphorylation similar to the global translational repressor 4E-BPs. Although one of

the phosphorylation sites located within the eIF4E-binding region of *Hs* 4E-BP1-3 (described above) is also present in *Dm* Mextli, there is no evidence that Mextli proteins are regulated by phosphorylation. Unlike the human 4E-BPs 1-3, *Dm* Mextli does neither contain a TOS motif that marks it as a substrate for the adaptor protein Raptor of the mTORC1 complex, nor does it contain potential phosphorylation sites that are suitable for a serine/threonine-specific protein kinase. Still, phosphoregulation of *Dm* eIF4E-Mextli complexes via a different pathway cannot be excluded, but more research is required to reveal which translational programs this complex controls and how it is regulated *in vivo*.

5.5 The specificity of GIGYF for 4EHP gives rise to a unique translational repressor complex that competes with eIF4E

The second eIF4E-family member, 4EHP, offers an alternative mechanism of translational regulation in addition to the binding of 4E-BPs to eIF4E. 4EHP is able to bind to the 5'mRNA cap but not to eIF4G, wherefore it regulates the translation of specific mRNAs (Cho et al., 2006; Cho et al., 2005; Fu et al., 2016; Villaescusa et al., 2009). The GIGYF proteins 1 and 2 are specific interaction partners of 4EHP (Morita et al., 2012). Interestingly, GIGYF1/2 proteins only bind to 4EHP and not to eIF4E (Rom et al., 1998). The molecular determinants required for the specificity of binding to eIF4E- and 4EHP were unclear, especially due to the similarity between the eIF4E- and 4EHP-binding motifs between the binding partners (Marcotrigiano et al., 1999; Morita et al., 2012; Rosettani et al., 2007).

My work on the 4EHP-GIGYF translational repressor complexes provided the first structural insight into a 4EHP-specific protein bound to 4EHP. With the molecular details we could learn that 4EHP binds preferentially to GIGYF1/2 and not to bipartite 4E-BPs. This specificity of 4EHP is provided through the recognition of 4EHP-specific surfaces by the auxiliary sequences of GIGYF1/2, which are absent in 4E-BPs. Therefore, the GIGYF-specific interaction sites of 4EHP might enable the specific formation of the 4EHP-GIGYF translational repressor complex *in vivo*, especially since other bipartite 4E-BPs like 4E-BP1 are present in much higher amount in the cell than GIGYF1/2 and could therefore occupy 4EHP and prevent it from being recruited to its specific targets.

However, the structural similarity of the canonical and non-canonical 4EHP-binding motifs of GIGYF1/2 with eIF4E-4E-BP complexes makes it difficult to rationalize why eIF4G does not bind to 4EHP and, *vice versa*, why GIGYF1/2 proteins do not bind to eIF4E. Specifically, the almost identical canonical binding motifs of GIGYF1/2 and eIF4G must confer binding to either eIF4E-protein. Indeed, GIGYF1/2 binds weakly to eIF4E *in vitro*. This also implies that the non-canonical binding regions play a role in the specific recognition of the correct eIF4E-protein. In agreement with this, a single amino acid substitution between eIF4E and 4EHP within their otherwise conserved binding region significantly affects the binding affinity to 4E-BP1. Thus, even small differences in the binding interfaces in either binding partner might affect their relative affinities towards eIF4E- or 4EHP *in vivo*. Nonetheless, it remains unclear whether the non-canonical motifs are the sole determinants of the exclusive binding of the GIGYF1/2 proteins and eIF4G to 4EHP and eIF4E, respectively. In this case it would be of interest to determine the

contributions of specific residues, which likely requires further extensive mutational approaches in combination with quantitative binding studies.

RNA-binding proteins (RBPs) are known to recruit the 4EHP-GIGYF translational repressor complex to specific mRNAs to repress translation (Fu et al., 2016; Morita et al., 2012; Tao and Gao, 2015). Since 4EHP has an approximately 200-fold lower affinity for an m^7GTP cap analog than eIF4E, its tethering to an mRNA via GIGYF and a recruiting RBP might increase its ability to compete with eIF4E for the 5' mRNA cap (Rosettani et al., 2007). In agreement with this model, my structural work provides insights on how 4EHP might replace eIF4E on the 5' mRNA cap.

First, the close proximity of the C-terminal auxiliary sequences of GIGYF to the cap-binding pocket of 4EHP in our structures could point to a role of the subsequent arginine/glycine-rich sequences of GIGYF in RNA-binding. Binding of GIGYF to RNA might stabilize 4EHP in close distance to the mRNA cap and therefore increase its local concentration to compete with eIF4E. However, arginine/glycine-rich low complexity regions usually bind non-specifically to RNAs, which brings several drawbacks to this idea (Thandapani et al., 2013). Non-specific RNA-binding does not guarantee the tethering of 4EHP in the 5'UTR of the mRNA close to the cap, as it potentially can allow binding to any sequence of the mRNA. Therefore, it is not clear how this RNA-binding capability could ultimately result in an advantage for 4EHP to compete with eIF4E present at the mRNA cap. Such an advantage, however, could be provided by yet unknown factors that facilitate the recognition of specific sequences in the mRNA. Alternatively, it seems more likely that binding of 4EHP to the cap is the initial step of the repression mechanism that subsequently locates the arginine/glycine-rich sequences in proximity of the 5' UTR

of the mRNA. Binding of GIGYF to the mRNA would then support the interaction of 4EHP with the cap structure of the mRNA. However, it remains to be determined if the arginine/glycine-rich sequences adjacent to the 4EHP-binding region of GIGYF indeed bind to RNA.

Second, the specific dimeric arrangement of the 4EHP-GIGYF complexes in solution might increase the binding of 4EHP on the target mRNA through avidity effects. Additional factors might also contribute to the avidity effect of the 4EHP-GIGYF complex. For instance, the C-terminus of GIGYF contains a predicted coiled-coil region, which might facilitate the multimerization of GIGYF proteins. However, the oligomeric state of GIGYF proteins *in vivo* is not known and needs to be further analyzed. Moreover, target mRNAs might contain several repeats of the recognition motif for a specific RBP, which in turn recruits the 4EHP-GIGYF complex. For example, the AU-rich binding sequence of TTP is repeated multiple times within the ARE (AU-rich element) of the target mRNAs, as in the tumor necrosis factor (TNF) mRNA (Brooks and Blackshear, 2013). Therefore, the cooperation of all of those factors could result in binding of multiple 4EHP-GIGYF complexes on target mRNAs, which would strongly increase the probability of 4EHP to occupy the mRNA cap.

Regarding the tight interaction between 4EHP and GIGYF proteins, an open question concerns the regulation of this complex. Interestingly, mouse 4EHP knockout cells have reduced expression of GIGYF, indicating that 4EHP and GIGYF are interdependent in terms of stability (Morita et al., 2012). So far, the molecular pathway involved in the control of the stability of 4EHP and GIGYF is unknown. Binding of GIGYF to 4EHP could preclude the exposure of potential sequences that target GIGYF for degradation, for instance via the proteasome. Therefore, the

expression levels of 4EHP might dictate the relative abundance of the 4EHP-GIGYF complex. Furthermore, it remains unclear if the interaction between GIGYF and 4EHP itself is regulated. Based on the phosphoregulation of the global translational repressor 4E-BPs, phosphorylation of GIGYF residues could also play a role in this mechanism. Indeed, several potential tyrosine phosphorylation sites have been suggested in mouse GIGYF proteins (Giovannone et al., 2003). Additionally, proteomic databases on posttranslational modifications indicate several serine and threonine phosphorylation sites. Although none of those sites are located within the 4EHP-binding region of GIGYF, two serine residues are located just N-terminal of the canonical 4EHP-binding motif of GIGYF proteins. These two sites might play a similar regulatory role as the N-terminal phosphorylation sites in 4E-BPs, which have been shown to induce a folding transition when phosphorylated that masks the canonical eIF4E-binding motif (Bah et al., 2015). Nevertheless, more binding assays as well as studies of protein dynamics *in vitro* are required to obtain a mechanistic understanding of those phosphorylation sites. The regulatory pathway that controls the activity of GIGYF proteins *in vivo* is also unclear and could be the object of future studies.

6. CONCLUSION

4E-BPs are translational repressors that act on the initiation step of translation by interfering with the formation of the eIF4F complex on the 5'cap structure of an mRNA. Mechanistically, this repression is based on a direct competition with the translation initiation factor eIF4G for binding to eIF4E. Since a variety of proteins in the cell follow this principle, 4E-BPs comprise a broad range of factors that are involved in different cellular functions. Therefore, translational repression can act in a global manner as a response to environmental cues, or at the level of specific mRNAs. At the beginning of my doctoral work, the underlying molecular details of the competitive binding mechanism were incomplete. For this reason, my doctoral work initially focused on the interaction of different 4E-BPs and eIF4G with eIF4E. My biophysical and structural studies on 4E-BPs elucidated common binding principles for different 4E-BPs (4E-T, Thor and 4E-BP1) to eIF4E and showed that the key interactions are conserved between 4E-BPs and across species. Moreover, my studies provided the first complete structural insight into the binding of the global translational repressor 4E-BP1 to eIF4E and rationalized previous studies on this factor that indicated additional eIF4E-binding regions (Lukhele et al., 2013; Paku et al., 2012). To complement our insights into the competitive binding mechanism, I initiated and participated in a collaborative project that aimed at understanding the binding mode of eIF4E-eIF4G complexes, since eIF4G was also reported to contain additional eIF4E-binding sequences (Umenaga et al., 2011). In this structural study, we visualized the extended interactions between eIF4G and eIF4E and revealed that the additional eIF4E-binding occurs on the same surfaces of eIF4E contacted by 4E-BPs. Therefore, the competition for eIF4E-binding is based on a complete molecular mimicry, beyond the conserved canonical eIF4E-binding motif. Moreover,

we could determine that the amino acid composition of the non-conserved additional eIF4E-binding regions of these proteins determine the ability of 4E-BPs to compete with eIF4G. Taken together, these studies provide a fundamental structural insight into the regulation of translation, which is of high relevance in therapeutic approaches to treat malignancies in which eIF4E activity is upregulated.

In another project of my PhD, I was working on a highly specialized invertebrate-specific 4E-BP called Mextli, which was reported to act as a translational activator during certain developmental stages and cell types in *Drosophila melanogaster* (Hernandez et al., 2013). I used a structural approach to compare Mextli proteins from *Dm* and *Ce*, and with this I revealed striking differences in their eIF4E-binding mode. *Ce* Mextli follows the bipartite binding mode similar to other 4E-BPs, while *Dm* Mextli evolved a complex, tripartite binding mode for eIF4E. Distinct structural differences to other 4E-BPs allow this tripartite arrangement of *Dm* Mextli, which confers specific functional properties to the complex not observed in any other eIF4E-4E-BP complex. By comparing the different binding modes of the *Ce* and *Dm* Mextli proteins, we could show that the *Dm* eIF4E-Mextli complex is very resistant to competition by bipartite 4E-BPs. This distinct functional property might enable *Dm* Mextli to fulfill its specific cellular function in the presence of other 4E-BPs. Therefore, *Dm* Mextli is the first example of a 4E-BP extending the bipartite binding mode. So far, there are no other examples of 4E-BPs that bind to eIF4E in a similar manner. However, the structural characteristics we could describe for *Dm* Mextli might help to identify similar binding modes and at the same time exclude that previously described 4E-BPs bind to eIF4E using a tripartite binding region.

In the last part of my doctoral work I focused on 4EHP translational repressor complexes. In contrast to eIF4E, 4EHP cannot bind to eIF4G and therefore acts as a translational repressor. Conversely, the GIGYF proteins only bind to 4EHP but not to eIF4E. In both cases, the molecular determinants for the discrimination between the eIF4E proteins are unclear, since the conserved canonical binding motif is highly similar between eIF4G and GIGYF proteins. My structural studies on 4EHP-GIGYF complexes demonstrated that GIGYF proteins gain affinity and achieve specificity because they contain additional auxiliary binding elements that contact 4EHP-specific surfaces. These binding sites on 4EHP are not used by 4E-BPs that also bind to eIF4E, and therefore provide the selectivity for the interaction with GIGYF. This unique binding mode might be especially important *in vivo*, where the low abundant 4EHP-GIGYF complex is facing an excess of other 4E-BPs. Nevertheless, due to the similarity in the overlapping bipartite binding interface between eIF4E- and 4EHP-complexes, it is difficult to rationalize why proteins like eIF4G and GIGYF are not able to bind to either eIF4E-protein. Future work to address this question might require comprehensive mutational and kinetic studies on the binding properties of those proteins.

In conclusion, my doctoral work revealed a common structural binding mode for 4E-BPs binding to eIF4E proteins. With the Mextli proteins and the 4EHP-GIGYF complexes I further showed specific exceptions, which extend the conserved bipartite binding to establish novel functional properties of this interaction. Therefore, my results contributed to the understanding of translational regulation by 4E-BPs, elucidated functional properties of eIF4E-complexes and provided insights into the selectivity and specificity for eIF4E-proteins (Figure 21).

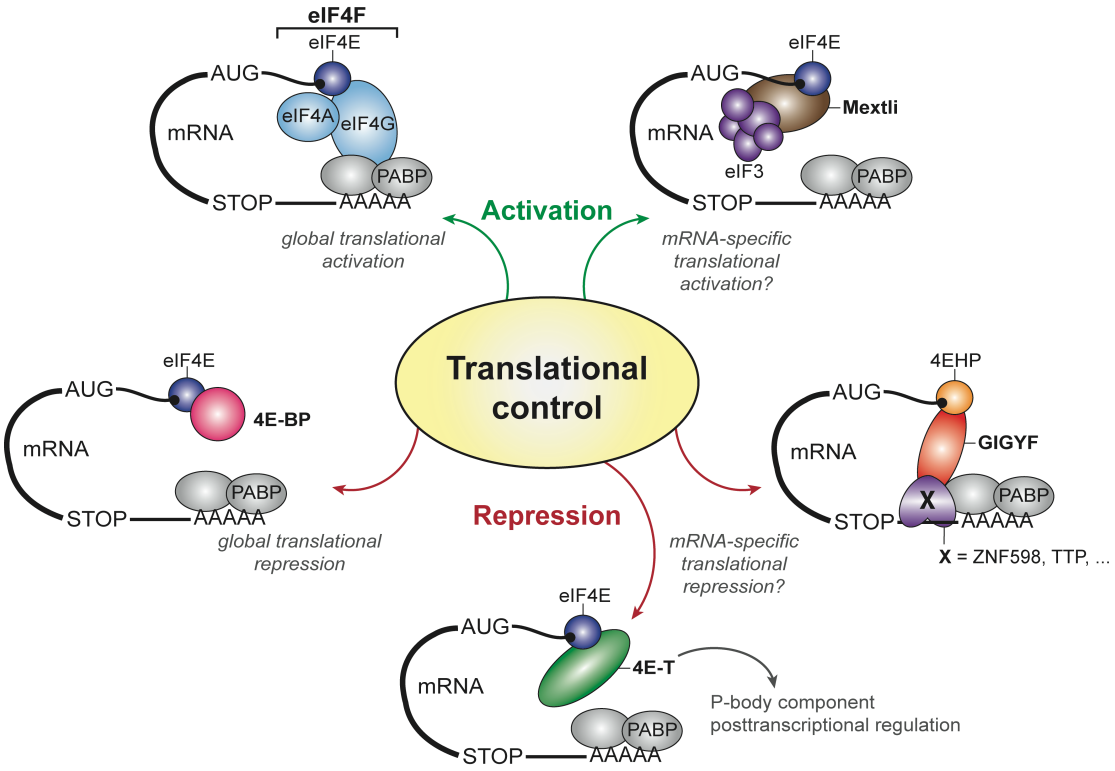


Figure 21: Translational control by eIF4E- and 4EHP-complexes. Different factors influence mRNA translation either on a global level or in a message specific way. eIF4G (light blue) and 4E-BPs (magenta) bound to eIF4E (dark blue) act as global translational activators and repressors, respectively. mRNA-specific eIF4E-complexes might be formed for example by Mextli proteins (brown) or 4E-T (green). Alternatively, the 4EHP-GIGYF (orange and red, respectively) complex is likely recruited to specific targets by different RNA-binding proteins (purple, labeled as 'X'). However, open questions remain about the translational programs that are controlled by those mRNA-specific complexes and about the mechanisms that are involved to control mRNA expression.

eIF4E- and eIF4E-related complexes are involved in diverse translational programs that act either globally or in a message-specific manner. My work contributed to the field of translational control by shedding light on structural and functional similarities and differences between these complexes (Figure 21). Many open questions remain regarding the cellular function of these complexes. To obtain a broader understanding about translation, future work should address whether there are specific targets that are regulated by certain 4E-BPs, and which additional factors are involved in their function (Figure 21).

7. AUTHOR CONTRIBUTIONS

This section describes my contribution to the papers discussed in this thesis.

1) **4E-BPs require non-canonical 4E-binding motifs and a lateral surface of eIF4E to repress translation**
   Igreja, C., Peter, D., Weiler, C., and Izaurralde, E. (2014)
   Nature Communications **5**, 4790

In this work, I performed the biophysical analysis of the interaction between 4E-BPs and eIF4E. I designed and cloned some of the constructs used in this study. Further I purified the proteins used in the *in vitro* competition assays (4E-BP peptides and eIF4E-eIF4G complexes) and performed some of the *in vitro* pulldown assays and competition assays myself. I conducted all of the ITC measurements and did the data analysis. Further I contributed to the preparation of the manuscript.

2) **Molecular Architecture of 4E-BP Translational Inhibitors Bound to eIF4E**
   Peter, D., Igreja, C., Weber, R., Wohlbold, L., Weiler, C., Ebertsch, L., Weichenrieder, O., and Izaurralde, E. (2015)
   Molecular Cell **57**(6), 1074-1087

My contribution to this project concerned the structural characterization of eIF4E-4E-BP complexes. I designed and cloned all of the constructs used in the *in vitro* experiments and crystallization, including the 4E-BP chimera. I expressed and purified all of the proteins that were used in pulldown assays, competition assays, ITC measurements and for crystallization trials. I performed some of the pulldown and competition assays. I conducted all of the ITC measurements and analyzed the data. I crystallized all of the eIF4E-complexes presented in this study. I determined and analyzed all of the crystal structures. I further contributed to the preparation of the manuscript.

3) Mextli proteins use both canonical bipartite and novel tripartite binding modes to form eIF4E complexes that display differential sensitivity to 4E-BP regulation

Peter, D., Weber, R., Köne, C., Chung, M.Y., Ebertsch, L., Truffault, V., Weichenrieder, O., Igreja, C., and Izaurralde, E. (2015)
Genes & Development **29**, 1835-1849

My major contribution to this study was the structural and biophysical characterization of eIF4E-Mextli complexes. I designed all of the constructs used in the *in vitro* experiments. I cloned many of the constructs used for protein expression in *E. coli* with the help of Carolin Köne. I expressed and purified proteins that were used in competition assays, ITC measurements, NMR measurements and for crystallization trials. I supervised Carolin Köne performing the expression and purification of some of the proteins. I performed some of the pulldown and competition assays together with Carolin Köne and Ramona Weber. I conducted all of the ITC measurements myself or supervised the measurements done by Ramona Weber. I analyzed the ITC data. I crystallized all of the eIF4E-Mextli complexes and determined the structures together with Carolin Köne under my supervision. I analyzed the crystal structures. Together with Vincent Truffault, I performed the NMR measurements and analyzed the data. I contributed to the writing of the manuscript and the preparation of the figures.

4) The Structures of eIF4E-eIF4G Complexes Reveal an Extended Interface to Regulate Translation Initiation

Grüner, S., Peter, D., Weber, R., Wohlbold, L., Chung, M.Y., Weichenrieder, O., Valkov, E., Igreja, C., and Izaurralde, E. (2016)
Molecular Cell **64**, 467-479

I initiated the project, designed and cloned some of the initial constructs. I further supported the project with the expression and purification of some of the proteins used in competition assays (*D. melanogaster* Thor peptide and different eIF4E-eIF4G complexes). I performed some of the pulldown assays and competitions assays. I contributed to the analysis of the structures and to the preparation of manuscript.

5) GIGYF1/2 proteins use auxiliary sequences to selectively bind to 4EHP and repress target mRNA expression

Peter, D., Weber, R., Sandmeir, F., Wohlbold, L., Helms, S., Bawankar, P., Valkov, E., Igreja, C., and Izaurralde, E. (2017)

Genes & Development accepted manuscript, *in press*

I contributed to this project with the structural and biophysical studies of 4EHP-GIGYF1/2 complexes. I designed and cloned the constructs used for *in vitro* experiments and crystallization trials. I expressed and purified the proteins used for pulldown assays, competition assays, ITC measurements, SAXS measurements and crystallization trials. I performed some of the pulldown assays and competition assays. I conducted all of the ITC measurements and the analysis. I crystallized all of the 4EHP-GIGYF complexes presented in the study and solved the structures. I supervised Felix Sandmeir in the expression, purification, crystallization and structure determination of the 4EHP-4E-BP1 complex. I analyzed all of the structures. I performed the SAXS experiments with the help of Eugene Valkov and analyzed the data. I also contributed to the writing of the manuscript and the preparation of the figures.

8. REFERENCES

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. J Biol Chem *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. J Mol Biol *385*, 491-506.

Altmann, M., Schmitz, N., Berset, C., and Trachsel, H. (1997). A novel inhibitor of cap-dependent translation initiation in yeast: p20 competes with eIF4G for binding to eIF4E. EMBO J *16*, 1114-1121.

Andrei, M.A., Ingelfinger, D., Heintzmann, R., Achsel, T., Rivera-Pomar, R., and Luhrmann, R. (2005). A role for eIF4E and eIF4E-transporter in targeting mRNPs to mammalian processing bodies. RNA *11*, 717-727.

Armengol, G., Rojo, F., Castellvi, J., Iglesias, C., Cuatrecasas, M., Pons, B., Baselga, J., and Ramon y Cajal, S. (2007). 4E-binding protein 1: a key molecular "funnel factor" in human cancer with clinical implications. Cancer Res *67*, 7551-7555.

Avdulov, S., Li, S., Michalek, V., Burrichter, D., Peterson, M., Perlman, D.M., Manivel, J.C., Sonenberg, N., Yee, D., Bitterman, P.B.*, et al.* (2004). Activation of translation complex eIF4F is essential for the genesis and maintenance of the malignant phenotype in human mammary epithelial cells. Cancer Cell *5*, 553-563.

Aylett, C.H., Sauer, E., Imseng, S., Boehringer, D., Hall, M.N., Ban, N., and Maier, T. (2016). Architecture of human mTOR complex 1. Science *351*, 48-52.

Bah, A., Vernon, R.M., Siddiqui, Z., Krzeminski, M., Muhandiram, R., Zhao, C., Sonenberg, N., Kay, L.E., and Forman-Kay, J.D. (2015). Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. Nature *519*, 106-109.

Bandyra, K.J., and Luisi, B.F. (2013). Licensing and due process in the turnover of bacterial RNA. RNA Biol *10*, 627-635.

Banko, J.L., Poulin, F., Hou, L., DeMaria, C.T., Sonenberg, N., and Klann, E. (2005). The translation repressor 4E-BP2 is critical for eIF4F complex formation, synaptic plasticity, and memory in the hippocampus. J Neurosci *25*, 9581-9590.

Barreau, C., Paillard, L., and Osborne, H.B. (2005). AU-rich elements and associated factors: are there unifying principles? Nucleic Acids Res *33*, 7138-7150.

Bartel, D.P. (2009). MicroRNAs: target recognition and regulatory functions. Cell *136*, 215-233.

Behm-Ansmant, I., Rehwinkel, J., Doerks, T., Stark, A., Bork, P., and Izaurralde, E. (2006a). mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. Genes Dev *20*, 1885-1898.

Behm-Ansmant, I., Rehwinkel, J., and Izaurralde, E. (2006b). MicroRNAs silence gene expression by repressing protein expression and/or by promoting mRNA decay. Cold Spring Harb Symp Quant Biol *71*, 523-530.

Bhat, M., Robichaud, N., Hulea, L., Sonenberg, N., Pelletier, J., and Topisirovic, I. (2015). Targeting the translation machinery in cancer. Nat Rev Drug Discov *14*, 261-278.

Brooks, S.A., and Blackshear, P.J. (2013). Tristetraprolin (TTP): interactions with mRNA and proteins, and current thoughts on mechanisms of action. Biochim Biophys Acta *1829*, 666-679.

Buffington, S.A., Huang, W., and Costa-Mattioli, M. (2014). Translational control in synaptic plasticity and cognitive dysfunction. Annu Rev Neurosci *37*, 17-38.

Burnett, P.E., Barrow, R.K., Cohen, N.A., Snyder, S.H., and Sabatini, D.M. (1998). RAFT1 phosphorylation of the translational regulators p70 S6 kinase and 4E-BP1. Proc Natl Acad Sci U S A *95*, 1432-1437.

Buskirk, A.R., and Green, R. (2017). Ribosome pausing, arrest and rescue in bacteria and eukaryotes. Philos Trans R Soc Lond B Biol Sci *372*.

Buttgereit, F., and Brand, M.D. (1995). A hierarchy of ATP-consuming processes in mammalian cells. Biochem J *312 (Pt 1)*, 163-167.

Cencic, R., Desforges, M., Hall, D.R., Kozakov, D., Du, Y., Min, J., Dingledine, R., Fu, H., Vajda, S., Talbot, P.J.*, et al.* (2011). Blocking eIF4E-eIF4G interaction as a strategy to impair coronavirus replication. J Virol *85*, 6381-6389.

Chekulaeva, M., Filipowicz, W., and Parker, R. (2009). Multiple independent domains of dGW182 function in miRNA-mediated repression in Drosophila. RNA *15*, 794-803.

Cheng, C.J., Bahal, R., Babar, I.A., Pincus, Z., Barrera, F., Liu, C., Svoronos, A., Braddock, D.T., Glazer, P.M., Engelman, D.M.*, et al.* (2015). MicroRNA silencing for cancer therapy targeted to the tumour microenvironment. Nature *518*, 107-110.

Cho, P.F., Gamberi, C., Cho-Park, Y.A., Cho-Park, I.B., Lasko, P., and Sonenberg, N. (2006). Cap-dependent translational inhibition establishes two opposing morphogen gradients in Drosophila embryos. Curr Biol *16*, 2035-2041.

Cho, P.F., Poulin, F., Cho-Park, Y.A., Cho-Park, I.B., Chicoine, J.D., Lasko, P., and Sonenberg, N. (2005). A new paradigm for translational control: inhibition via 5'-3' mRNA tethering by Bicoid and the eIF4E cognate 4EHP. Cell *121*, 411-423.

Cosentino, G.P., Schmelzle, T., Haghighat, A., Helliwell, S.B., Hall, M.N., and Sonenberg, N. (2000). Eap1p, a novel eukaryotic translation initiation factor 4E-associated protein in Saccharomyces cerevisiae. Mol Cell Biol *20*, 4604-4613.

Cridge, A.G., Castelli, L.M., Smirnova, J.B., Selley, J.N., Rowe, W., Hubbard, S.J., McCarthy, J.E., Ashe, M.P., Grant, C.M., and Pavitt, G.D. (2010). Identifying eIF4E-binding protein translationally-controlled transcripts reveals links to mRNAs bound by specific PUF proteins. Nucleic Acids Res *38*, 8039-8050.

Culjkovic, B., Topisirovic, I., Skrabanek, L., Ruiz-Gutierrez, M., and Borden, K.L. (2005). eIF4E promotes nuclear export of cyclin D1 mRNAs via an element in the 3'UTR. J Cell Biol *169*, 245-256.

De Benedetti, A., and Graff, J.R. (2004). eIF-4E expression and its role in malignancies and metastases. Oncogene *23*, 3189-3199.

Decker, C.J., and Parker, R. (2012). P-bodies and stress granules: possible roles in the control of translation and mRNA degradation. Cold Spring Harb Perspect Biol *4*, a012286.

Dever, T.E., Dar, A.C., and Sicheri, F. (2007). The eIF2α Kinases. In Translational Control in Biology and Medicine, M.B. Mathews, N. Sonenberg, and J.W.B. Hershey, eds. (New York, USA: Cold Spring Harbor Laboratory Press), pp. 319-344.

Dinkova, T.D., Keiper, B.D., Korneeva, N.L., Aamodt, E.J., and Rhoads, R.E. (2005). Translation of a small subset of Caenorhabditis elegans mRNAs is dependent on a specific eukaryotic translation initiation factor 4E isoform. Mol Cell Biol *25*, 100-113.

Donnelly, N., Gorman, A.M., Gupta, S., and Samali, A. (2013). The eIF2alpha kinases: their structures and functions. Cell Mol Life Sci *70*, 3493-3511.

Dostie, J., Ferraiuolo, M., Pause, A., Adam, S.A., and Sonenberg, N. (2000). A novel shuttling protein, 4E-T, mediates the nuclear import of the mRNA 5' cap-binding protein, eIF4E. EMBO J *19*, 3142-3156.

Dowling, R.J., Topisirovic, I., Alain, T., Bidinosti, M., Fonseca, B.D., Petroulakis, E., Wang, X., Larsson, O., Selvaraj, A., Liu, Y.*, et al.* (2010). mTORC1-mediated cell proliferation, but not cell growth, controlled by the 4E-BPs. Science *328*, 1172-1176.

Duval, M., Simonetti, A., Caldelari, I., and Marzi, S. (2015). Multiple ways to regulate translation initiation in bacteria: Mechanisms, regulatory circuits, dynamics. Biochimie *114*, 18-29.

Eulalio, A., Behm-Ansmant, I., and Izaurralde, E. (2007). P bodies: at the crossroads of post-transcriptional pathways. Nat Rev Mol Cell Biol *8*, 9-22.

Eulalio, A., Huntzinger, E., and Izaurralde, E. (2008). GW182 interaction with Argonaute is essential for miRNA-mediated translational repression and mRNA decay. Nat Struct Mol Biol *15*, 346-353.

Eyler, D.E., Wehner, K.A., and Green, R. (2013). Eukaryotic release factor 3 is required for multiple turnovers of peptide release catalysis by eukaryotic release factor 1. J Biol Chem *288*, 29530-29538.

Ferraiuolo, M.A., Basak, S., Dostie, J., Murray, E.L., Schoenberg, D.R., and Sonenberg, N. (2005). A role for the eIF4E-binding protein 4E-T in P-body formation and mRNA decay. J Cell Biol *170*, 913-924.

Fosgerau, K., and Hoffmann, T. (2015). Peptide therapeutics: current status and future directions. Drug Discov Today *20*, 122-128.

Frank, J., and Spahn, C.M.T. (2006). The ribosome and the mechanism of protein synthesis. Rep Prog Phys *69*, 1383-1417.

Frolova, L., Le Goff, X., Zhouravleva, G., Davydova, E., Philippe, M., and Kisselev, L. (1996). Eukaryotic polypeptide chain release factor eRF3 is an eRF1- and ribosome-dependent guanosine triphosphatase. RNA *2*, 334-341.

Fu, R., Olsen, M.T., Webb, K., Bennett, E.J., and Lykke-Andersen, J. (2016). Recruitment of the 4EHP-GYF2 cap-binding complex to tetraproline motifs of tristetraprolin promotes repression and degradation of mRNAs with AU-rich elements. RNA *22*, 373-382.

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-Cieszynska, 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.

Giovannone, B., Lee, E., Laviola, L., Giorgino, F., Cleveland, K.A., and Smith, R.J. (2003). Two novel proteins that are linked to insulin-like growth factor (IGF-I) receptors by the Grb10 adapter and modulate IGF-I signaling. J Biol Chem *278*, 31564-31573.

Giovannone, B., Tsiaras, W.G., de la Monte, S., Klysik, J., Lautier, C., Karashchuk, G., Goldwurm, S., and Smith, R.J. (2009). GIGYF2 gene disruption in mice results in

neurodegeneration and altered insulin-like growth factor signaling. Hum Mol Genet *18*, 4629-4639.

Gkogkas, C.G., Khoutorsky, A., Ran, I., Rampakakis, E., Nevarko, T., Weatherill, D.B., Vasuta, C., Yee, S., Truitt, M., Dallaire, P.*, et al.* (2013). Autism-related deficits via dysregulated eIF4E-dependent translational control. Nature *493*, 371-377.

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.

Gosselin, P., Martineau, Y., Morales, J., Czjzek, M., Glippa, V., Gauffeny, I., Morin, E., Le Corguille, G., Pyronnet, S., Cormier, P.*, et al.* (2013). Tracking a refined eIF4E-binding motif reveals Angel1 as a new partner of eIF4E. Nucleic Acids Res *41*, 7783-7792.

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

Grzmil, M., and Hemmings, B.A. (2012). Translation regulation as a therapeutic target in cancer. Cancer Res *72*, 3891-3900.

Gualerzi, C.O., and Pon, C.L. (2015). Initiation of mRNA translation in bacteria: structural and dynamic aspects. Cell Mol Life Sci *72*, 4341-4367.

Ha, M., and Kim, V.N. (2014). Regulation of microRNA biogenesis. Nat Rev Mol Cell Biol *15*, 509-524.

Hernandez, G., Altmann, M., Sierra, J.M., Urlaub, H., Diez del Corral, R., Schwartz, P., and Rivera-Pomar, R. (2005). Functional analysis of seven genes encoding eight translation initiation factor 4E (eIF4E) isoforms in Drosophila. Mech Dev *122*, 529-543.

Hernandez, G., Diez del Corral, R., Santoyo, J., Campuzano, S., and Sierra, J.M. (1997). Localization, structure and expression of the gene for translation initiation factor eIF-4E from Drosophila melanogaster. Mol Gen Genet *253*, 624-633.

Hernández, G., Gillespie, K.M., Bachvaroff, T.R., Jagus, R., Igreja, C., Peter, D., Bulfoni, M., and Cosson, B. (2016). Evolution of eIF4E-interacting proteins. In Evolution of the protein synthesis machinery and its regulation, G. Hernández, and R. Jagus, eds. (Heidelberg, London, New York: Springer), pp. 207-234.

Hernandez, G., Miron, M., Han, H., Liu, N., Magescas, J., Tettweiler, G., Frank, F., Siddiqui, N., Sonenberg, N., and Lasko, P. (2013). Mextli is a novel eukaryotic translation initiation factor 4E-binding protein that promotes translation in Drosophila melanogaster. Mol Cell Biol *33*, 2854-2864.

Hernandez, G., and Vazquez-Pianzola, P. (2005). Functional diversity of the eukaryotic translation initiation factors belonging to eIF4 families. Mech Dev *122*, 865-876.

Hershey, J.W., Sonenberg, N., and Mathews, M.B. (2012). Principles of translational control: an overview. Cold Spring Harb Perspect Biol *4*.

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

Hsieh, A.C., Liu, Y., Edlind, M.P., Ingolia, N.T., Janes, M.R., Sher, A., Shi, E.Y., Stumpf, C.R., Christensen, C., Bonham, M.J., *et al.* (2012). The translational landscape of mTOR signalling steers cancer initiation and metastasis. Nature *485*, 55-61.

Huntzinger, E., and Izaurralde, E. (2011). Gene silencing by microRNAs: contributions of translational repression and mRNA decay. Nat Rev Genet *12*, 99-110.

Ibrahimo, S., Holmes, L.E., and Ashe, M.P. (2006). Regulation of translation initiation by the yeast eIF4E binding proteins is required for the pseudohyphal response. Yeast *23*, 1075-1088.

Igreja, C., and Izaurralde, E. (2011). CUP promotes deadenylation and inhibits decapping of mRNA targets. Genes Dev *25*, 1955-1967.

Jackson, R.J., Hellen, C.U., and Pestova, T.V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. Nat Rev Mol Cell Biol *11*, 113-127.

Jonas, S., and Izaurralde, E. (2015). Towards a molecular understanding of microRNA-mediated gene silencing. Nat Rev Genet *16*, 421-433.

Joshi, B., Cameron, A., and Jagus, R. (2004). Characterization of mammalian eIF4E-family members. Eur J Biochem *271*, 2189-2203.

Joshi, B., Lee, K., Maeder, D.L., and Jagus, R. (2005). Phylogenetic analysis of eIF4E-family members. BMC Evol Biol *5*, 48.

Jung, M.Y., Lorenz, L., and Richter, J.D. (2006). Translational control by neuroguidin, a eukaryotic initiation factor 4E and CPEB binding protein. Mol Cell Biol *26*, 4277-4287.

Kafri, M., Metzl-Raz, E., Jona, G., and Barkai, N. (2016). The Cost of Protein Production. Cell Rep *14*, 22-31.

Kamenska, A., Lu, W.T., Kubacka, D., Broomhead, H., Minshall, N., Bushell, M., and Standart, N. (2014a). Human 4E-T represses translation of bound mRNAs and enhances microRNA-mediated silencing. Nucleic Acids Res *42*, 3298-3313.

Kamenska, A., Simpson, C., and Standart, N. (2014b). eIF4E-binding proteins: new factors, new locations, new roles. Biochem Soc Trans *42*, 1238-1245.

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.

Keiper, B.D., Lamphear, B.J., Deshpande, A.M., Jankowska-Anyszka, M., Aamodt, E.J., Blumenthal, T., and Rhoads, R.E. (2000). Functional characterization of five eIF4E isoforms in Caenorhabditis elegans. J Biol Chem *275*, 10590-10596.

Kim, V.N., Han, J., and Siomi, M.C. (2009). Biogenesis of small RNAs in animals. Nat Rev Mol Cell Biol *10*, 126-139.

Kinkelin, K., Veith, K., Grunwald, M., and Bono, F. (2012). Crystal structure of a minimal eIF4E-Cup complex reveals a general mechanism of eIF4E regulation in translational repression. RNA *18*, 1624-1634.

Ko, S.Y., Guo, H., Barengo, N., and Naora, H. (2009). Inhibition of ovarian cancer growth by a tumor-targeting peptide that binds eukaryotic translation initiation factor 4E. Clin Cancer Res *15*, 4336-4347.

Kohler, R., Mooney, R.A., Mills, D.J., Landick, R., and Cramer, P. (2017). Architecture of a transcribing-translating expressome. Science *356*, 194-197.

Kozak, M. (1983). Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. Microbiol Rev *47*, 1-45.

Kozak, M. (1991). Structural features in eukaryotic mRNAs that modulate the initiation of translation. J Biol Chem *266*, 19867-19870.

Krishnamoorthy, T., Pavitt, G.D., Zhang, F., Dever, T.E., and Hinnebusch, A.G. (2001). Tight binding of the phosphorylated alpha subunit of initiation factor 2 (eIF2alpha) to the regulatory subunits of guanine nucleotide exchange factor eIF2B is required for inhibition of translation initiation. Mol Cell Biol *21*, 5018-5030.

Kubacka, D., Kamenska, A., Broomhead, H., Minshall, N., Darzynkiewicz, E., and Standart, N. (2013). Investigating the consequences of eIF4E2 (4EHP) interaction with 4E-transporter on its cellular distribution in HeLa cells. PLoS One *8*, e72761.

Kuzuoğlu-Öztürk, D., Bhandari, D., Huntzinger, E., Fauser, M., Helms, S., and Izaurralde, E. (2016). miRISC and the CCR4-NOT complex silence mRNA targets independently of 43S ribosomal scanning. EMBO J *35*, 1186-1203.

Lall, S., Friedman, C.C., Jankowska-Anyszka, M., Stepinski, J., Darzynkiewicz, E., and Davis, R.E. (2004). Contribution of trans-splicing, 5' -leader length, cap-poly(A) synergism, and initiation factors to nematode translation in an Ascaris suum embryo cell-free system. J Biol Chem *279*, 45573-45585.

Landick, R., Carey, J., and Yanofsky, C. (1985). Translation activates the paused transcription complex and restores transcription of the trp operon leader region. Proc Natl Acad Sci U S A *82*, 4663-4667.

Lasda, E.L., and Blumenthal, T. (2011). Trans-splicing. Wiley Interdiscip Rev RNA *2*, 417-434.

Lazaris-Karatzas, A., Montine, K.S., and Sonenberg, N. (1990). Malignant transformation by a eukaryotic initiation factor subunit that binds to mRNA 5' cap. Nature *345*, 544-547.

Leonard, S., Plante, D., Wittmann, S., Daigneault, N., Fortin, M.G., and Laliberte, J.F. (2000). Complex formation between potyvirus VPg and translation eukaryotic initiation factor 4E correlates with virus infectivity. J Virol *74*, 7730-7737.

Lukhele, S., Bah, A., Lin, H., Sonenberg, N., and Forman-Kay, J.D. (2013). Interaction of the Eukaryotic Initiation Factor 4E with 4E-BP2 at a Dynamic Bipartite Interface. Structure.

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.

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.

Mamane, Y., Petroulakis, E., Rong, L., Yoshida, K., Ler, L.W., and Sonenberg, N. (2004). eIF4E--from translation to transformation. Oncogene *23*, 3172-3179.

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.

Miron, M., Lasko, P., and Sonenberg, N. (2003). Signaling from Akt to FRAP/TOR targets both 4E-BP and S6K in Drosophila melanogaster. Mol Cell Biol *23*, 9117-9126.

Miron, M., Verdu, J., Lachance, P.E., Birnbaum, M.J., Lasko, P.F., and Sonenberg, N. (2001). The translational inhibitor 4E-BP is an effector of PI(3)K/Akt signalling and cell growth in Drosophila. Nat Cell Biol *3*, 596-601.

Miyoshi, H., Dwyer, D.S., Keiper, B.D., Jankowska-Anyszka, M., Darzynkiewicz, E., and Rhoads, R.E. (2002). Discrimination between mono- and trimethylated cap structures by two isoforms of Caenorhabditis elegans eIF4E. EMBO J *21*, 4680-4690.

Moerke, N.J., Aktas, H., Chen, H., Cantel, S., Reibarkh, M.Y., Fahmy, A., Gross, J.D., Degterev, A., Yuan, J., Chorev, M.*, et al.* (2007). Small-molecule inhibition of the interaction between the translation initiation factors eIF4E and eIF4G. Cell *128*, 257-267.

Moore, M.J. (2005). From birth to death: the complex lives of eukaryotic mRNAs. Science *309*, 1514-1518.

Morisaki, T., Lyon, K., DeLuca, K.F., DeLuca, J.G., English, B.P., Zhang, Z., Lavis, L.D., Grimm, J.B., Viswanathan, S., Looger, L.L.*, et al.* (2016). Real-time quantification of single RNA translation dynamics in living cells. Science *352*, 1425-1429.

Morita, M., Ler, L.W., Fabian, M.R., Siddiqui, N., Mullin, M., Henderson, V.C., Alain, T., Fonseca, B.D., Karashchuk, G., Bennett, C.F.*, et al.* (2012). A novel 4EHP-GIGYF2 translational repressor complex is essential for mammalian development. Mol Cell Biol *32*, 3585-3593.

Murphy, J.F., Klein, P.G., Hunt, A.G., and Shaw, J.G. (1996). Replacement of the tyrosine residue that links a potyviral VPg to the viral RNA is lethal. Virology *220*, 535-538.

Nakamura, A., Sato, K., and Hanyu-Nakamura, K. (2004). Drosophila cup is an eIF4E binding protein that associates with Bruno and regulates oskar mRNA translation in oogenesis. Dev Cell *6*, 69-78.

Napoli, I., Mercaldo, V., Boyl, P.P., Eleuteri, B., Zalfa, F., De Rubeis, S., Di Marino, D., Mohr, E., Massimi, M., Falconi, M.*, et al.* (2008). The fragile X syndrome protein represses activity-dependent translation through CYFIP1, a new 4E-BP. Cell *134*, 1042-1054.

Nelson, M.R., Leidal, A.M., and Smibert, C.A. (2004). Drosophila Cup is an eIF4E-binding protein that functions in Smaug-mediated translational repression. EMBO J *23*, 150-159.

Nishimura, T., Padamsi, Z., Fakim, H., Milette, S., Dunham, W.H., Gingras, A.C., and Fabian, M.R. (2015). The eIF4E-Binding Protein 4E-T Is a Component of the mRNA Decay Machinery that Bridges the 5' and 3' Termini of Target mRNAs. Cell Rep *11*, 1425-1436.

Ozgur, S., Basquin, J., Kamenska, A., Filipowicz, W., Standart, N., and Conti, E. (2015). Structure of a Human 4E-T/DDX6/CNOT1 Complex Reveals the Different

Interplay of DDX6-Binding Proteins with the CCR4-NOT Complex. Cell Rep *13*, 703-711.

Paku, K.S., Umenaga, Y., Usui, T., Fukuyo, A., Mizuno, A., In, Y., Ishida, T., and Tomoo, K. (2012). A conserved motif within the flexible C-terminus of the translational regulator 4E-BP is required for tight binding to the mRNA cap-binding protein eIF4E. Biochem J *441*, 237-245.

Papadopoulos, E., Jenni, S., Kabha, E., Takrouri, K.J., Yi, T., Salvi, N., Luna, R.E., Gavathiotis, E., Mahalingam, P., Arthanari, H.*, et al.* (2014). Structure of the eukaryotic translation initiation factor eIF4E in complex with 4EGI-1 reveals an allosteric mechanism for dissociating eIF4G. Proc Natl Acad Sci U S A *111*, E3187-3195.

Pause, A., Belsham, G.J., Gingras, A.C., Donze, O., Lin, T.A., Lawrence, J.C., Jr., and Sonenberg, N. (1994). Insulin-dependent stimulation of protein synthesis by phosphorylation of a regulator of 5'-cap function. Nature *371*, 762-767.

Pelletier, J., Graff, J., Ruggero, D., and Sonenberg, N. (2015). Targeting the eIF4F translation initiation complex: a critical nexus for cancer development. Cancer Res *75*, 250-263.

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

Pisarev, A.V., Skabkin, M.A., Pisareva, V.P., Skabkina, O.V., Rakotondrafara, A.M., Hentze, M.W., Hellen, C.U., and Pestova, T.V. (2010). The role of ABCE1 in eukaryotic posttermination ribosomal recycling. Mol Cell *37*, 196-210.

Pisareva, V.P., Skabkin, M.A., Hellen, C.U., Pestova, T.V., and Pisarev, A.V. (2011). Dissociation by Pelota, Hbs1 and ABCE1 of mammalian vacant 80S ribosomes and stalled elongation complexes. EMBO J *30*, 1804-1817.

Proshkin, S., Rahmouni, A.R., Mironov, A., and Nudler, E. (2010). Cooperation between translating ribosomes and RNA polymerase in transcription elongation. Science *328*, 504-508.

Regnier, P., and Hajnsdorf, E. (2013). The interplay of Hfq, poly(A) polymerase I and exoribonucleases at the 3' ends of RNAs resulting from Rho-independent termination: A tentative model. RNA Biol *10*, 602-609.

Rendl, L.M., Bieman, M.A., Vari, H.K., and Smibert, C.A. (2012). The eIF4E-binding protein Eap1p functions in Vts1p-mediated transcript decay. PLoS One *7*, e47121.

Rhoads, R.E. (2009). eIF4E: new family members, new binding partners, new roles. J Biol Chem *284*, 16711-16715.

Ricci, E.P., Limousin, T., Soto-Rifo, R., Rubilar, P.S., Decimo, D., and Ohlmann, T. (2013). miRNA repression of translation in vitro takes place during 43S ribosomal scanning. Nucleic Acids Res *41*, 586-598.

Rinker-Schaeffer, C.W., Graff, J.R., De Benedetti, A., Zimmer, S.G., and Rhoads, R.E. (1993). Decreasing the level of translation initiation factor 4E with antisense RNA causes reversal of ras-mediated transformation and tumorigenesis of cloned rat embryo fibroblasts. Int J Cancer *55*, 841-847.

Rom, E., Kim, H.C., Gingras, A.C., Marcotrigiano, J., Favre, D., Olsen, H., Burley, S.K., and Sonenberg, N. (1998). Cloning and characterization of 4EHP, a novel mammalian eIF4E-related cap-binding protein. J Biol Chem *273*, 13104-13109.

Rosettani, P., Knapp, S., Vismara, M.G., Rusconi, L., and Cameron, A.D. (2007). Structures of the human eIF4E homologous protein, h4EHP, in its m7GTP-bound and unliganded forms. J Mol Biol *368*, 691-705.

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.

Ruggero, D. (2013). Translational control in cancer etiology. Cold Spring Harb Perspect Biol *5*.

Ruggero, D., Montanaro, L., Ma, L., Xu, W., Londei, P., Cordon-Cardo, C., and Pandolfi, P.P. (2004). The translation factor eIF-4E promotes tumor formation and cooperates with c-Myc in lymphomagenesis. Nat Med *10*, 484-486.

Salvi, N., Papadopoulos, E., Blackledge, M., and Wagner, G. (2016). The Role of Dynamics and Allostery in the Inhibition of the eIF4E/eIF4G Translation Initiation Factor Complex. Angew Chem Int Ed Engl *55*, 7176-7179.

Sarkar, N. (1997). Polyadenylation of mRNA in prokaryotes. Annu Rev Biochem *66*, 173-197.

Schaffer, F.L., Ehresmann, D.W., Fretz, M.K., and Soergel, M.I. (1980). A protein, VPg, covalently linked to 36S calicivirus RNA. J Gen Virol *47*, 215-220.

Schalm, S.S., Fingar, D.C., Sabatini, D.M., and Blenis, J. (2003). TOS motif-mediated raptor binding regulates 4E-BP1 multisite phosphorylation and function. Curr Biol *13*, 797-806.

Sekiyama, N., Arthanari, H., Papadopoulos, E., Rodriguez-Mias, R.A., Wagner, G., and Leger-Abraham, M. (2015). Molecular mechanism of the dual activity of 4EGI-1: Dissociating eIF4G from eIF4E but stabilizing the binding of unphosphorylated 4E-BP1. Proc Natl Acad Sci U S A *112*, E4036-4045.

Sengupta, M.S., Low, W.Y., Patterson, J.R., Kim, H.M., Traven, A., Beilharz, T.H., Colaiacovo, M.P., Schisa, J.A., and Boag, P.R. (2013). ifet-1 is a broad-scale translational repressor required for normal P granule formation in C. elegans. J Cell Sci *126*, 850-859.

Shoemaker, C.J., and Green, R. (2011). Kinetic analysis reveals the ordered coupling of translation termination and ribosome recycling in yeast. Proc Natl Acad Sci U S A *108*, E1392-1398.

Sonenberg, N., and Hinnebusch, A.G. (2009). Regulation of translation initiation in eukaryotes: mechanisms and biological targets. Cell *136*, 731-745.

Steege, D.A. (2000). Emerging features of mRNA decay in bacteria. RNA *6*, 1079-1090.

Steinberger, J., Chu, J., Maiga, R.I., Sleiman, K., and Pelletier, J. (2016). Developing anti-neoplastic biotherapeutics against eIF4F. Cell Mol Life Sci.

Steitz, T.A. (2008). A structural understanding of the dynamic ribosome machine. Nat Rev Mol Cell Biol *9*, 242-253.

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 (elF4A) in translation is in direct proportion to the degree of mRNA 5' secondary structure. RNA *7*, 382-394.

Tait, S., Dutta, K., Cowburn, D., Warwicker, J., Doig, A.J., and McCarthy, J.E. (2010). Local control of a disorder-order transition in 4E-BP1 underpins regulation of translation via eIF4E. Proc Natl Acad Sci U S A *107*, 17627-17632.

Tao, X., and Gao, G. (2015). Tristetraprolin Recruits Eukaryotic Initiation Factor 4E2 To Repress Translation of AU-Rich Element-Containing mRNAs. Mol Cell Biol *35*, 3921-3932.

Thandapani, P., O'Connor, T.R., Bailey, T.L., and Richard, S. (2013). Defining the RGG/RG motif. Mol Cell *50*, 613-623.

Truitt, M.L., Conn, C.S., Shi, Z., Pang, X., Tokuyasu, T., Coady, A.M., Seo, Y., Barna, M., and Ruggero, D. (2015). Differential Requirements for eIF4E Dose in Normal Development and Cancer. Cell *162*, 59-71.

Tucker, M., Staples, R.R., Valencia-Sanchez, M.A., Muhlrad, D., and Parker, R. (2002). Ccr4p is the catalytic subunit of a Ccr4p/Pop2p/Notp mRNA deadenylase complex in Saccharomyces cerevisiae. EMBO J *21*, 1427-1436.

Umenaga, Y., Paku, K.S., In, Y., Ishida, T., and Tomoo, K. (2011). Identification and function of the second eIF4E-binding region in N-terminal domain of eIF4G:

comparison with eIF4E-binding protein. Biochem Biophys Res Commun *414*, 462-467.

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.

Valkov, E., Sharpe, T., Marsh, M., Greive, S., and Hyvonen, M. (2012). Targeting protein-protein interactions and fragment-based drug discovery. Top Curr Chem *317*, 145-179.

Van Doren, K., and Hirsh, D. (1990). mRNAs that mature through trans-splicing in Caenorhabditis elegans have a trimethylguanosine cap at their 5' termini. Mol Cell Biol *10*, 1769-1772.

Villaescusa, J.C., Buratti, C., Penkov, D., Mathiasen, L., Planaguma, J., Ferretti, E., and Blasi, F. (2009). Cytoplasmic Prep1 interacts with 4EHP inhibiting Hoxb4 translation. PLoS One *4*, e5213.

von Manteuffel, S.R., Gingras, A.C., Ming, X.F., Sonenberg, N., and Thomas, G. (1996). 4E-BP1 phosphorylation is mediated by the FRAP-p70s6k pathway and is independent of mitogen-activated protein kinase. Proc Natl Acad Sci U S A *93*, 4076-4080.

Wallace, A., Filbin, M.E., Veo, B., McFarland, C., Stepinski, J., Jankowska-Anyszka, M., Darzynkiewicz, E., and Davis, R.E. (2010). The nematode eukaryotic translation initiation factor 4E/G complex works with a trans-spliced leader stem-loop to enable efficient translation of trimethylguanosine-capped RNAs. Mol Cell Biol *30*, 1958-1970.

Wells, S.E., Hillner, P.E., Vale, R.D., and Sachs, A.B. (1998). Circularization of mRNA by eukaryotic translation initiation factors. Mol Cell *2*, 135-140.

Wilhelm, J.E., Hilton, M., Amos, Q., and Henzel, W.J. (2003). Cup is an eIF4E binding protein required for both the translational repression of oskar and the recruitment of Barentsz. J Cell Biol *163*, 1197-1204.

Wu, B., Eliscovich, C., Yoon, Y.J., and Singer, R.H. (2016). Translation dynamics of single mRNAs in live cells and neurons. Science *352*, 1430-1435.

Yanagiya, A., Suyama, E., Adachi, H., Svitkin, Y.V., Aza-Blanc, P., Imataka, H., Mikami, S., Martineau, Y., Ronai, Z.A., and Sonenberg, N. (2012). Translational homeostasis via the mRNA cap-binding protein, eIF4E. Mol Cell *46*, 847-858.

Youtani, T., Tomoo, K., Ishida, T., Miyoshi, H., and Miura, K. (2000). Regulation of human eIF4E by 4E-BP1: binding analysis using surface plasmon resonance. IUBMB Life *49*, 27-31.

Zappavigna, V., Piccioni, F., Villaescusa, J.C., and Verrotti, A.C. (2004). Cup is a nucleocytoplasmic shuttling protein that interacts with the eukaryotic translation initiation factor 4E to modulate Drosophila ovary development. Proc Natl Acad Sci U S A *101*, 14800-14805.

Zhou, W., Quah, S.T., Verma, C.S., Liu, Y., Lane, D.P., and Brown, C.J. (2012). Improved eIF4E binding peptides by phage display guided design: plasticity of interacting surfaces yield collective effects. PLoS One *7*, e47235.

9. ABBREVIATIONS

4E-BM	eIF4E-binding motif
4E-BP1-3	eIF4E-binding proteins 1-3
4E-BPs	eIF4E-binding proteins
4E-T	eIF4E-transporter protein
4EHP	eIF4E-homologous protein
A sequences	auxiliary sequences
α1-3	sequence region with α-helical propensity or fold number 1-3
ABCE1	ABC-family ATPase ABCE1
AGO1-4	Argonaute proteins 1-4
ARE	AU-rich element
ASD	autism spectrum disorders
ASU	asymmetric unit
ATP	Adenosine triphosphate
C motif	canonical motif
CCR4	Carbon catabolite repressor protein 4
Ce	*Caenorhabditis elegans*
CYFIP1	Cytoplasmic FMR1-interacting protein 1
Dm	*Drosophila melanogaster*
eEF1A	eukaryotic elongation factor 1-alpha
eEF2	eukaryotic elongation factor 2
EEP	endo/exonuclease phosphatase domain
eIF2-α, -β, -γ	eukaryotic translation initiation factor 2 subunits α, β or γ
eIF2B	eukaryotic translation initiation factor 2B
eIF4A	eukaryotic translation initiation factor 4A
eIF4E	eukaryotic translation initiation factor 4E
eIF4F	eukaryotic translation initiation factor 4F
eIF4G	eukaryotic translation initiation factor 4G
eIFs	eukaryotic translation initiation factors
eRF1	eukaryotic translation termination factor 1
eRF3	eukaryotic translation termination factor 3
FMR	fragile X mental retardation

GB1	B1 domain of protein G
GIGYF	Grb10-interacting GYF protein
GnRH	gonadotropin-releasing hormone
GnRH-RI	gonadotropin-releasing hormone receptor
GTP	guanosine triphosphate
Hs	*Homo sapiens*
HSQC	heteronuclear single quantum coherence
IFE1-5	initiation factor of *elegans* 1-5
IP	immunoprecipitation
ITC	isothermal titration calorimetry
K_D	dissociation constant
KH	hnRNP K homology domain
m^7GpppN	7-methylguanosine 5′-triphosphate dinucleotide
m^7GTP	7-methylguanosine 5′-triphosphate
Met-tRNA$^{Met}_i$	initiator methionyl-tRNA
MIF4G	middle domain of eukaryotic translation initiation factor 4G
miRISC	microRNA-induced silencing complex
MMG	monomethylguanosine
mRNA	messenger RNA
mRNP	messenger ribonucleoprotein
NC motif	non-canonical motif
NMR	nuclear magnetic resonance
NOT	Negative on TATA
nt	nucleotide
p4E-BP1	phosphorylated 4E-BP1
PABP	poly(A)-binding protein
PDB	Protein databank
pHLIP	pH-induced transmembrane structure
PI3K	phosphoinositide 3-kinase
PIC	preinitiation complex
poly(A)	stretch of multiple adenosine monophosphates
RBP	RNA-binding protein
RISC	RNA-induced silencing complex

RNA	ribonucleic acid
SAXS	small-angle X-ray scattering
Sc	*Saccharomyces cerevisiae*
SD	Shine-Dalgarno
SL	spliced leader
TC	ternary complex
TMG	trimethylguanosine
TNF	tumor necrosis factor
tRNA	transfer RNA
TTP	Tristetraprolin
UTR	untranslated region of an mRNA
VPg	viral protein genome-linked factor

10. APPENDIX

10.1 List of publications

10.1.1 Discussed publications

The original manuscripts of the publications listed in this section can be found in section 10.3. The manuscripts are either published or accepted for publication.

1) **4E-BPs require non-canonical 4E-binding motifs and a lateral surface of eIF4E to repress translation**
   Igreja, C., Peter, D., Weiler, C. and Izaurralde, E. (2014)
   Nature Communications **5**, 4790

2) **Molecular Architecture of 4E-BP Translational Inhibitors Bound to eIF4E**
   Peter, D., Igreja, C., Weber, R., Wohlbold, L., Weiler, C., Ebertsch, L., Weichenrieder, O. and Izaurralde, E. (2015)
   Molecular Cell **57**(6), 1074-1087

3) **Mextli proteins use both canonical bipartite and novel tripartite binding modes to form eIF4E complexes that display differential sensitivity to 4E-BP regulation**
   Peter, D., Weber, R., Köne, C., Chung, M.Y., Ebertsch, L., Truffault, V., Weichenrieder, O., Igreja, C. and Izaurralde, E. (2015)
   Genes & Development **29**(17), 1835-1849

4) **The Structures of eIF4E-eIF4G Complexes Reveal an Extended Interface to Regulate Translation Initiation**
   Grüner, S., Peter, D., Weber, R., Wohlbold, L., Chung, M.Y., Weichenrieder, O., Valkov, E., Igreja, C. and Izaurralde, E. (2016)
   Molecular Cell **64**(3), 467-479

5) GIGYF1/2 proteins use auxiliary sequences to selectively bind to 4EHP and repress target mRNA expression

Peter, D., Weber, R., Sandmeir, F., Wohlbold, L., Helms, S., Bawankar, P., Valkov, E., Igreja,C. and Izaurralde, E. (2017)

Genes & Development accepted manuscript, *in press*

10.1.2 Additional publications during the PhD

6) An asymmetric PAN3 dimer recruits a single PAN2 exonuclease to mediate mRNA deadenylation and decay

Jonas, S.*, Christie, M.*, Peter, D., Bhandari, D., Loh, B., Huntzinger, E., Weichenrieder, O. and Izaurralde, E. (2014)

Nature Structural and Molecular Biology **21**(7), 599-608

*equal contributions

7) Evolution of eIF4E-interacting proteins

In: *Evolution of the protein synthesis machinery and its regulation*

Hernández, G., Gillespie, K.M., Bachvaroff, T.R., Jagus, R., Igreja, C., Peter, D., Bulfoni, M. and Cosson B. (2016)

Eds. Greco Hernández and Rosemary Jagus.

pp. 207-234; Springer. Heidelberg, London, New York.

10.1.3 Additional publications before the PhD

8) Structural requirements for enzymatic activities of foamy virus protease-reverse transcriptase

Schneider, A., Peter, D., Schmitt, J., Leo, B., Richter, F., Rösch, P., Wöhrl, B.M. and Hartl, M.J. (2014)

Proteins **82**(3), 375-385

10.2 List of crystal structures

This section contains an overview of the crystal structures I solved and published during my doctoral work. The crystal structures listed in the following are referenced to the corresponding accession code of the worldwide protein data bank (http://www.wwpdb.org).

protein / protein complex	PDB accession code
N. crassa Pan2 WD40 domain	4CZV
D. melanogaster eIF4E-Thor complex	4UE8
D. melanogaster eIF4E-4E-T complex	4UE9
D. melanogaster eIF4E-eIF4G complex	4UEC
H. sapiens eIF4E-4E-BP1 complex	4UED
D. melanogaster eIF4E-4E-BP chimera (crystal form 1)	4UEA
D. melanogaster eIF4E-4E-BP chimera (crystal form 2)	4UEB
D. melanogaster eIF4E-Mextli complex (cap-bound state)	5ABU
D. melanogaster eIF4E-Mextli complex (cap-free state)	5ABV
C. elegans eIF4E-Mextli complex (cap-bound state)	5ABX
C. elegans eIF4E-Mextli complex (cap-free state)	5ABY
H. sapiens 4EHP-GIGYF1 complex	5NVK
H. sapiens 4EHP-GIGYF2 complex	5NVL
H. sapiens 4EHP-GIGYF2 complex (lacking the auxiliary sequences)	5NVM
H. sapiens 4EHP-4E-BP1 complex	5NVN

10.3 Original manuscripts of the discussed publications

nature COMMUNICATIONS

ARTICLE

Received 15 May 2014 | Accepted 24 Jul 2014 | Published 2 Sep 2014

DOI: 10.1038/ncomms5790 OPEN

4E-BPs require non-canonical 4E-binding motifs and a lateral surface of eIF4E to repress translation

Cátia Igreja[1], Daniel Peter[1], Catrin Weiler[1] & Elisa Izaurralde[1]

eIF4E-binding proteins (4E-BPs) are a widespread class of translational regulators that share a canonical (C) eIF4E-binding motif (4E-BM) with eIF4G. Consequently, 4E-BPs compete with eIF4G for binding to the dorsal surface on eIF4E to inhibit translation initiation. Some 4E-BPs contain non-canonical 4E-BMs (NC 4E-BMs), but the contribution of these motifs to the repressive mechanism—and whether these motifs are present in all 4E-BPs—remains unknown. Here, we show that the three annotated *Drosophila melanogaster* 4E-BPs contain NC 4E-BMs. These motifs bind to a lateral surface on eIF4E that is not used by eIF4G. This distinct molecular recognition mode is exploited by 4E-BPs to dock onto eIF4E–eIF4G complexes and effectively displace eIF4G from the dorsal surface of eIF4E. Our data reveal a hitherto unrecognized role for the NC 4E-BMs and the lateral surface of eIF4E in 4E-BP-mediated translational repression, and suggest that bipartite 4E-BP mimics might represent efficient therapeutic tools to dampen translation during oncogenic transformation.

[1] Department of Biochemistry, Max Planck Institute for Developmental Biology, Spemannstrasse 35, 72076 Tübingen, Germany. Correspondence and requests for materials should be addressed to E.I. (email: elisa.izaurralde@tuebingen.mpg.de).

The regulation of protein synthesis at the initiation step is a widespread and reversible mechanism to control gene expression in eukaryotes[1,2]. During translation initiation, the small ribosomal subunit is recruited to mRNA by the eukaryotic initiation factor 4F (eIF4F) complex, which comprises the cap-binding protein eIF4E, the scaffolding protein eIF4G and the DEAD-box RNA helicase eIF4A. The eIF4E protein recognizes the mRNA m^7GpppN cap structure and interacts with eIF4G, which promotes translation initiation via the recruitment of the 43S pre-initiation complex[1]. eIF4G binds eIF4E through a conserved motif (or canonical eIF4E-binding motif, C 4E-BM) of sequence $TyrX_4Leu\Phi$, where Φ is hydrophobic, and X is any amino acid[3–5].

The assembly of the eIF4F complex is regulated by a diverse group of eIF4E-binding proteins (4E-BPs), which share a similar C $TyrX_4Leu\Phi$ motif with eIF4G. Therefore, 4E-BPs bind to the same surface on eIF4E, sterically blocking its interaction with eIF4G and preventing translation initiation[4–7]. The association of 4E-BPs with eIF4E is reversible and regulated by phosphorylation. Unphosphorylated or hypophosphorylated 4E-BPs exhibit a high affinity for eIF4E and repress translation, whereas hyper-phosphorylated 4E-BPs lose their affinity for eIF4E[2,8,9].

At a functional level, 4E-BPs play essential roles in the control of translation during development and regulate neuronal plasticity by repressing translation at a global or message-specific level[9–14]. Through their inhibitory effect on translation, 4E-BPs negatively regulate cell proliferation and act as tumor suppressors[9,11]. However, the 4E-BP anti-oncogenic function is compromised in many tumors, resulting in increased eIF4E activity and protein synthesis, which is required for tumorigenic transformation[9]. Consequently, a detailed molecular understanding of the interaction between eIF4E and 4E-BPs is crucial to design or improve drugs that may be useful in pathological conditions in which eIF4E activity and global translation are upregulated[9,15,16].

The C motifs of eIF4G and 4E-BPs adopt similar α-helical structures on binding to a conserved patch of hydrophobic residues on the dorsal side of the eIF4E cap-binding pocket[5,7,17]. Additional surfaces on eIF4E also contribute to the interaction with eIF4G as well as with a subset of 4E-BPs by binding to residues that are carboxy terminal to the C motifs, which contain NC 4E-BMs[17–20]. To date, NC motifs have only been identified and characterized in eIF4G, vertebrate 4E-BP1–3 and *D. melanogaster* CUP[17,19,21–23]. The NC motifs of 4E-BPs are not conserved between orthologous proteins across the animal kingdom. Therefore, it is not known whether all 4E-BPs contain NC motifs. Functionally, NC motifs have been proposed to play an auxiliary role by cooperating with their cognate C motifs to increase the binding affinity for eIF4E[17,19,20,22].

The protein CUP is an insect-specific 4E-BP that controls the translation of maternal messenger RNAs during oogenesis and embryogenesis[21,24–26]. The crystal structure of *Dm* eIF4E bound to a CUP peptide containing the C and NC 4E-BMs revealed that both motifs adopt an α-helical conformation and contact two orthogonal surfaces on eIF4E[27]. The C 4E-BM binds to the conserved dorsal surface of eIF4E, as observed for the C motifs of eIF4G and 4E-BP1,2. The NC motif docks in an antiparallel fashion onto a lateral and conserved surface of eIF4E[27].

A comparison of the *Dm* eIF4E–CUP complex with the structure of yeast eIF4E in complex with a fragment of eIF4G indicates that the NC motif of CUP and yeast eIF4G bind to partially overlapping surfaces on the lateral side of eIF4E[17,27]. Consequently, NC motifs could also contribute to the steric incompatibility with eIF4G and participate in the competition process. However, the contribution of NC motifs to the ability of 4E-BPs to displace eIF4G has not yet been elucidated.

To shed light on the role of NC motifs in 4E-BP-mediated translational repression, we investigated whether different *Dm* 4E-BPs contain NC motifs and how these motifs contribute to the displacement of eIF4G from eIF4E. We show that similar to CUP, Thor (ortholog of 4E-BP1–3) and 4E-T (4E-transporter) bind to eIF4E through a bipartite sequence that contains a C motif and a NC motif. The newly identified NC motifs in Thor and 4E-T share no sequence similarity with their vertebrate counterparts or with CUP. Nevertheless, these motifs share an overlapping lateral binding surface on eIF4E with the NC motif of CUP, which is required for the binding of 4E-BPs but not of eIF4G. The binding to an eIF4E surface that is not used by eIF4G allows 4E-BPs to dock onto preexisting eIF4E–eIF4G complexes to begin to displace eIF4G from the dorsal surface. Our data reveal a hitherto unrecognized diversity of NC motifs and establish the relevance of these motifs in the mechanism by which 4E-BPs repress translation. More generally, our data indicate that bipartite 4E-BP mimics have a competitive advantage over eIF4G and might represent potent repressors for the treatment of malignancies, in which eIF4E activity is upregulated.

Results

4E-BPs bind to a lateral surface of eIF4E. To gain insight into the binding mode of different 4E-BPs to eIF4E, we compared the interaction of *Dm* CUP, Thor and 4E-T with *Dm* eIF4E (Fig. 1a). In coimmunoprecipitation and pull-down assays, we confirmed that all the proteins interacted with endogenous eIF4E in *Dm* Schneider (S2) cells (Supplementary Fig. 1a–e).

Dm CUP interacts with eIF4E through C and NC motifs[27]. In particular, the CUP residues Tyr327, Leu332, Met333 and Arg336 in the C motif interact with residues on the dorsal surface of eIF4E, including Trp106 and Leu167 (Fig. 1b,c and Supplementary Fig. 2a,b). In addition, the CUP residues Leu364, Leu368, Met371 and Ile373 in the NC motif contact a eIF4E lateral surface that is centered at residues Ile96 and Ile112 (Fig. 1b,d and Supplementary Fig. 2a,b)[27].

To determine whether Thor, 4E-T and eIF4G also recognize the lateral surface of eIF4E, we substituted residues Ile96 and Ile112 with Ala (eIF4E mutant II-AA) and performed coimmunoprecipitation assays in S2 cells. As a control, we used an eIF4E mutant with a Trp106Ala substitution (W106A) on the dorsal binding surface, because this substitution abolishes the binding of CUP and eIF4G to eIF4E[21,28,29]. As expected, the W106A substitution strongly reduced the binding of eIF4E to endogenous eIF4G and to all three of the 4E-BPs (Fig. 1e–g, lanes 7). By contrast, the II-AA mutations disrupted the association of eIF4E with CUP, Thor and 4E-T but not with eIF4G (Fig. 1e–g, lanes 8). Thus, in contrast to eIF4G, 4E-BPs recognize and depend on the lateral surface to efficiently bind to eIF4E in cell lysates, in which eIF4G (or other 4E-BPs) is also present.

Identification of NC 4E-BMs in Thor and 4E-T. The immunoprecipitation assays shown in Fig. 1e–g indicate that similar to CUP, Thor and 4E-T contain NC motifs that interact with the lateral binding surface of eIF4E. In human 4E-BP1,2, the NC IPGVTS/T motif (located C-terminally to the C motif), increases the binding affinity of the proteins for eIF4E by approximately three orders of magnitude[19,22]. However, the IPGVTS/T motif is not conserved across the animal kingdom (Supplementary Fig. 2c). Nevertheless, several hydrophobic residues are present in the corresponding region in *Dm* Thor (residues Pro76–Pro84; Supplementary Fig. 2c).

To determine whether the Thor residues 76–84 constitute a NC 4E-BMs, we substituted Cys78, Leu79 and Leu80 with alanine (NC*) or deleted the motif (ΔNC, Supplementary Table 1). In

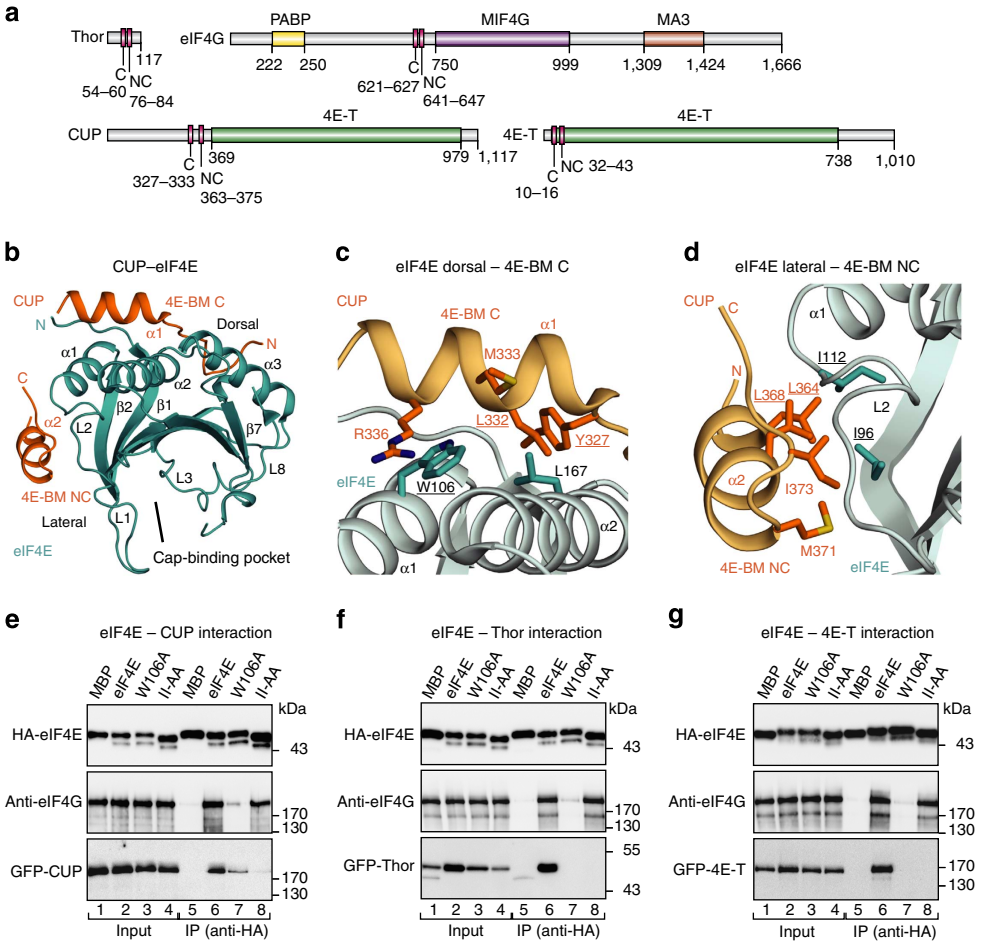


Figure 1 | 4E-BPs bind to a lateral surface of eIF4E. (**a**) Schematic representation of eIF4G and the 4E-BPs that were analyzed in this study. All the proteins contain a canonical (C) and a non-canonical (NC) 4E-BM. CUP and 4E-T contain a region with similarity to human 4E-transporter (4E-T) region. eIF4G contains a PABP-interacting region and MIF4G and MA3 domains. The amino-acid positions at the domain/motif boundaries are indicated below the protein outlines. (**b**) Cartoon representation of the overall structure of the eIF4E-CUP complex. eIF4E is shown in cyan and CUP in orange (PDB code 4AXG)[27]. Selected secondary structure elements are labeled in black for eIF4E and in orange for CUP. (**c**,**d**) Close-up views of the dorsal (**c**) and lateral (**d**) interfaces between eIF4E and CUP. Selected interface residues are shown as cyan and orange sticks for eIF4E and CUP, respectively. The eIF4E and CUP residues are labeled in black and orange, respectively, and underlined if they are mutated. (**e**–**g**) WB showing the interaction of HA-eIF4E (either WT or mutated) with GFP-tagged full-length 4E-BPs (CUP, Thor and 4E-T) and endogenous eIF4G. The size markers (kDa) are shown to the right of each panel. The original WB shown in this figure can be found in Supplementary Fig. 8.

the coimmunoprecipitation assays, the deletion of the Thor residues 76–84 abolished the interaction with eIF4E (Fig. 2a, lane 12), whereas the alanine substitutions decreased the eIF4E binding (Supplementary Fig. 1d, lane 9, NC*). By contrast, the substitution of the flanking residues Arg81, Gly82 and Thr83 by alanine was ineffective (Fig. 2a, lane 10). As a control, amino-acid substitutions in the C motif (C*, Supplementary Table 1) also disrupted the interaction with eIF4E (Fig. 2a and Supplementary Fig. 1d). Thus, the interaction of Thor with eIF4E requires both a C and a downstream NC motif in cell lysates.

In human 4E-T, sequences downstream of the C motif also contribute to the interaction with eIF4E[30]. Again, these sequences are not conserved in insects (Supplementary Fig. 2d). Nevertheless, based on the observation that in CUP and Thor, the NC motifs are located ~15–29 residues from the C motifs, are hydrophobic and, in the case of CUP, exhibit helical propensity, we inspected the *Dm* 4E-T sequence for motifs that fulfill these criteria. We identified a region in the insect 4E-T (residues 32–43) that could contain a potential NC motif and is located at a similar position as is the motif in the human protein (Supplementary Fig. 2d). In the coimmunoprecipitation assays, alanine substitutions or deletions of various residues in this motif (Supplementary Table 1) caused a drastic reduction in the 4E-T binding to eIF4E (Fig. 2b, lanes 10 and 11, and Supplementary Fig. 1e), similar to the disruption of the C motif (C*, Fig. 2b, lane 9, and Supplementary Fig. 2d). Thus, a NC 4E-BMs is also present in the *Dm* 4E-T that is conserved in *Drosophila* species.

4E-BPs and eIF4G display similar affinities for eIF4E. Next, we compared the binding efficiencies of the minimal eIF4E-binding regions of the 4E-BPs (C + NC, Supplementary Table 1) in pull-down assays. These regions were expressed with an amino-terminal MBP-tag and a C-terminal GB1-tag[31]. In parallel, we analyzed the minimal eIF4E-binding fragment of eIF4G (residues 578–650), which includes the C motif and the SDVVL motif that was identified in *Hs* eIF4G (corresponding to *Dm* VKNVSI, Supplementary Fig. 2e), which plays an auxiliary function in stabilizing the eIF4G interaction with eIF4E[23]. The bipartite C + NC regions of the three 4E-BPs and the eIF4G fragment pulled down the purified eIF4E at comparable levels (Fig. 2c).

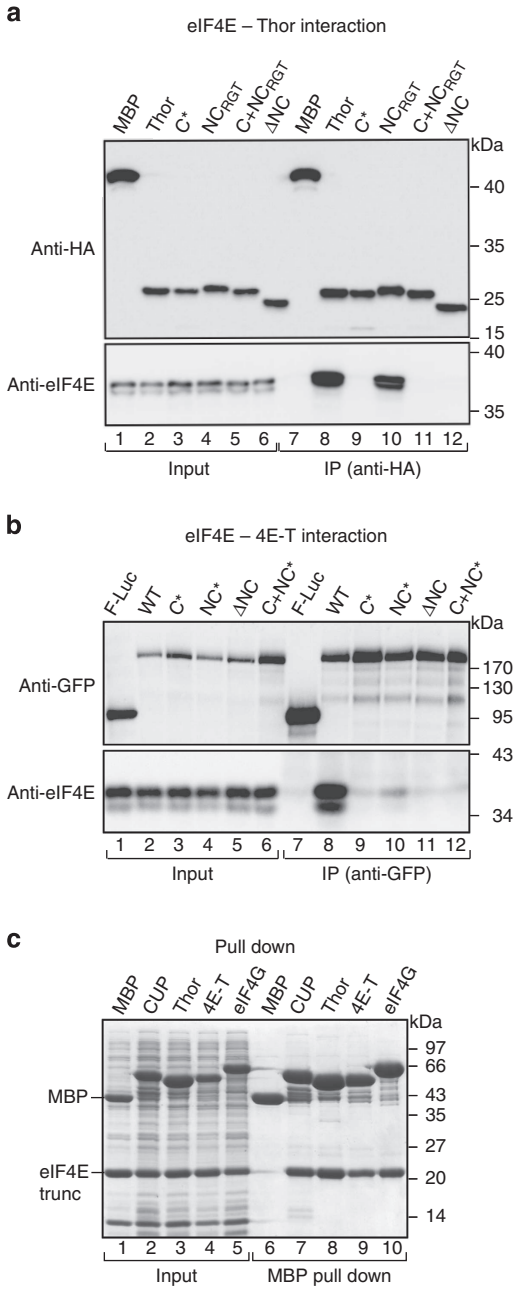


Figure 2 | Identification of non-canonical motifs in Thor and 4E-T. (**a**) WB showing the interaction of HA-Thor (full length, either WT or mutated) with endogenous eIF4E in S2 cells. The proteins were immunoprecipitated using anti-HA antibodies. The inputs (1%) and immunoprecipitates (30%) were analyzed by WB using anti-HA and anti-eIF4E antibodies. The original WB shown in this panel can be found in Supplementary Fig. 8. (**b**) WB showing the interaction of GFP–4E-T (full length, WT or mutated) with endogenous eIF4E. The proteins were immunoprecipitated using anti-GFP antibodies. The inputs (1%) and immunoprecipitates (30%) were analyzed by WB using anti-GFP and anti-eIF4E antibodies. The original WBs can be found in Supplementary Fig. 8. (**c**) MBP pull down showing the interaction of His_6-tagged eIF4E (residues 69–248, trunc) with MBP-tagged 4E-BP C + NC fragments (see Supplementary Table 1) and eIF4G (residues 578–650). All the fragments contained, in addition, a C-terminal GB1 tag. The input samples (10%) and bound fractions (15%) were analyzed using SDS-PAGE followed by Coomassie blue staining. The size markers (kDa) are shown to the right of each panel.

To obtain information on the affinities and thermodynamic parameters, we performed isothermal titration calorimetry (ITC) experiments. The bipartite regions of all three of the 4E-BPs and eIF4G exhibited comparable binding affinities for eIF4E, with dissociation constants (K_Ds) in the nanomolar range (Table 1 and Supplementary Fig. 3). The K_D values obtained for *Dm* Thor and eIF4G are comparable to those that have been reported for the human proteins[5,19,22,23,32]. Notably, although the binding of all proteins to eIF4E is enthalpically driven, the entropic penalties differ between these proteins, suggesting differences in the binding mechanisms. In particular, the interaction between CUP and eIF4E displayed the highest entropic penalty, which is indicative of a lower degree of conformational freedom in the bound state. Thus, CUP may undergo larger disorder-to-order transitions on binding, which is consistent with the formation of two α-helices[27]. 4E-T and Thor exhibited lower entropic penalties, suggesting a more dynamic conformation in the bound state.

To understand the contribution of the NC 4E-BMs to the affinity of 4E-BPs for eIF4E, we analyzed the binding of 4E-BP peptides containing only the C motifs or the complementary sequences comprising the linker (L) region between the two motifs and the NC motif (L + NC, Supplementary Fig. 4 and Supplementary Table 1). The affinities of the C motifs in isolation were one to three orders of magnitude lower than the C + NC peptides, indicating that the NC motifs contribute significantly to the overall affinity. Interestingly, the C motifs of all three 4E-BPs exhibited significant differences in binding affinities, with the affinity of the 4E-T peptide being approximately one and two orders of magnitude higher than those of the CUP and Thor peptides, respectively (Table 1 and Supplementary Fig. 4). The differences between 4E-BPs were more pronounced for the L + NC peptides, because only the CUP peptide interacted with eIF4E at detectable levels. The binding of the CUP peptide (L + NC) was enthalpically driven, with a K_D comparable to that of the C motif. These results indicate a similar contribution to the energetics of binding by the C and NC motifs of CUP.

Finally, we determined the affinities of the bipartite peptides (C + NC) for the eIF4E II-AA mutant. The affinities of CUP and Thor peptides were reduced by one and two orders of magnitude, respectively (Table 1 and Supplementary Fig. 5). In contrast, 4E-T binding was not significantly affected perhaps reflecting the higher affinity of its C 4E-BM. Similarly, the mutations in the lateral surface of eIF4E did not affect eIF4G binding.

We conclude that although 4E-BPs and eIF4G display similar affinities for eIF4E, they use different binding modes. These differences can be mainly attributed to the linker regions and the NC motifs, consistent with their sequence diversity, although differences in affinities for the C motifs were also detected. Moreover, the results of the ITC experiments also indicate that the affinity of 4E-BPs for eIF4E results from synergistic effects between the C and NC motifs.

4E-BP NC motifs are sufficient to bind eIF4E. To further analyze the binding modes of the 4E-BPs and eIF4G to eIF4E, we performed pull-down assays with recombinant proteins that were expressed in *Escherichia coli*. In contrast to the experiments in cell lysates, the *in vitro* pull-down assays allowed us to investigate the interactions of the individual proteins in the absence of other 4E-BPs, which could compete for binding and could obscure the interpretation of the results. We tested recombinant fragments of eIF4G, CUP and 4E-T and full-length Thor for binding to either the eIF4E wild-type (WT) or II-AA mutant (that is, with a disrupted lateral surface). eIF4G and the 4E-BPs pulled down

Table 1 | Thermodynamic parameters for the interaction of eIF4E with eIF4G and 4E-BP peptides.

	K_D (M)	ΔH (kcal mol^{-1})	TΔS (kcal mol^{-1})	ΔG (kcal mol^{-1})	Molar ratio
Peptide + eIF4E					
eIF4G 578-680	$17 \pm 13 \times 10^{-9}$	−18.5 ± 2	7.97	−10.51	0.99 ± 0.08
CUP C + NC	$9.1 \pm 0.5 \times 10^{-9}$	−34.5 ± 0.35	23.86	−10.78	1.05 ± 0.02
CUP C	$1.6 \pm 0.1 \times 10^{-7}$	−16.84 ± 0.04	7.70	−9.13	1.01 ± 0.01
CUP L + NC	$1.03 \pm 0.03 \times 10^{-7}$	−18.5 ± 0.2	9.10	−9.38	0.98 ± 0.01
Thor C + NC	$1.4 \pm 0.3 \times 10^{-9}$	−16.8 ± 1	4.90	−11.87	0.95 ± 0.02
Thor C	$2.26 \pm 0.06 \times 10^{-6}$	−12.1 ± 2.8	2.24	−9.82	1.06 ± 0.01
Thor L + NC	nb	nb	nb	nb	nb
4E-T C + NC	$5.6 \pm 2.4 \times 10^{-9}$	−22.8 ± 3	11.73	−11.11	0.95 ± 0.01
4E-T C	$1.6 \pm 0.2 \times 10^{-8}$	−18.6 ± 0.7	8.16	−10.45	0.95 ± 0.01
4E-T L + NC	nb	nb	nb	nb	nb
Peptide + eIF4E (II-AA)					
eIF4G 578-680	$40 \pm 9.5 \times 10^{-9}$	−16.24 ± 0.04	6.32	−9.93	1.03 ± 0.01
CUP C + NC	$5.0 \pm 0.8 \times 10^{-8}$	−18.6 ± 0.7	8.79	−9.84	0.98 ± 0.01
Thor C + NC	$4.7 \pm 0.3 \times 10^{-7}$	−7.6 ± 0.3	−0.91	−8.46	0.97 ± 0.01
4E-T C + NC	$8.8 \pm 2 \times 10^{-9}$	−12.9 ± 0.3	2.16	−10.80	1.03 ± 0.04

C, canonical; eIF4, eukaryotic initiation factor 4; L, linker; nb, no binding; NC, non-canonical.
See Supplementary Figs 3–5.

comparable amounts of WT and mutant eIF4E (Fig. 3a, lanes 9 and 10; Fig. 3b, lanes 10 and 11; and Fig. 3c,d, lanes 13 and 14), indicating that these proteins interact with the eIF4E mutant lacking a functional lateral binding surface *in vitro*. The results obtained *in vitro* contrast with the observation that the 4E-BPs did not interact with the II-AA mutant in cell lysates (Fig. 1e–g). One possible explanation for this difference is that cell lysates contain eIF4G, which blocks the dorsal surface of eIF4E, leaving only the lateral surface available for 4E-BPs. If the lateral surface is in addition mutated, then 4E-BPs may not be able to interact with eIF4E and displace bound eIF4G (see below).

The interaction of eIF4G and 4E-BPs with the eIF4E II-AA mutant is most likely mediated by their C motifs that bind to the dorsal surface of eIF4E. To confirm this assumption, we introduced mutations in the C motifs (C* mutants, Supplementary Table 1). Substitutions in the C motif of eIF4G abolished its interaction with either WT or mutant eIF4E (Fig. 3a, lanes 11 and 12). By contrast, the equivalent substitutions in the C motifs of 4E-BPs did not prevent their binding to eIF4E, reflecting a truly bipartite-binding mode (Fig. 3b, lane 12, and Fig. 3c,d, lane 15). However, the CUP and Thor C* mutants were strongly impaired in their binding to the eIF4E II-AA mutant (Fig. 3b, lane 13; Fig. 3c, lane 16), indicating that the C* mutants use the lateral surface of eIF4E. The 4E-T C* mutant showed reduced binding to both WT and mutant eIF4E (Fig. 3d, lanes 15 and 16).

Substitutions in the NC motifs (NC*) did not prevent the interaction of 4E-BPs with either WT or mutant eIF4E, most likely because the C motifs are sufficient for binding (Fig. 3b, lanes 14 and 15; Fig. 3c,d, lanes 17 and 18). The interaction of the three 4E-BPs with WT eIF4E was strongly reduced when the two motifs were mutated (C + NC*, Fig. 3b, lane 16, Fig. 3c,d, lanes 19). Remarkably, some residual binding to eIF4E was observed. These results suggest that the linker regions between the motifs in CUP and 4E-T and additional residues in Thor (which was full length) contact eIF4E and contribute to the interaction. The results obtained for the Thor NC* and C + NC* mutants were confirmed using a mutant with a deleted NC motif (ΔNC, Supplementary Fig. 6a).

Collectively, our results indicate that 4E-BPs interact with eIF4E using a bipartite-binding mode and recognize a lateral surface on eIF4E that is not used by eIF4G. Two main observations support these conclusions. First, mutations in the C motifs abolish the interaction of eIF4G but not of 4E-BPs with eIF4E. Second, mutations on the lateral surface of eIF4E abolish or reduce the binding of 4E-BPs to eIF4E when their binding to the dorsal surface is also compromised. Our results further indicate that the eIF4G residues downstream of the C motif, including the VKNVSI motif, do not use the binding surface centered at residues Ile96 and Ile112 and are not sufficient for binding to eIF4E when the C motif is mutated, which is in agreement with the proposed auxiliary role of these sequences[23]. Finally, it is important to note that although mutations in the C motifs of Thor and 4E-T do not disrupt binding to eIF4E, a deletion of the C motif prevents binding (L + NC peptides, see ITC experiments). These results suggest that mutations in the C motifs of these proteins do not completely abolish binding to the eIF4E dorsal surface, or that the formation of an α-helical structure (which is likely maintained in the mutants) is indirectly required to facilitate the binding of the linker region and NC motifs.

4E-BPs use the eIF4E lateral surface to compete with eIF4G. The observation that 4E-BPs can bind to the eIF4E II-AA mutant *in vitro* (that is, in the absence of competition) but not in cell lysates (that is, in the presence of eIF4G) suggests that 4E-BPs are not able to compete with eIF4G for binding to eIF4E when the lateral binding surface is disrupted.

To further investigate the role of the lateral binding surface of eIF4E in the competition mechanism, we performed competition assays using preassembled eIF4E–eIF4G complexes containing either eIF4E WT or the II-AA mutant and GST-tagged eIF4G (residues 578–650). eIF4G formed stable complexes both with WT and mutant eIF4E (Fig. 4a–c, lanes 4 and 5, respectively). These preassembled eIF4E–eIF4G complexes were challenged with increasing amounts of peptides containing the C and NC (C + NC) motifs of 4E-BPs or the same eIF4G fragment. Proteins that were associated with eIF4E were recovered by eIF4E pull-down assays.

The CUP, 4E-T and Thor C + NC peptides displaced eIF4G from the complex and associated with eIF4E (Fig. 4a–c, lanes 7–10 versus 6, Supplementary Figs 6b–d and 7a–c). The CUP and 4E-T peptides were able to effectively displace eIF4G when present at

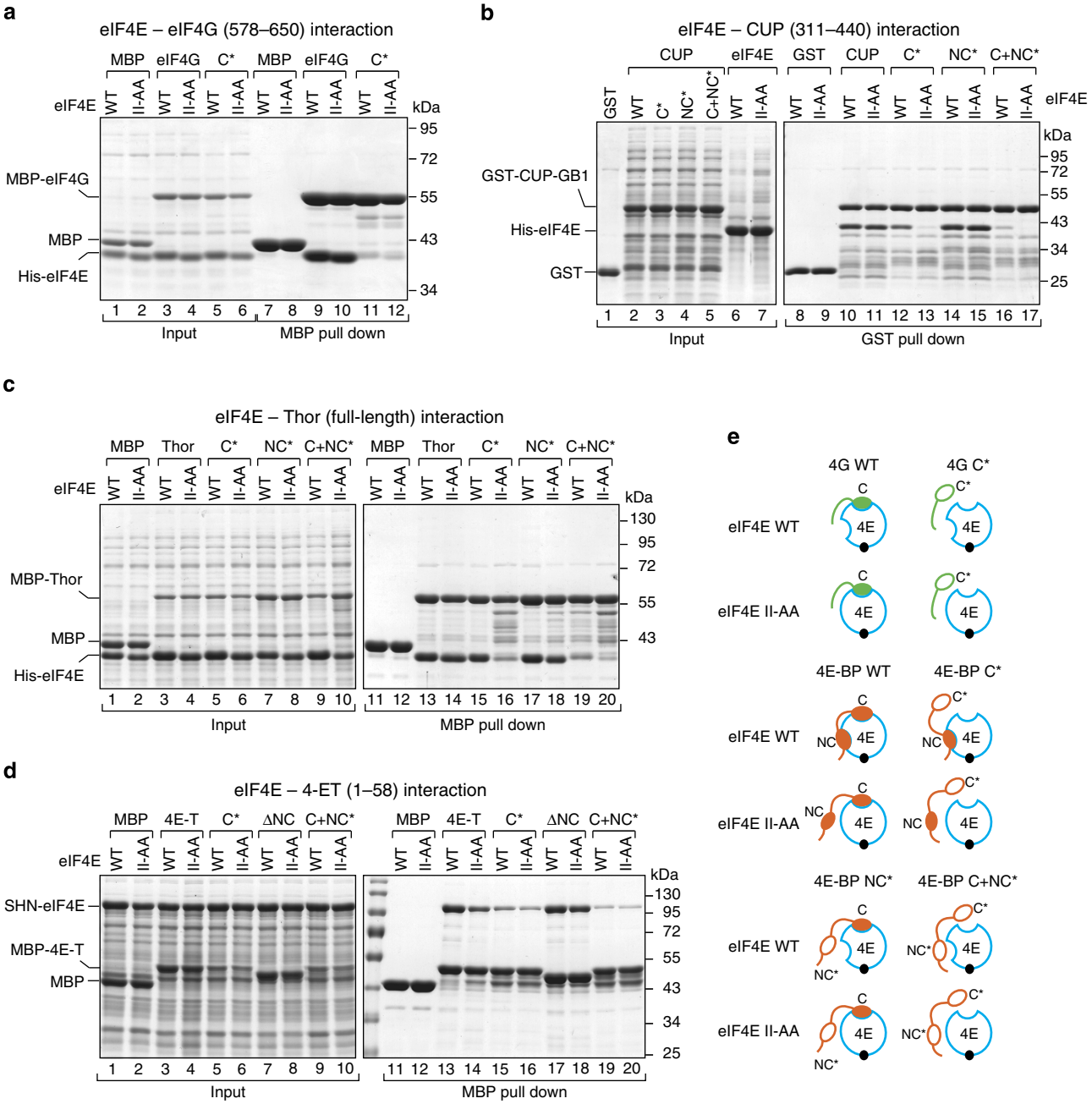
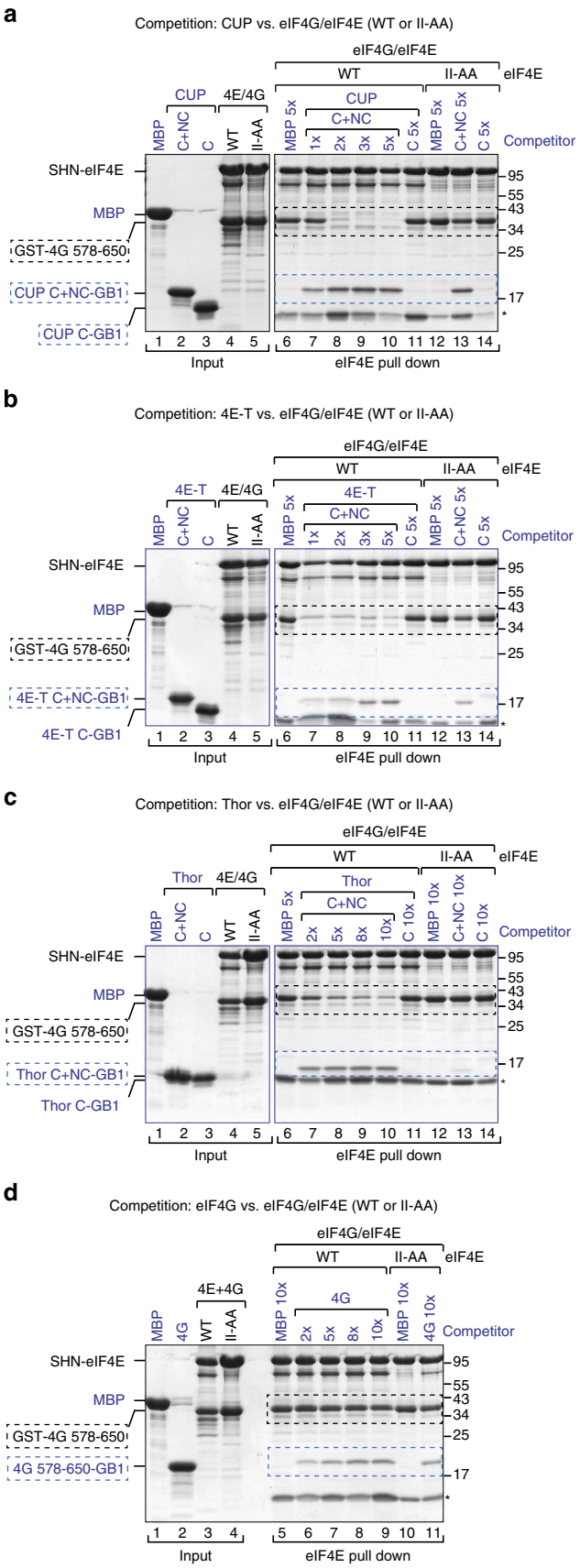


Figure 3 | 4E-BPs interact with eIF4E using a bipartite-binding mechanism. (**a**) MBP pull-down assay showing the interaction of His_6-eIF4E (full length, either WT or II-AA mutant) and MBP-eIF4G (residues 578–650; either WT or canonical 4E-BM mutant (C*)). The input (25%) and bound fractions (50%) were analyzed by SDS-PAGE followed by Coomassie blue staining. (**b**) GST pull-down assay showing the interaction of His_6-eIF4E (WT or II-AA mutant) and GST-CUP-GB1 (residues 311–440, WT or mutated in the canonical motif (C*), non-canonical motif (NC*) or both motifs (C + NC*)). (**c**) MBP pull down showing the association of MBP-Thor (full length, WT or 4E-BM mutants) with His_6-eIF4E (WT or II-AA mutant). The samples were analyzed as described in **a**. (**d**) MBP pull-down assay showing the interaction of MBP–4E-T (fragment 1–58, WT or the indicated mutants) with eIF4E (WT or II-AA mutant). eIF4E was expressed with a tag consisting of the streptavidin-binding peptide (strep), His_6 and the NusA protein (SHN tag). (**e**) Schematic representation of the different eIF4E–4E-BPs and eIF4E–eIF4G complexes that were analyzed in the pull-down assays. 4E, eIF4E (blue circle). Black circle, m^7GTP-cap structure. The 4E-BPs are shown in orange and the eIF4G in green. The asterisks indicate mutations in the corresponding motifs. These mutations are described in Supplementary Table 1.

two- and onefold molar excess, respectively. Under the same conditions, the 4E-BP C + NC peptides did not efficiently displace eIF4G from complexes that contained the eIF4E II-AA mutant (Fig. 4a–c, lanes 13 versus 12, Supplementary Figs 6b–d and 7a–c). Thus, binding to the lateral surface is required for 4E-BPs to effectively compete with eIF4G. In agreement with this conclusion, peptides containing only the 4E-BP C motifs did not displace eIF4G from eIF4E, although they were tested at the highest molar concentration (Fig. 4a–c, lanes 11 versus 10 and Supplementary Fig. 7a–c).

In striking contrast to the 4E-BP peptides, the eIF4G peptide hardly competed with GST-eIF4G for binding to eIF4E, irrespective of whether eIF4E was WT or mutated (Fig. 4d, lanes 5–11, Supplementary Figs 6b and 7d). Mechanistically, our results

indicate that 4E-BPs are more efficient competitors than is eIF4G and must bind to the lateral surface of eIF4E to effectively displace eIF4G from preassembled eIF4E–eIF4G complexes.



4E-BPs use the NC motifs to compete with eIF4G. Given that binding of 4E-BPs to the lateral surface of eIF4E is required for competition with eIF4G and that peptides containing only the 4E-BP C motifs cannot compete with eIF4G (Fig. 4), we next investigated the requirement for NC motifs. To this end, we performed competition assays using preassembled eIF4E–eIF4G complexes and excess 4E-BP peptides lacking either the C or NC motifs. The WT CUP C + NC peptide interacted with eIF4E and efficiently displaced preassembled eIF4G (Fig. 5a, lane 9 versus 6). Peptides containing either the C or the NC motifs of CUP did not compete with eIF4G (Fig. 5a, lanes 7 and 8), although these peptides bind to eIF4E in the absence of eIF4G (Fig. 5b, lanes 6 and 7), which is in agreement with the ITC experiments.

Similar results were obtained for Thor. Notably, deleting the non-canonical motif in the context of full-length Thor was sufficient to abolish its ability to compete with eIF4G (Fig. 5c, lane 10 versus 7), although in the absence of eIF4G this deletion mutant interacted with eIF4E (Supplementary Fig. 6a, lanes 17 and 18). Mutations in the canonical motif also abolished competition, as expected (Fig. 5c, lane 9). We conclude that 4E-BPs require both canonical and non-canonical motifs to compete with eIF4G for eIF4E binding. Thus, the non-canonical motifs play an essential role in the competition mechanism.

4E-BPs exhibit a kinetic competitive advantage over eIF4G. Given that the 4E-BPs and eIF4G display similar affinities for eIF4E, the differences in the ability to efficiently displace pre-bound eIF4G in competition assays are likely explained by the binding kinetics and the bipartite-binding mode. To obtain additional information on the ability of 4E-BPs to compete with eIF4G, we challenged preassembled eIF4E–eIF4G complexes with five- to tenfold molar excess of 4E-BP and eIF4G peptides and monitored the amount of eIF4G remaining bound to eIF4E over time.

In the absence of competitors, eIF4G remained bound to eIF4E, as expected (Fig. 5d, lane 4). In the presence of a tenfold molar excess of eIF4G peptide, we observed a 50% dissociation of prebound eIF4G after 4 h at 4 °C (Fig. 5d and Supplementary Fig. 7e). In the presence of a fivefold molar excess of CUP and 4E-T peptides (C + NC), we observed a 50% eIF4G dissociation in 2.5 ± 0.5 and 22 min, respectively, whereas the half-life of the eIF4E–eIF4G complexes in the presence of tenfold molar excess of Thor was 37 ± 9 min. (Fig. 5e–g, and Supplementary Fig. 7f). The simplest explanation of these results is that the bipartite-binding mode and the binding to an eIF4E surface that is not used by eIF4G confer on 4E-BPs a kinetic competitive advantage because they can bind preassembled eIF4E–eIF4G complexes without the need for prior eIF4G dissociation.

Figure 4 | 4E-BPs require binding to the lateral surface of eIF4E to compete with eIF4G. (**a**–**d**) Purified eIF4E-eIF4G complexes (2 μM) containing SHN-tagged eIF4E (full length, WT or II-AA mutant) and GST-eIF4G (residues 578–650) were incubated with increasing amounts of CUP (**a**), 4E-T (**b**), Thor (**c**) and eIF4G (**d**) peptides fused C terminally to GB1. The 4E-BP peptides contained the canonical and non-canonical motifs (C + NC) or only the C and are described in Supplementary Table 1. MBP served as a negative control. The proteins that were bound to eIF4E were pulled down using Strep-Tactin beads. The competitor peptides are labeled in blue, and their positions are highlighted by blue, dashed boxes. The black dashed boxes indicate the position of preassembled GST-eIF4G. Numbers above the lanes indicate fold molar excess of the competitor peptides. The corresponding quantification of the competition assays is shown in Supplementary Fig. 7a–d ($n = 2$). Asterisks indicated a contaminant protein.

eIF4G competes with 4E-BPs bound to the eIF4E II-AA mutant. Next, we asked whether eIF4G could compete with 4E-BPs when their binding to the lateral surface of eIF4E was disrupted. For this purpose, preassembled complexes containing eIF4E (WT or II-AA mutant) bound to GST-4E-BP fragments were challenged with excess amounts of MBP-eIF4G (residues 578–650). Proteins that were bound to eIF4E were recovered via eIF4E pull down. MBP-eIF4G did not displace CUP, Thor, 4E-T or eIF4G bound to WT eIF4E (Fig. 6a, lane 6 versus 5, and Fig. 6b, lanes 8, 10 and 14). In contrast, MBP-eIF4G partially displaced CUP (Fig. 6a, lane 8 versus 7) and completely displaced full-length Thor (Fig. 6b, lane 12 versus 11) bound to the eIF4E II-AA mutant. These observations indicate that eIF4G can compete with 4E-BPs for binding to eIF4E only when their interaction with the lateral surface of eIF4E is impaired. Thus, the dissociation of 4E-BPs from the lateral surface of eIF4E (for instance, on phosphorylation) may be sufficient for their dissociation from eIF4E to allow eIF4G to resume translation (Fig. 6c).

The non-canonical motifs mediate translational repression. To determine the role of non-canonical motifs in translational repression, we tested whether 4E-BPs repressed the expression of a firefly luciferase (F-Luc) reporter when coexpressed in S2 cells. A short uncapped and unadenylated RNA served as a transfection control (control RNA). To rule out the possibility that the inhibition of F-Luc expression resulted from changes in the F-Luc mRNA levels, we analyzed these levels by northern blotting and determined translation efficiencies (Fig. 7a,b).

The CUP N-terminal fragment or full-length Thor inhibited the expression of the F-Luc reporter in a dose-dependent manner (Fig. 7a–d). 4E-T caused mRNA degradation when overexpressed and was excluded from the analysis (C.I. and E.I., unpublished results). Mutations in either the canonical or non-canonical motifs as well as the combined mutations suppressed CUP- and Thor-mediated repression (Fig. 7a,b). The mutant proteins were expressed at levels that were comparable to the highest tested level for the WT protein (Fig. 7c,d, WB). Thus, both the canonical and non-canonical motifs are required for Thor and CUP to repress translation in a cellular context, which is in agreement with the competition assays.

The non-canonical motifs regulate eIF4E localization. 4E-BPs are nucleocytoplasmic shuttling proteins that transport eIF4E to the nucleus[26,33–35]. Although eIF4E nuclear functions are not clearly understood, the nuclear retention/import of eIF4E could contribute to the efficient inhibition of cap-dependent translation. In addition, human 4E-T can also induce the accumulation of eIF4E in mRNA processing bodies or P-bodies[36]. To determine whether the non-canonical motifs contribute to the regulation of eIF4E subcellular distribution mediated by 4E-BPs, we analyzed the localization of endogenous eIF4E by immunofluorescence in S2 cells expressing WT or mutant 4E-BPs (Fig. 8).

At a steady-state, CUP and Thor distributed evenly throughout the cytoplasm (Fig. 8a,e). By contrast, 4E-T accumulated in cytoplasmic foci, which correspond to P-bodies as judged by the colocalization with the P-body marker Trailer hitch (Fig. 8i and Supplementary Fig. 7g). Endogenous eIF4E was also evenly distributed in the cytoplasm in cells overexpressing WT CUP and Thor as well as the mutant versions of these proteins (Fig. 8a–h, middle panels). In contrast, in cells expressing 4E-T, eIF4E was detected in P-bodies (Fig. 8i). Thus, 4E-T can drag eIF4E into P-bodies. Accordingly, the number of eIF4E-positive P-bodies was reduced in cells overexpressing 4E-T mutants (C*, NC*, C + NC*; Fig. 8j–l), although the mutants still localized to P-bodies. Thus, both the canonical and non-canonical motifs of 4E-T are required to induce the accumulation of eIF4E in P-bodies.

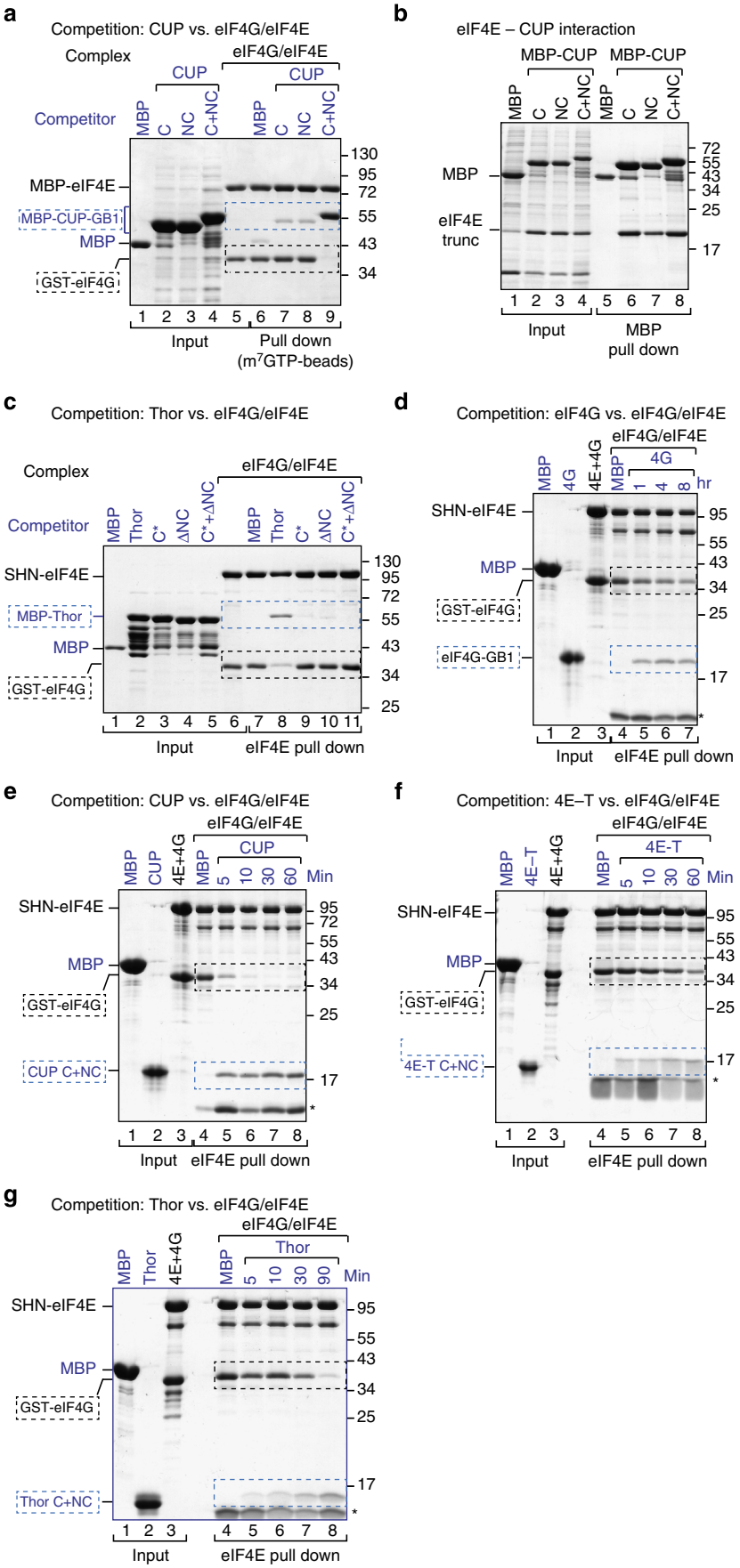
Next, we treated S2 cells with Leptomycin B (LMB), a drug that inhibits nuclear export by CRM1, which has been shown to export 4E-BPs[26,33,37]. The LMB treatment induced the nuclear accumulation of CUP and 4E-T proteins (Fig. 8m,u) and a partial nuclear accumulation of Thor (Fig. 8q). Concomitantly, endogenous eIF4E accumulated in the nucleus (Fig. 8m,q,u, middle panels). eIF4E nuclear accumulation was dependent on binding to the 4E-BPs because this accumulation was strongly reduced in cells expressing the 4E-BP mutants (Fig. 8n–p,r–t,v–x). None of the 4E-BPs required binding to eIF4E to translocate to the nucleus in the LMB-treated cells (Fig. 8m–x, left panels). Taken together, our data indicates that both the canonical and non-canonical motifs are required for 4E-BPs to regulate eIF4E subcellular distribution.

Discussion

In this study, we show that similar to CUP, Thor and 4E-T employ a bipartite interface that is composed of canonical and non-canonical motifs to bind to the dorsal and lateral surfaces of eIF4E, respectively. While the dorsal binding surface of eIF4E is also used by eIF4G[5,7,17], the lateral binding surface is only used by 4E-BPs and is required for 4E-BPs to displace eIF4G from preassembled eIF4E–eIF4G complexes. Based on these results, we propose that the lateral surface of eIF4E provides an exclusive docking surface for 4E-BPs on eIF4E–eIF4G complexes. After docking, 4E-BPs can begin to displace eIF4G by establishing interactions with the eIF4E dorsal surface via their own canonical motifs, further stabilizing their association with eIF4E (Fig. 6c).

The ability to bind laterally to the side of eIF4E that is not used by eIF4G enable 4E-BPs to displace eIF4G even when their binding affinities are similar and under conditions in which 4E-BPs are not in great excess compared with eIF4G. Indeed, by docking to preassembled eIF4E–eIF4G complexes, the 4E-BPs increase their local concentration and can rapidly dissociate bound eIF4G, inhibiting ongoing translation. Our model also provides one possible explanation for why eIF4G is a poor competitor compared with 4E-BPs. Indeed, eIF4G will not bind eIF4E unless the prebound eIF4G or 4E-BPs dissociate. In this

Figure 5 | 4E-BPs require the non-canonical motifs to compete with eIF4G. (**a**) Purified eIF4E–eIF4G complexes were incubated with fivefold molar excess amounts of CUP peptides containing either the C, the NC or both motifs (C + NC). The peptides were fused N terminally to MBP and C terminally to GB1. The eIF4E–eIF4G complexes contained MBP-tagged eIF4E and GST-tagged eIF4G (residues 578–650). The eIF4E-bound proteins were pulled down using m^7GTP-sepharose beads and analyzed by SDS–PAGE. (**b**) MBP pull down showing the interaction of purified eIF4E (69–248) with the CUP fragments shown in **a**. (**c**) Purified eIF4E–eIF4G complexes were incubated with fivefold molar excess of MBP-Thor (full length, either WT or the indicated mutants). The eIF4E-bound proteins were pulled down using Strep-Tactin beads and analyzed as described in **a**. (**d**–**g**) Purified eIF4E–eIF4G complexes (1 μM) containing SHN-eIF4E (full length) and GST-eIF4G (residues 578–650) prebound to Strep-Tactin beads were incubated with a 10-fold molar excess of eIF4G (residues 578–650, **d**) and Thor (**g**) or a fivefold molar excess of CUP (**e**) and 4E-T (**f**) peptides fused C terminally to GB1. The 4E-BP peptides contained the C + NC motifs. Proteins bound to eIF4E were recovered at the indicated time points. In all of the panels, the competitor proteins are labeled in blue and highlighted by blue, dashed boxes. The black, dashed boxes mark the position of preassembled GST-eIF4G. Quantification of the dissociation assays is shown in Supplementary Fig. 7e,f. Each experiment was repeated at least twice.

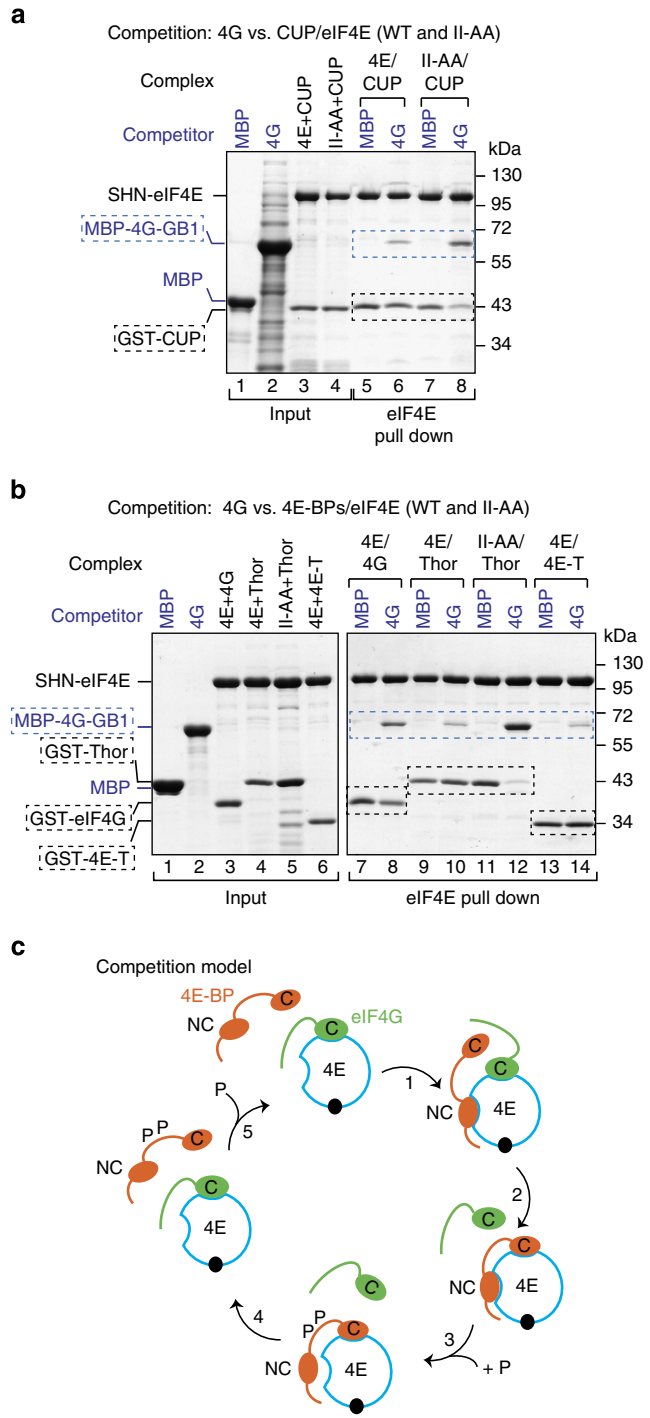


context, it will be of interest to determine the contribution of the canonical and non-canonical motifs to the association (K_{on}) and dissociation (K_{off}) rate constants of 4E-BP proteins.

How can eIF4G bind back to eIF4E to resume translation? We show that eIF4G can displace 4E-BPs when their binding to the lateral surface of eIF4E is impaired. Although in our studies this interaction was impaired by mutations, *in vivo* this impairment could be achieved by posttranslational modifications such as phosphorylation. Indeed, it is well established that the phosphorylation of 4E-BPs reduces their affinity for eIF4E[2,8]. Thus, it will be of interest to dissect the impact of phosphorylation on the interaction of 4E-BPs with either the lateral or dorsal surfaces of eIF4E.

Owing to their lack of conservation, it has remained unclear whether non-canonical motifs are present in all 4E-BPs. Our data indicate that the non-canonical motifs are intrinsic to the ability of 4E-BPs to compete with eIF4G and thus are likely to be present in all 4E-BPs that repress translation. At the functional level, non-canonical motifs have been proposed to play an auxiliary role and have been mainly implicated in the regulation of the affinity of eIF4E for the mRNA cap structure through allosteric effects[7,17,27,28,38]. Specifically, the binding of the 4E-BP1,2 non-canonical motifs to eIF4E increases the affinity for the cap structure[19,22,38]. Here, we show that the non-canonical motifs are essential, not auxiliary, for 4E-BP function in inhibiting translation. Given the diversity of non-canonical motifs and their different modes of interaction with eIF4E, it is possible that their binding to the lateral surface of eIF4E modulates the affinity for the cap in different ways, thereby mediating different effects. For example, by increasing the affinity of eIF4E for the cap structure, 4E-BPs may stabilize translationally repressed mRNA targets as observed for CUP[39]. Alternatively, by decreasing the affinity of eIF4E for the mRNA cap, 4E-BPs may destabilize the repressed mRNA target through decapping and subsequent decay.

In summary, our current understanding of 4E-BPs role in translational repression is predominantly based on the study of the low-molecular-weight 4E-BPs of the 4E-BP1–3 family. The identification of additional, high-molecular-weight 4E-BPs together with the characterization of their interaction mode with eIF4E reveals an unexpected sequence diversity of the eIF4E-binding regions and of the functional mechanisms. The functional diversity of 4E-BPs is further enhanced by the presence of additional domains in the high-molecular-weight 4E-BPs. These additional domains link eIF4E binding with other cellular processes, such as mRNA decay, as described for CUP and 4E-T[36,39]. Understanding the molecular basis for the interaction of diverse 4E-BPs with eIF4E will provide valuable insight into the variety of mechanisms that are employed by these proteins to regulate gene expression. These studies promise to uncover novel therapeutic strategies to selectively target dysregulated translation in cancer.



Methods

DNA constructs. The plasmids expressing the luciferase reporters, control RNA and GFP- or HA-tagged eIF4E, eIF4G, Tral and CUP (WT or mutated) have been previously described[39–42]. The plasmids expressing HA-Thor-V5 and GFP-Thor were obtained by inserting the corresponding DNA into the EcoRV and XhoI sites of the pAc5.1-λN-HA or pAc5.1-GFP vectors, respectively. A plasmid expressing HA-4E-T was obtained by inserting the corresponding DNA (CG32016 isoform B)

Figure 6 | eIF4G competes with 4E-BPs when their binding to the lateral surface of eIF4E is impaired. (**a**,**b**) Purified complexes containing SHN-tagged eIF4E bound to either eIF4G or the indicated 4E-BPs were incubated with MBP or MBP-eIF4G-GB1. The amount of eIF4G or 4E-BP proteins that were associated with eIF4E was determined by pull down using Strep-Tactin beads. The complexes contained GST-eIF4G (residues 578–650), GST-CUP (residues 311–440), GST-Thor (full length) or GST-4E-T (residues 1–58). (**c**) Competition model: eIF4E (blue circle) contains a dorsal and a lateral surface that bind to the C and NC motifs of 4E-BPs (shown in orange), respectively. The dorsal surface also binds to the canonical motif of eIF4G (shown in green). The eIF4E lateral binding surface provides a docking site for the non-canonical motifs of 4E-BPs when eIF4G is bound to the dorsal surface of eIF4E via its canonical motif (1). After docking, 4E-BPs displace eIF4G from the dorsal surface of eIF4E and repress translation (2). Phosphorylation (P) of 4E-BPs destabilizes their association with eIF4E (3). Therefore, eIF4G can bind to eIF4E and translation resumes (4). In humans, 4E-BP1–3 the phosphorylation sites are located in the linker region between the 4E-BMs and in the sequences N-terminal to the canonical motif (not shown). Dephosphorylation of 4E-BPs is required for binding to eIF4E (5). Symbols are as in Fig. 3e.

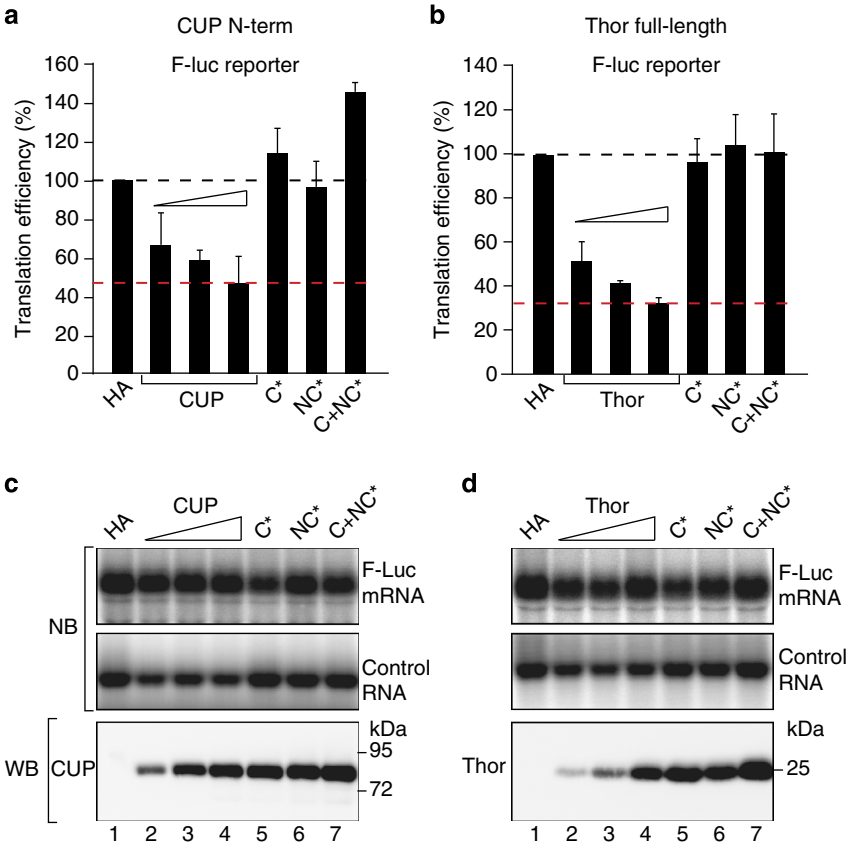


Figure 7 | Role of non-canonical motifs in translational repression. (**a**,**b**) S2 cells were transfected with plasmids expressing a F-Luc reporter (F-Luc) and HA-CUP (N-Term, residues 1–402) or HA-Thor (full length, either WT or the indicated mutants). The F-Luc mRNA levels were analyzed by northern blotting and were normalized to those of the control RNA. The normalized F-Luc mRNA levels were used to normalize the F-Luc activity to obtain the translation efficiencies, which were set to 100 in cells expressing the HA peptide. The mean values ± s.d. from three independent experiments are shown. The red, dashed lines indicate the F-Luc levels for the maximum repressive activity that was exhibited by the 4E-BPs. The black, dashed lines indicate the F-Luc levels expected in the absence of repression. (**c**,**d**) Upper panels show northern blot (NB) analyses of representative RNA samples corresponding to the experiments that are shown in **a**,**b**, respectively. The lower panels show the expression of the 4E-BP proteins analyzed by WB. The original northern and WB can be found in Supplementary Fig. 9.

into the EcoRI and NotI restriction sites of the pAc5.1-λN-HA vector. For expression in *E. coli*, the DNA encoding Thor (full length) and 4E-T (residues 1–58) was inserted into the XhoI-MfeI and AflII-NotI sites, respectively of the pnEA-NvM vector[43] (which provides an N-terminal MBP tag followed by a TEV protease cleavage site). A DNA fragment coding eIF4G (residues 578–650) was inserted into the XhoI and BamHI restriction sites of the pnEA-NvM or pnEA-NvG (which provides an N-terminal GST tag) vector[43].

A DNA fragment encoding full-length *Dm* eIF4E was inserted into the NdeI-BamHI restriction sites of the pnEK-NvH vector (which provides an N-terminal hexa-histidine (His_6) tag) or the pnEK-NvSHN vector (which provides an N-terminal Strep-NusA-His-tag). The DNA fragments encoding CUP, Thor and 4E-T minimal eIF4E-binding fragments (C + NC), the individual canonical (C) and non-canonical motifs (NC, in the case of CUP), and the L + NC peptides were cloned into the NdeI-NheI restriction sites of the pnEA-NpM vector with an N-terminal MBP tag followed by an HRV3C protease cleavage site[43]. The DNA encoding the B1 domain of immunoglobulin-binding protein G (GB1)[31] was inserted C terminally into the described fragments by site-directed insertion using the QuikChange mutagenesis kit (Stratagene). The DNA encoding a truncated eIF4E protein (residues 69–248) was cloned into the NdeI-NheI restriction sites of the pnEA-NpH vector (which provides an N-terminal His_6-tag followed by a HRV3C protease cleavage site)[43]. All the mutants were generated by site-directed mutagenesis using the QuikChange mutagenesis kit (Stratagene) and the oligonucleotide sequences provided in Supplementary Table 3. All the constructs and mutations were confirmed by sequencing and are listed in Supplementary Table 1. The plasmids expressing full-length His_6-eIF4E and GST-CUP (residues 311–440) were kindly provided by F. Bono[27].

Coimmunoprecipitation assays and western blotting. The coimmunoprecipitations assays were performed as previously described[41]. For the pull downs using m^7GTP beads, 25 µl of immobilized γ-aminophenyl-m^7GTP (C_{10}-spacer—Jena Bioscience) beads was added to the cell lysates and the mixtures were rotated for 1 h at 4 °C. The beads were washed three times with NET buffer (50 mM Tris (pH 7.4), 150 mM NaCl, 1 mM EDTA and 0.1% Triton X-100). The bound proteins were eluted with 2 × SDS–polyacrylamide gel electrophoresis (SDS–PAGE) sample buffer and analyzed by western blotting (WB). All of the WB experiments were developed with the ECL western blotting detection system (GE Healthcare) as recommended by the manufacturer. The antibodies used in this study are listed in Supplementary Table 2.

Protein expression and purification. Unless indicated otherwise, all the proteins were expressed in *E. coli* BL21 Star (DE3) cells (Invitrogen) that were grown in LB medium overnight at 20 °C. The lysis buffers were supplemented with DNaseI (5 µg ml^{-1}), lysozyme (1 mg ml^{-1}) and protease inhibitor cocktail (Roche). The truncated His_6-eIF4E (residues 69–248) that was used in the ITC experiments and in Figs 2c and 5b was purified in lysis buffer containing 50 mM HEPES (pH 7.2), 300 mM NaCl, 20 mM imidazole and 2 mM β-mercaptoethanol using Ni^{2+}-affinity chromatography (HisTrap HP 5 ml, GE Healthcare) and eluted with a gradient of 20–500 mM imidazole. After the cleavage of the His_6-tag with HRV3C protease (homemade), the protein was further purified using a heparin column (HiTrap Heparin HP 5 ml, GE Healthcare), followed by size exclusion chromatography (Superdex 75 16/60, GE Healthcare) in 20 mM Na-phosphate (pH 7.0), 300 mM NaCl and 2 mM dithiothreitol (DTT).

To obtain the preassembled eIF4E–eIF4G complexes used in Figs 4 and 5d–g, full-length eIF4E (WT or the II-AA mutant) containing an N-terminal SNH-tag was coexpressed with an N-terminal GST-tagged eIF4G (residues 578–650). The cells were lysed by sonication in lysis buffer containing 50 mM HEPES (pH 7.2), 300 mM NaCl and 2 mM DTT. The complexes were purified from the cleared lysates using Protino Glutathione Agarose 4B beads (Machery-Nagel). The complex was further purified using a heparin column (HiTrap Heparin HP 5 ml, GE Healthcare) and a

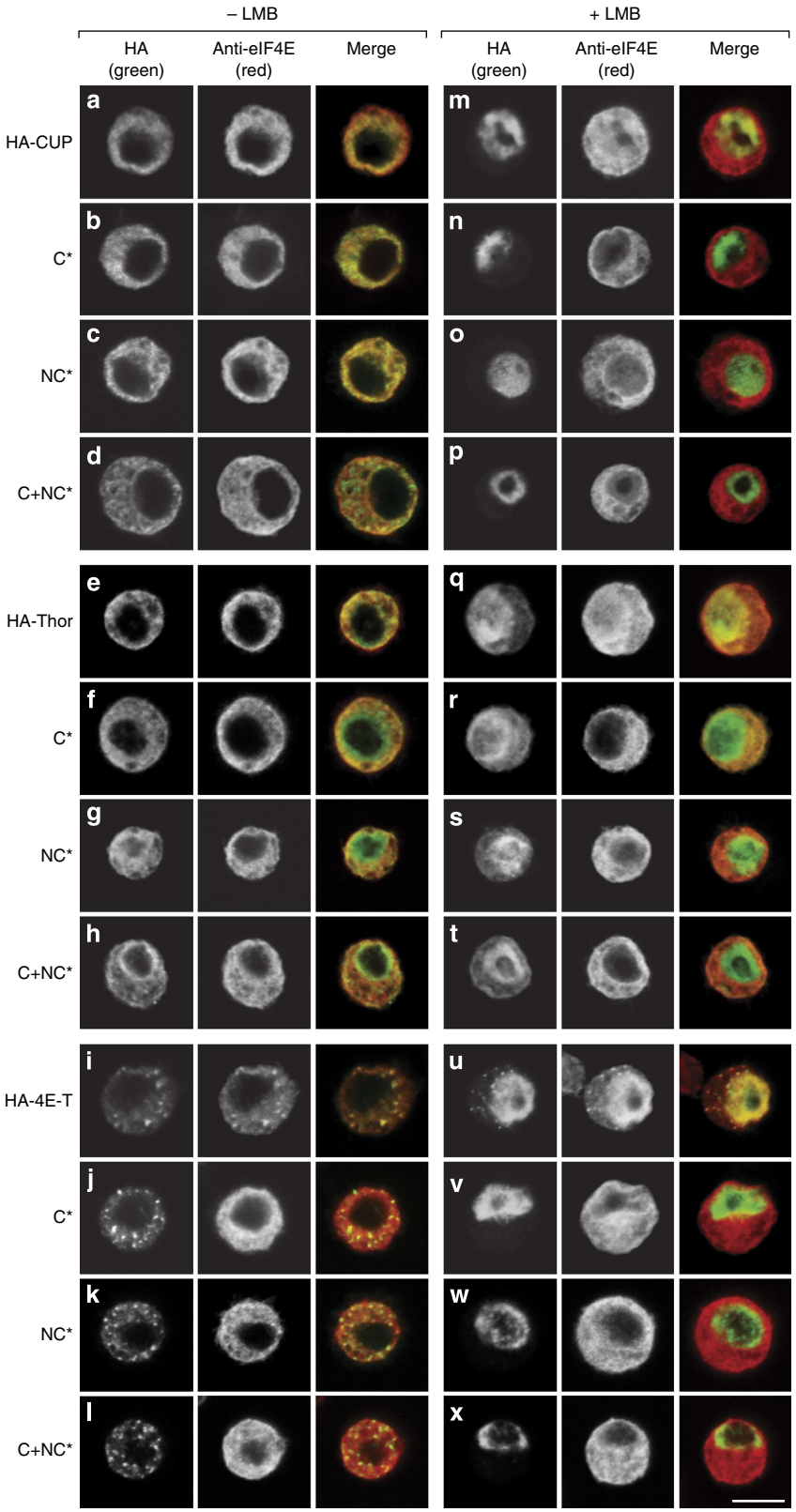


Figure 8 | The non-canonical motifs are required for 4E-BP-mediated eIF4E nuclear import. (**a**–**x**) S2 cells expressing HA-tagged versions of CUP (**a**–**d**; **m**–**p**), Thor (**e**–**h**; **q**–**t**) or 4E-T (**i**–**l**; **u**–**x**) or the corresponding 4E-BMs mutants (indicated on the left) were treated with LMB for 12 h (+ LMB) or with methanol as control (− LMB). The cells were fixed and the localization of the HA-tagged proteins and endogenous eIF4E was determined by indirect immunofluorescence using anti-HA and anti-eIF4E antibodies. The merged pictures show the HA signal in green and the eIF4E signal in red. Scale bar, 5 μm.

final size exclusion chromatography (Superdex 200 16/60, GE Healtcare) in 20 mM HEPES (pH 7.2), 200 mM NaCl and 2 mM DTT.

For the ITC measurements and the competition assays shown in Figs 4 and 5d–g, the 4E-BP peptides corresponding to the canonical (C) motifs, the combined (C + NC) motifs or L + NC were expressed with an HRV3C cleavable N-terminal MBP-tag and a non-cleavable C-terminal GB1 domain. The cells were lysed by sonication in lysis buffer containing 50 mM HEPES (pH 7.2), 300 mM NaCl and 2 mM DTT. The proteins were purified from the cleared lysates using amylose resin (New England Biolabs) followed by the cleavage of the MBP tag with HRV3C protease overnight at 4 °C. The proteins were further purified by size exclusion chromatography (Superdex 75 16/60, GE Healthcare) in 20 mM Na-phosphate (pH 7.0), 300 mM NaCl and 2 mM DTT. For the eIF4G fragment (residues 578–650) the MBP was removed after cleavage with TEV protease through an additional anion exchange chromatography (HiTrap Q HP 5 ml, GE Healthcare) before the final gel filtration.

Protein pull-down assays. For the pull-down assays shown in Fig. 3 and Supplementary Fig. 6a, the full-length His_6- or SHN-tagged eIF4E (WT or II-AA mutant) was coexpressed with GST or MBP-tagged protein fragments, including CUP (residues 311–440), Thor (full length), 4E-T (residues 1–58) or eIF4G (residues 578–650) in *E. coli* BL21 (DE3) STAR cells in autoinducing medium[44] overnight at 20 °C. The cells were resuspended in lysis buffer (20 mM Tris–HCl (pH 7.5), 100 mM NaCl, 1 mM DTT) supplemented with EDTA-free protease inhibitor (Roche) and lysed by sonication. The cleared lysates were incubated for 1 h with 20 µl of Protino Glutathione Agarose 4B beads (Macherey-Nagel) or amylose resin (New England Biolabs). The beads were washed three times with lysis buffer, and the bound proteins were eluted with lysis buffer containing 25 mM L-glutathione or 25 mM maltose for 15 min. The proteins were analyzed by 10–15% SDS–PAGE followed by Coomassie blue staining.

In Fig. 2c, eIF4G (residues 578–650) and the C + NC peptides of CUP, Thor and 4E-T were expressed with a N-terminal MBP and a C-terminal GB1 tag. The bacterial cells were resuspended in 5 ml of lysis buffer containing 50 mM HEPES (pH 7.2), 200 mM NaCl and 2 mM DTT and lysed by sonication. Purified eIF4E (residues 69–248) was added to the cleared lysates (40–80 µl), adjusted to 0.3 ml with lysis buffer and incubated with 30 µL of amylose resin for 1 h at 4 °C. The beads were washed three times with lysis buffer and eluted with 64 µl of the same buffer containing 25 mM maltose. The proteins were analyzed by 15% SDS–PAGE followed by Coomassie blue staining.

Competition assays. For the competition assays shown in Figs 5a,c and 6a,b, complexes containing SHN- or MBP-tagged eIF4E (WT or II-AA mutant) bound to GST-eIF4G (residues 578–650), GST-Thor (full length), GST-4E-T (residues 1–58) or GST-CUP (311–440) were obtained by coexpressing the corresponding proteins in *E. coli* BL21 (DE3) STAR cells (30 ml culture). The cells were resuspended in lysis buffer (5 ml) that was supplemented with EDTA-free protease inhibitor (Roche) and 1 mg ml^{-1} lysozyme and lysed by sonication. The cleared lysates were incubated with 400 µl of Protino Glutathione Agarose 4B (Macherey-Nagel) for 1 h. The beads were washed three times with lysis buffer, and the proteins were eluted after 10 min of incubation with lysis buffer containing 25 mM L-glutathione. The protein complexes were stored at − 20 °C or used in competition assays.

The purified recombinant complexes were mixed with excess amounts of the indicated purified competitor proteins (Fig. 5c) or with bacterial lysates expressing the competitor proteins (Figs 5a and 6) and incubated for 30 min at 4 °C. After incubation, 20 µl of immobilized γ-Aminophenyl-m^7GTP or Strep-Tactin Sepharose (IBA), were added to the samples and incubated for another 40 min at 4 °C. The beads were washed three times with lysis buffer and eluted with lysis buffer containing 2.5 mM desthiobiotin (Strep-Tactin Sepharose) or with 20 µl of SDS–PAGE loading buffer (Aminophenyl-m^7GTP beads). The proteins were analyzed by 10–15% SDS–PAGE followed by staining with Coomassie blue staining.

For the titration experiments shown in Fig. 4, 2 µM of purified complexes containing SHN-eIF4E (WT or II-AA mutant) bound to GST-eIF4G (residues 578–650) were incubated with increasing amounts (2–20 µM) of purified competitor proteins for 20 min at 4 °C. The eIF4E-bound proteins were recovered via Strep-Tactin Sepharose pull down and eluted with lysis buffer containing 2.5 mM desthiobiotin. The proteins were analyzed by 15% SDS–PAGE followed by staining with Coomassie blue staining.

In the 'kinetic assays' shown in Fig. 5d–g, the purified complexes containing SHN-eIF4E and GST-eIF4G (578–650; 1 µM) were incubated with Strep-Tactin beads for 20 min. The prebound complex was then challenged with 5 µM (CUP and 4E-T) or 10 µM (Thor and eIF4G) of competitor proteins for the indicated time points. The eIF4E-associated proteins were pulled down, eluted and analyzed as described above.

ITC analysis. The ITC experiments were performed using a VP-ITC microcalorimeter (MicroCal) at 20 °C. The solution of eIF4E (residues 69–248, WT or mutant: 1–20 µM) in the calorimetric cell was titrated with tenfold concentrated solutions of GB1-stabilized peptides corresponding to 4E-BPs C + NC (10 µM), C (50 µM), L + NC (100 µM CUP, 200 µM Thor and 4E-T) or eIF4G (residues 578–650, 20 µM) that were dissolved in the same buffer (20 mM Na-phosphate (pH 7.0) and 150 mM NaCl). The titration experiments consisted of an initial injection of 2 µl followed by 28 injections of 10 µl at an interval of 240 s. Each binding experiment was repeated twice. The thermodynamic parameters were estimated using a one-site binding model (Origin version 7.0), whereby the datapoint of the first injection was removed for the analysis[45].

Translation repression assays. S2 cells were transfected in 6-well plates using Effectene transfection reagent (Qiagen) according to the manufacturers protocol. The transfection mixtures contained: 0.1 µg of F-Luc reporter plasmid (F-Luc-V5), 0.3 µg of control RNA reporter, and increasing amounts of plasmids expressing HA-CUP (fragment 1–402; 0.05–0.2 µg) and HA-Thor (full length, 0.1–0.5 µg). The plasmids expressing the corresponding mutants or the HA peptide control were transfected at the highest concentration. In all the experiments, the cells were collected three days after transfection. The F-Luc activity was measured using the Dual-Luciferase reporter assay system (Promega). The northern blotting was performed as previously described[42]. The F-Luc mRNA levels were determined by northern blotting and were normalized to those of the control RNA. The normalized F-Luc mRNA levels were then used to normalize the F-Luc activity, to obtain translation efficiencies.

Immunofluorescence. S2 cells expressing HA-tagged versions of CUP, Thor and 4E-T or the indicated mutants were treated with Leptomycin B (100 nM; Sigma) or methanol as a control for 12 h. After the LMB treatment, the cells were allowed to adhere to poly-D-lysine–coated coverslips for 15 min and were fixed with 2% paraformaldehyde for 10 min. The cells were then permeabilized with 0.1% Triton X-100 in PBS (10 min) and stained with affinity-purified monoclonal anti-HA (Covance 1:1,000) and polyclonal anti-eIF4E (1:2,000) antibodies in PBS containing 1% BSA for 1 h. Alexa Fluor 594-labeled goat anti-rabbit and 488-labeled anti-mouse antibodies (Invitrogen) were used at dilutions of 1:1,000 and 1:2,000, respectively. The cells were mounted using Fluoromount-G (Southern Biotech). The images were acquired at room temperature using a confocal microscope (TCS SP2; Leica) that was fitted with a Plan-Apochromat × 100 NA 1.40 oil immersion objective and a series of three photomultipliers (Hamamatsu Photonics) controlled with the Leica confocal software (version 2.61). The images were prepared using Photoshop (Adobe).

References

1. Jackson, R. J., Hellen, C. U. & Pestova, T. V. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat. Rev. Mol. Cell Biol.* **11,** 113–127 (2010).
2. Richter, J. D. & Sonenberg, N. Regulation of cap-dependent translation by eIF4E inhibitory proteins. *Nature* **433,** 477–480 (2005).
3. Hershey, P. E. *et al.* The Cap-binding protein eIF4E promotes folding of a functional domain of yeast translation initiation factor eIF4G1. *J. Biol. Chem.* **274,** 21297–21304 (1999).
4. Mader, S., Lee, H., Pause, A. & Sonenberg, N. 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 (1995).
5. Marcotrigiano, J., Gingras, A. C., Sonenberg, N. & Burley, S. K. Cap-dependent translation initiation in eukaryotes is regulated by a molecular mimic of eIF4G. *Mol. Cell* **3,** 707–716 (1999).
6. Altmann, M., Schmitz, N., Berset, C. & Trachsel, H. A novel inhibitor of cap-dependent translation initiation in yeast: p20 competes with eIF4G for binding to eIF4E. *EMBO J.* **16,** 1114–1121 (1997).
7. Matsuo, H. *et al.* Structure of translation factor eIF4E bound to m^7GDP and interaction with 4E-binding protein. *Nat. Struct. Biol.* **4,** 717–724 (1997).
8. Gingras, A. C. *et al.* Regulation of 4E-BP1 phosphorylation: a novel two-step mechanism. *Genes Dev.* **13,** 1422–1437 (1999).
9. Martineau, Y., Azar, R., Bousquet, C. & Pyronnet, S. Anti-oncogenic potential of the eIF4E-binding proteins. *Oncogene* **32,** 671–677 (2013).
10. Bidinosti, M. *et al.* Postnatal deamidation of 4E-BP2 in brain enhances its association with raptor and alters kinetics of excitatory synaptic transmission. *Mol. Cell* **37,** 797–808 (2010).
11. Dowling, R. J. *et al.* mTORC1-mediated cell proliferation, but not cell growth, controlled by the 4E-BPs. *Science* **328,** 1172–1176 (2010).
12. Gkogkas, C. G. *et al.* Autism-related deficits via dysregulated eIF4E-dependent translational control. *Nature* **493,** 371–377 (2013).
13. Hsieh, A. C. *et al.* The translational landscape of mTOR signalling steers cancer initiation and metastasis. *Nature* **485,** 55–61 (2012).
14. Lasko, P., Cho, P., Poulin, F. & Sonenberg, N. Contrasting mechanisms of regulating translation of specific *Drosophila* germline mRNAs at the level of 5'-cap structure binding. *Biochem. Soc. Trans.* **33,** 1544–1546 (2005).
15. Faivre, S., Kroemer, G. & Raymond, E. Current development of mTOR inhibitors as anticancer agents. *Nat. Rev. Drug Discov.* **5,** 671–688 (2006).

16. Jia, Y., Polunovsky, V., Bitterman, P. B. & Wagner, C. R. Cap-dependent translation initiation factor eIF4E: an emerging anticancer drug target. *Med. Res. Rev.* **32,** 786–814 (2012).
17. Gross, J. D. *et al.* Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E. *Cell* **115,** 739–750 (2003).
18. Gosselin, P. *et al.* The translational repressor 4E-BP called to order by eIF4E: new structural insights by SAXS. *Nucleic Acids Res.* **39,** 3496–3503 (2011).
19. Lukhele, S., Bah, A., Lin, H., Sonenberg, N. & Forman-Kay, J. D. Interaction of the eukaryotic initiation factor 4E with 4E-BP2 at a dynamic bipartite interface. *Structure* **21,** 2186–2196 (2013).
20. Mizuno, A. *et al.* Importance of C-terminal flexible region of 4E-binding protein in binding with eukaryotic initiation factor 4E. *FEBS Lett.* **582,** 3439–3444 (2008).
21. Nelson, M. R., Leidal, A. M. & Smibert, C. A. *Drosophila* cup is an eIF4E-binding protein that functions in Smaug-mediated translational repression. *EMBO J.* **23,** 150–159 (2004).
22. Paku, K. S. *et al.* A conserved motif within the flexible C-terminus of the translational regulator 4E-BP is required for tight binding to the mRNA cap-binding protein eIF4E. *Biochem. J.* **441,** 237–245 (2012).
23. Umenaga, Y., Paku, K. S., In, Y., Ishida, T. & Tomoo, K. Identification and function of the second eIF4E-binding region in N-terminal domain of eIF4G: comparison with eIF4E-binding protein. *Biochem. Biophys. Res. Commun.* **414,** 462–467 (2011).
24. Nakamura, A., Sato, K. & Hanyu-Nakamura, K. *Drosophila* cup is an eIF4E binding protein that associates with Bruno and regulates oskar mRNA translation in oogenesis. *Dev. Cell* **6,** 69–78 (2004).
25. Wilhelm, J. E., Hilton, M., Amos, Q. & Henzel, W. J. Cup is an eIF4E binding protein required for both the translational repression of oskar and the recruitment of Barentsz. *J. Cell Biol.* **163,** 1197–1204 (2003).
26. Zappavigna, V., Piccioni, F., Villaescusa, J. C. & Verrotti, A. C. Cup is a nucleocytoplasmic shuttling protein that interacts with the eukaryotic translation initiation factor 4E to modulate *Drosophila* ovary development. *Proc. Natl Acad. Sci. USA* **101,** 14800–14805 (2004).
27. Kinkelin, K., Veith, K., Grunwald, M. & Bono, F. Crystal structure of a minimal eIF4E-Cup complex reveals a general mechanism of eIF4E regulation in translational repression. *RNA* **18,** 1624–1634 (2012).
28. Ptushkina, M. *et al.* Cooperative modulation by eIF4G of eIF4E-binding to the mRNA 5' cap in yeast involves a site partially shared by p20. *EMBO J.* **17,** 4798–4808 (1998).
29. Pyronnet, S. *et al.* Human eukaryotic translation initiation factor 4G (eIF4G) recruits mnk1 to phosphorylate eIF4E. *EMBO J.* **18,** 270–279 (1999).
30. Kubacka, D. *et al.* Investigating the consequences of eIF4E2 (4EHP) interaction with 4E-transporter on its cellular distribution in HeLa cells. *PLoS ONE* **8,** e72761 (2013).
31. Cheng, Y. & Patel, D. J. An efficient system for small protein expression and refolding. *Biochem. Biophys. Res. Commun.* **317,** 401–405 (2004).
32. Slepenkov, S. V., Korneeva, N. L. & Rhoads, R. E. Kinetic mechanism for assembly of the m^7GpppG.eIF4E.eIF4G complex. *J. Biol. Chem.* **283,** 25227–25237 (2008).
33. Dostie, J., Ferraiuolo, M., Pause, A., Adam, S. A. & Sonenberg, N. A novel shuttling protein, 4E-T, mediates the nuclear import of the mRNA 5' cap-binding protein, eIF4E. *EMBO J.* **19,** 3142–3156 (2000).
34. Rong, L. *et al.* Control of eIF4E cellular localization by eIF4E-binding proteins, 4E-BPs. *RNA* **14,** 1318–1327 (2008).
35. Sukarieh, R., Sonenberg, N. & Pelletier, J. The eIF4E-binding proteins are modifiers of cytoplasmic eIF4E relocalization during the heat shock response. *Am. J. Physiol. Cell Physiol.* **296,** C1207–C1217 (2009).
36. Ferraiuolo, M. A. *et al.* A role for the eIF4E-binding protein 4E-T in P-body formation and mRNA decay. *J. Cell Biol.* **170,** 913–924 (2005).
37. Wolff, B., Sanglier, J. J. & Wang, Y. Leptomycin B is an inhibitor of nuclear export: inhibition of nucleo-cytoplasmic translocation of the human immunodeficiency virus type 1 (HIV-1) Rev protein and Rev-dependent mRNA. *Chem. Biol.* **4,** 139–147 (1997).
38. Siddiqui, N. *et al.* Structural insights into the allosteric effects of 4EBP1 on the eukaryotic translation initiation factor eIF4E. *J. Mol. Biol.* **415,** 781–792 (2012).
39. Igreja, C. & Izaurralde, E. CUP promotes deadenylation and inhibits decapping of mRNA targets. *Genes Dev.* **25,** 1955–1967 (2011).
40. Zekri, L., Kuzuoglu-Ozturk, D. & Izaurralde, E. GW182 proteins cause PABP dissociation from silenced miRNA targets in the absence of deadenylation. *EMBO J.* **32,** 1052–1065 (2013).
41. Tritschler, F. *et al.* Similar modes of interaction enable Trailer Hitch and EDC3 to associate with DCP1 and Me31B in distinct protein complexes. *Mol. Cell Biol.* **28,** 6695–6708 (2008).
42. Behm-Ansmant, I. *et al.* mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. *Genes Dev.* **20,** 1885–1898 (2006).
43. Diebold, M. L., Fribourg, S., Koch, M., Metzger, T. & Romier, C. Deciphering correct strategies for multiprotein complex assembly by co-expression: application to complexes as large as the histone octamer. *J. Struct. Biol.* **175,** 178–188 (2011).
44. Studier, F. W. Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **41,** 207–234 (2005).
45. Mizoue, L. S. & Tellinghuisen, J. The role of backlash in the "first injection anomaly" in isothermal titration calorimetry. *Anal Biochem.* **326,** 125–127 (2004).

Acknowledgements

This work was supported by the Max Planck Society and the Gottfried Wilhelm Leibniz Program of the Deutsche Forschungsgemeinschaft (DFG, awarded to E.I). We are grateful to J. Ristau for generating some of the constructs used in this study, to F. Bono for providing the His-eIF4E and GST-CUP 311–440 constructs, and to Min-Yi Chung for excellent technical assistance. We thank S. Würtenberger for assistance with the ITC measurements and S. Jonas, S. Grüner, V. Truffault and L. Zekri for insightful discussions.

Author contributions

C.I. performed most of the experiments and supervised assays performed by C.W.; D.P. contributed to some experiments, performed ITC measurements and contributed to the writing of the manuscript; C.W. performed pull-down assays and prepared several DNA constructs; E.I. was the principal investigator who coordinated the project; C.I. and E.I. designed the research project and drafted the manuscript.

Additional information

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

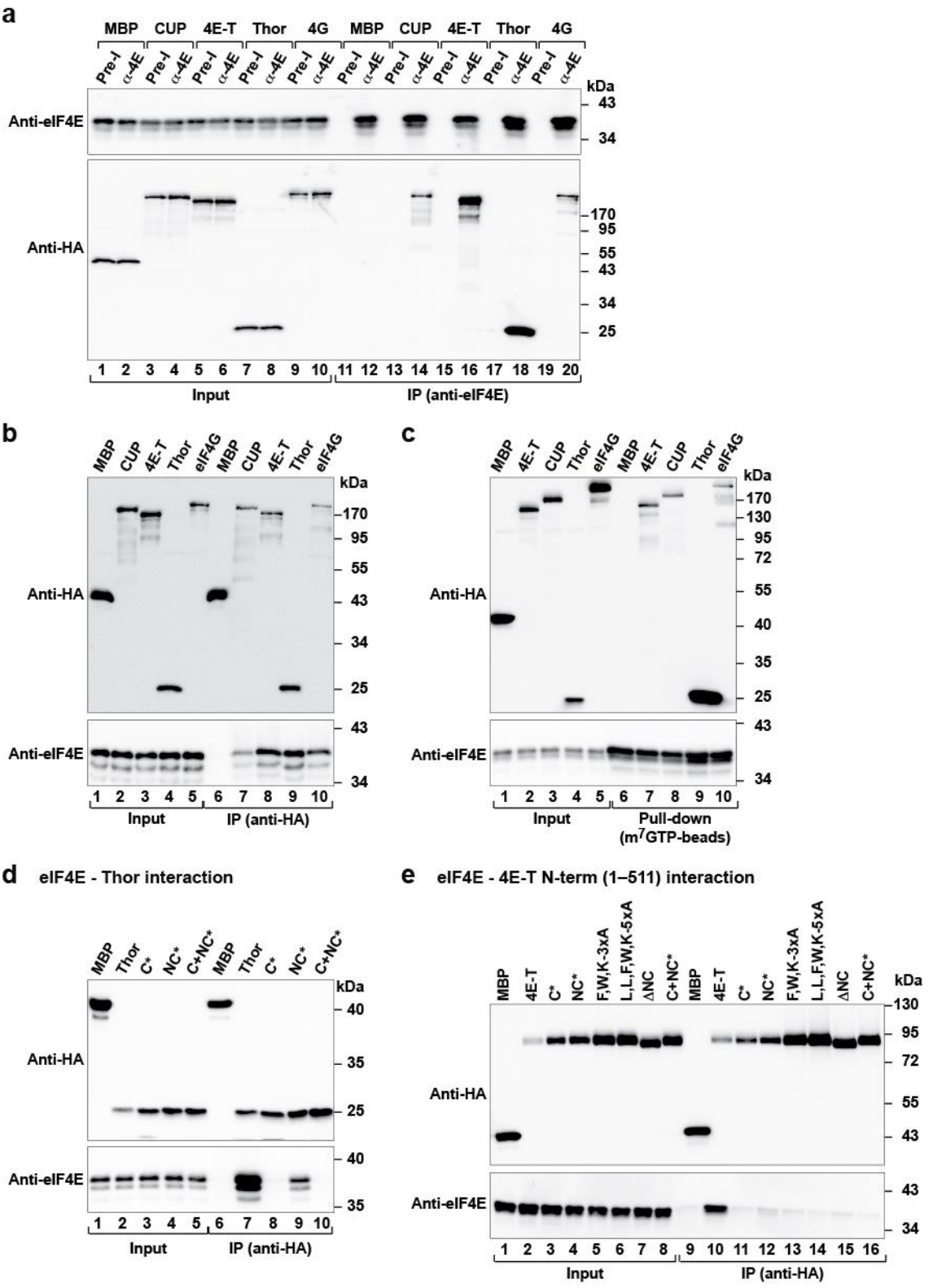
Competing financial interests: The authors declare no competing financial interests.

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

How to cite this article: Igreja, C. *et al.* 4E-BPs require non-canonical 4E-binding motifs and a lateral surface of eIF4E to repress translation. *Nat. Commun.* 5:4790 doi: 10.1038/ncomms5790 (2014).

Supplementary Figures

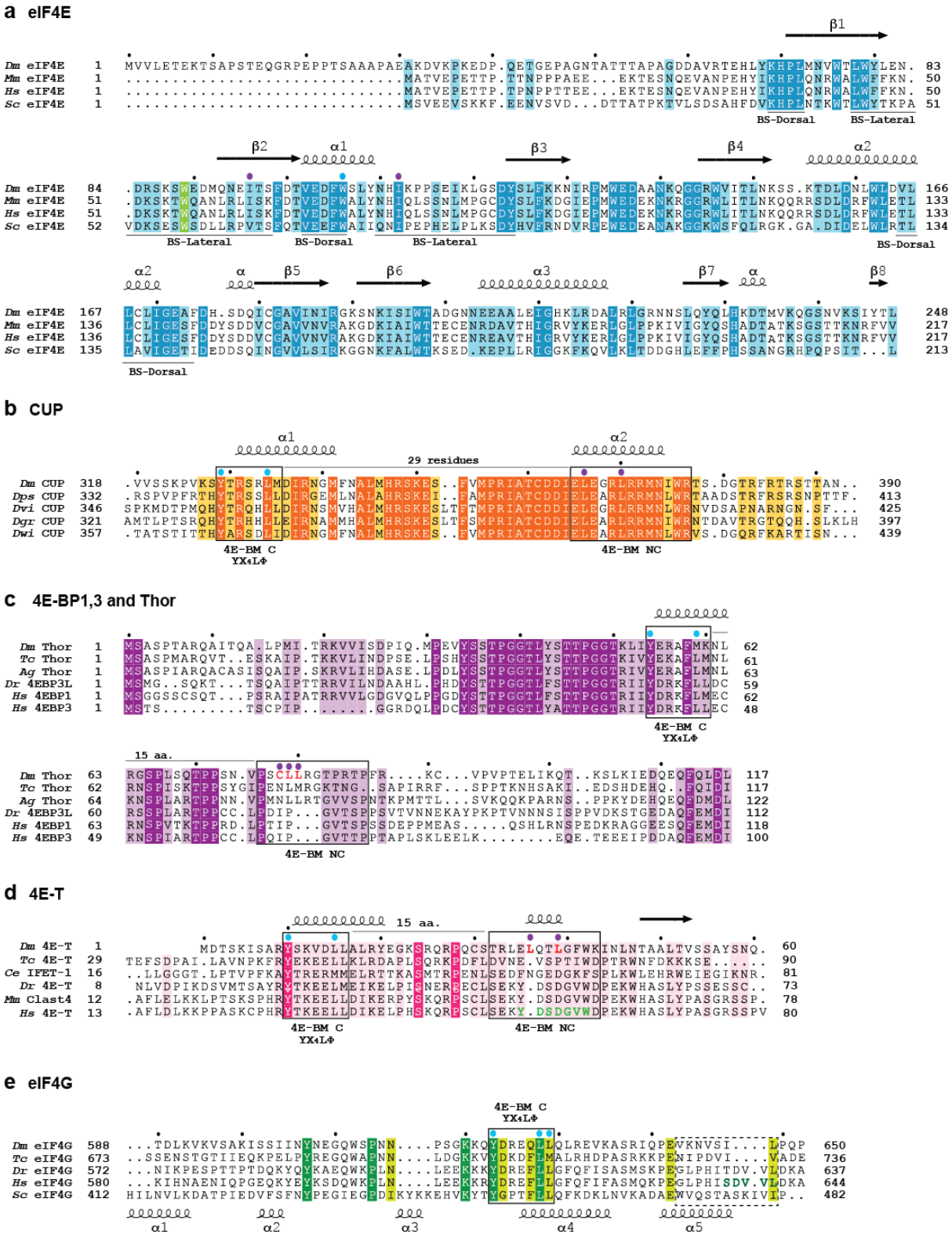
Supplementary Fig. 1



Supplementary Figure 1. The interaction of eIF4E with 4E-BPs and eIF4G in S2 cell lysates. (**a**) Western blot showing the interaction of HA-tagged full length 4E-BPs (CUP,

Thor and 4E-T) or HA-eIF4G and endogenous eIF4E in S2 cell lysates. The proteins were immunoprecipitated using a polyclonal anti-eIF4E antibody (α-4E). The pre-immune (Pre-I) serum and an HA-tagged version of MBP served as negative controls. The inputs (1%) and immunoprecipitates (10%) were analyzed by western blotting using anti-HA and anti-eIF4E antibodies. (**b,c**) The interaction of HA-tagged full length 4E-BPs (CUP, Thor and 4E-T) and eIF4G with endogenous eIF4E was analyzed by coimmunoprecipitation using anti-HA antibodies (**b**) or by pull-down using m^7GTP-Sepharose beads (**c**). HA-MBP served as a negative control. The inputs (10%) and immunoprecipitates (30%) were analyzed by western blotting using anti-HA and anti-eIF4E antibodies. (**d**) Western blot showing the interaction of HA-Thor (full-length either wild-type or mutants) with endogenous eIF4E in S2 cells. The proteins were immunoprecipitated using anti-HA antibodies and analyzed as described in (**b**,**c**). (**e**) Western blot showing the interaction of HA-4E-T (residues 1–511, wild-type or mutants) with endogenous eIF4E. The proteins were immunoprecipitated using anti-HA antibodies and analyzed as described in (**b**,**c**). The size markers (kDa) are shown to the right of each panel. Mutants are described in Supplementary Table 1.

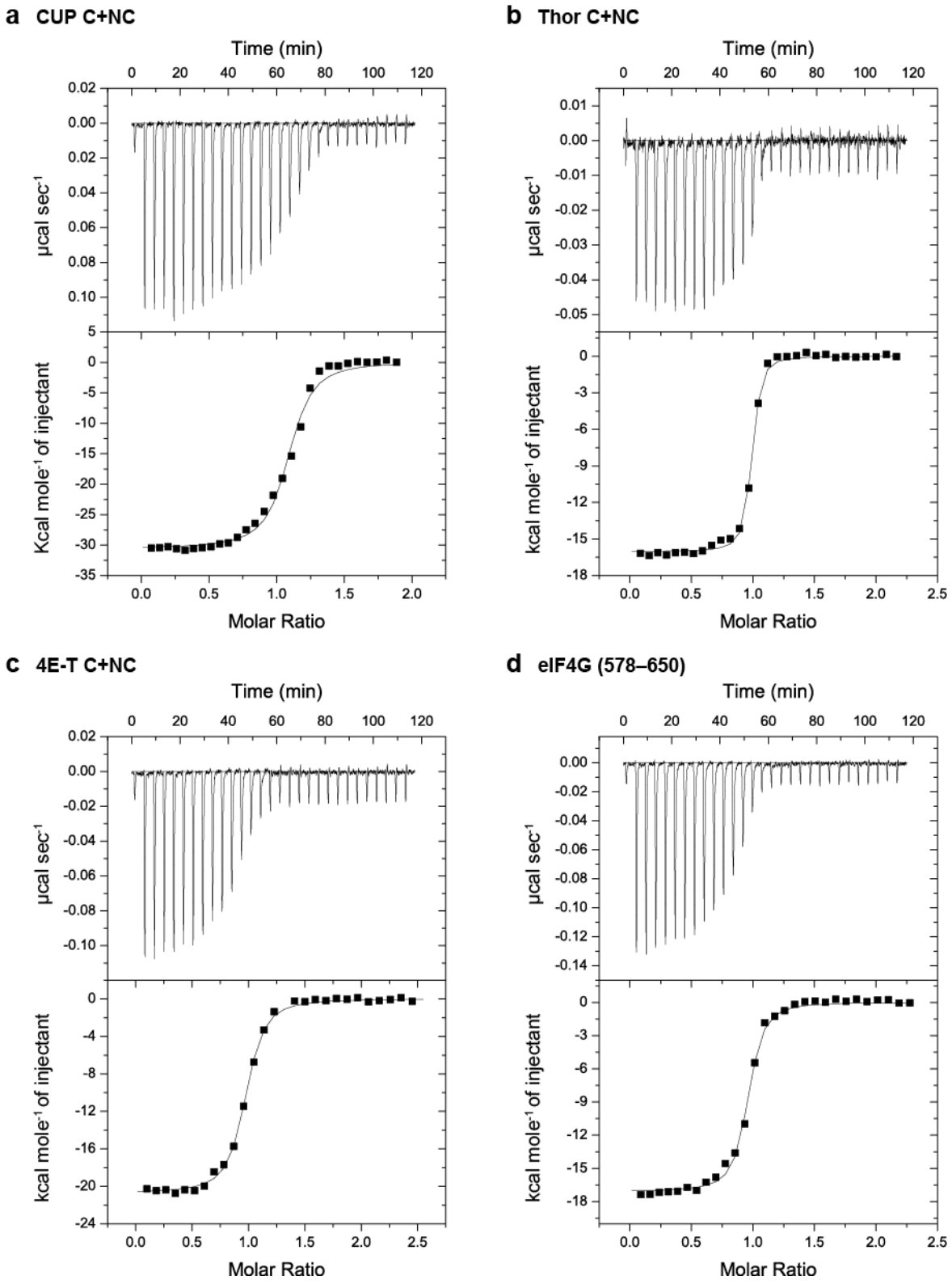
Supplementary Fig. 2



Supplementary Figure 2. Sequence alignments. Residues conserved in all aligned sequences are highlighted with a dark color background and printed in white. Residues with >70% similarity are shown with a light color background. (**a**) Structure-based sequence

alignment of eIF4E proteins from *Drosophila melanogaster* (*Dm*), *Mus musculus* (*Mm*), *Homo sapiens* (*Hs*) and *Saccharomyces cerevisae* (*Sc*). Secondary structure elements of the *Dm* eIF4E (PDB code: 4AXG)[27] are indicated above the sequences. The lateral and dorsal binding surfaces (BS) are indicated with a line below the sequences. Residues mutated in this study are indicated by filled circles above the sequence. Circles are colored in cyan or magenta for the dorsal and lateral surfaces, respectively. (**b**) Sequence alignment of the eIF4E-interacting regions of CUP orthologous proteins from insects. Species are as follows: *Drosophila melanogaster* (*Dm*), *Drosophila pseudoobscura* (*Dps*), *Drosophila virilis* (*Dvi*), *Drosophila grimshawi* (*Dgr*), *Drosophila willistoni* (*Dwi*). Secondary structure elements of the *Dm* CUP (PDB code: 4AXG) protein are indicated above the sequences. The canonical (C) and non-canonical (NC) 4E-BMs are boxed in black. Residues mutated in this study are indicated with filled circles above the sequence. Circles are colored in cyan or magenta for the canonical and non-canonical 4E-BMs, respectively. (**c**) Sequence alignment of Thor orthologous proteins. Species are as follows: *Drosophila melanogaster* (*Dm*), *Tribolium castaneum* (*Tc*), *Anopheles gambiae* (*Ag*), *Danio rerio* (*Dr*) and *Homo sapiens* (*Hs*). Additional symbols are as described in panel (**b**). (**d**) Alignment of the eIF4E-interacting region of 4E-T orthologous proteins. Species are *Caenorhabditis elegans* (*Ce*) and as in (a,**c**). Predicted secondary structure elements are indicated above the sequences. A putative non-canonical 4E-BM identified in the *Hs* 4E-T protein is highlighted in green characters. Additional symbols are as described in panel (**b**). (**e**) Alignment of eIF4E-interacting regions of eIF4G orthologous proteins. Species are as in panels (**a,d**). A putative non-canonical 4E-BM identified in the *Hs* eIF4G protein and the corresponding motif in *Dm* are indicated by a dashed box. Secondary structure elements of the *Sc* eIF4G (PDB 1RF8) protein are indicated below the sequences. Additional symbols are as described in panel (**b**).

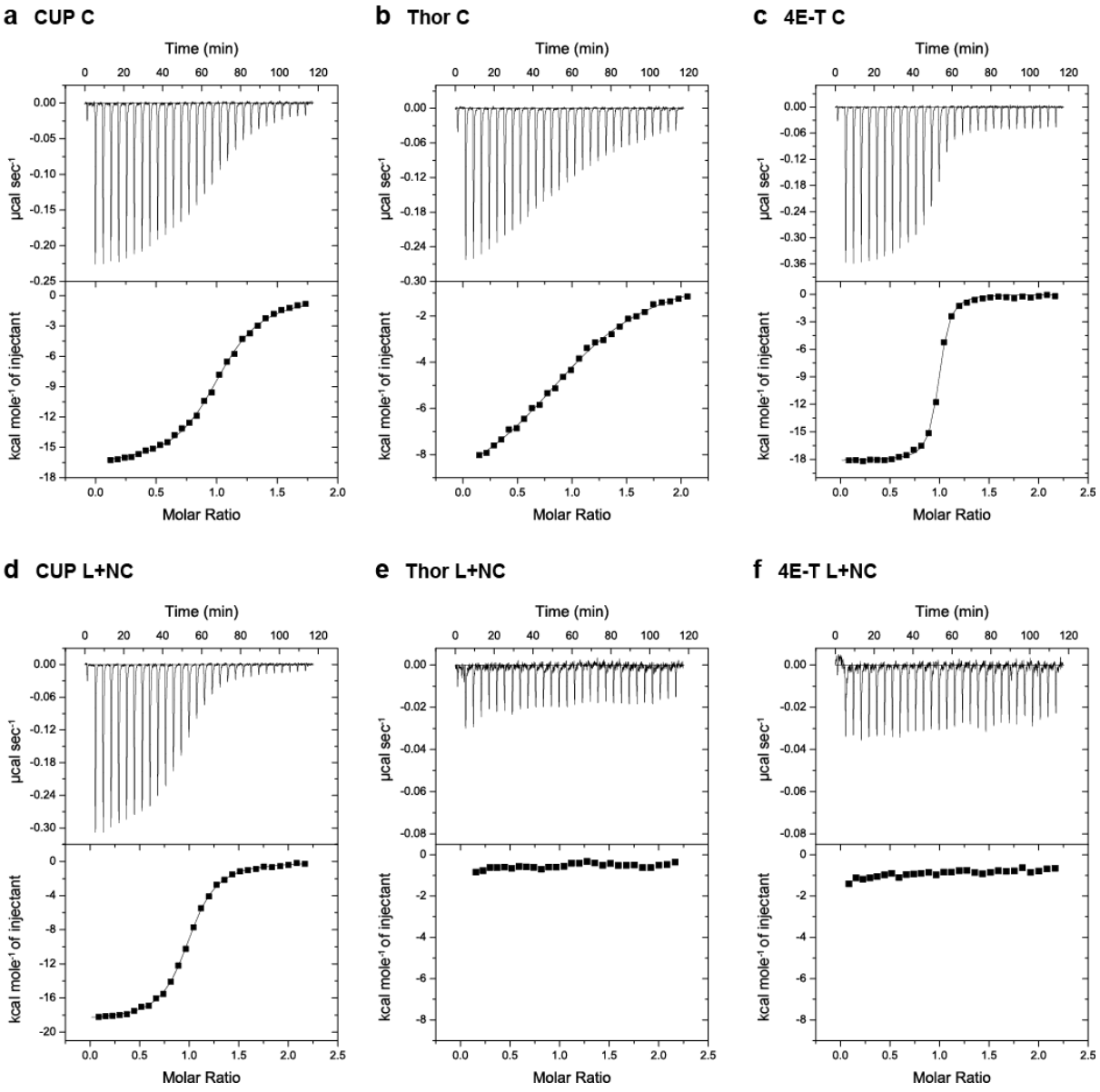
Supplementary Fig. 3



Supplementary Figure 3. Thermodynamic parameters for the interaction of eIF4E with eIF4G and 4E-BP peptides. (**a–d**) ITC profiles of CUP (**a**), Thor (**b**), 4E-T (**c**) and eIF4G

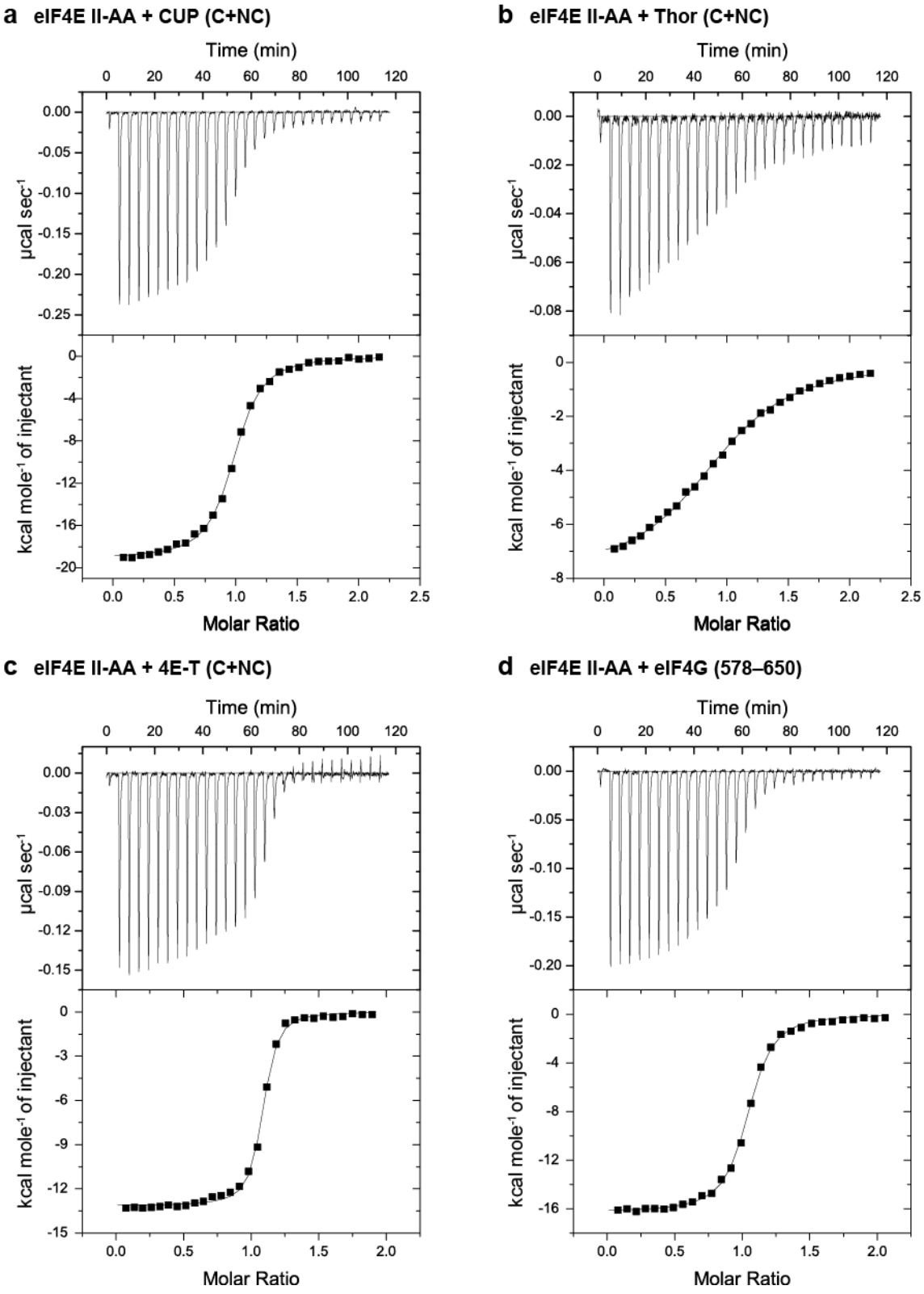
(**d**) minimal regions for binding to eIF4E (wild-type (WT), residues 69–248). The top panels represent the raw data (μcal sec^{-1}) whereas the bottom panels show the integrated data (kcal $mole^{-1}$ of injectant) of heat changes that were best-fitted using a one-site binding model. The thermodynamic parameters correspond to heat changes that were detected when 4E-BPs (10 μM) or eIF4G (20 μM) were injected into the calorimetric cell (1.4 mL) containing 1 μM (**a,b,c**) or 2 μM (**d**) eIF4E.

Supplementary Fig. 4



Supplementary Figure 4. Thermodynamic parameters for the interaction of eIF4E with 4E-BP peptides. (**a–f**) ITC profiles for the interaction of eIF4E (WT, residues 69-248) and CUP (**a,d**), Thor (**b,e**) and 4E-T (**c,f**) peptides containing only the canonical motifs (**a–c**) or the L+NC regions (**d–f**). The measurements were performed as described in Supplementary Fig. 3. The solutions contained the peptides corresponding to the canonical motifs (50 µM) or the L+NC regions (100 µM for CUP and 200 µM for Thor and 4E-T). The eIF4E concentration in the calorimetric cell was 5 µM (for the C peptides), 10 µM (for CUP L+NC) and 20 µM (for Thor and 4E-T L+NC peptides).

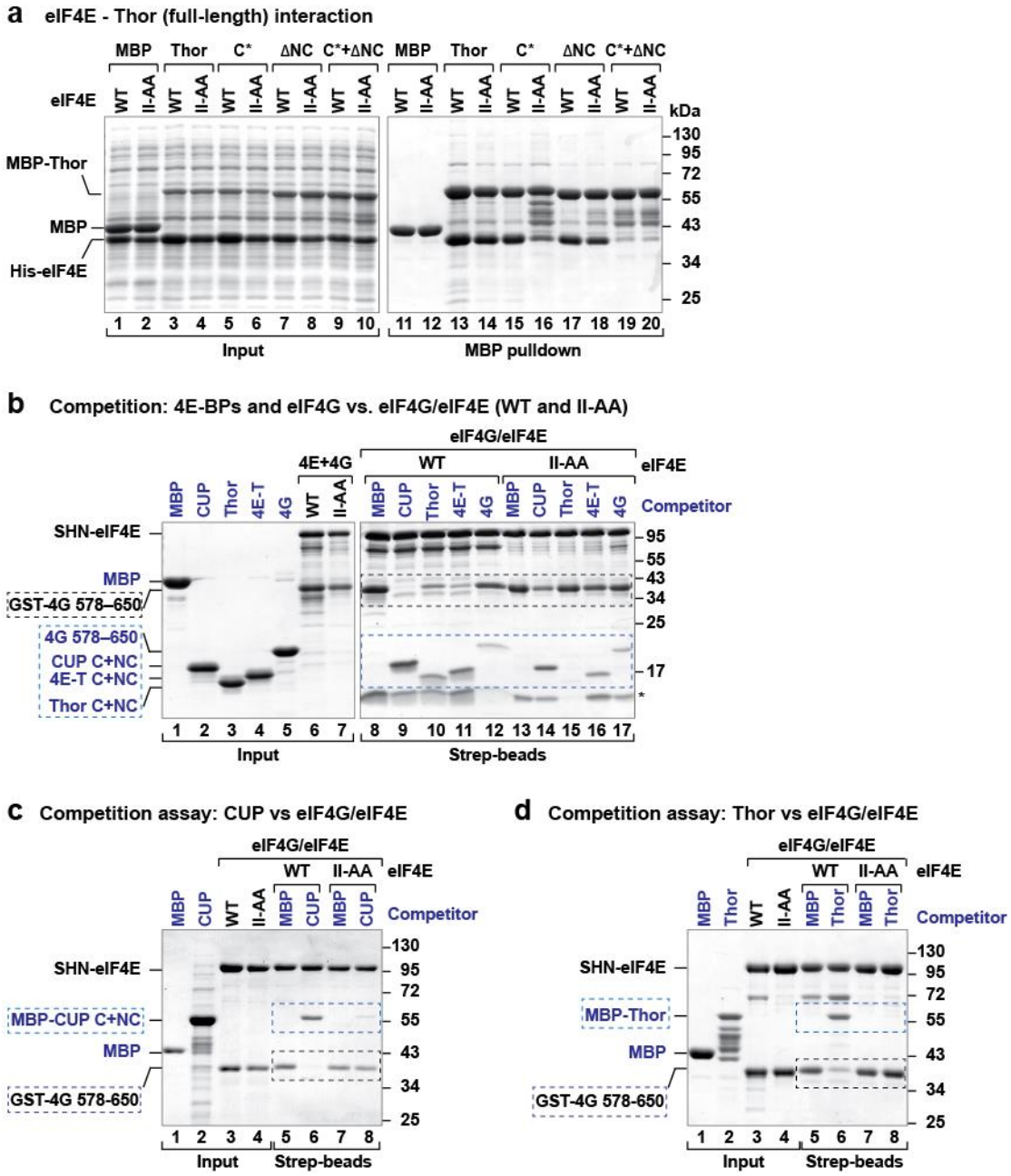
Supplementary Fig. 5



Supplementary Figure 5. Thermodynamic parameters for the interaction of eIF4E II-AA mutant with 4E-BP peptides. (**a–d**) ITC profiles for the interaction of eIF4E IIAA

mutant (residues 69-248) and CUP (**a**), Thor (**b**), 4E-T (**c**) and eIF4G (**d**) peptides. The measurements were performed as described in **Supplementary Fig. 3**. The solutions contained the 4E-BP C+NC peptides (10 μM) and 20 μM for the eIF4G peptide. The eIF4E II-AA mutant concentration in the calorimetric cell was 5 μM.

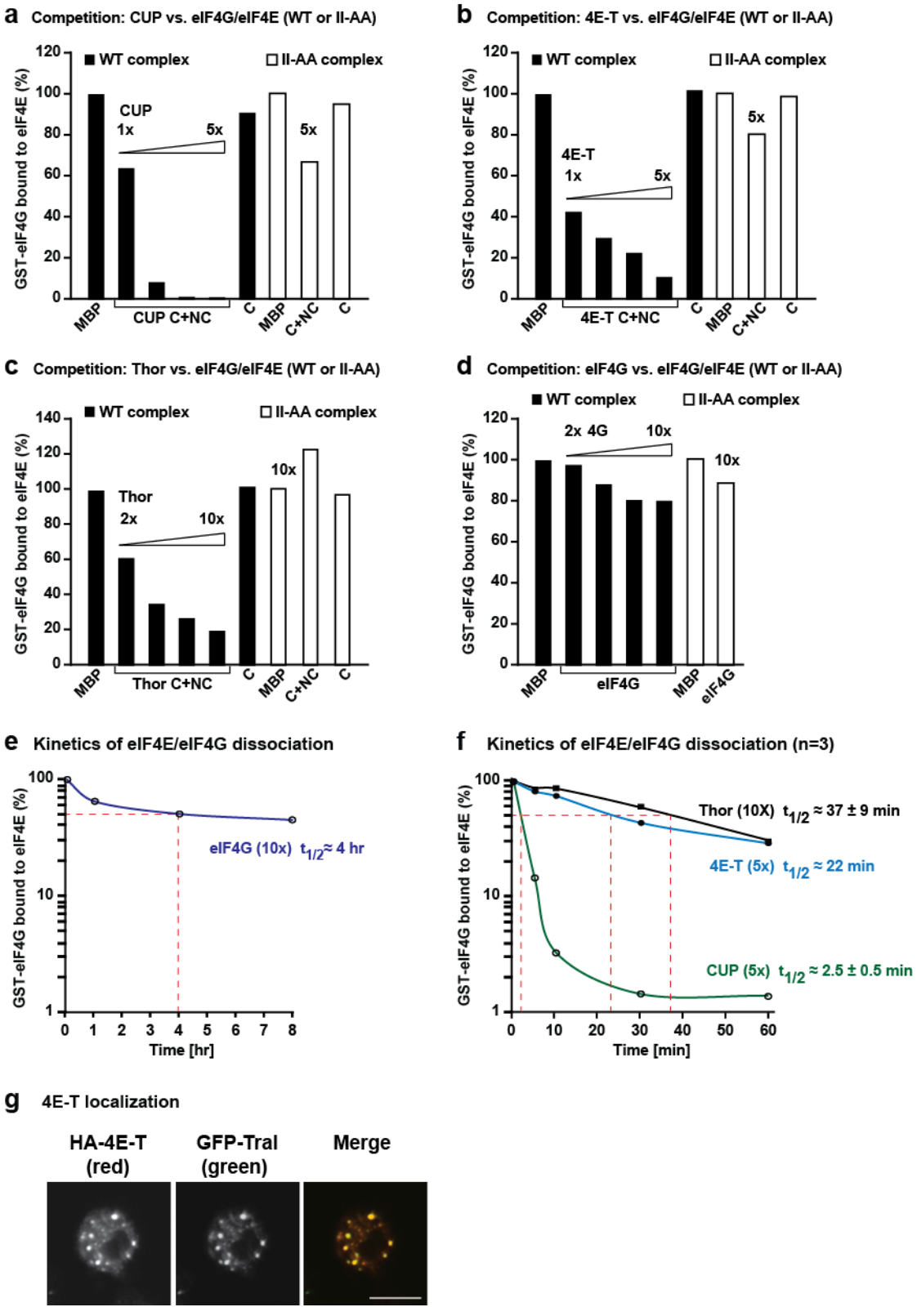
Supplementary Fig. 6



Supplementary Figure 6. Identification of a non-canonical 4E-BM in Thor and competition assays. (**a**) The association of MBP-Thor (full length, wild-type or 4E-BM mutant) with His_6-eIF4E (WT or II-AA mutant) was analyzed by MBP pull-down as described in Fig. 3c. (**b**) Purified eIF4E–eIF4G complexes containing SHN-eIF4E (wild-type or II-AA mutant) and GST-eIF4G (578–650) were incubated with CUP, Thor and 4E-T fragments (C+NC) fused C-terminally to GB1. The proteins that were bound to eIF4E were

pulled down using Strep-Tactin beads and analyzed by SDS-PAGE. The competitor peptides are labeled in blue, and their positions are highlighted by blue, dashed boxes. The black, dashed boxes indicate the position of preassembled GST-eIF4G. (**c,d**) The purified eIF4E–eIF4G complexes described in panel (**b**) were incubated with a 5-fold molar excess of CUP (**c**) (C+NC fragment) or MBP-Thor full length (**d**). The eIF4E-bound proteins were pulled down using Strep-Tactin beads and analyzed by SDS-PAGE. The size markers (kDa) are shown to the right of each panel.

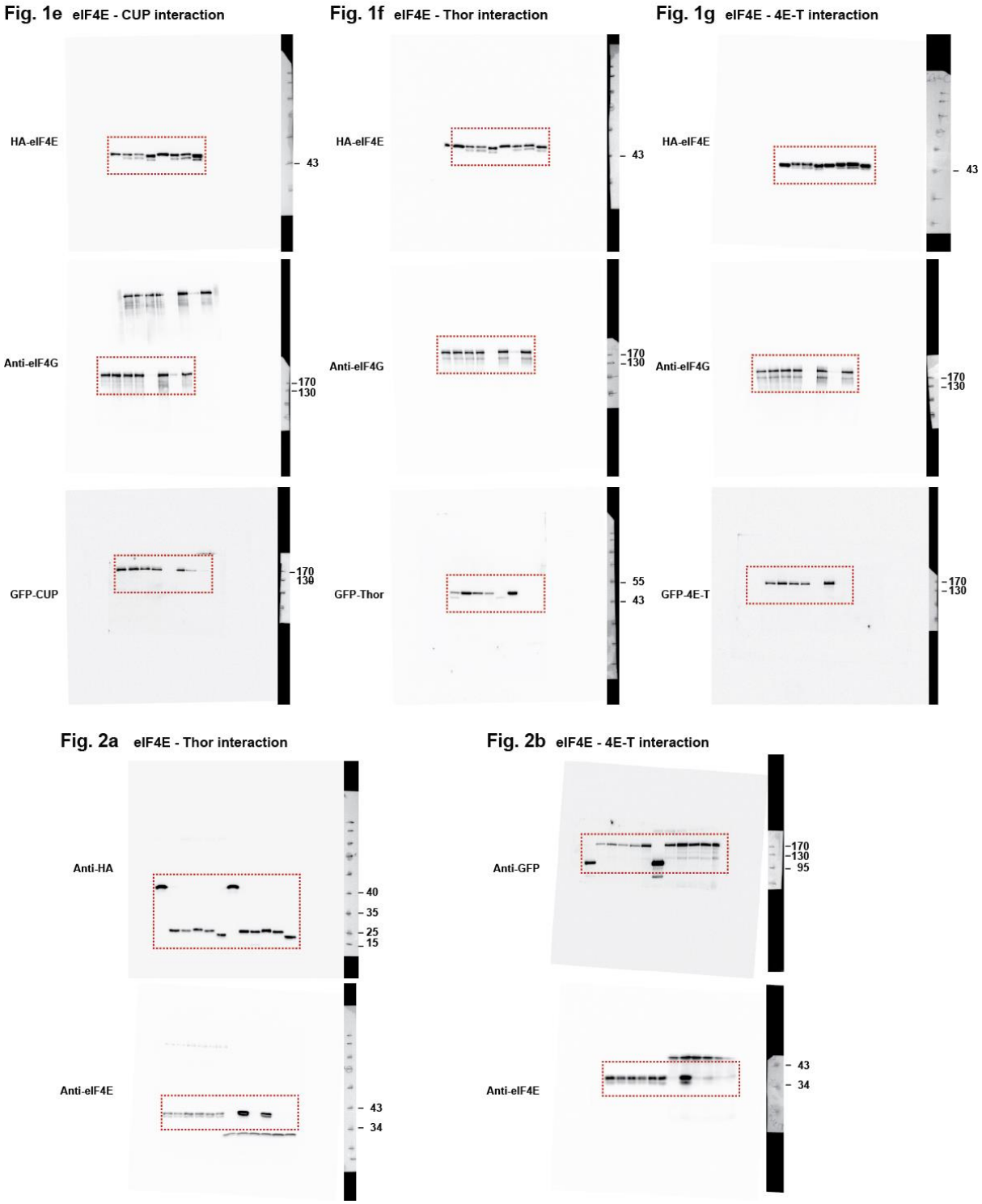
Supplementary Fig. 7



Supplementary Figure 7. Quantitative analysis of the competition assays. (**a–d**) Quantification of GST-eIF4G remaining bound to eIF4E (WT or II-AA mutant) in the

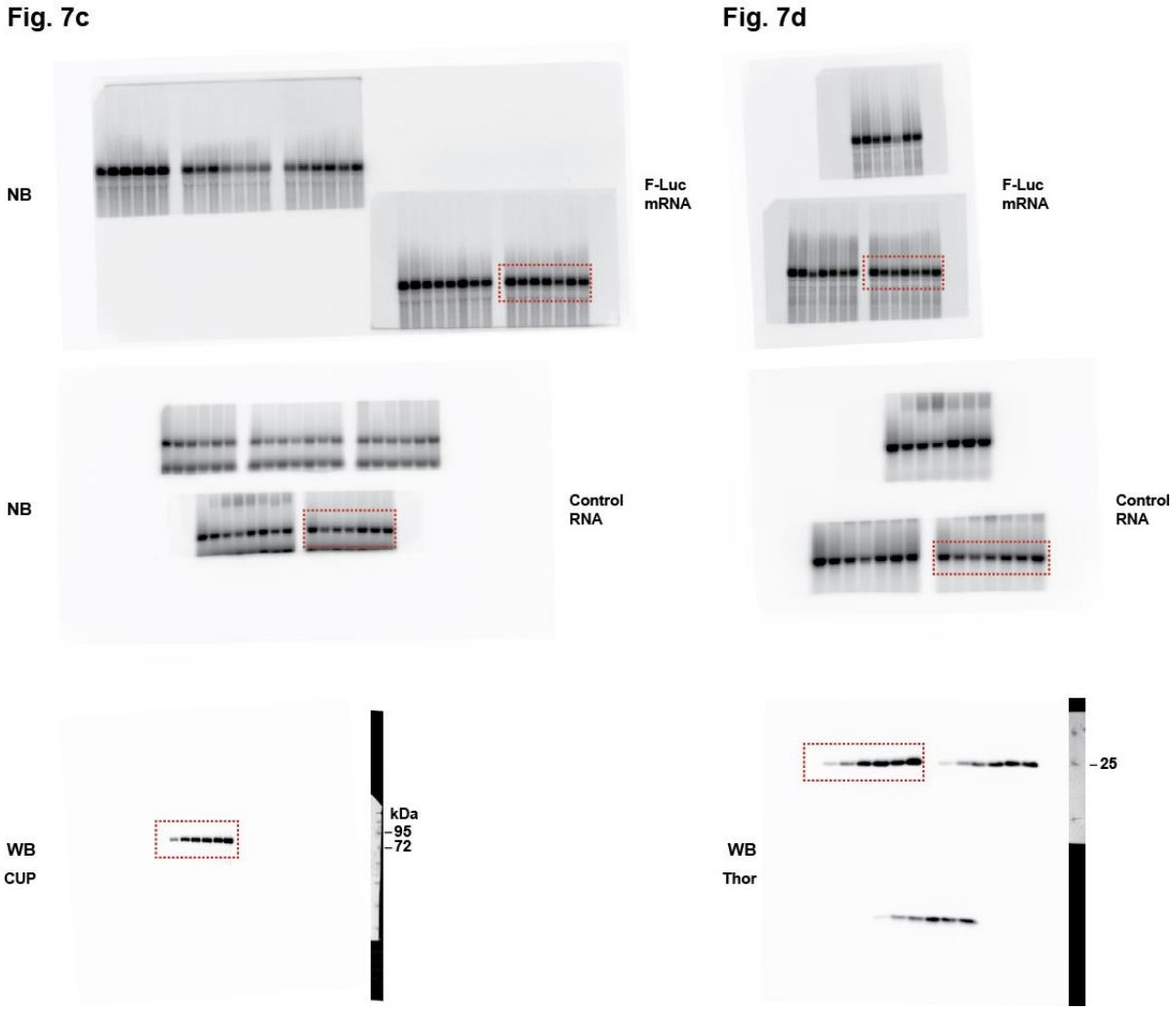
presence of MBP (negative control), increasing amounts of 4E-BP peptides or eIF4G-GB1 in the experiments shown in Fig. 4a–d. The amounts of GST-eIF4G bound to eIF4E in each experimental condition were determined using the Image J software after Coomassie blue staining. To rule out the possibility that changes in the amount of GST-eIF4G bound to eIF4E resulted from variations in the loading volume, all values were normalized to the levels of SHN-eIF4E present in each condition. These values were set to 100 in the presence of MBP. Each experiment was repeated at least twice. (**e,f**) Quantification of the experiments shown in Fig. 5d–g. The amounts of GST-eIF4G remaining bound to eIF4E in the presence of MBP, 4E-BP peptides or eIF4G-GB1 were normalized to those of eIF4E and set to 100 in the presence of MBP. The data represent averages of two (eIF4G and 4E-T) or three (CUP and Thor) independent experiments. (**g**) *Dm* 4E-T localizes to P-bodies. Confocal fluorescent micrographs of fixed S2 cells expressing HA-4E-T and GFP-Tral. The localization of HA-4E-T was determined by indirect immunofluorescence using an anti-HA antibody. The merged pictures show the HA signal in red and the GFP-Tral signal in green. Bar: 5 μm.

Supplementary Fig. 8



Supplementary Figure 8. Original images of western blots used in Fig. 1e–g and Fig. 2a,b.

Supplementary Fig. 9



Supplementary Figure 9. Original images of western and northern blots used in Fig. 7c,d.

Supplementary Table 1. Mutants and constructs used in this study

Protein	Name of the construct	Mutations / fragments	Binding site / motif
eIF4E (1–248) (isoform C) P48598-2	W106A	W106A	Dorsal surface
	II-AA	I96A, I112A	Lateral surface
	trunc	69–248	
eIF4G (1–1666) (isoform A) O61380	4G	578–650	eIF4E-binding region
	C*	Y621A, L626A, L627A	Canonical
CUP (1–1117) Q9VMA3	CUP	311–440	eIF4E-binding region
	C+NC	325–376	eIF4E-binding region short
	C	325–341	Canonical
	NC	362–376	Non-canonical
	L+NC	342–376	Linker+non-canonical
	C*	Y327A, L332A	Canonical
	NC*	L364A, L368A	Non-canonical
	C+NC*	Y327A, L332A, L364A, L368A	Double mutant
	N-term	1–402	N-terminus
Thor (1–117) Q9XZ56	Thor	Full length	
	C+NC	50–83	eIF4E-binding region short
	C	50–63	Canonical
	NC	77–83	Non-canonical
	L+NC	64–83	Linker+non-canonical
	C*	Y54A, M59A	Canonical
	NC*	C78A, L79A, L80A	Non-canonical
	R,G,T	R81A, G82A, T83A	Non-canonical
	ΔNC	Δ76–84	Non-canonical
	C+NC*	Y54A, M59A, C78A, L79A, L80A	Double mutant
	C*+ΔNC	Y54A, M59A, Δ76-85	Double mutant
4E-T (1–1010) Q8IH18	4E-T	1–58	eIF4E-binding region
	C+NC	9–44	eIF4E-binding region short
	C	9–21	Canonical
	NC	31–44	Non-canonical
	L+NC	22–44	Linker+non-canonical
	C*	Y10A, L15A	Canonical
	NC*	L36A, L39A	Non-canonical
	F,W,K-3xA	F41A, W42A, K43A	Non-canonical
	L,L,F,W,K-5xA	L36A, L39A, F41A, W42A, K43A	Non-canonical
	ΔNC	Δ28–43	Non-canonical
	C+NC*	Y10A, L15A, L36A, L39A	Double mutant
	C*+ΔNC	Y10A, L15A, Δ28-43	Double mutant
	N-term	1–511	N-terminus

Supplementary Table 2. Antibodies used in this study

Antibody	Source	Catalog Number	Dilution	Monoclonal/ Polyclonal
Anti-HA-HRP (for Western blot)	Roche	12 013 819 001	1:5,000	Monoclonal
Anti-eIF4E	In house		1:3,000	Rabbit polyclonal
Anti-GFP	In house		1:2,000	Rabbit polyclonal
Anti-eIF4G	In house		1:3,000	Rabbit polyclonal
Anti-HA (for Immunofluorescence)	Covance	MMS-101P	1:1,000	Monoclonal
Anti-rabbit-HRP	GE Healthcare	NA934V	1:10,000	Polyclonal
Alexa Fluor 594–labelled goat anti–rabbit	Invitrogen	A-11080	1:1,000	Polyclonal

Supplementary Table 3. Mutagenesis primers used in this study

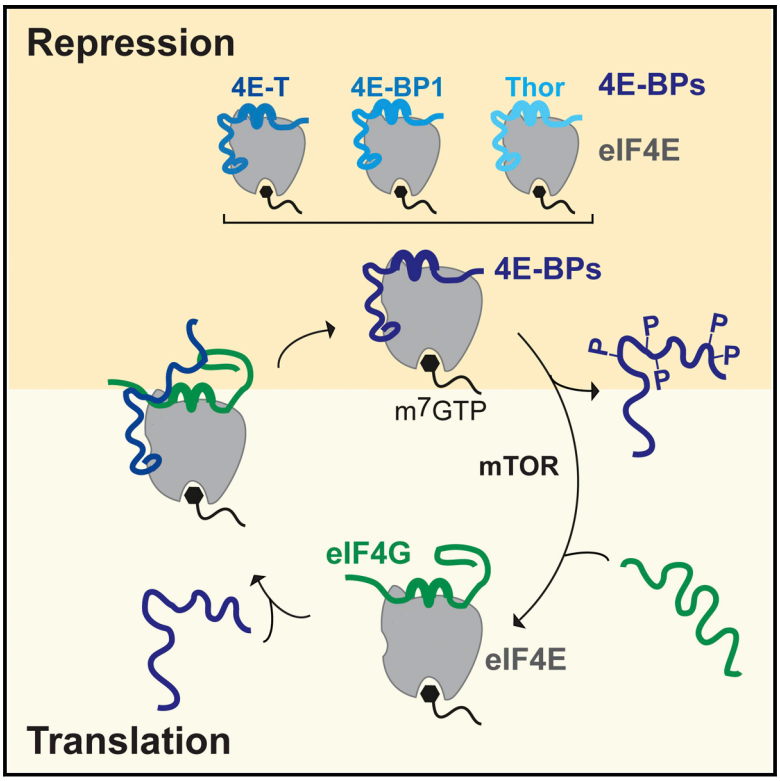
Protein	Mutations	Primer sequence 5'- 3'(Forward)
eIF4E	W106A	cttcgataccgtcgaggacttcgcgagcctatacaaccacatcaagccc
	II-AA	1 - gggaggacatgcaaaacgaggccaccagcttcgataccgtcga 2 - tctggagcctatacaaccacgccaagcccccatcagagatcaa
eIF4G	Y621A, L626A, L627A	cggaaaaaaacaagctgaccgagaacaggctgctcagttacgcga
CUP	Y327A, L332A	aagccggtcaagagcgctacccgctcccgcgcgatggacattcgcaac
	L364A, L368A	ccacctgcgacgacattgaggcggagggaagggctcgccgcatgaatatttggcgc
Thor	Y54A, M59A	ctcctggaggcaccaaacttatcgccgagcgggctttcgcgaagaatctccgtggctccccatt
	C78A, L79A, L80A	ctccgccgtccaacgtgcccagtgctgcggcgaggggcactccgcgtactccctt
	R81A, G82A, T83A	ccaacgtgcccagttgcttgctggctgccgctccgcgtactcccttccgcaagtg
	Δ76–84	gccaaactccgccgtccaacgtgcgtactcccttccgcaagtgcgt
4E-T	Y10A, L15A	caaagattagtgccagggcctcgaaagtagacgcgttagctctaagatatgaag
	L36A, L39A	cgcaatgttcaacacgacttgaagcgcagacggcgggtttttggaaaattaatctgaa
	F41A, W42A, K43A	gacttgaattgcagacgctaggtgcggcggcgattaatctgaacacagctgcgtt
	Δ28–43	gccaaactccgccgtccaacgtgcgtactcccttccgcaagtgcgt

Article

Molecular Cell

Molecular Architecture of 4E-BP Translational Inhibitors Bound to eIF4E

Graphical Abstract



Authors

Daniel Peter, Cátia Igreja, ...,
Oliver Weichenrieder, Elisa Izaurralde

Correspondence

oliver.weichenrieder@tuebingen.mpg.de (O.W.),
elisa.izaurralde@tuebingen.mpg.de (E.I.)

In Brief

The eIF4E-binding proteins (4E-BPs) repress translation by competing with eIF4G for binding to eIF4E. Peter et al. crystallized complexes of several 4E-BPs bound to eIF4E, revealing common binding principles and providing a molecular explanation for competition efficiency and for the negative effect of phosphorylation on complex formation.

Highlights

- Structures of 4E-BPs bound to eIF4E reveal common binding principles
- A conserved elbow loop allows 4E-BPs to bind two orthogonal surfaces on eIF4E
- Non-canonical motifs use different conformations to engage a lateral eIF4E pocket
- Phosphorylation of 4E-BPs impairs the competition with eIF4G for eIF4E binding

Accession Numbers

4UE8
4UE9
4UEA
4UEB
4UEC
4UED

Peter et al., 2015, Molecular Cell *57*, 1074–1087
March 19, 2015
http://dx.doi.org/10.1016/j.molcel.2015.01.017

CellPress

CellPress

Molecular Cell
Article

Molecular Architecture of 4E-BP Translational Inhibitors Bound to eIF4E

Daniel Peter,[1] Cátia Igreja,[1] Ramona Weber,[1] Lara Wohlbold,[1] Catrin Weiler,[1] Linda Ebertsch,[1] Oliver Weichenrieder,[1,*] and Elisa Izaurralde[1,*]

[1]Department of Biochemistry, Max Planck Institute for Developmental Biology, Spemannstrasse 35, 72076 Tübingen, Germany

*Correspondence: oliver.weichenrieder@tuebingen.mpg.de (O.W.), elisa.izaurralde@tuebingen.mpg.de (E.I.)

http://dx.doi.org/10.1016/j.molcel.2015.01.017

SUMMARY

The eIF4E-binding proteins (4E-BPs) represent a diverse class of translation inhibitors that are often deregulated in cancer cells. 4E-BPs inhibit translation by competing with eIF4G for binding to eIF4E through an interface that consists of canonical and non-canonical eIF4E-binding motifs connected by a linker. The lack of high-resolution structures including the linkers, which contain phosphorylation sites, limits our understanding of how phosphorylation inhibits complex formation. Furthermore, the binding mechanism of the non-canonical motifs is poorly understood. Here, we present structures of human eIF4E bound to 4E-BP1 and fly eIF4E bound to Thor, 4E-T, and eIF4G. These structures reveal architectural elements that are unique to 4E-BPs and provide insight into the consequences of phosphorylation. Guided by these structures, we designed and crystallized a 4E-BP mimic that shows increased repressive activity. Our studies pave the way for the rational design of 4E-BP mimics as therapeutic tools to decrease translation during oncogenic transformation.

INTRODUCTION

The eIF4E-binding proteins (4E-BPs) inhibit protein synthesis at a global or message-specific level and play important roles in diverse cellular processes such as development and synaptic plasticity (Banko et al., 2005; Bidinosti et al., 2010; Martineau et al., 2013). Through their repressive effect on translation, 4E-BPs inhibit cell proliferation and act as tumor suppressors, and their activity is frequently dysregulated in cancer (Boussemart et al., 2014; Dowling et al., 2010; Hsieh et al., 2012; Martineau et al., 2013). Furthermore, dysregulation of the 4E-BP-eIF4E interaction is thought to contribute to autism spectrum disorders (ASDs) (Banko et al., 2005, 2007; Gkogkas et al., 2013). Consequently, a detailed molecular understanding of the interaction between eIF4E and 4E-BPs is of primary importance because of the increasing interest in designing eIF4E inhibitors as therapeutic tools to reduce translation in diseases associated with increased protein synthesis (Jia et al., 2012; Martineau et al., 2013).

The 4E-BPs exert their inhibitory effect on protein synthesis by interfering with the assembly of the eukaryotic initiation factor eIF4F, which consists of the cap-binding protein eIF4E, the scaffolding protein eIF4G, and the RNA helicase eIF4A (Jackson et al., 2010). eIF4E binds to the mRNA 5′ cap structure and interacts with eIF4G, which in turn initiates translation by recruiting the 43S preinitiation complex (Jackson et al., 2010). eIF4G interacts with eIF4E through a conserved motif (termed the canonical 4E-BM) of sequence $YX_4L\Phi$ (where Y denotes Tyr, X denotes any amino acid, L denotes Leu, and Φ denotes a hydrophobic residue), which is also present in 4E-BPs (Mader et al., 1995; Marcotrigiano et al., 1999). Because of the presence of this shared binding motif, 4E-BPs compete with eIF4G for binding to the same conserved patch of hydrophobic residues on the dorsal side of eIF4E, thereby blocking translation initiation (Gross et al., 2003; Mader et al., 1995; Marcotrigiano et al., 1999; Matsuo et al., 1997).

Downstream of the canonical (C) motifs, 4E-BPs contain non-canonical (NC) 4E-binding motifs (4E-BMs) connected by a linker of 15–30 residues. Although the non-canonical motifs do not share sequence similarity and are not conserved between orthologs across species, they all seem to bind to the same lateral surface of eIF4E, which is conserved and is not required for eIF4G binding (Gosselin et al., 2011; Igreja et al., 2014; Kinkelin et al., 2012; Lukhele et al., 2013; Mizuno et al., 2008; Paku et al., 2012). The non-canonical motifs increase the affinity of 4E-BPs for eIF4E by three orders of magnitude and are required for 4E-BPs to be able to compete with eIF4G and repress translation (Igreja et al., 2014; Lukhele et al., 2013; Paku et al., 2012).

The interaction of 4E-BPs with eIF4E is inhibited by the phosphorylation of 4E-BPs at multiple sites (Gingras et al., 1999, 2001). In vertebrate 4E-BP1–3 and its *Drosophila melanogaster* (*Dm*) ortholog Thor, two phosphorylation sites are located in the linker region. In the absence of high-resolution structures of eIF4E-4E-BP complexes in which the linker regions can be visualized, it has been difficult to rationalize how phosphorylation negatively regulates complex formation. Furthermore, despite the essential role of the non-canonical motifs of 4E-BPs in translational repression, the molecular details of the interaction of these motifs with eIF4E are known only for the *Dm* 4E-BP protein CUP (Kinkelin et al., 2012). However, because of the limited sequence identity and the lack of the regulatory linker region, the CUP model could not be used to derive general principles regarding the interaction of 4E-BPs with eIF4E.

To obtain general and predictive insight into the binding mode of 4E-BPs to eIF4E and its regulation, we adopted a comparative approach, solving the structures of three 4E-BPs bound to eIF4E

and of eIF4G bound to eIF4E. The structures reveal the binding mode of the linker regions and the non-canonical motifs and provide insight into the mechanism by which phosphorylation prevents 4E-BPs from repressing translation. Most important, our studies provide a hitherto missing structural framework for the rational design of translational inhibitors to treat malignancies in which eIF4E activity and protein synthesis are upregulated.

RESULTS

Structural Overview

To understand the binding mode of 4E-BPs to eIF4E, we crystallized and determined the structures of *Dm* eIF4E in complex with the minimal binding fragments of two 4E-BPs (Thor and eIF4E-transporter or 4E-T) at 1.10 and 2.15 Å resolution, respectively (Figures 1A–1G; Tables 1 and S1). The Thor and 4E-T peptides used in this study contained the canonical and non-canonical 4E-BMs connected by the linker region and retained the ability of the full-length proteins to bind to eIF4E (Igreja et al., 2014). To better understand the mechanism of the competition of 4E-BPs with eIF4G, we also crystallized *Dm* eIF4E bound to a *Dm* eIF4G peptide containing the canonical motif and the downstream auxiliary VKNVSI motif (*Hs* SDVVL; Figure S1), which stabilizes the interaction of eIF4G with eIF4E (Umenaga et al., 2011). The eIF4E-eIF4G complex was crystallized in the presence of an m^7GpppG cap analog, and the structure was determined at 2.4 Å resolution (Figures 1H, 1I, and S2; Table 1). Finally, we determined the crystal structure of human (*Hs*) 4E-BP1 in complex with *Hs* eIF4E at 1.75 Å resolution (Figure 1D; Table 1), which demonstrated conservation of the key structural features observed in the *Dm* complexes.

The structures of *Dm* eIF4E and *Hs* eIF4E are similar to the structures of free or bound eIF4E from diverse organisms, and no major conformational changes in the proteins occur upon 4E-BP binding (Figures 1B–1H and S2A) (Gross et al., 2003; Kinkelin et al., 2012; Mizuno et al., 2008; Paku et al., 2012; Papadopoulos et al., 2014; Siddiqui et al., 2012; Umenaga et al., 2011; Volpon et al., 2006). Briefly, eIF4E adopts a crescent-shaped conformation formed by a strongly bent β sheet of eight antiparallel β strands. The cap-binding cavity is located on the concave, ventral surface (Figure 1H). The convex surface is decorated by three α helices that form the dorsal binding surface of the protein. The major difference between the various eIF4E structures is observed in loops L1 and L3, which harbor the Trp residues that coordinate the cap structure. These loops undergo structural rearrangements upon binding to the cap structure (Volpon et al., 2006), as observed in the eIF4E-eIF4G complex bound to a cap analog (Figure 1H). In the absence of the cap analog, loop L1 adopts different conformations, and loop L3 is mainly disordered (Figures 1B–1D versus Figure 1H).

In contrast to eIF4E, the 4E-BPs differ in their structural details. Nevertheless, the interacting peptide sequences can be divided into three common structural elements, each of which binds to a defined and common surface on eIF4E (Figures 1B–1J and 2): (1) an N-terminal α helix formed by the canonical motifs and binding to the dorsal surface of eIF4E; (2) an elbow loop immediately following the canonical α helix that bends the peptide backbone by approximately 90°, orienting the linker regions downward to engage the lateral surface of eIF4E (Figure 1G); and (3) a C-terminal loop formed by the non-canonical motifs (NC loop). The NC loop shows the largest structural differences among the 4E-BPs (Figure 1E) and contacts a hydrophobic pocket on the lateral surface of eIF4E, providing an additional anchoring point. The hydrophobic pocket on the lateral surface of eIF4E is also contacted by the non-canonical motif of CUP, which adopts an α helical conformation (Figure 1J) (Kinkelin et al., 2012). However, the elbow loop and the linker region are not visible in the structure of the CUP-eIF4E complex (Figure 1J).

In the eIF4G-eIF4E complex, only the canonical motif and additional short N- and C-terminal extensions were observed (Figures 1H and S2B–S2F). The auxiliary VKNVSI was not visible in our structure, suggesting that it may form transient interactions or may be susceptible to proteolysis. The solution structure of the *Saccharomyces cerevisiae* eIF4E-eIF4G complex indicated that the eIF4G residues C-terminal to the canonical motif wrap around the unstructured N-terminal extension of eIF4E (Figures S2E and S2F) (Gross et al., 2003). However, it is not known whether a similar interaction forms between the metazoan proteins, as *Dm* eIF4G exhibited a similar affinity for eIF4E independently of whether the N-terminal extension was included (Table S2) (Igreja et al., 2014). Consequently, the eIF4E N-terminal extension was deleted in the construct used for crystallization (Table S1).

The 4E-BP interface on eIF4E overlaps with, yet is more extensive than, the eIF4G-eIF4E interface (Figures 2A–2D). It comprises a dorsal surface shared by the canonical helices of 4E-BPs and eIF4G, a lateral surface shared by all 4E-BPs but not used by eIF4G, and small surface patches unique to each of the 4E-BPs (Figure 2E).

The Canonical Motifs

As expected, the canonical motifs of 4E-BPs and eIF4G fold into an α helix, which is held in position by interactions analogous to the ones previously reported for canonical peptides in complex with eIF4E (Figures 3A–3H, S2G, and S2H) (Gross et al., 2003; Kinkelin et al., 2012; Mizuno et al., 2008; Paku et al., 2012; Umenaga et al., 2011). Specifically, residues corresponding to LΦ in the consensus sequence $YX_4L\Phi$ form similar hydrophobic contacts with the conserved *Dm* eIF4E residues V102, W106, and L167 (V69, W73, and L135 in *Hs* eIF4E, respectively; Figures 3A–3D). The hydroxyl group of the invariant Tyr (Y) side chain contacts the backbone of the conserved H70-P71-L72 motif in eIF4E (H37-P38-L39 in *Hs* eIF4E) and is in van der Waals contact with V102 of *Dm* eIF4E (*Hs* V69) in all structures (Figures 3E–3H, S2G, and S2H).

The canonical α helices are flanked on both sides by conserved Arg/Lys residues that adopt a similar orientation in all structures (Figures 2A–2D, 3A–3D, S1, and S2G), suggesting an extended consensus sequence for the canonical motif ($YX[R/K]X_2L\Phi X_2R/K$). Accordingly, alignment and sequence profile of the canonical motifs of all available eIF4G and 4E-BP sequences indicates that the Arg/Lys residues at positions 2 and 9 are well conserved (Figure S2I). The side chains of the C-terminal Arg residues ($R63^{Thor}$, $R19^{4E\text{-}T}$, $R63^{BP1}$, $R628^{4G}$, and $R336^{CUP}$) stack onto $W106^{4E}$ (*Hs* $W73^{4E}$), shielding its hydrophobic surface from solvent exposure (Figures 3A–3D and S2G).

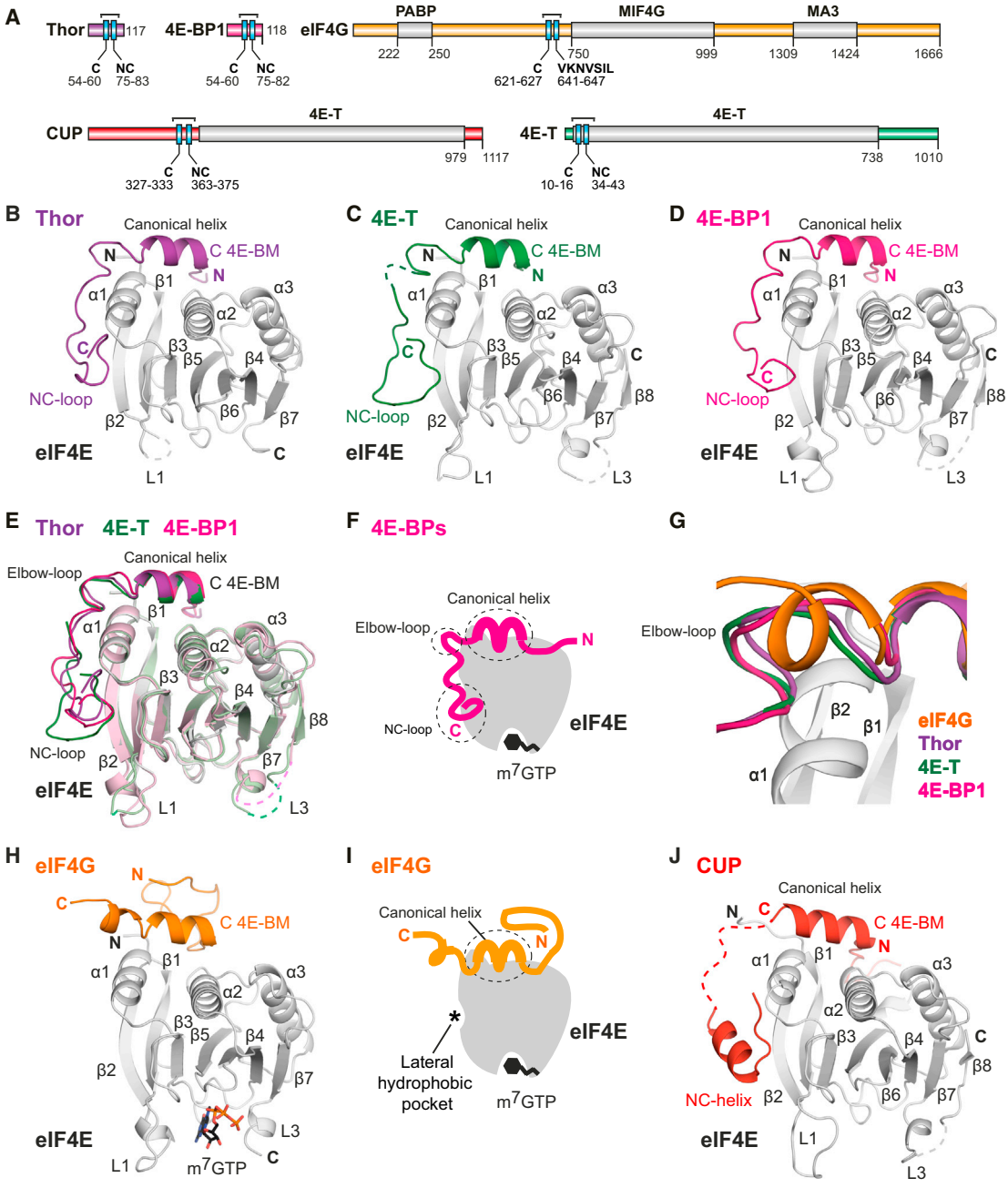


Figure 1. Overall Structures of 4E-BPs and eIF4G Bound to eIF4E

(A) Domain organization of eIF4G and the 4E-BPs analyzed in this study. The 4E-BPs contain a canonical (C) and a non-canonical (NC) 4E-BM. CUP and 4E-T contain a region with similarity to human 4E-T; eIF4G contains a PABP-interacting region and MIF4G and MA3 domains. See also Figure S1.

(B–D) Overview of the structures of eIF4E bound to the indicated 4E-BPs. Selected secondary structure elements are labeled in black for eIF4E and in color for 4E-BPs.

(E) Superposition of the structures of eIF4E bound to 4E-T, Thor, and 4E-BP1 reveals differences in the conformation of the NC loops.

(F) Schematic representation of eIF4E bound to 4E-BPs.

(G) Alignment of the eIF4G and 4E-BP structures by superposition of their canonical helices indicates that the elbow loop is a common structural feature of 4E-BPs.

(H) Overview of the structure of eIF4E bound to eIF4G and cap analog. See also Figure S2.

(I) Schematic representation of eIF4E bound to eIF4G and m^7GTP.

(J) Overview of the structure of the eIF4E-CUP complex (PDB accession number 4AXG; Kinkelin et al., 2012).

Table 1. Data Collection and Refinement Statistics

	Dm eIF4E-Thor Complex	*Dm* eIF4E-4E-T Complex	*Dm* eIF4E-eIF4G Complex	*Hs* eIF4E-4E-BP1 Complex	*Dm* eIF4E-Triple Chimera Complex 1	*Dm* eIF4E-Triple Chimera Complex 2
Space group	$P2_1$	$P6_522$	$P2_12_12_1$	$P2_1$	$P6_1$	$P2_12_12_1$
Unit cell						
Dimensions (Å)						
a, b, c	35.5, 73.3, 38.3	40.0, 40.0, 443	54.9, 68.1, 104	42.1, 66.6, 43.1	94.5, 94.5, 132	38.7, 94.5, 167
Angles (°)						
α, β, γ	90, 112.5, 90	90, 90, 120	90, 90, 90	90, 118.3, 90	90, 90, 120	90, 90, 90
Data collection						
Wavelength (Å)	1.000	1.000	1.038	1.070	1.000	0.999
Resolution (Å)	36.6–1.10	36.9–2.15	48.6–2.40	37.9–1.75	47.3–2.62	47.9–2.52
	(1.13–1.10)	(2.21–2.15)	(2.46–2.40)	(1.80–1.75)	(2.69–2.62)	(2.59–2.52)
R_{sym}	0.032 (0.324)	0.091 (0.707)	0.095 (0.790)	0.041 (0.309)	0.068 (0.846)	0.110 (0.582)
Mean $I/\sigma I$	23.7 (4.0)	25.6 (4.5)	15.4 (2.3)	13.4 (2.5)	14.4 (2.1)	9.4 (2.1)
Completeness (%)	95.8 (88.7)	100 (100)	99.9 (99.8)	96.9 (98.1)	99.6 (99.5)	96.2 (97.1)
Multiplicity	6.0 (4.2)	34.7 (31.4)	6.5 (6.6)	2.9 (2.8)	4.9 (5.0)	4.7 (4.1)
Refinement						
Resolution (Å)	36.5–1.10	36.9–2.15	48.5–2.40	37.9–1.75	47.3–2.62	41.8–2.52
No. of reflections	70,121	12,796	15,777	20,588	19,990	20,794
R_{work}/R_{free}	0.145/0.161	0.235/0.271	0.210/0.253	0.192/0.237	0.205/0.262	0.205/0.245
No. of atoms	1,852	1,739	3,176	1,878	5,001	5,054
Protein	1,648	1,701	3,066	1,759	5,001	5,044
Ligand/ion	13	/	66	/	/	/
Water	191	38	44	119	/	10
B factors (Å^2)	19.8	50.0	53.6	32.8	70.1	44.4
Protein	18.6	50.1	53.7	32.4	70.1	44.4
Ligand/ion	33.5	/	56.9	/	/	/
Water	29.3	45.4	43.4	38.4	/	32.8
Ramachandran plot						
Favored (%)	98.5	97.0	98.4	98.6	98.0	99.0
Disallowed (%)	0	0	0	0	0	0
Root-mean-square deviation						
Bond lengths (Å)	0.007	0.002	0.003	0.004	0.003	0.004
Bond angles (°)	1.168	0.564	0.667	0.780	0.647	0.671

Values in parentheses are for highest-resolution shell. Ligands: *Dm* eIF4E-Thor complex: tetraethylene glycol, two Na atoms; *Dm* eIF4E-eIF4G complex: one m^7GTP per eIF4E molecule.

Furthermore, in the 4E-BP structures, these Arg residues contact the side chain of $N110^{4E}$ (*Hs* $N77^{4E}$), which is not formed in the eIF4G structure (see Figure 4).

On the N-terminal portion of the canonical α-helix, the Arg/Lys residues ($R56^{Thor}$, $K12^{4E\text{-}T}$, $R56^{BP1}$, $R622^{4G}$, and $R329^{CUP}$) contact $L167^{4E}$ (*Hs* L135). Furthermore, in all structures except for Thor, these residues form a salt bridge to a conserved acidic residue in eIF4E (*Dm* $D164^{4E}$, *Hs* $E132^{4E}$; Figures 3B–3D). In Thor, this salt bridge is formed by K60 at the Φ position of the consensus motif (Figure 3A), which is a specific feature of the Thor interaction.

The Elbow Loops

A key distinguishing feature of 4E-BPs is the linker region, which extends the 4E-BP binding interfaces beyond the canonical α helices and forms the characteristic elbow loop that enables 4E-BPs to engage two orthogonal surfaces on eIF4E (Figure 1G). Despite the lack of sequence conservation, the elbow loops are organized around common principles (Figures 4A–4F). They begin immediately after the Arg residues flanking the canonical helices on the C-terminal side ($R63^{Thor,BP1}$ and $R19^{4E\text{-}T}$); they comprise a helical half-turn that begins with a Pro ($P66^{Thor,BP1}$) or a Gly ($G22^{4E\text{-}T}$) residue located at the tip of the elbow; and they end with similarly arranged Ser or Thr residues ($S68^{Thor}$, $S24^{4E\text{-}T}$, and $T68^{BP1}$). The side chains of these Ser/Thr residues contact the carbonyl oxygens of preceding residues ($L62^{Thor}$, $L18^{4E\text{-}T}$, $R63^{BP1}$) and fix the backbone, as do the internal interactions within the helical half-turn (Figures 4A–4C).

The elbow loops are anchored to the lateral surface of eIF4E primarily through interactions with the conserved Asn (*Dm*

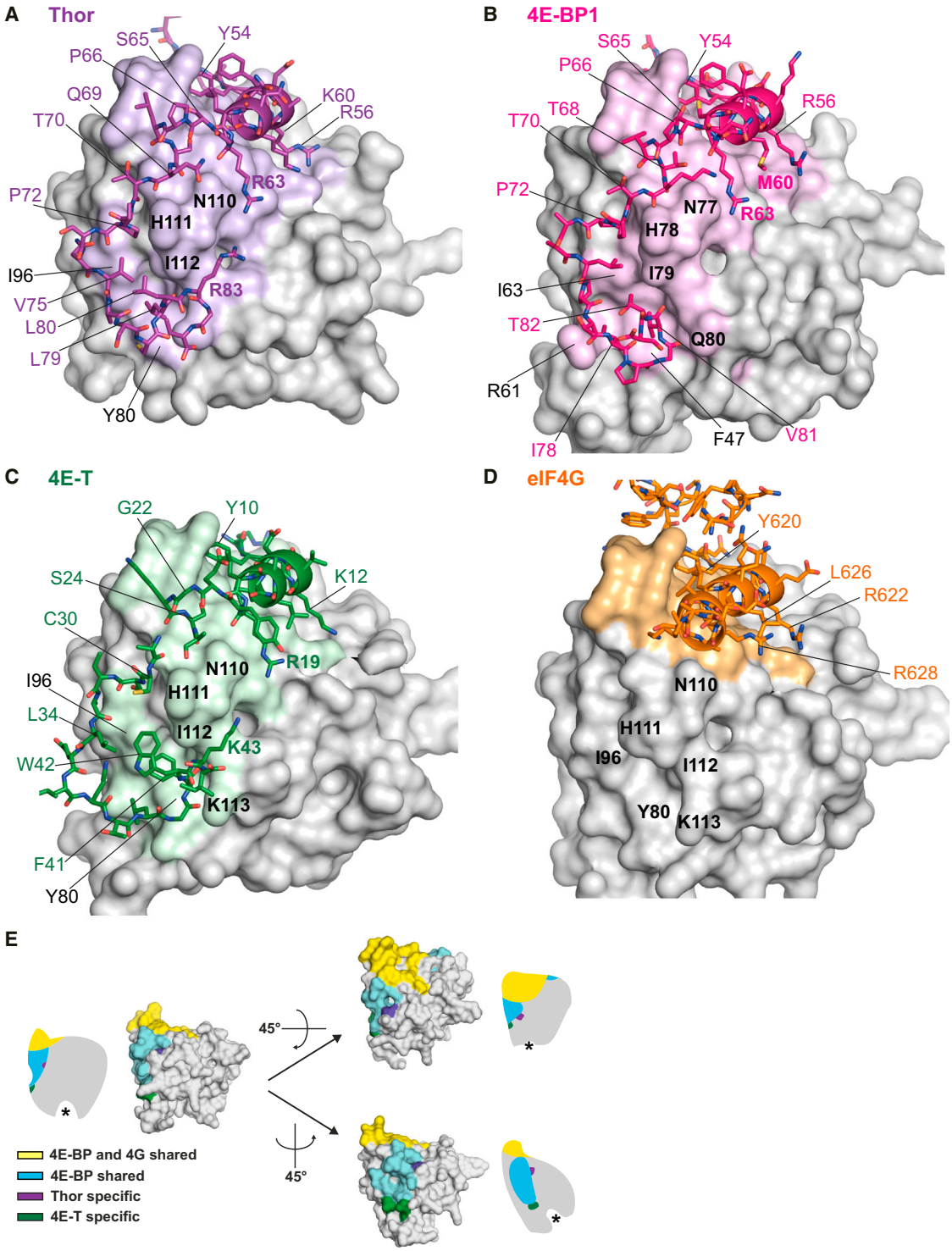


Figure 2. Surface Representation of the 4E-BP and eIF4G Binding Interfaces on eIF4E

(A–D) The eIF4E residues that are within 4 Å of the bound 4E-BP or eIF4G peptides are shown in color. Selected residues lining the binding surface are labeled in black for eIF4E and in color for 4E-BPs and eIF4G. The 4E-BP and the eIF4G peptides are shown as sticks.

(E) Superposition of the eIF4E surfaces that bind eIF4G and 4E-BPs. The region of overlap between the eIF4G and 4E-BP binding surfaces is shown in yellow. The surface shared by 4E-BPs but not used by eIF4G is colored in cyan. Surfaces used exclusively by Thor and 4E-T are shown in magenta and green, respectively.

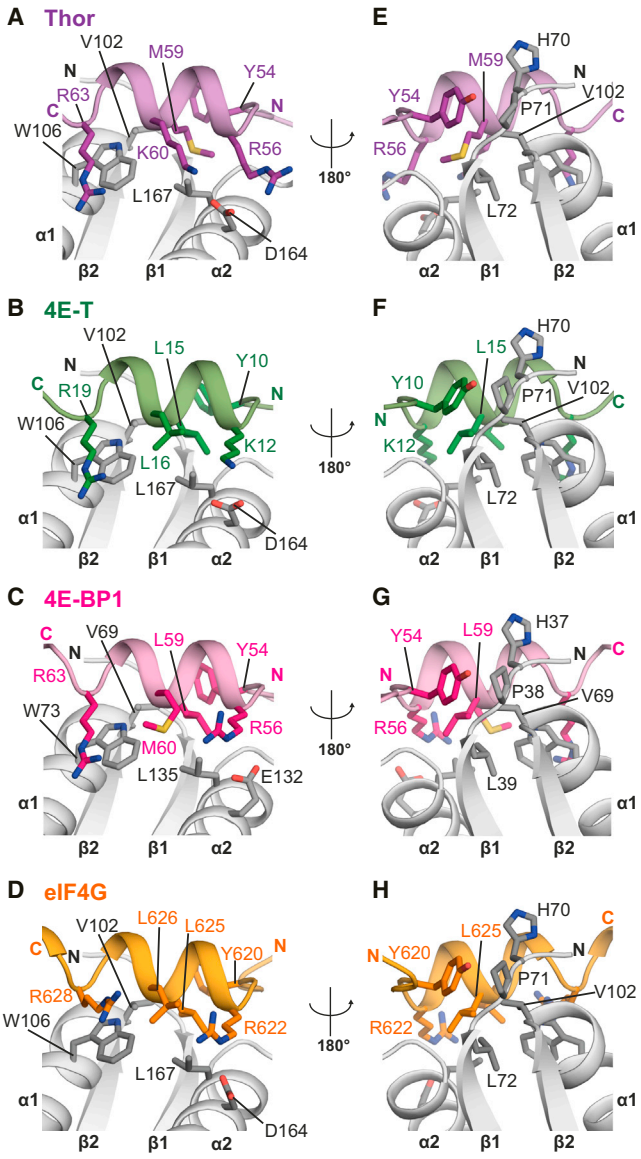


Figure 3. Interaction of the Canonical Motifs of 4E-BPs and eIF4G with the Dorsal Surface of eIF4E

(A–H) Close-up views of the interaction between the dorsal surface of eIF4E and the canonical helices of 4E-BPs and eIF4G. Selected interface residues are shown as gray sticks for eIF4E and as colored sticks for 4E-BPs and eIF4G. See also Figures S2–S4.

N110^{4E} and *Hs* N77^{4E}) and His residues (*Dm* H111^{4E}, *Hs* H78^{4E}). The Asn residue contacts the backbone carbonyl group of the structurally equivalent residues S68^{Thor}, T68^{BP1}, and S24^{4E-T}, as well as the side chains of the structurally conserved R63^{Thor,BP1} and R19^{4E-T} located C-terminal to the canonical helices (Figures 2 and 4D–4F). The aromatic ring of the conserved His residue is locked in position upon ligand binding. In the 4E-BP structures, this ring stacks on the peptide bonds preceding and following the conserved T70^{Thor,BP1} phosphorylation site (perpendicular and parallel π stacking, Figures 4D and 4E) or is in van der Waals contact with the C30^{4E-T} side chain (Figure 4F, corresponding to P71^{Thor,BP1}).

The Non-Canonical Loops

A second anchoring point on the lateral surface of eIF4E is provided by a conserved hydrophobic pocket. This pocket is lined by residues Y80^{4E}, I96^{4E}, and I112^{4E} (F47^{4E}, I63^{4E}, and I79^{4E} in *Hs* eIF4E, respectively) and is engaged by the residues that form the non-canonical motifs of the 4E-BPs. Strikingly, however, the various 4E-BPs, including CUP, adopt very different backbone conformations as they approach the pocket (Figures 1E, 1J, and 2). In particular, the pocket is engaged by the hydrophobic side chains of L79^{Thor}, V81^{BP1}, F41^{4E-T}, or I373^{CUP} (Figures 4G–4I and S2J). Furthermore, the carbonyl oxygens of these residues are fixed by a main-chain contact to the nitrogen of *Dm* K113^{4E} (*Hs* Q80^{4E}), which is internally buttressed at its carbonyl oxygen by *Dm* Y124^{4E} (*Hs* Y91^{4E}; Figures 4J–4L). The three *Dm* 4E-BPs also form one additional common main-chain contact from the nitrogen of R81^{Thor} (K43^{4E-T}, R375^{CUP}) to the carbonyl oxygen of *Dm* H111^{4E}, whereas the corresponding Arg side chain locks the backbone of *Dm* I112^{4E} in position (Figures 4G, 4I, 4J, and 4L). These Arg/Lys residues are not conserved in vertebrate 4E-T and 4E-BP1–3 (Figure S1).

Importantly, although the peptides of 4E-BP1, Thor, and 4E-T show some internal stabilization by backbone hydrogen bonds or hydrophobic contacts, they do not form defined secondary structures. This property differs for CUP, in which the residues that engage the eIF4E pocket and its vicinity are located on an α helix (Figures 1J and S2J) (Kinkelin et al., 2012). However, this helix seems to form only upon binding, as indicated by the large entropic penalty associated with the binding of CUP to eIF4E (Igreja et al., 2014).

Structure Validation

To validate the observed interactions, we substituted key interface residues and tested complex formation in pull-down assays in vitro using recombinant proteins and in coimmunoprecipitation assays in human and *Dm* cells. Substitutions of residues in either the canonical or the non-canonical motif of Thor or 4E-T did not prevent complex formation in vitro in the absence of eIF4G, indicating that the remaining sequences are sufficient for binding to eIF4E (Figures S3A and S3B) (Igreja et al., 2014). However, these mutations did abolish or impair the ability of 4E-BPs to compete with endogenous eIF4G for binding to eIF4E in human and *Dm* cell lysates (Figures S4A–S4D) (Igreja et al., 2014). Binding to eIF4E was abrogated when mutations in the canonical and non-canonical motifs were combined (Figures S3A, S3B, and S4A–S4D C*+NC*).

Substitutions of each of the Arg/Lys residues in the canonical or non-canonical motifs by glutamic acid (or alanine in the case of Thor) did not affect the binding of *Dm* 4E-BPs to eIF4E in vitro (Figures S3C–S3F). Binding to eIF4E was impaired only when the three Arg/Lys residues (i.e., two in the canonical and one in the non-canonical motif) were substituted (Figures S3C–S3E), indicating that these residues contribute to the affinity of the interaction. Furthermore, although the double mutations in the two Arg/Lys residues in the canonical motif did not affect the binding of Thor to free eIF4E, they did impair its ability to compete with

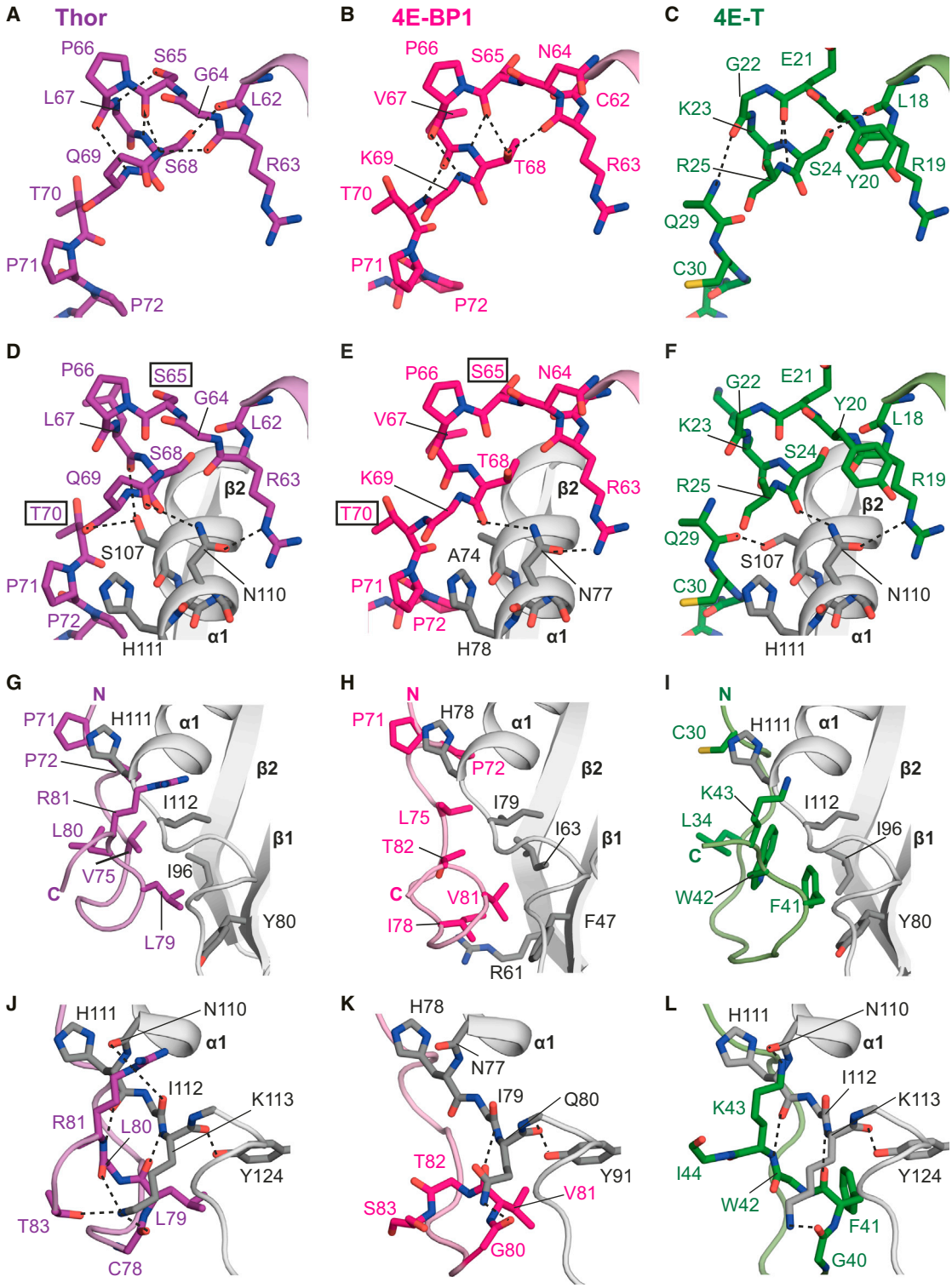


Figure 4. The Elbow and the Non-Canonical Loops

(A–C) Close-up views of the elbow loops and interactions within the loops. The side chains of L67^{Thor}, K69^{BP1} and K23^{4ET} have been omitted for clarity.

(D–F) Close-up views of the interaction between eIF4E and the 4E-BP linker regions C-terminal to the elbow loops. Selected interface residues are shown as gray and colored sticks for eIF4E and 4E-BPs, respectively. The side chain of K69^{BP1} has been omitted for clarity; R25^{4ET} and Q29^{4ET} are not visible in the structure. Conserved phosphorylated residues in Thor and 4E-BP1 are boxed.

(legend continued on next page)

eIF4G for binding to eIF4E (Figure S3G). We therefore conclude that the Arg/Lys residues in the canonical motif participate in complex formation and in the competition mechanism.

Phosphorylation of Linker Residues Interferes with the Competition Mechanism

Our high-resolution crystal structures provide insight into the conformation of the two conserved phosphorylation sites (S65 and T70) in the linker regions of Thor and 4E-BP1 (Figures 4D and 4E). In both structures, the side chains of these residues do not contribute to the binding interface, indicating that phosphorylation does not directly interfere with complex formation. Furthermore, phosphomimetic mutations at S65 and T70 (S65D, T70D) did not affect the interaction of a Thor fragment (32–83) with eIF4E (Figure S5 and Table S2), as reported previously (Gingras et al., 2001; Karim et al., 2001). However, the mutations disrupted the interaction with eIF4E when combined with mutations in the canonical motif (Figure S4E), suggesting that phosphorylation interferes with the 4E-BP binding mechanism.

We therefore tested the effect of phosphomimetic mutations on the ability of the Thor fragment to displace eIF4G from preassembled eIF4E-eIF4G complexes in competition assays in vitro. Preassembled eIF4E-eIF4G complexes were challenged with 2-fold molar excess of Thor fragment. The amount of eIF4G remaining bound to eIF4E over time was determined (Figures 5A and 5B). In the presence of MBP (maltose-binding protein), eIF4G remained bound to eIF4E, as expected (Figure 5B, lanes 5 and 10). The wild-type Thor fragment displaced eIF4G from preassembled eIF4E-eIF4G complexes more efficiently than the fragment containing phosphomimetic mutations at S65 and T70 (Figures 5A and 5B). Indeed, the half-life of the eIF4E-eIF4G complexes was 18 min in the presence of wild-type Thor, compared with 60 min for the S65D, T70D mutant (Figures 5A and 5B). Conversely, eIF4G displaced the Thor phosphomimetic mutant from preassembled eIF4E-Thor complexes but failed to efficiently displace Thor wild-type (Figures S4F and S4G).

To further validate the conclusions obtained using the phosphomimetic mutant, we performed isothermal titration calorimetry (ITC) measurements and competition assays with Thor synthetic peptides (residues 50–83), either wild-type or phosphorylated at positions S65 and T70. In agreement with the results obtained with the phosphomimetic mutant, the wild-type and phosphorylated peptides exhibited nanomolar affinities for eIF4E (Figure S5 and Table S2). However, the phosphorylated peptide was strongly impaired in its ability to displace eIF4G from preassembled eIF4E-eIF4G complexes (Figures 5C and 5D). Note that similar amounts of wild-type and phosphorylated peptide were used in the competition assay but the phosphorylated peptide stains less efficiently with Coomassie (Figure 5D). We conclude that the phosphomimetic mutations faithfully recapitulate the impact of phosphorylation of the linker residues.

To determine the effect of phosphorylation in translational repression, we tested whether the S65D, T70D mutant repressed the expression of a *Renilla* luciferase (R-Luc) reporter when coexpressed in S2 cells. A short uncapped and unadenylated RNA served as a transfection control (control RNA). To rule out the possibility that the inhibition of R-Luc expression resulted from changes in mRNA levels, we analyzed these levels by northern blotting and determined the translation efficiencies (i.e., the R-Luc activity normalized to R-Luc mRNA levels).

Remarkably, phosphomimetic mutations at S65 and T70 were sufficient to prevent Thor from repressing translation of the R-Luc reporter in *Dm* cells (Figures 5E and S4H). The proteins were expressed to comparable levels (Figure 5F). We conclude that the phosphorylation of linker residues is sufficient to reduce the potency of 4E-BPs as translational repressors. Taken together, our results indicate that phosphorylation at S65 and T70 residues does not affect binding to free eIF4E but does interfere with the ability of 4E-BPs to displace eIF4G from preexisting eIF4E-eIF4G complexes and repress translation.

Design of a Potent Bipartite 4E-BP Mimic

The binding mode observed in our structures and the observation that mutations in the non-canonical motifs prevent 4E-BPs from competing with eIF4G (Figures S4A–S4D) (Igreja et al., 2014) support a model in which 4E-BPs use the lateral surface of eIF4E to dock on the eIF4E-eIF4G complex before eIF4G dissociates and then compete for binding to the dorsal surface through the canonical motifs. Conceptually, and as supported by the present structures, this two-step competition mechanism does not require conformational changes in eIF4E and is therefore distinct from the mechanism described for 4EGI-1 (Papadopoulos et al., 2014). 4EGI-1 is a small molecule inhibitor of the eIF4E-eIF4G interaction that binds in close proximity to the hydrophobic pocket on the lateral surface of eIF4E and presumably displaces eIF4G from the dorsal surface by an allosteric mechanism (Figure 6A) (Moerke et al., 2007; Papadopoulos et al., 2014). The observable rearrangements in eIF4E upon 4EGI-1 binding include an extension of helix α1 by one turn, which disrupts the backbone stabilization by Y91^{4E} (*Dm* Y124^{4E}), and a reorientation of the side chain of H78^{4E} (*Dm* H111^{4E}) toward the hydrophobic pocket that is not possible in the presence of 4E-BPs (Figures 6B–6E and S4I) (Papadopoulos et al., 2014).

Compared with the nanomolar affinities of the 4E-BPs, the affinity of 4EGI-1 for eIF4E (10-20 μM) is relatively low, and higher concentrations of 4EGI-1 are required to displace eIF4G from eIF4E (Moerke et al., 2007; Papadopoulos et al., 2014). We therefore propose that bipartite 4E-BP mimics are likely to be more potent inhibitors of protein synthesis than monopartite inhibitors.

To test this prediction, we engineered a 4E-BP mimic based on our structural information and on previously reported biophysical properties of the 4E-BP peptides (Igreja et al., 2014). The rationally designed inhibitor comprised the high-affinity canonical motif of 4E-T, the non-canonical motif of CUP (which forms a helix and binds to eIF4E in isolation; Kinkelin et al., 2012; Igreja et al., 2014), and the linker region of Thor, which

(G–I) Interaction of the non-canonical loops with the lateral surface of eIF4E. See also Figure S2J.

(J–L) Close-up views of the non-canonical loops and interactions with residues lining the lateral hydrophobic pocket of eIF4E. Selected 4E-BP and eIF4E residues are shown as sticks (C78–R81^{Thor}, G80–S83^{BP1}, and G40–I44^{4ET}). eIF4E secondary structure elements (β1 and β2) and the side chains of *Dm* I112^{4E} (*Hs* I79^{4E}), C78^{Thor}, L80^{Thor}, T82^{BP1}, W42^{4ET}, and I44^{4ET} have been omitted for clarity.

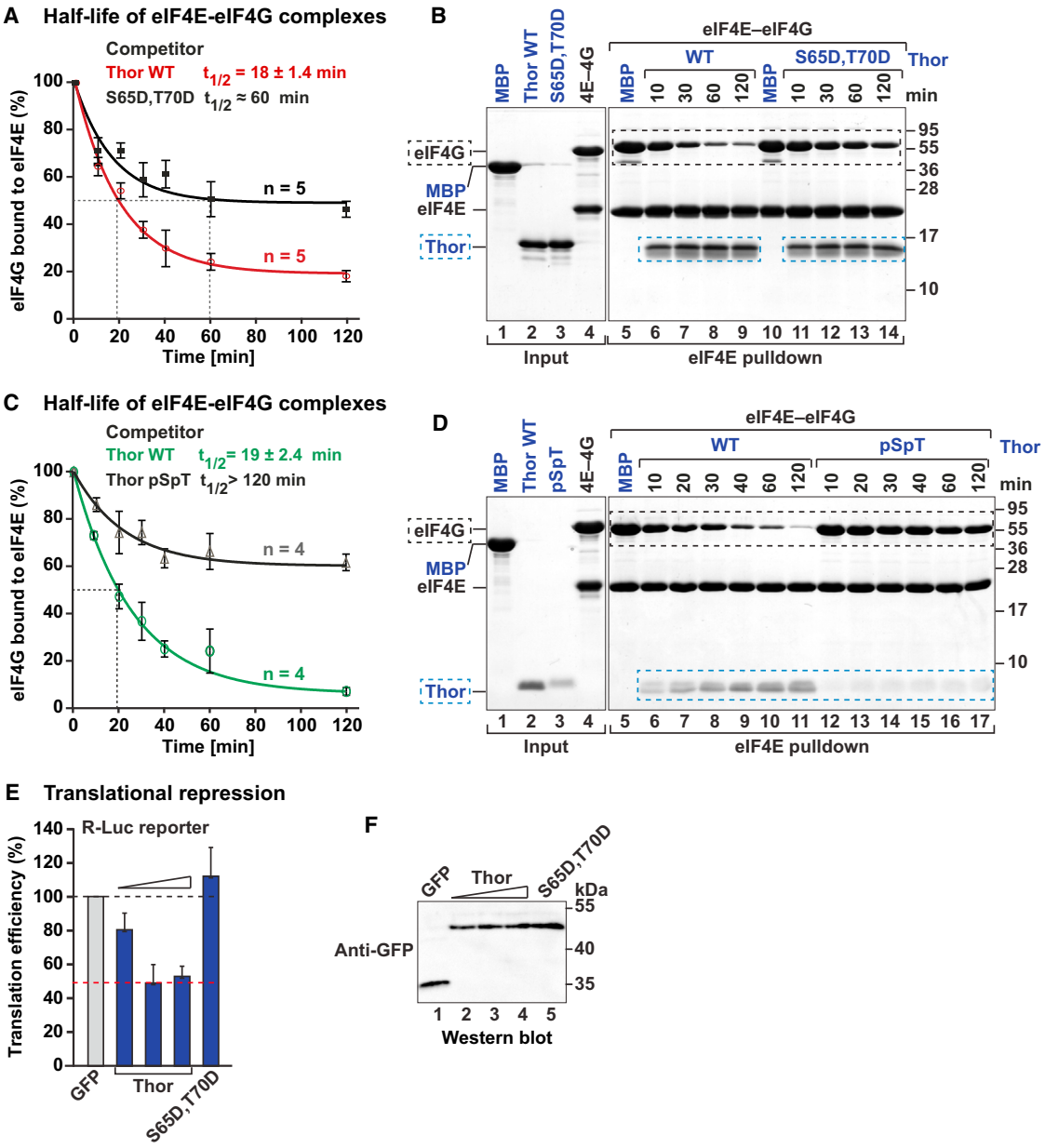


Figure 5. Effects of Phosphorylation of Thor Linker Residues

(A and B) Half-life of eIF4E-eIF4G complexes. Purified eIF4E-eIF4G complexes containing His_6-tagged eIF4E and MBP-eIF4G (residues 578–650) were incubated with a 2-fold molar excess of Thor (residues 32–83, either wild-type [WT] or the S65D, T70D phosphomimetic mutant). MBP served as a negative control. The proteins bound to eIF4E were pulled down using Ni-NTA beads at the indicated time points. Mean values ± SD from five independent experiments (n = 5) are shown. A representative SDS-PAGE gel is shown in (B). The competitor peptides are labeled in blue, and their positions are highlighted by blue dashed boxes. The black dashed boxes indicate the position of MBP-eIF4G. See also Figures S4F and S4G.

(C and D) The half-life of eIF4E-eIF4G complexes in the presence of a 2-fold molar excess of Thor synthetic peptides (residues 50–83, either WT or phosphorylated at S65 and T70) was determined as described in (A) and (B). Note that the phosphorylated peptide does not stain well with Coomassie. See also Figure S5 and Table S2.

(E and F) Translation efficiency of a R-Luc reporter in *Dm* cells expressing GFP or GFP-tagged Thor (WT or the S65D, T70D phosphomimetic mutant). Mean values ± SD from three independent experiments are shown. The red dashed line indicates the translation efficiency for the maximum repressive activity exhibited by Thor WT, whereas the black dashed line indicates the translation efficiency expected in the absence of repression. Protein expression was analyzed by western blotting (F). The RNA levels were determined by northern blotting (Figure S4H).

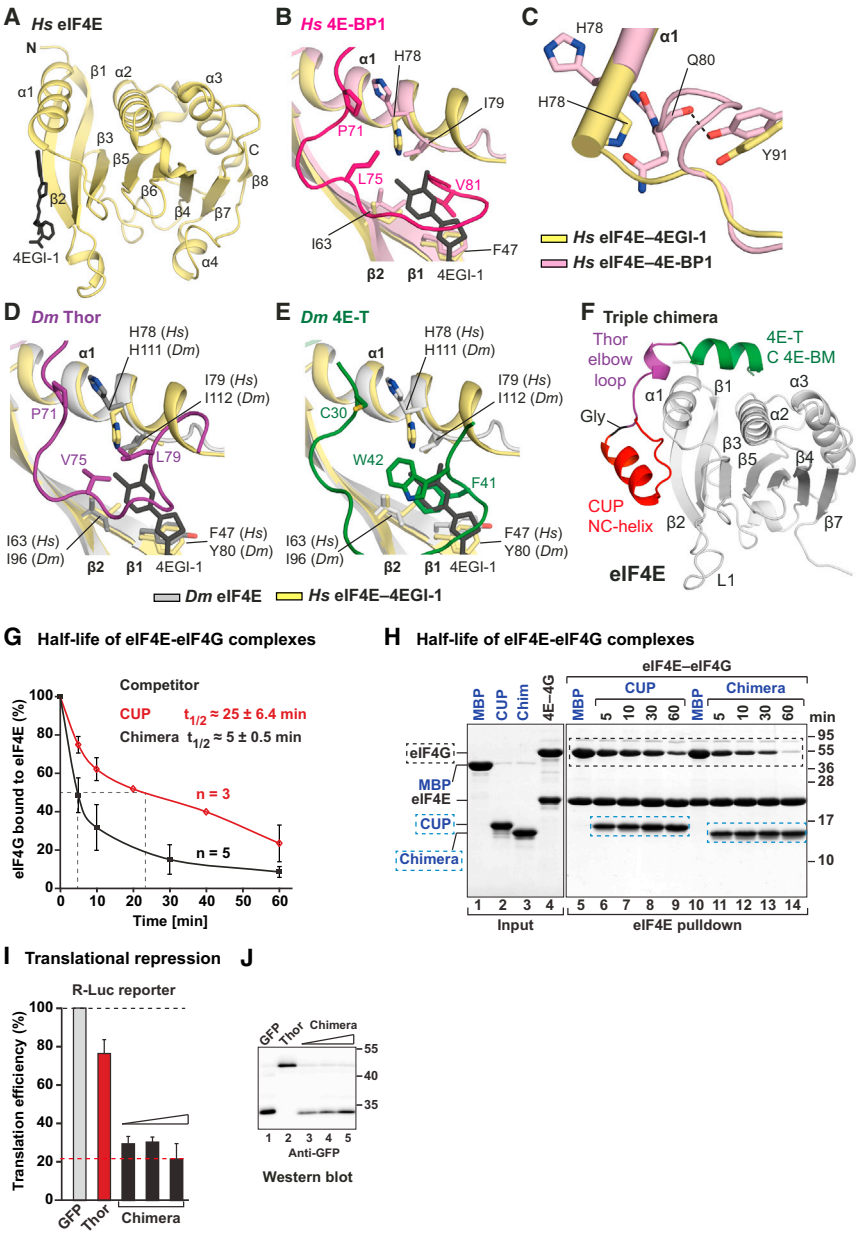


Figure 6. Translational Repression by a 4E-BP Mimic Peptide

(A–E) Overview of the structure of eIF4E bound to the 4EGI-1 inhibitor (PDB accession number 4TPW; Papadopoulos et al., 2014) and superposition with the indicated 4E-BP structures. Selected secondary structure elements and residues are labeled in black for eIF4E and in color for the 4E-BPs. See also Figure S4I.

(F) Overview of the structure of eIF4E bound to the triple chimera peptide (crystal form 2). Selected secondary structure elements are labeled in black for eIF4E and in color for the peptide. See also Figures S6 and S7.

(G and H) The half-life of eIF4E-eIF4G complexes in the presence of 1.5-fold molar excess of CUP or of the triple chimera peptide was determined as described in Figures 5A and 5B. See also Figures S7A and S7B.

(I and J) The translation efficiency of a R-Luc reporter in *Dm* cells expressing GFP or GFP-tagged Thor (WT) or the triple chimera was analyzed as described in Figures 5E and 5F. Mean values ± SD from three independent experiments are shown. The expression of the GFP-tagged proteins was analyzed by western blotting (J). The RNA levels were determined by northern blotting (Figure S7C). See also Figures S7D and S7E.

exhibits the most extensive intra- and intermolecular interaction network. ITC experiments on this triple chimera showed that it possesses 2.6-fold higher binding affinity for eIF4E than the CUP or Thor peptides (Figure S5F and Table S2).

We also determined the structure of the triple chimera peptide bound to *Dm* eIF4E in two crystal forms at 2.62 and 2.52 Å resolution (Figures 6F and S6A–S6E and Table 1). These structures demonstrated that the designed peptide binds to eIF4E as expected and in different crystal packing environments. In particular, the canonical helix of 4E-T, the linker region of Thor, and the non-canonical helix of CUP adopted conformations similar to those observed in the structures of the parental proteins, with root-mean-square deviations no higher than 0.67 Å (Figures S6F–S6H).

Next, we tested the ability of the chimeric peptide to displace eIF4G in competition assays in vitro. We observed that the triple chimera displaced eIF4G from preassembled eIF4E-eIF4G complexes more efficiently than CUP or Thor peptides when added at 1.5-fold molar excess. Indeed, the half-life of the eIF4E-eIF4G complexes was 5- and 3-fold shorter in the presence of triple chimera peptide compared with the CUP or Thor peptides, respectively, at the same concentration (Figures 6G, 6H, S7A, and S7B).

Importantly, the chimeric peptide fused to GFP repressed the expression of an R-Luc reporter in *Dm* cells more efficiently than wild-type Thor, even when expressed at a lower level (Figures 6I, 6J, and S7C). Additionally, the chimeric peptide fused to GFP interacted with *Hs* eIF4E (Figure S7D) and competed with endogenous eIF4G for binding to eIF4E in human cells (Figure S7E). These results provide a basis for the rational design of potent 4E-BP mimics for therapeutic applications.

DISCUSSION

Our study reveals common conserved principles of the interaction between 4E-BPs and eIF4E as well as individual features that guide the binding of the non-canonical motifs of 4E-BPs. Furthermore, the structures presented here provide insight into the mechanism by which phosphorylation prevents 4E-BPs

from repressing translation. Finally, our structural analyses provide a basis for the rational design of eIF4E inhibitors for the treatment of diseases in which eIF4E activity is upregulated.

4E-BP Binding Mechanisms

In combination with previous studies, our structures show that the canonical motifs of diverse 4E-BPs and eIF4G use a similar mechanism to bind eIF4E and that this binding mode is conserved (Gross et al., 2003; Kinkelin et al., 2012; Mizuno et al., 2008; Paku et al., 2012; Umenaga et al., 2011). In contrast, the binding mode of the linker regions and non-canonical motifs of 4E-BPs differ in the molecular details as expected on the basis of the differences in sequence and length of these regions. Nevertheless, these regions exhibit common structural features, such as the elbow and the non-canonical loops, which are unique to 4E-BPs.

Although the linker regions and non-canonical loops are largely devoid of secondary structure elements, such as α helices and β strands, they adopt a defined conformation in the bound state that is stabilized by intra- and intermolecular interactions. Previous studies indicated that the corresponding regions of *Hs* 4E-BP2 and sea urchin 4E-BP form "fuzzy complexes" in solution, sampling alternative conformations in the bound state and contacting larger surfaces on eIF4E (Gosselin et al., 2011; Lukhele et al., 2013). However, the fact that we could crystallize defined states and conformations that remain the same in a different crystal packing environment in the context of the designed chimera, indicates that these regions also have a preferred conformation in the bound state, probably reflecting the most stable state.

Competition with eIF4G

The 4E-BP interface on eIF4E overlaps with the interface of eIF4G but is more extensive and buries a total surface area of 1,450 $Å^2$ (Thor), 1,370 $Å^2$ (4E-T), or 1,235 $Å^2$ (4E-BP1) relative to 655 $Å^2$ for eIF4G in our structure (Figure 2). The 4E-BP interface comprises the dorsal surface of eIF4E shared with eIF4G and the lateral surface of eIF4E shared by all 4E-BPs but not used by eIF4G. Although we cannot exclude that metazoan eIF4G also contacts the lateral surface of eIF4E, as observed in the solution structure of *Sc* eIF4G (Gross et al., 2003), several lines of evidence indicate that these potential additional contacts may be transient in nature and do not require the lateral hydrophobic pocket of eIF4E. First, mutations in the canonical motif of eIF4G abolish binding to eIF4E, indicating that the sequences flanking the canonical motif are not sufficient to mediate eIF4E-binding. In contrast, 4E-BPs exhibit a truly bipartite binding mode, and mutations in the canonical motifs do not disrupt binding to eIF4E in vitro, because the linker and non-canonical regions are sufficient for binding (Igreja et al., 2014). Second, mutations in the lateral hydrophobic pocket of eIF4E prevent 4E-BPs from competing with eIF4G and from repressing translation, but these mutations do not affect eIF4G binding to eIF4E (Igreja et al., 2014). These observations, together with the binding mode observed in our structures, suggest that 4E-BPs can bind to eIF4E-eIF4G complexes before eIF4G dissociates and then benefit from the local enrichment to efficiently compete for binding to the dorsal surface using the canonical motifs.

Nevertheless, eIF4G most likely establishes interactions with eIF4E beyond the canonical motif that are not visible in our structure. Indeed, *Hs* and *Dm* eIF4G peptides containing the canonical motif and flanking sequences including the auxiliary motif exhibit nanomolar affinities for eIF4E, in contrast to the micromolar affinities observed for the peptides containing only the canonical motif (Umenaga et al., 2011). These additional interactions may differ across species, as the auxiliary (VKNVSI) motif is not conserved (Figure S1E).

Phosphorylation Interferes with the Competition Mechanism

The binding of 4E-BPs to eIF4E is regulated by coordinated and ordered phosphorylation events. In human 4E-BPs, phosphorylation at Thr residues upstream of the canonical motif constitutes a priming event that allows, upon external stimuli, the phosphorylation of the Thr and Ser residues in the linker region. Consequently, two waves of phosphorylation events (independent and dependent on external stimuli) are necessary to prevent 4E-BP1 from binding to eIF4E (Gingras et al., 1999, 2001). In the absence of high-resolution structures in which these residues could be visualized, it has been difficult to determine the impact of phosphorylation on the interaction of 4E-BPs with either the lateral or dorsal surfaces of eIF4E.

The structures presented here indicate that the target sites in the linker region (S65, T70) are not part of the interface with eIF4E and hence are not directly involved in eIF4E binding. In agreement with the solvent exposure of the respective side chains, we show that phosphorylation of these residues (or phosphomimetic mutations) does not significantly affect Thor binding to eIF4E in vitro (Figure S5 and Table S2). Similar observations were reported in previous studies for human 4E-BP-1,2 (Bah et al., 2014; Gingras et al., 2001; Karim et al., 2001). However, we show that phosphorylation of linker residues does impair the ability of Thor to displace eIF4G from preassembled eIF4E-eIF4G complexes (Figures 5A–5D). More important, phosphomimetic mutations at the phosphorylation sites in the linker region prevent Thor from repressing translation in vivo (Figures 5E and 5F). On the basis of these observations and of the proximity of the phosphorylated residues to the elbow loop, it is likely that the primary effect of phosphorylation is interference with the competition mechanism. This interference could occur by a conformational stabilization of the peptide in the free state and/or a destabilization of bound intermediates that decrease the competitive advantage. Consistent with this model, nuclear magnetic resonance studies using short *Hs* 4E-BP1 peptides containing the canonical motif and linker region but lacking the non-canonical motifs (residues 51–67) demonstrated a reduction of helix propensity upon phosphorylation (Tait et al., 2010).

Nevertheless, the mechanism by which phosphorylation inhibits different 4E-BPs from binding to eIF4E remains to be fully elucidated. In the case of 4E-BP2, phosphorylation of residues N-terminal to the canonical motif induces the formation of a four-stranded β domain that sequesters the canonical motif, reducing the affinity for eIF4E by two orders of magnitude (Bah et al., 2014). This event primes the phosphorylation of residues in the linker region, which further reduces affinity by

10-fold, suggesting that there may be additional order-to-disorder transitions upon phosphorylation that also involve the linker residues. In the case of *Dm* Thor, phosphorylation at residues upstream of the canonical motif is sufficient to prevent binding to eIF4E (Miron et al., 2003), suggesting that the impact of phosphorylation of specific residues is likely to differ for the different 4E-BPs.

4E-BP Peptide Mimics as Therapeutic Tools

The dysregulation of cap-dependent translation is causally linked to human diseases including cancer, ASDs, and fragile X syndrome (Banko et al., 2005, 2007; Boussemart et al., 2014; Gkogkas et al., 2013; Sharma et al., 2010). In cancer cells, increased eIF4E activity facilitates the translation of mRNAs coding for growth factors and oncogenic proteins (Dowling et al., 2010). In ASD and fragile X syndrome, increased synthesis of synaptic proteins results in augmented connectivity (Gkogkas et al., 2013; Sharma et al., 2010). In this context, the development of eIF4E inhibitors for therapeutic applications has been the object of intensive research (Jia et al., 2012). The validity of this approach is underscored by the observation that 4EGI-1 inhibits tumor growth in human breast and melanoma cancer xenografts and reverses disease-associated phenotypes in mouse models of ASD (Gkogkas et al., 2013; Moerke et al., 2007).

However, the affinity of 4EGI-1 for eIF4E is approximately three orders of magnitude lower than the affinity of 4E-BPs (Moerke et al., 2007). We therefore speculate that bipartite 4E-BP mimics are likely to represent more potent inhibitors of translation, first, because of their higher affinity for eIF4E, and second, because 4E-BPs sterically block the eIF4G-binding surface, whereas 4EGI-1 interferes with eIF4G binding indirectly using an allosteric mechanism (Papadopoulos et al., 2014).

As a proof of principle, we designed a peptide-based 4E-BP mimic by combining the canonical motif of 4E-T with the non-canonical motif of CUP (i.e., the motifs with the highest respective affinity for eIF4E). These motifs were connected by the linker region of Thor, which establishes the most extensive contacts with eIF4E among all linkers. The triple chimeric peptide showed an increased affinity compared to the equivalent peptides of Thor and CUP and faster eIF4G displacement kinetics (Figure 6). More remarkably, the chimeric peptide inhibited translation in S2 cells at a concentration at which Thor was ineffective (Figures 6I and 6J). The peptide also efficiently disrupted endogenous eIF4E-eIF4G complexes in human cells. These observations illustrate the feasibility of designing high-affinity eIF4E binding peptides as tools for therapeutic applications.

EXPERIMENTAL PROCEDURES

DNA Constructs

The plasmids used for the expression of *Dm* proteins (full length or fragments) in *Escherichia coli* or S2 cells have been described previously (Igreja et al., 2014). The plasmids for the expression of V5-SBP or GFP-tagged 4E-BP1 and 4E-BP3 in human cells were obtained by inserting the corresponding DNA between the XhoI and BamHI sites of the pT7-V5-SBP and pT7-EGFP-C1 vectors, respectively. *Hs* eIF4E cDNA sequence was introduced in the EcoRI and BamHI restriction sites of the pλN-HA-C1 vector. The plasmids for expression in *E. coli* are described in detail in the Supplemental Information. A synthetic DNA fragment coding for the triple chimera peptide was cloned between the EcoRV and XhoI restriction sites of the pAc5.1B-EGFP vector and the pT7-EGFP-C1 vector for expression in *Dm* and *Hs* cells, respectively. The triple chimera includes the canonical motif of 4E-T (residues 9–20), the linker region of Thor (residues 65–72), the non-canonical motif of CUP (residues 362–376), and an additional Gly residue between the linker and the non-canonical portion of the peptide. All the mutants used in this study were generated by site-directed mutagenesis using the QuikChange mutagenesis kit (Stratagene). All the constructs and mutations were confirmed by sequencing and are listed in Table S1.

Protein Expression and Purification

All recombinant proteins were expressed in *E. coli* BL21 Star (DE3) cells (Invitrogen) grown in Luria broth medium overnight at 20°C. The cells were lysed by sonication in lysis buffer containing 50 mM HEPES (pH 7.2), 300 mM NaCl, and 2 mM dithiothreitol (DTT) supplemented with DNaseI (5 μg/ml), lysozyme (1 mg/ml), and protease inhibitor cocktail (Roche). To purify the complexes containing *Dm* eIF4E, His_6-tagged *Dm* eIF4E (residues 69–248) and the MBP-tagged 4E-BP or eIF4G, were expressed separately. The complexes were formed after cell lysis by mixing the corresponding lysates. To obtain complexes containing *Hs* eIF4E, His_6-tagged *Hs* eIF4E (residues 36–217) was co-expressed with MBP-tagged *Hs* 4E-BP1 (residues 50–83). All complexes were purified from cleared cell lysates using amylose resin (New England Biolabs), followed by cleavage of MBP- and His_6-tags with HRV3C protease overnight at 4°C and additional cleavage with TEV protease for the complex with *Dm* eIF4G. After cleavage of the tags, the complexes were separated from MBP and His_6 using a heparin column (HiTrap Heparin HP 5 ml; GE Healthcare), further purified on a Superdex 75 column (GE Healthcare) and stored in 10 mM HEPES (pH 7.2), 200 mM NaCl, 2 mM DTT, and 5% glycerol at −80°C.

For the ITC measurements and competition assays, the GB1-stabilized peptides were purified as described previously (Igreja et al., 2014). The complex consisting of His_6-tagged *Dm* eIF4E (residues 69–248) and MBP-tagged *Dm* eIF4G (residues 578–650) used for the competition assays was purified from cleared cell lysates using amylose resin (New England Biolabs), followed by a heparin column (HiTrap Heparin HP 5 ml; GE Healthcare) and a final purification on a Superdex 200 column (GE Healthcare).

Crystallization, Data Collection, and Structure Determination

A detailed description of the crystallization conditions and the structure determination process can be found in Supplemental Experimental Procedures. All diffraction data sets were recorded on a PILATUS 6M detector at the PXII beamline of the Swiss Light Source at a temperature of 100 K. The diffraction data and refinement statistics are summarized in Table 1.

Coimmunoprecipitation Assays and Western Blotting

Coimmunoprecipitation assays in S2 cells and western blotting were performed as described previously (Igreja et al., 2014). A detailed protocol for the coimmunoprecipitation assays in human HEK293T cells can be found in the Supplemental Information. Immunoprecipitated proteins were analyzed by western blotting. All western blots were developed using the ECL western blotting detection system (GE Healthcare). The antibodies used in this study are listed in Table S3.

Pull-Downs, ITC Analysis, Competition and Translation Repression Assays

The in vitro pull-down and competition assays, ITC analysis, and translation repression assays were performed as described previously (Igreja et al., 2014) and are described in the Supplemental Information.

ACCESSION NUMBERS

Coordinates for the structures have been deposited in the Protein Data Bank (PDB) under accession codes 4UE8 (*Dm* eIF4E-Thor), 4UE9 (*Dm* eIF4E-4E-T), 4UEA (*Dm* eIF4E-chimeric peptide, crystal form 1), 4UEB (*Dm* eIF4E-chimeric peptide, crystal form 2), 4UEC (*Dm* eIF4E-eIF4G), and 4UED (*Hs* eIF4E-4E-BP1).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and three tables and can be found with this article online at http://dx.doi.org/10.1016/j.molcel.2015.01.017.

AUTHOR CONTRIBUTIONS

D.P. purified, crystallized, and solved the structures of all eIF4E complexes presented in this study and performed ITC measurements. D.P. and O.W. collected and analyzed diffraction data. D.P., L.E., R.W., and C.W. performed pull-downs and competition assays. L.W. performed binding assays in human cells. C.I. performed coimmunoprecipitation assays and translation inhibition assays in S2 cells. L.E. and L.W. generated many of the DNA constructs used in this study. E.I. was the principal investigator who coordinated the project. C.I. and E.I. supervised the project. D.P., C.I., E.I., and O.W. wrote the manuscript. All authors corrected the manuscript.

ACKNOWLEDGMENTS

We are grateful to Fulvia Bono for providing the His-eIF4E and GST-CUP 311–440 constructs and to M.-Y. Chung for excellent technical assistance. We thank R. Büttner and T. Raisch for the setup of crystallization screens and the staff at the PX beamline of the Swiss Light Source for assistance with data collection. This work was supported by the Max Planck Society and the Gottfried Wilhelm Leibniz Program of the Deutsche Forschungsgemeinschaft (awarded to E.I).

Received: October 28, 2014
Revised: December 22, 2014
Accepted: January 7, 2015
Published: February 19, 2015

REFERENCES

Bah, A., Vernon, R.M., Siddiqui, Z., Krzeminski, M., Muhandiram, R., Zhao, C., Sonenberg, N., Kay, L.E., and Forman-Kay, J.D. (2014). Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. Nature. Published online December 24, 2014. http://dx.doi.org/10.1038/nature13999.

Banko, J.L., Poulin, F., Hou, L., DeMaria, C.T., Sonenberg, N., and Klann, E. (2005). The translation repressor 4E-BP2 is critical for eIF4F complex formation, synaptic plasticity, and memory in the hippocampus. J. Neurosci. *25*, 9581–9590.

Banko, J.L., Merhav, M., Stern, E., Sonenberg, N., Rosenblum, K., and Klann, E. (2007). Behavioral alterations in mice lacking the translation repressor 4E-BP2. Neurobiol. Learn. Mem. *87*, 248–256.

Bidinosti, M., Ran, I., Sanchez-Carbente, M.R., Martineau, Y., Gingras, A.C., Gkogkas, C., Raught, B., Bramham, C.R., Sossin, W.S., Costa-Mattioli, M., et al. (2010). Postnatal deamidation of 4E-BP2 in brain enhances its association with raptor and alters kinetics of excitatory synaptic transmission. Mol. Cell *37*, 797–808.

Boussemart, L., Malka-Mahieu, H., Girault, I., Allard, D., Hemmingsson, O., Tomasic, G., Thomas, M., Basmadjian, C., Ribeiro, N., Thuaud, F., et al. (2014). eIF4F is a nexus of resistance to anti-BRAF and anti-MEK cancer therapies. Nature *513*, 105–109.

Dowling, R.J., Topisirovic, I., Alain, T., Bidinosti, M., Fonseca, B.D., Petroulakis, E., Wang, X., Larsson, O., Selvaraj, A., Liu, Y., et al. (2010). mTORC1-mediated cell proliferation, but not cell growth, controlled by the 4E-BPs. Science *328*, 1172–1176.

Gingras, A.C., Gygi, S.P., Raught, B., Polakiewicz, R.D., Abraham, R.T., Hoekstra, M.F., Aebersold, R., and Sonenberg, N. (1999). 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-Cieszynska, A., Aebersold, R., and Sonenberg, N. (2001). Hierarchical phosphorylation of the translation inhibitor 4E-BP1. Genes Dev. *15*, 2852–2864.

Gkogkas, C.G., Khoutorsky, A., Ran, I., Rampakakis, E., Nevarko, T., Weatherill, D.B., Vasuta, C., Yee, S., Truitt, M., Dallaire, P., et al. (2013). Autism-related deficits via dysregulated eIF4E-dependent translational control. Nature *493*, 371–377.

Gosselin, P., Oulhen, N., Jam, M., Ronzca, J., Cormier, P., Czjzek, M., and Cosson, B. (2011). The translational repressor 4E-BP called to order by eIF4E: new structural insights by SAXS. Nucleic Acids Res. *39*, 3496–3503.

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

Hsieh, A.C., Liu, Y., Edlind, M.P., Ingolia, N.T., Janes, M.R., Sher, A., Shi, E.Y., Stumpf, C.R., Christensen, C., Bonham, M.J., et al. (2012). The translational landscape of mTOR signalling steers cancer initiation and metastasis. Nature *485*, 55–61.

Igreja, C., Peter, D., Weiler, C., and Izaurralde, E. (2014). 4E-BPs require non-canonical 4E-binding motifs and a lateral surface of eIF4E to repress translation. Nat. Commun. *5*, 4790.

Jackson, R.J., Hellen, C.U., and Pestova, T.V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. Nat. Rev. Mol. Cell Biol. *11*, 113–127.

Jia, Y., Polunovsky, V., Bitterman, P.B., and Wagner, C.R. (2012). Cap-dependent translation initiation factor eIF4E: an emerging anticancer drug target. Med. Res. Rev. *32*, 786–814.

Karim, M.M., Hughes, J.M., Warwicker, J., Scheper, G.C., Proud, C.G., and McCarthy, J.E. (2001). A quantitative molecular model for modulation of mammalian translation by the eIF4E-binding protein 1. J. Biol. Chem. *276*, 20750–20757.

Kinkelin, K., Veith, K., Grünwald, M., and Bono, F. (2012). Crystal structure of a minimal eIF4E-Cup complex reveals a general mechanism of eIF4E regulation in translational repression. RNA *18*, 1624–1634.

Lukhele, S., Bah, A., Lin, H., Sonenberg, N., and Forman-Kay, J.D. (2013). Interaction of the eukaryotic initiation factor 4E with 4E-BP2 at a dynamic bipartite interface. Structure *21*, 2186–2196.

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.

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.

Martineau, Y., Azar, R., Bousquet, C., and Pyronnet, S. (2013). Anti-oncogenic potential of the eIF4E-binding proteins. Oncogene *32*, 671–677.

Matsuo, H., Li, H., McGuire, A.M., Fletcher, C.M., Gingras, A.C., Sonenberg, N., and Wagner, G. (1997). Structure of translation factor eIF4E bound to m^7GDP and interaction with 4E-binding protein. Nat. Struct. Biol. *4*, 717–724.

Miron, M., Lasko, P., and Sonenberg, N. (2003). Signaling from Akt to FRAP/TOR targets both 4E-BP and S6K in Drosophila melanogaster. Mol. Cell. Biol. *23*, 9117–9126.

Mizuno, A., In, Y., Fujita, Y., Abiko, F., Miyagawa, H., Kitamura, K., Tomoo, K., and Ishida, T. (2008). Importance of C-terminal flexible region of 4E-binding protein in binding with eukaryotic initiation factor 4E. FEBS Lett. *582*, 3439–3444.

Moerke, N.J., Aktas, H., Chen, H., Cantel, S., Reibarkh, M.Y., Fahmy, A., Gross, J.D., Degterev, A., Yuan, J., Chorev, M., et al. (2007). Small-molecule inhibition of the interaction between the translation initiation factors eIF4E and eIF4G. Cell *128*, 257–267.

Paku, K.S., Umenaga, Y., Usui, T., Fukuyo, A., Mizuno, A., In, Y., Ishida, T., and Tomoo, K. (2012). A conserved motif within the flexible C-terminus of the translational regulator 4E-BP is required for tight binding to the mRNA cap-binding protein eIF4E. Biochem. J. *441*, 237–245.

Papadopoulos, E., Jenni, S., Kabha, E., Takrouri, K.J., Yi, T., Salvi, N., Luna, R.E., Gavathiotis, E., Mahalingam, P., Arthanari, H., et al. (2014). Structure of the eukaryotic translation initiation factor eIF4E in complex with 4EGI-1 reveals an allosteric mechanism for dissociating eIF4G. Proc. Natl. Acad. Sci. U S A *111*, E3187–E3195.

Sharma, A., Hoeffer, C.A., Takayasu, Y., Miyawaki, T., McBride, S.M., Klann, E., and Zukin, R.S. (2010). Dysregulation of mTOR signaling in fragile X syndrome. J. Neurosci. *30*, 694–702.

Siddiqui, N., Tempel, W., Nedyalkova, L., Volpon, L., Wernimont, A.K., Osborne, M.J., Park, H.W., and Borden, K.L. (2012). Structural insights into the allosteric effects of 4EBP1 on the eukaryotic translation initiation factor eIF4E. J. Mol. Biol. *415*, 781–792.

Tait, S., Dutta, K., Cowburn, D., Warwicker, J., Doig, A.J., and McCarthy, J.E. (2010). Local control of a disorder-order transition in 4E-BP1 underpins regulation of translation via eIF4E. Proc. Natl. Acad. Sci. U S A *107*, 17627–17632.

Umenaga, Y., Paku, K.S., In, Y., Ishida, T., and Tomoo, K. (2011). Identification and function of the second eIF4E-binding region in N-terminal domain of eIF4G: comparison with eIF4E-binding protein. Biochem. Biophys. Res. Commun. *414*, 462–467.

Volpon, L., Osborne, M.J., Topisirovic, I., Siddiqui, N., and Borden, K.L. (2006). Cap-free structure of eIF4E suggests a basis for conformational regulation by its ligands. EMBO J. *25*, 5138–5149.

Molecular Cell

Supplemental Information

Molecular Architecture of 4E-BP Translational Inhibitors Bound to eIF4E

Daniel Peter, Cátia Igreja, Ramona Weber, Lara Wohlbold, Catrin Weiler, Linda Ebertsch, Oliver Weichenrieder, and Elisa Izaurralde

SUPPLEMENTAL FIGURES

Figure S1

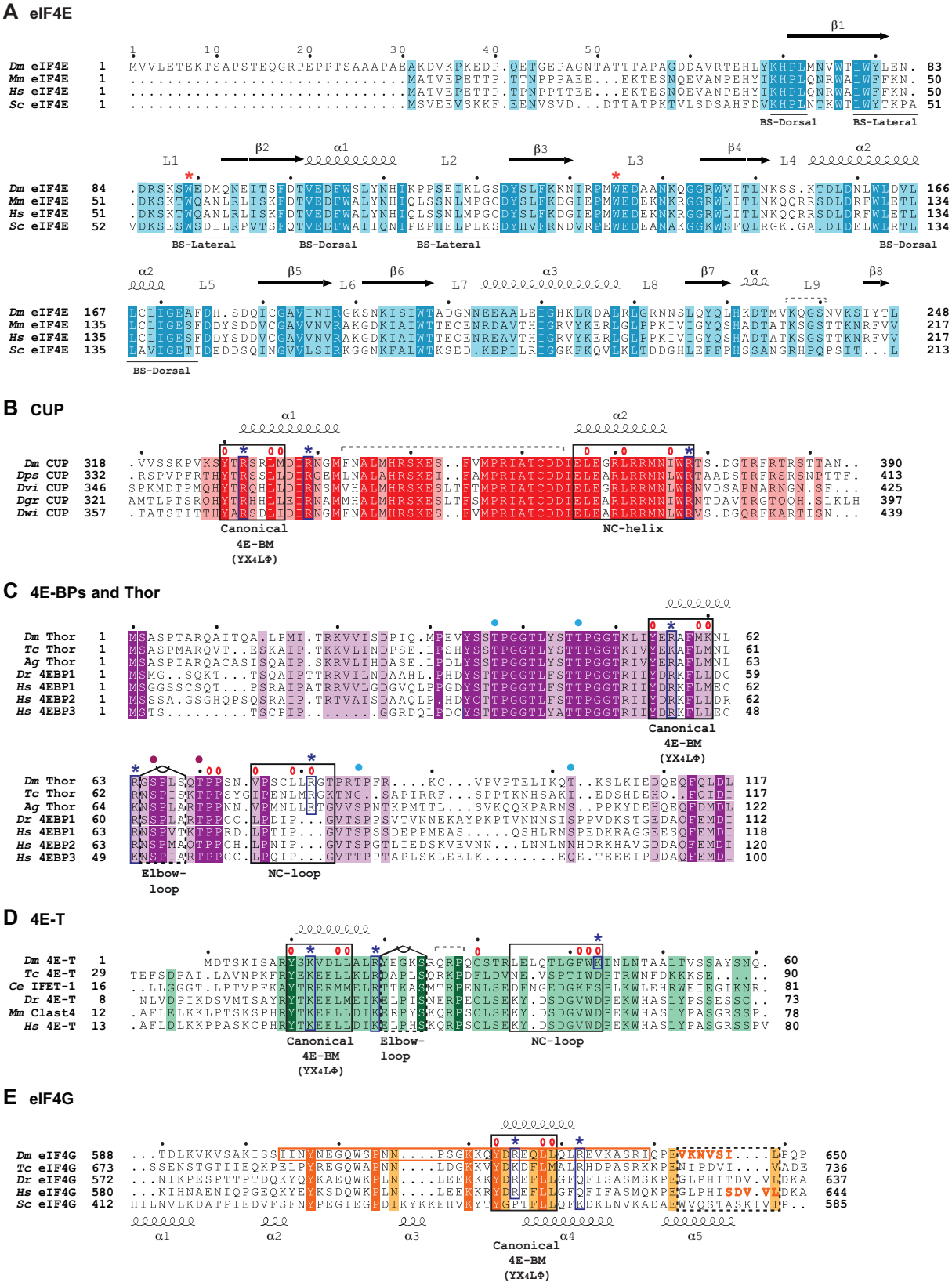


Figure S1, related to Figure 1. Sequence alignments

In all aligned sequences, conserved residues are highlighted with a dark-colored background and printed in white. Residues with >70% similarity are shown with a light color background. Secondary structure elements are indicated above the sequences for the *Dm* proteins and below the sequence for *Saccharomyces cerevisiae* eIF4G. The structurally conserved Arg/Lys residues in the canonical motifs of 4E-BPs and eIF4G and in the non-canonical motifs of *Dm* 4E-BPs are indicated by blue asterisks and are boxed in blue. In the alignments of 4E-BP proteins, the canonical (C) and non-canonical (NC) motifs are boxed in black. 4E-BP and eIF4G residues within the motifs contacting eIF4E are indicated by red open circles; phosphorylated residues in Thor and 4E-BPs are indicated by circles colored in magenta for the sites visible in our structures or in cyan for additional known sites.

(**A**) Structure-based sequence alignment of eIF4E orthologous proteins from *Drosophila melanogaster* (*Dm*), *Mus musculus* (*Mm*), *Homo sapiens* (*Hs*) and *Saccharomyces cerevisiae* (*Sc*). The Trp residues coordinating the m^7GTP nucleotide are indicated by red asterisks. The lateral and dorsal binding surfaces (BS) are indicated by a line below the sequences.

(**B**) Sequence alignment of the eIF4E-interacting regions of CUP orthologous proteins from insects. The species are as follows: *Drosophila melanogaster* (*Dm*), *Drosophila pseudoobscura* (*Dps*), *Drosophila virilis* (*Dvi*), *Drosophila grimshawi* (*Dgr*) and *Drosophila willistoni* (*Dwi*). Secondary structural elements of the *Dm* CUP are according to the structure of CUP in complex with eIF4E (PDB ID: 4AXG; Kinkelin et al., 2012). Residues that were not visible in the CUP structure are highlighted by a gray dashed bracket above the alignment.

(**C**) Sequence alignment of Thor orthologous proteins. The species are as follows: *Drosophila melanogaster* (*Dm*), *Tribolium castaneum* (*Tc*), *Anopheles gambiae* (*Ag*), *Danio rerio* (*Dr*) and *Homo sapiens* (*Hs*). The residues forming the elbow loop are indicated by a triangle above the alignment.

(**D**) Alignment of the eIF4E-interacting region of 4E-T orthologous proteins. The species are *Caenorhabditis elegans* (*Ce*) and *Mus musculus* (*Mm*); other species are as in (C).

(**E**) Alignment of the eIF4E-interacting region of eIF4G orthologous proteins. The species are as in panels (A) and (D). The secondary structural elements for *Sc* eIF4G are based on the NMR structure (PDB ID: 1RF8; Gross et al., 2003) and are indicated below the sequences. The residues visible in our structure are boxed in orange. The auxiliary VKNVSI motif (*Hs* SDVVL) is printed in bold orange characters.

Figure S2

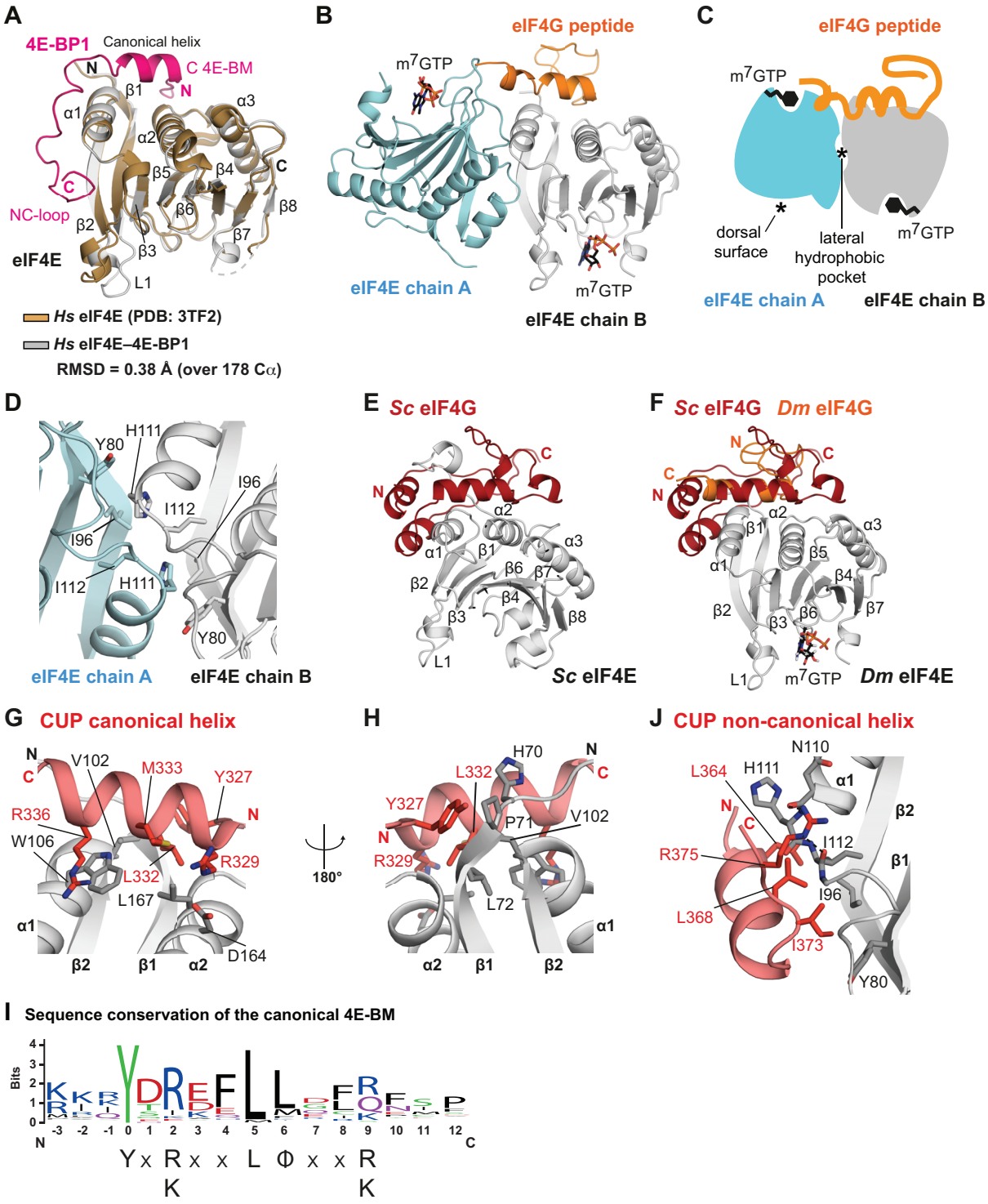


Figure S2, related to Figures 1 and 3. Structures of 4E-BPs and eIF4G bound to eIF4E

(**A**) Superposition of the structure of free human eIF4E (brown, PDB ID: 3TF2; Siddiqui et al., 2012) with the structure of *Hs* eIF4E (gray) bound to 4E-BP1 (magenta).

(**B**) Cartoon representation showing the crystal packing in the structure of the *Dm* eIF4E-eIF4G complex bound to m^{7}GTP.

(**C**) Schematic representation of eIF4E dimerization in the structure of the *Dm* eIF4E-eIF4G complex bound to m^7GTP.

(**D**) Close-up view of the interaction between two eIF4E molecules in the asymmetric unit of the *Dm* eIF4E-eIF4G complex bound to m^7GTP.

(**E**) Overview of the structure of the *Saccharomyces cerevisiae* eIF4E-eIF4G complex (PDB ID: 1RF8; Gross et al., 2003). The solution structure of the *Saccharomyces cerevisiae* eIF4G-eIF4E complex indicates that eIF4G residues C-terminal to the canonical motif wrap around the unstructured N-terminal extension of eIF4E.

(**F**) Superposition of the canonical motif of *Sc* eIF4G onto the canonical motif of *Dm* eIF4G in the *Dm* eIF4E–eIF4G complex. *Sc* eIF4E has been ommitted for clarity.

(**G** and **H**) Close-up views of the interaction between the dorsal surface of eIF4E and the canonical motif of CUP (PDB ID: 4AXG; Kinkelin et al., 2012). Selected interface residues are shown as gray sticks for eIF4E and as colored sticks for CUP.

(**I**) Sequence Logo of the $YX_4L\Phi$ canonical 4E-BM. The alignment underlying the motif includes sequences of different characterized 4E-BPs (4E-BP1–3/Thor, 4E-T, CUP, Angel, Mextli, p20, Eap1 and Maskin) as well as eIF4G from different species. The motif was created using WebLogo (http://weblogo.berkeley.edu/logo.cgi; Crooks et al., 2004).

(**J**) Interaction of the non-canonical α-helix of CUP with the lateral hydrophobic pocket of eIF4E (PDB ID: 4AXG).

Figure S3

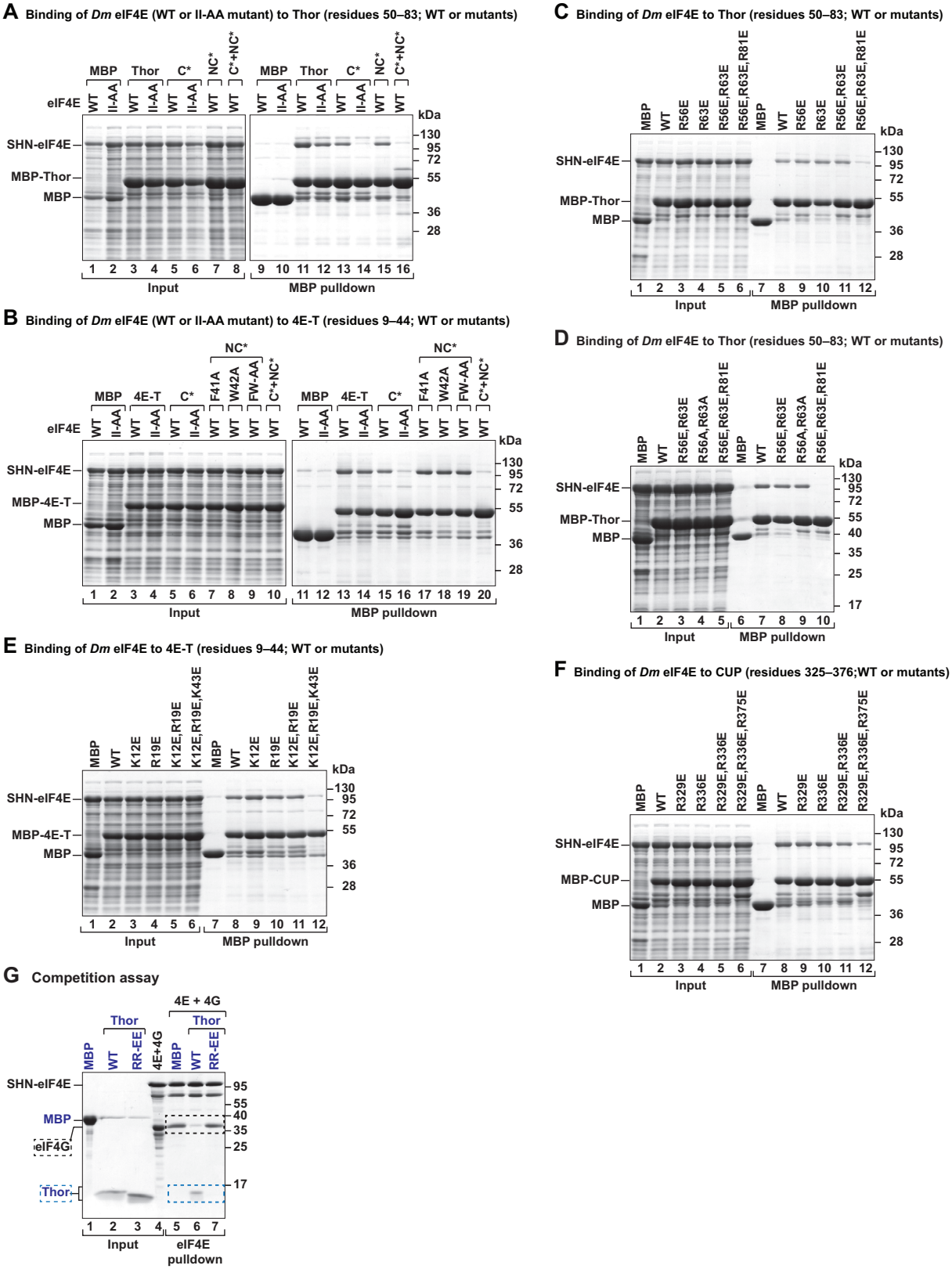


Figure S3, related to Figure 3 and Table S1. Interaction of *Dm* 4E-BPs and eIF4E *in vitro*

(**A** to **F**) MBP pulldowns showing the interaction of MBP-tagged 4E-BPs (wild-type or the indicated mutants) with full-length eIF4E (wild-type or II-AA mutant). eIF4E was expressed with a SHN tag consisting of the streptavidin-binding peptide (strep), His_6 and the NusA protein. The eIF4E II-AA mutant carries alanine substitutions in residues that line the lateral hydrophobic pocket (I96A and I112). The MBP-tagged 4E-BP fragments used in the pulldown assays contained Thor residues 50–83, 4E-T residues 9–44 and CUP residues 325–376 and were all C-terminally fused to GB1. Protein mutants are described in Table S1. The C and NC mutants are as follows; Thor C* (Y54A, M59A), Thor NC* (V75D, L79D), 4E-T C* (Y10A, L15A) and 4E-T NC* (F41D, W42D).

(**G**) Purified complexes containing SHN-eIF4E bound to GST-eIF4G were incubated with 10 fold molar excess of Thor fragments (residues 50–83; either wild-type or the R56E, R63E mutant) C-terminally fused to GB1. The eIF4E-bound proteins were pulled down using Ni-NTA beads and analyzed by SDS-PAGE. The competitor proteins are labeled in blue and highlighted by a blue dashed box. The black dashed box marks the position of GST-eIF4G.

Figure S4

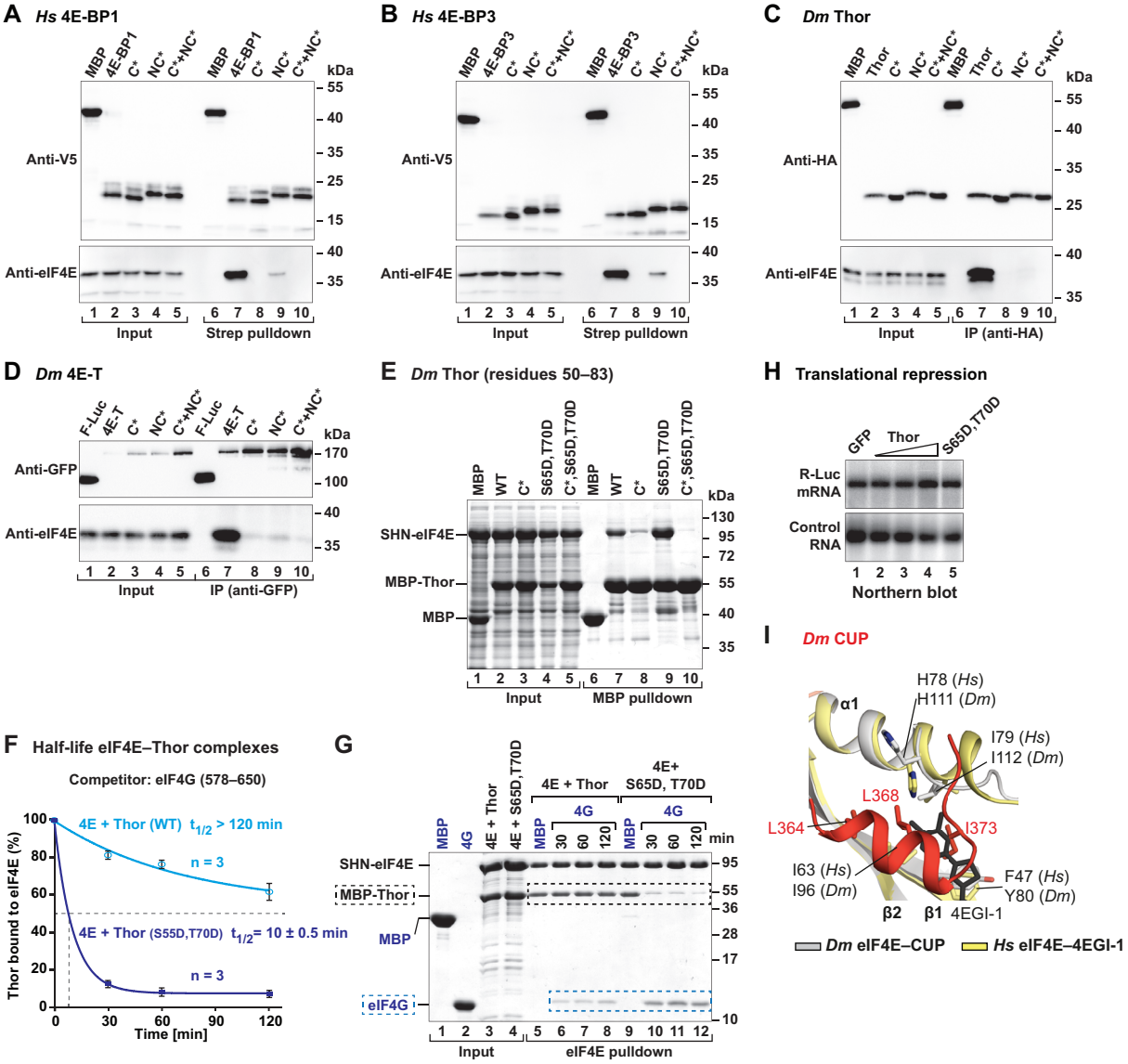


Figure S4, related to Figure 3, 5, Table S1 and Table S3. Interaction of 4E-BPs with eIF4E in *Dm* and human cells

(**A** and **B**) Western blot showing the interaction of 4E-BP1 and 4E-BP3 (full-length, either wild-type or mutated) with endogenous eIF4E in human HEK293T cells. The full-length 4E-BP1 and 4E-BP3 proteins were expressed with an N-terminal tag containing the streptavidin peptide followed by a V5 tag. The proteins were pulled down using streptavidin (Strep) beads. The inputs (0.75% for the V5 tagged proteins and 1.6% for eIF4E) and immunoprecipitates (1.5% for the V5 tagged proteins and 3% for eIF4E) were analyzed by western blotting using anti-V5 and anti-eIF4E antibodies. Alanine substitutions of the Tyr and Leu residues in the

canonical motifs (C*) or of Leu and Val residues in the non-canonical motifs (NC*) abolish or reduce binding of 4E-BPs to eIF4E in human cells (in the presence of endogenous eIF4G).

(**C**) Western blot showing the interaction of HA-Thor (full-length, either wild-type or mutated) with endogenous eIF4E in S2 cells. The proteins were immunoprecipitated using anti-HA antibodies. The inputs (1.5% for Thor and 0.3% for eIF4E) and immunoprecipitates (15% for Thor and 30% for eIF4E) were analyzed by western blotting using anti-HA and anti-eIF4E antibodies. Substitutions of residues in the canonical (C*: Y54A, M59A) or the non-canonical (NC*: V75D, L79D) motif abolish Thor interaction with eIF4E in S2 cells (in the presence of eIF4G). Thus the mutations prevent Thor from displacing eIF4G from preassembled endogenous eIF4E-eIF4G complexes.

(**D**) Western blot showing the interaction of GFP-tagged 4E-T (full-length, either wild-type or mutated) with endogenous eIF4E in S2 cells. The proteins were immunoprecipitated using anti-GFP antibodies. The inputs (1.5% for 4E-T and 0.3% for eIF4E) and immunoprecipitates (20%) were analyzed by western blotting using anti-GFP and anti-eIF4E antibodies. Substitutions in the canonical (C*: Y10A, L15A) or non-canonical (NC*: F41A, W42A) motif abolish binding of 4E-T to eIF4E in S2 cells (in the presence of eIF4G). Size markers (kDa) are shown to the right of each panel. Protein mutants are described in Table S1.

(**E**) MBP pulldown assay showing the interaction of MBP-Thor (residues 32–83 wild-type or the indicated mutants) and full-length SHN-eIF4E. Thor fragments were all C-terminally fused to GB1.

(**F** and **G**) Half-life of eIF4E–Thor complexes. Purified eIF4E–Thor complexes (4 μM) containing SNH-tagged eIF4E (residues 69–248) and MBP-tagged Thor (full length, either wild-type or the S65D, T70D phosphomimetic mutant) immobilized on Ni-NTA beads were incubated with a 5-fold molar excess of eIF4G (residues 578–650) or MBP as a negative control. The eIF4G peptide was fused C-terminally to GB1. Proteins bound to eIF4E were recovered at the indicated time points, analyzed by SDS-PAGE (G) and quantify in three

independent experiments (F). Error bars in panel (F) represent standard deviations. In panel (G), the competitor eIF4G protein is labeled in blue and highlighted by a blue dashed box. The black dashed box marks the position of Thor.

(**H**) Northern blot (NB) analyses of representative RNA samples corresponding to the experiment shown in Figure 5E,F.

(**I**) Superposition of the structure of eIF4E (yellow) bound to the 4EGI-1 inhibitor (black; PDB ID: 4TPW) with the structure of the *Dm* eIF4E (gray) bound to CUP (red; PDB ID: 4AXG).

Figure S5

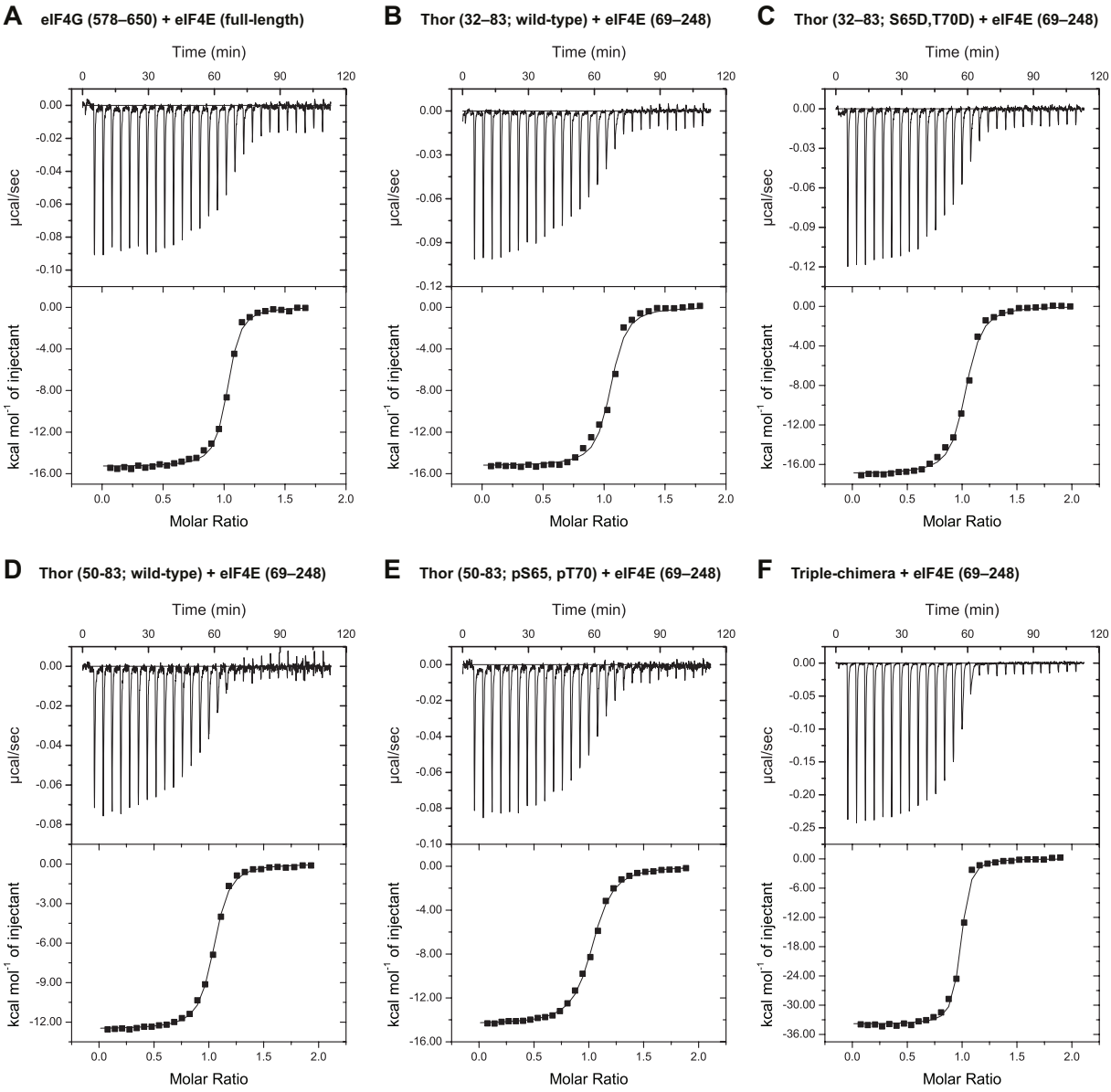


Figure S5, related to Figures 5, 6 and Table S2. Calorimetric titration data for the interaction of eIF4E with eIF4G and 4E-BP-derived peptides

(**A** to **F**) ITC profiles for the interaction of *Dm* eIF4E (full-length or residues 69–248) with the following peptides: (A) eIF4G (residues 578–650); (B and C) Thor (residues 32–83) either wild-type (B) or the S65D, T70D phosphomimetic mutant (C); (D and E) Thor synthetic peptide (residues 50–83) either wild-type (D) or phosphorylated at S65D, T70D (E); and (F) triple chimera peptide. The thermodynamic parameters are shown in Table S2.

Figure S6

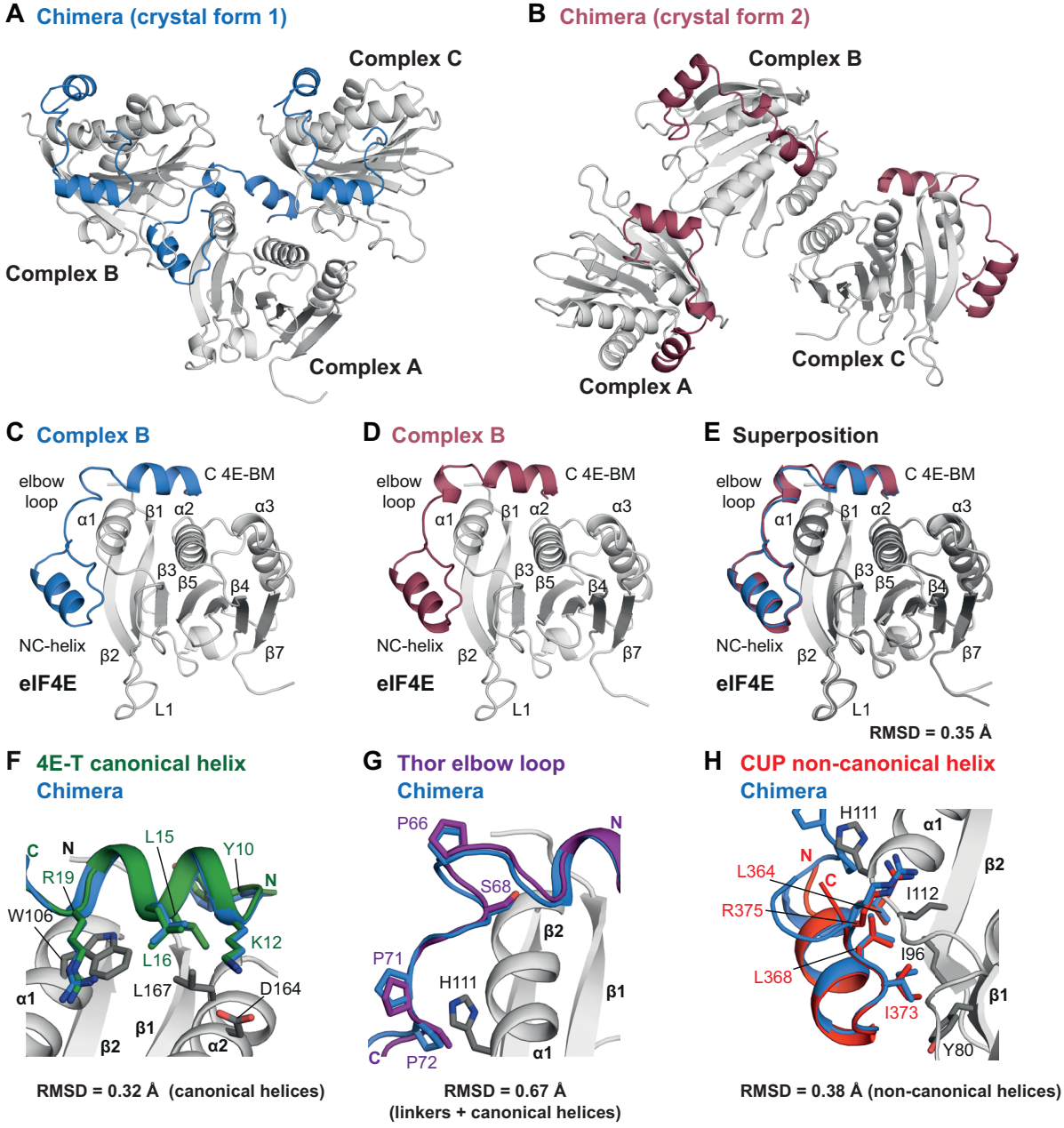


Figure S6, related to Figure 6. Structure of the triple chimera peptide in complex with eIF4E

(**A** and **B**) Cartoon representation showing the crystal packing in the structures of two crystal forms of *Dm* eIF4E bound to the chimeric peptide.

(**C–E**) Overview of the structures of *Dm* eIF4E bound to the chimeric peptide and superposition of the structures derived from the two crystal forms. Selected secondary structure elements are labeled in black for eIF4E and in color for the peptide. The structures of the two crystal forms superpose with an RMSD of 0.35 Å over 205 Cα.

(**F–H**) Superposition of the structure of eIF4E bound to the triple chimera peptide with the structures of eIF4E bound to 4E-T (F), Thor (G) and CUP (H; PDB ID: 4AXG).

Figure S7

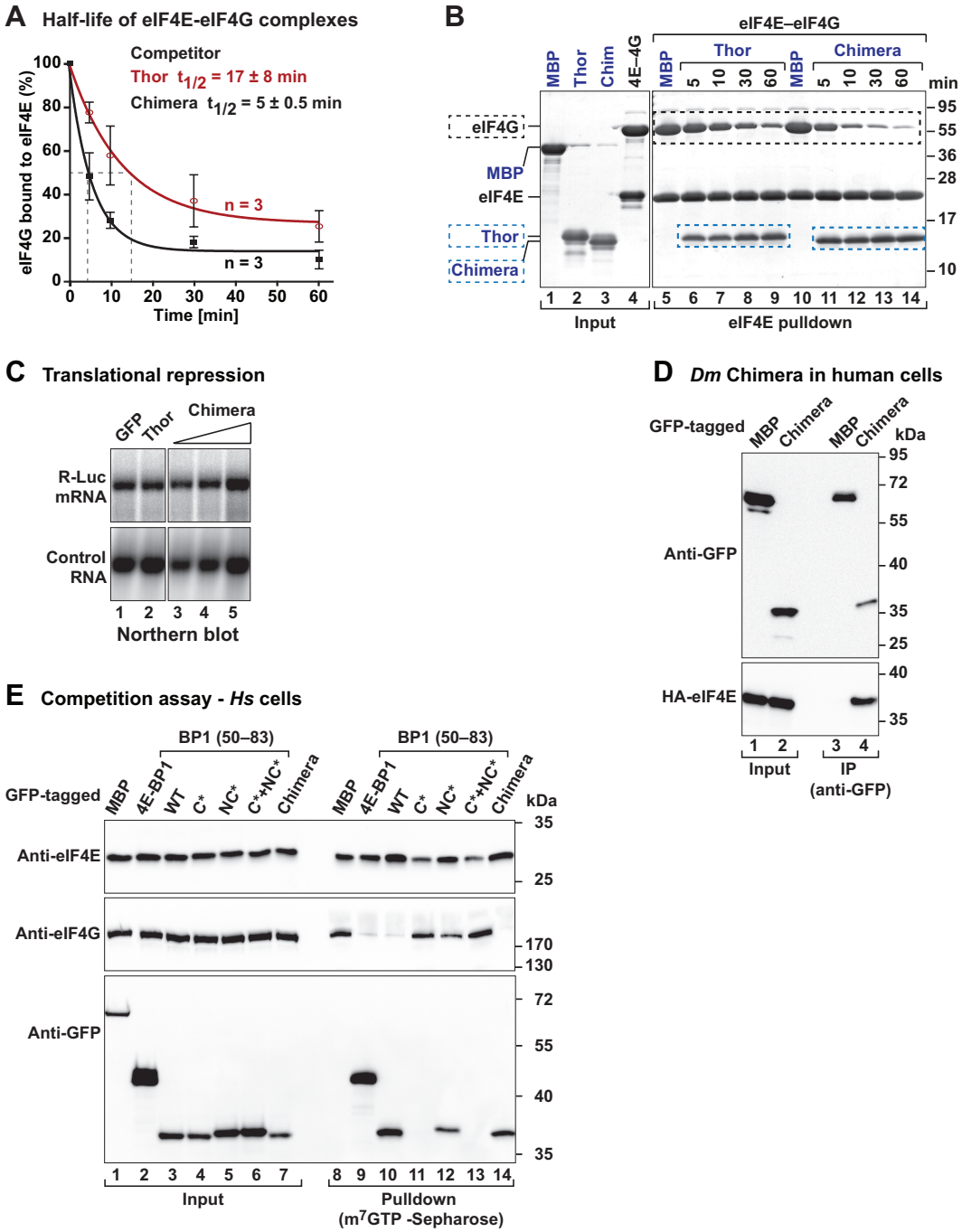


Figure S7, related to Figure 6 and Table S3. Competition and translation repression assays using the triple chimera

(**A** and **B**) Purified eIF4E–eIF4G complexes (2 μM) containing His-tagged eIF4E (residues 69–248) and MBP-tagged eIF4G (residues 578–650) immobilized on Ni-NTA beads were incubated with a 1.5-fold molar excess of Thor peptide (residues 32–83) or the triple chimera peptide. The peptides were fused C-terminally to GB1. Proteins bound to eIF4E were recovered at the indicated time points, analyzed by SDS-PAGE (B) and quantified. The competitor proteins are labeled in blue and highlighted by a blue, dashed box. The black, dashed box marks the position of MBP-eIF4G. The half-life of eIF4G-eIF4E complexes is shown in panel (A). Error bars represent standard deviations from 3 independent experiments (n=3).

(**C**) Northern blot analyses of representative RNA samples corresponding to the experiment shown in Figure 6I,J.

(**D**) Western blot showing the interaction of GFP-tagged chimeric peptide with endogenous HA-eIF4E in human HEK293T cells. The proteins were coimmunoprecipitated using anti-GFP antibodies. GFP-MBP served as negative control. The inputs (2.5% for the GFP-tagged proteins and 1% for eIF4E) and immunoprecipitates (10% for the GFP-tagged proteins and 25% for eIF4E) were analyzed by western blotting using anti-GFP and anti-HA antibodies.

(**E**) Human HEK293T cells were transfected with GFP-tagged proteins including MBP, a 4E-BP1 fragment (residues 50–83; either wild-type or the indicated C* and NC* mutants) and the chimeric peptide. Proteins bound to endogenous eIF4E were pulled down using m^7GTP-Sepharose and analyzed by western blotting using anti-eIF4E, anti-eIF4G and anti-GFP antibodies. See also Tables S1 and S3.

Table S1 related to Figure 1. Mutants and constructs used in this study

Protein	Name of the construct	Fragments / mutations	Binding site / motif
***Dm* eIF4E** (1–248) (isoform C) P48598-2	4E	Full-length	
	II-AA	I96A, I112A	Lateral surface
	trunc	69–248	
***Dm* eIF4G** (1–1666) (isoform A) O61380	4G	578–650	Extended eIF4E-binding region
CUP (1–1117) Q9VMA3	C+NC	325–376	eIF4E-binding region
	R1	R329E	Structurally conserved Arg/Lys residues
	R2	R336E	
	R1,R2	R329E, R336E	
	R1,R2,R3	R329E, R336E, R375E	
Thor (1–117) Q9XZ56	Thor	Full-length	
	C+NC	50–83	eIF4E-binding region
	C*	Y54A, M59A	Canonical
	NC*	V75D, L79D	Non-canonical
	C*+NC*	Y54A, M59A, V75D, L79D	Canonical + non-canonical
	R1	R56E	Structurally conserved Arg/Lys residues
	R2	R63E	
	R1,R2	R56E, R63E	
	RR-AA	R56A, R63A	
	R1,R2,R3	R56E, R63E, R81E	
	Thor extended	32–83	Extended eIF4E-binding region
	S65D, T70D	32–83; S65D, T70D	Phosphomimetic
	C* + S65D, T70D	32–83; Y54A, M59A, S65D, T70D	Canonical + phosphomimetic
	peptide	50–83	Wild-type peptide
	phosphopeptide	50–83; pS65, pT70	Phosphorylated peptide
4E-T (1–1010) Q8IH18	C+NC	9–44	eIF4E-binding region
	C*	Y10A, L15A	Canonical
	NC*	F41D, W42D	Non-canonical
	C*+NC*	Y10A, L15A, F41D, W42D	Canonical + non-canonical
	K1	K12E	Structurally conserved Arg/Lys residues
	R2	R19E	
	R1,R2	K12E, R19E	
	R1,R2,K3	K12E, R19E, K43E	
***Hs* eIF4E** (1-217)	4E	Full-length	
	trunc	36–217	
***Hs* 4E-BP1** (1-118)	C+NC	50–83	eIF4E binding region
	C*	Y54A, L59A	Canonical
	NC*	L75A, V81A	Non-canonical
	C*+NC*	Y54A, L59A, L75A, V81A	Canonical + non-canonical
***Hs* 4E-BP3** (1-100)	C+NC	36–69	eIF4E binding region
	C*	Y40A, L45A	Canonical
	NC*	L61A, V67A	Non-canonical
	C*+NC*	Y40A, L45A, L61A, V67A	Canonical + non-canonical
Triple chimera	Chimeric peptide	4E-T residues 9–20 Thor residues 65–72 + Gly CUP residues 362–376	Canonical Linker + Gly Non-canonical

Table S2 related to Figures 1 and S5. Thermodynamic parameters for the interaction of eIF4G and 4E-BP derived peptides with eIF4E (full-length or 69–248)

Peptide	K_D (M)	ΔH (kcal mol^{-1})	-TΔS (kcal mol^{-1})	ΔG (kcal mol^{-1})	Molar ratio
Triple chimera	$3.5 \pm 0.2 \cdot 10^{-9}$	-32.8 ± 0.9	21.7	-11.1	0.98 ± 0.05
eIF4G 578–650* (eIF4E full-length)	$5.7 \pm 0.9 \cdot 10^{-9}$	-15.0 ± 0.7	4.1	-10.9	0.99 ± 0.01
Thor (32–83) Wild-type	$9.4 \pm 1.4 \cdot 10^{-9}$	-14.7 ± 0.7	4.0	-10.7	1.02 ± 0.01
Thor (32–83) S65D,T70D	$11.9 \pm 1.5 \cdot 10^{-9}$	-17.2 ± 1.9	6.6	-10.6	0.99 ± 0.01
Thor (50–83) wild-type	$10.5 \pm 2.5 \cdot 10^{-9}$	-12.8 ± 0. 7	2.1	-10.7	1.01 ± 0.01
Thor (50–83) pS65, pT70	$16.9 \pm 1.3 \cdot 10^{-9}$	-15.0 ± 0.9	4.5	-10.5	1.00 ± 0.01

The triple chimera exhibits nanomolar affinity for eIF4E, similar to the individual *Dm* 4E-BPs (Igreja et al., 2014). The high entropic penalty on the interaction is similar to that observed for the CUP peptide but different to Thor and 4E-T, in agreement with the observation that the non-canonical portion of the chimeric peptide also folds into an α-helix (Figure 6F). Values were determined as described in the Supplemental Experimental Procedures.

Table S3 related to Figures 5 and 6. Antibodies used in this study

Antibody	Source	Catalog Number	Dilution	Monoclonal/ Polyclonal
Anti-V5	Invitrogen	554205	1:5,000	Monoclonal
Anti-HA-HRP (western blot)	Roche	12 013 819 001	1:5,000	Monoclonal
Anti-HA (Immunoprecipitation)	Covance	MMS-101P	1:1,000	Monoclonal
Anti- *Dm* eIF4E	In house		1:3,000	Rabbit polyclonal
Anti- *Hs* eIF4E	BETHYL laboratories	A301-154A	1:2,000	Rabbit polyclonal
Anti- *Hs* eIF4G	BETHYL laboratories	A301-776A	1:2,000	Rabbit polyclonal
Anti-GFP	In house		1:2,000	Rabbit polyclonal
Anti-rabbit-HRP	GE Healthcare	NA934V	1:10,000	Polyclonal
Anti-mouse-HRP	GE Healthcare	RPN4201	1:10,000	Polyclonal

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

DNA constructs

For expression in *E. coli*, a DNA fragment coding for *Hs* eIF4E (residues 36–217) was inserted into the NdeI and NheI restriction sites of the pnYC-NpH vector (Diebold et al., 2011), resulting in expressed polypeptides with an N-terminal His_6-tag cleavable by HRV3C protease. DNA fragments coding for *Hs* 4E-BP1 (residues 50–83) or *Dm* Thor (residues 32–83) were inserted into the NdeI and NheI restriction sites of the pnEA-NpM vector (Diebold et al., 2011), producing an N-terminal fusion with MBP cleavable by HRV3C protease. The DNA sequence coding for the B1 domain of immunoglobulin-binding protein G (GB1; Cheng et al., 2004) was inserted C-terminal to the 4E-BP fragments by site-directed mutagenesis. A synthetic DNA fragment coding for the triple chimera peptide was cloned between the XhoI and NheI restriction sites of the pnEA-NvM vector (Diebold et al., 2011). The GB1 tag was inserted C-terminally into the pnEA-NvM-triple chimera construct as described above. All the constructs and mutations are listed in Table S1.

Coimmunoprecipitation assays in human cells

For coimmunoprecipitations in human cells, HEK293T cells were grown in 10 cm dishes and transfected using TurboFect transfection reagent (Thermo Scientific) according to the manufacturer's recommendations. The transfection mixtures contained a total of 10 µg plasmid including the SBP-V5 tagged proteins. Two days after transfection, the cells were washed with PBS and lysed in NET lysis buffer [50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.1% Triton-X 100, 1 mM EDTA] supplemented with 10% glycerol and protease inhibitors (Complete Protease Inhibitor Mix, Roche) for 10 min on ice. The cells were centrifuged at 18,000 *g* for 15 min at 4°C. The cleared cell lysates were treated with RNase A for 30 min and centrifuged again at 18,000 ×*g* for 15 min. For SBP pulldowns, 50 µl streptavidin-Sepharose (GE Healthcare) was directly added to the samples and the samples were rotated

for 1 h at 4°C. For m^7GTP pulldowns, 25 µl of γ-Aminophenyl-m^7GTP (Jena Biosciences) were added to the cell lysates and incubated for 1h at 4°C. The beads were washed three times with NET wash buffer [50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 0.1% Triton-X 100, 1 mM EDTA] and once with NET buffer without detergent. Bound proteins were eluted with 100 µl of SDS-PAGE sample buffer and analyzed by western blotting.

Crystallization

Crystals of *Dm* eIF4E (residues 69–248) in complex with Thor (residues 50–83) were obtained at 18°C using the hanging-drop vapor diffusion method one day after mixing the protein solution (12 mg/ml, 0.2 µl) with reservoir (0.2 µl) containing 18% PEG 3350, 0.2 M KNO_3 and 0.5% n-octyl-β-D-glucoside. Crystals of *Dm* eIF4E (residues 68–248) in complex with 4E-T (residues 9–44; 10 mg/ml) were obtained at 18°C using the hanging-drop vapor diffusion method. Crystals grew in 2–3 days after mixing the protein solution with reservoir solution containing 0.1 M Tris-HCl (pH 8.5), 28% PEG 400 and 0.05 M $LiSO_4$. Crystals of *Dm* eIF4E (residues 68–248) in complex with *Dm* eIF4G (residues 578–650; 14 mg/ml) and m^7GpppG (New England BioLabs) were obtained in 0.1 M MES (pH 6.5) and 25% PEG 2000 monomethyl ether (MME) at 20°C using the sitting-drop vapor diffusion method. The crystals were improved by microseeding in related conditions [0.1 M MES (pH 6.0) and 20% PEG 2000 MME at 18°C]. A single crystal of *Hs* eIF4E (residues 36–217) in complex with *Hs* 4E-BP1 (residues 50–83) was obtained in the Morpheus 96-well screen condition E12 (Molecular Dimensions) at 20°C using the sitting-drop vapor diffusion method. Crystals of *Dm* eIF4E (residues 69–248) in complex with the triple-chimera (crystal form 1, Table 1) were obtained at 20°C in the Morpheus 96-well screen condition E10 (Molecular Dimensions) using the sitting-drop vapor diffusion method. A second crystal form of this complex (crystal form 2, Table 1) was obtained in 0.1 M MES (pH 6.5) and 22% PEG 300 using the hanging drop

vapor diffusion method. When necessary, crystals were soaked in mother liquor supplemented with 10–15% glycerol for cryoprotection before flash-freezing in liquid nitrogen.

Data collection and structure determination

Data for all crystals were collected at 100K on a PILATUS 6M detector at the PXII beamline of the Swiss Light Source. Diffraction images of the *Dm* eIF4E–Thor complex were recorded at a wavelength of 1.000 Å. Diffraction data extended to a resolution of 1.10 Å. The data were processed with XDS and scaled using XSCALE (Kabsch, 2010). The structure was solved by molecular replacement using PHASER (McCoy et al., 2007) and one copy of the *Dm* eIF4E (PDB-ID: 4AXG; Kinkelin et al., 2012) as a model. To avoid model bias, the molecular replacement solution of the *Dm* eIF4E molecule was used to rebuild an initial model of eIF4E using ARP/wARP autobuilding from the CCP4 package (Langer et al., 2008; Winn et al., 2011). To complete the structure, iterative cycles of model building and refinement were carried out with COOT (Emsley et al., 2010) and PHENIX (Afonine et al., 2012). The Thor peptide chain was built manually into the difference density in COOT, and the structure of the complex was further refined with PHENIX. In the final model, 98.5% of all residues were in the favored regions of the Ramachandran plot and there were no outliers.

Crystals of the *Dm* eIF4E–4E-T complex diffracted to 2.15 Å resolution, and data were collected at a wavelength of 1.000 Å. Data processing, model building and refinement were performed as described above except that the structure of the *Dm* eIF4E–Thor complex was used as a search model. In the final model, 97.0% of all residues were in the favored regions of the Ramachandran plot and there were no outliers.

Diffraction data for crystals of the *Dm* eIF4E–eIF4G complex were collected at a wavelength of 1.038 Å and extended to 2.40 Å resolution. Data processing, model building and refinement were performed as described above except that the asymmetric unit contained two copies of the search model (*Dm* eIF4E from the eIF4E-Thor complex). The final model of the

Dm eIF4E–eIF4G complex contains 98.4% of all residues in the favored regions of the Ramachandran plot and no outliers.

Diffraction data for the crystals of the *Hs* eIF4E–4E-BP1 complex were collected at a wavelength of 1.070 Å and extended to 1.75 Å resolution. Data processing, model building and refinement were performed as described above, with the exception that the structure of *Hs* eIF4E (PDB-ID: 4TPW; Papadopoulos et al., 2014) was used as a search model. The final model contains 98.6% of all residues in the favored regions of the Ramachandran plot and no outliers.

Diffraction data for the crystals of *Dm* eIF4E in complex with the triple-chimera peptide (crystal form 1) were collected at a wavelength of 1.000 Å and extended to 2.62 Å resolution. A second crystal form (crystal form 2) of the *Dm* eIF4E-triple chimera complex diffracted to 2.52 Å resolution and data were collected at a wavelength of 0.999 Å. For both crystal forms, data processing, model building and refinement was done as described above, with the exception that the asymmetric unit contained three copies of the search model (*Dm* eIF4E from the eIF4E–Thor complex) and that the initial model of the complex was re-build using the PHENIX AutoBuild wizard (Terwilliger et al., 2008). The final models of the *Dm* eIF4E-triple-chimera (crystal forms 1 and 2) contain 98.0% and 99.0% of all residues in the favored regions of the Ramachandran plot, respectively, and no outliers.

The stereochemical properties of all structures were verified with MOLPROBITY (Chen et al., 2010), and structural images were prepared with PyMOL (http://www.pymol.org). The diffraction data and refinement statistics are summarized in Table 1.

ITC analysis

The ITC experiments were performed using a VP-ITC microcalorimeter (MicroCal) at 20°C as described previously (Igreja et al., 2014). The solution of eIF4E (residues 69–248, 2 μM) in the calorimetric cell was titrated with tenfold concentrated solutions of GB1-stabilized

peptides [eIF4G, Thor fragment, or triple chimera peptide (20 μM)] that were dissolved in the same buffer (20 mM Na-phosphate (pH 7.0) and 150 mM NaCl). Alternatively, synthetic Thor peptides (wild-type or phosphorylated; EMC microcollections) were used. The titration experiments consisted of an initial injection of 2 μL followed by 28 injections of 10 μL at intervals of 240 s. The thermodynamic parameters were estimated using a one-site binding model (Origin version 7.0); the datapoint of the first injection was removed for the analysis (Mizoue and Tellinghuisen, 2004). Errors represent standard errors from three independent experiments.

Competition and pulldown assays

The *in vitro* pulldown and competition assays were performed as described previously (Igreja et al., 2014), with the exception that the purified complexes containing *Dm* His_6-eIF4E (residues 69–248) and *Dm* MBP-eIF4G (residues 578–650; 2 μM) used in the competition assays were incubated with Ni-NTA beads for 20 min. The bound complexes were challenged with the indicated amounts of competitor for the indicated time points. After incubation the beads were pelleted and washed three times with 750 μL buffer containing 50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 20 mM Imidazole and 2 mM DTT. Proteins bound to the Ni-NTA beads were eluted with the same buffer containing 500 mM Imidazole. The proteins were analyzed by 15% SDS-PAGE followed by Coomassie blue staining. For the competition assay shown in Figure S4F,G, complexes consisting of MBP-tagged Thor (full length, wild type or mutant) and SHN-tagged eIF4E (full length) were obtained by coexpression. The proteins were purified from cleared lysates using amylose resin (New England Biolabs) and either stored at -80 or immobilized on Ni-NTA beads (4 μM). The immobilized complexes were incubated with 5-fold molar excess of GB1-stabilized eIF4G fragment (residues 578–650). The proteins were pulled down at the indicated time points, eluted and analyzed as described above.

Translation repression assays

S2 cells were transfected in 6-well plates using Effectene transfection reagent (Qiagen) according to the manufacturer's protocol. The transfection mixtures contained 1 μg of R-Luc reporter plasmid, 0.1 μg of a plasmid expressing a control RNA (uncapped, derived from 7SL RNA), and increasing amounts of plasmids expressing GFP-tagged triple chimera peptide or full-length Thor (wild-type or phosphomimetic mutant, 0.05–0.5 μg). The plasmid expressing the GFP control was transfected at the concentration necessary to obtain expression at a level similar to the highest expression level of GFP-Thor. In all of the experiments, the cells were collected 3 days after transfection. R-Luc activity was measured using the Dual-Luciferase reporter assay system (Promega) and northern blotting was performed as described previously (Igreja et al., 2014). The R-Luc mRNA levels were determined by northern blotting and were normalized to those of the control RNA. The normalized R-Luc mRNA levels were then used to normalize the R-Luc activity to obtain translation efficiencies.

SUPPLEMENTAL REFERENCES

Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., and Adams, P.D. (2012). Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr. D. Biol. Crystallogr. *68*, 352-367.

Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D. Biol. Crystallogr. *66*, 12-21.

Cheng, Y., and Patel, D.J. (2004). An efficient system for small protein expression and refolding. Biochem. Biophys. Res. Commun. *317*, 401-405.

Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004). WebLogo: A Sequence Logo Generator. Genome Res. *14*, 1188-1190.

Diebold, M.L., Fribourg, S., Koch, M., Metzger, T., and Romier, C. (2011). Deciphering correct strategies for multiprotein complex assembly by co-expression: application to complexes as large as the histone octamer. J. Struct. Biol. *175*, 178-188.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D. Biol. Crystallogr. *66*, 486-501.

Kabsch, W. (2010). Xds. Acta Crystallogr. D. Biol. Crystallogr. *66*, 125-132.

Langer, G., Cohen, S.X., Lamzin, V.S., and Perrakis, A. (2008). Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. Nat. Protoc. *3*, 1171-1179.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J. Appl. Crystallogr. *40*, 658-674.

Mizoue, L.S., and Tellinghuisen, J. (2004). The role of backlash in the "first injection anomaly" in isothermal titration calorimetry. Anal. Biochem. *326*, 125-127.

Terwilliger, T.C., Grosse-Kunstleve, R.W., Afonine, P.V., Moriarty, N.W., Zwart, P.H., Hung, L.-W., Read, R.J., and Adams, P.D. (2008). Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. Acta Crystallogr. D64, 61–69.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A.*, et al.* (2011). Overview of the CCP4 suite and current developments. Acta Crystallogr. D. Biol. Crystallogr. *67*, 235-242.

Mextli proteins use both canonical bipartite and novel tripartite binding modes to form eIF4E complexes that display differential sensitivity to 4E-BP regulation

Daniel Peter, Ramona Weber, Carolin Köne, Min-Yi Chung, Linda Ebertsch, Vincent Truffault, Oliver Weichenrieder, Cátia Igreja, and Elisa Izaurralde

Department of Biochemistry, Max Planck Institute for Developmental Biology, 72076 Tübingen, Germany

The eIF4E-binding proteins (4E-BPs) are a diverse class of translation regulators that share a canonical eIF4E-binding motif (4E-BM) with eIF4G. Consequently, they compete with eIF4G for binding to eIF4E, thereby inhibiting translation initiation. Mextli (Mxt) is an unusual 4E-BP that promotes translation by also interacting with eIF3. Here we present the crystal structures of the eIF4E-binding regions of the *Drosophila melanogaster* (*Dm*) and *Caenorhabditis elegans* (*Ce*) Mxt proteins in complex with eIF4E in the cap-bound and cap-free states. The structures reveal unexpected evolutionary plasticity in the eIF4E-binding mode, with a classical bipartite interface for *Ce* Mxt and a novel tripartite interface for *Dm* Mxt. Both interfaces comprise a canonical helix and a noncanonical helix that engage the dorsal and lateral surfaces of eIF4E, respectively. Remarkably, *Dm* Mxt contains a C-terminal auxiliary helix that lies anti-parallel to the canonical helix on the eIF4E dorsal surface. In contrast to the eIF4G and *Ce* Mxt complexes, the *Dm* eIF4E–Mxt complexes are resistant to competition by bipartite 4E-BPs, suggesting that *Dm* Mxt can bind eIF4E when eIF4G binding is inhibited. Our results uncovered unexpected diversity in the binding modes of 4E-BPs, resulting in eIF4E complexes that display differential sensitivity to 4E-BP regulation.

[*Keywords*: CUP; eIF4F; eIF4G; 4E-BP; Mextli; translational regulation]

Supplemental material is available for this article.

Received July 20, 2015; revised version accepted August 7, 2015.

The inhibition of cap-dependent translation initiation is a widespread and reversible mechanism for regulating gene expression in eukaryotes (Kong and Lasko 2012). This type of regulation is mediated by a diverse family of eIF4E-binding proteins (4E-BPs), which play essential roles in diverse biological processes, including cell proliferation, development, and neuronal plasticity (Banko et al. 2005; Dowling et al. 2010; Kong and Lasko 2012; Gkogkas et al. 2013).

Canonical cap-dependent translation initiation requires the assembly of the ternary eukaryotic initiation factor eIF4F, which consists of the cap-binding protein eIF4E, the adaptor protein eIF4G, and the RNA helicase eIF4A (Jackson et al. 2010). The eIF4G protein serves as a scaffold to mediate multiple protein–protein interactions that are essential for cap-dependent translation initiation. Specifically, eIF4G interacts with eIF4E bound to the mRNA 5′ cap structure and recruits the 43S preinitiation complex (PIC) through interactions with eIF3, thereby coupling translation initiation with the recognition of the mRNA cap structure (Jackson et al. 2010).

The interaction of eIF4G with eIF4E is mediated by a conserved motif, termed the canonical eIF4E-binding motif (4E-BM) of sequence $YX_4L\Phi$ (where Y is Tyr, X is any amino acid, L is Leu, and Φ is a hydrophobic residue) (Mader et al. 1995; Marcotrigiano et al. 1999). This interaction is competitively inhibited by the 4E-BPs, which, like eIF4G, contain a canonical 4E-BM (Mader et al. 1995; Marcotrigiano et al. 1999). The canonical 4E-BMs found in eIF4G and 4E-BPs adopt similar α-helical conformations and compete for binding to a conserved patch of hydrophobic residues on the dorsal surface of eIF4E that is opposite the cap-binding pocket (Mader et al. 1995; Matsuo et al. 1997; Marcotrigiano et al. 1999; Gross et al. 2003).

Corresponding authors: elisa.izaurralde@tuebingen.mpg.de, catia.igreja@tuebingen.mpg.de

Article published online ahead of print. Article and publication date are online at http://www.genesdev.org/cgi/doi/10.1101/gad.269068.115. Freely available online through the *Genes & Development* Open Access option.

C-terminal to the canonical motif, 4E-BPs contain noncanonical 4E-BMs that are connected by variable linkers. Structural studies have revealed that, despite a lack of sequence similarity in their linkers and noncanonical motifs, all 4E-BPs bind to a conserved lateral surface of eIF4E (Mizuno et al. 2008; Gosselin et al. 2011; Kinkelin et al. 2012; Paku et al. 2012; Lukhele et al. 2013; Igreja et al. 2014; Peter et al. 2015). Through this interaction, the noncanonical motifs increase the affinity of 4E-BPs for eIF4E by three orders of magnitude (Paku et al. 2012; Lukhele et al. 2013; Igreja et al. 2014; Peter et al. 2015) and are necessary for 4E-BPs to be able to compete with eIF4G and repress translation (Igreja et al. 2014; Peter et al. 2015).

Mextli (Mxt) is an invertebrate-specific 4E-BP involved in germline stem cell maintenance and early embryogenesis in *Drosophila melanogaster* (Hernández et al. 2013). Like all 4E-BPs, Mxt contains a canonical 4E-BM and competes with eIF4G for binding to eIF4E (Fig. 1A; Hernández et al. 2013). However, in contrast to other 4E-BPs, Mxt was reported to promote cap-dependent translation initiation through interactions with eIF3 (Hernández et al. 2013).

Mxt is currently the only known 4E-BP besides eIF4G that has been reported to stimulate translation. Therefore, understanding its eIF4E-binding mode will shed light on alternative mechanisms for initiating cap-dependent translation. To understand the binding mode of Mxt to eIF4E, we solved the crystal structures of the eIF4E-binding regions of *D. melanogaster* (*Dm*) and *Caenorhabditis elegans* (*Ce*) Mxt bound to their respective eIF4Es in either the presence or absence of an m^7GpppG cap analog. The canonical and noncanonical motifs of Mxt adopt an α-helical conformation upon binding to the dorsal and lateral surfaces of eIF4E, respectively, forming a bipartite binding interface. This type of interface has been observed for other eIF4E–4E-BP complexes, including eIF4E complexes with human 4E-BP1 and *Dm* CUP, Thor, and 4E-T

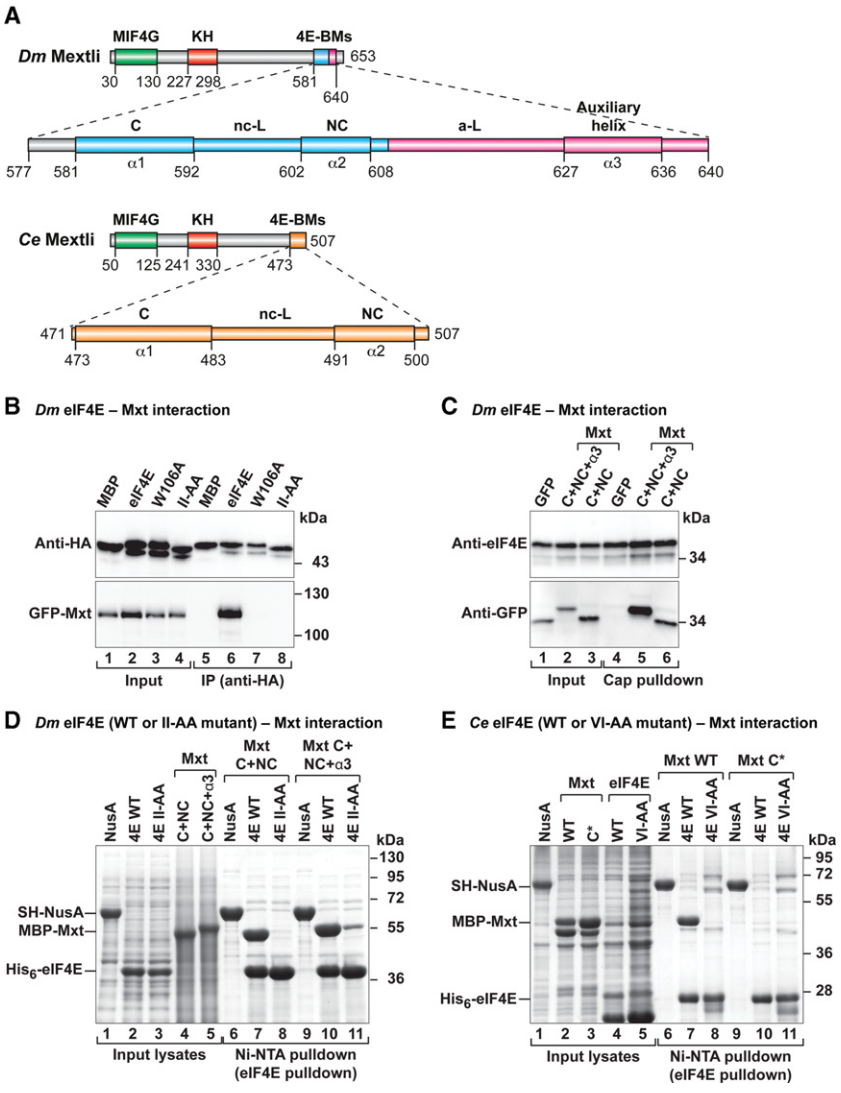


Figure 1. Mxt proteins associate with eIF4E through bipartite or tripartite binding sequences. (*A*) Mxt contains a MIF4G-like domain, an hnRNP K-homology (KH) domain, and C-terminal 4E-BMs. The eIF4E-binding region of *D. melanogaster* (*Dm*) Mxt contains three 4E-BMs (canonical [C], noncanonical [NC], and auxiliary [A] helix [α3]) connected by a noncanonical (nc-L) and auxiliary (a-L) linkers. The eIF4E-binding region of *Caenorhabditis elegans* (*Ce*) Mxt is bipartite and composed of canonical and noncanonical 4E-BMs connected by a noncanonical linker. The amino acid positions at the domain/motif boundaries are indicated *below* the protein outlines and are based on the structures presented in the present study. (*B*) Interaction of HA-tagged *Dm* eIF4E (either wild type or mutant) with GFP-tagged full-length *Dm* Mxt. The inputs (3%) and immunoprecipitates (30% for Mxt and 20% for eIF4E) were analyzed by Western blotting using anti-HA and anti-GFP antibodies. (*C*) Interaction of endogenous *Dm* eIF4E with the tripartite (C + NC + α3) and bipartite (C + NC) *Dm* Mxt peptides N-terminally fused to GFP. The proteins were pulled down with m^7GTP-Sepharose beads. The inputs (5% for Mxt fragments and 1% for eIF4E) and immunoprecipitates (20% for Mxt and 10% for eIF4E) were analyzed by Western blotting. (*D*) Ni-NTA pull-down assay showing the association of hexahistidine [His_6]-tagged *Dm* eIF4E (full-length; either wild type or the II-AA mutant) with maltose-binding protein (MBP)-tagged *Dm* Mxt fragments C-terminally fused to GB1. The inputs (10%) and bound fractions (25%) were analyzed by SDS-PAGE followed by Coomassie blue staining. SH-NusA served as negative control. (*E*) Ni-NTA pull-down assay showing the association of His_6-tagged *Ce* eIF4E (1–215; wild type or VI-AA mutant) with the MBP-tagged *Ce* Mxt eIF4E-binding region (471–507; either wild type or carrying mutations in the canonical motif [C*]). Samples were analyzed as described in *D*. The size markers (in kilodaltons) are shown at the *right* of each panel.

(Kinkelin et al. 2012; Lukhele et al. 2013; Peter et al. 2015). However, in *Dm* Mxt, the bipartite binding surface extends beyond the noncanonical motif and forms a third interface through an auxiliary linker and helix that contact eIF4E in a novel binding mode. This novel tripartite binding mode confers resistance to competition by other 4E-BPs, potentially providing a mechanism to sustain cap-dependent translation under conditions in which the interaction of eIF4G with eIF4E is inhibited. More generally, our study also identifies diverse binding modes of 4E-BPs that are valuable for the rational design of new translation inhibitors for the treatment of malignancies—such as cancer, autism spectrum disorders (ASDs), and fragile X syndrome—that are associated with the up-regulation of eIF4E activity and protein synthesis (Banko et al. 2005; Dowling et al. 2010; Gkogkas et al. 2013; Martineau et al. 2013).

Results

Mxt binds to the dorsal and lateral surfaces of eIF4E

The 4E-BPs bind to the dorsal and lateral surfaces of eIF4E using canonical and noncanonical 4E-BMs, respectively (Kinkelin et al. 2012; Paku et al. 2012; Lukhele et al. 2013; Igreja et al. 2014; Peter et al. 2015). To determine whether *Dm* Mxt contains a noncanonical motif that binds to the lateral surface of eIF4E, we used coimmunoprecipitation assays in *Dm* Schneider 2 (S2) cells and examined the effects of Ile96Ala and Ile112Ala substitutions (II-AA) in the lateral hydrophobic pocket of eIF4E. The II-AA mutations abolished the interaction of *Dm* eIF4E with Mxt and CUP but not with eIF4G (Fig. 1B; Supplemental Fig. S1A,B), as has been observed previously for other 4E-BPs (Igreja et al. 2014). In contrast, a Trp106Ala substitution (W106A) on the dorsal binding surface of eIF4E abolishes or strongly reduces binding to Mxt, CUP, and eIF4G, as would be expected for proteins containing canonical 4E-BMs (Fig. 1B, lane 7; Supplemental Fig. S1A, lane 7; Igreja et al. 2014; Peter et al. 2015). These results indicate that, in contrast to eIF4G but similar to other 4E-BPs, *Dm* Mxt requires both the dorsal and lateral surfaces to efficiently bind to eIF4E in cell lysates that contain eIF4G (or other 4E-BPs) and may compete with Mxt for binding to eIF4E.

Dm *Mxt contains a noncanonical 4E-BM and auxiliary binding sequences*

The immunoprecipitation assay shown in Figure 1B indicates that *Dm* Mxt contains a noncanonical motif that interacts with the lateral binding surface of eIF4E. Although the noncanonical motifs are not conserved among different 4E-BPs, they contain common features: (1) They are located ~15–30 residues downstream from the canonical motifs, (2) they contain hydrophobic residues, and (3) in the case of CUP, they exhibit helical propensity. We inspected the *Dm* Mxt sequence for motifs that fulfilled these criteria and identified a motif downstream from the canonical motif that exhibited helical propensity and contained hydrophobic residues. This motif was termed a noncanonical motif based on the structural studies presented below (NC, residues 602–608) (Fig. 1A; Supplemental Fig. S2A,B). We also identified a second patch of residues with helical propensity (residues 627–636) that was 38 residues downstream from the canonical motif, which was termed an auxiliary helix (α3) (Fig. 1A; Supplemental Fig. S2A,B). The residues connecting the noncanonical motif and auxiliary helix are the auxiliary linker (Fig. 1A).

To determine whether the auxiliary helix contributes to eIF4E binding, we expressed C-terminal Mxt fragments with or without this auxiliary helix (fragments C + NC and C + NC + α3, termed bipartite and tripartite, respectively) (Supplemental Table S1) and tested their binding to endogenous eIF4E in pull-down assays using m^7GTP-sepharose beads. The two *Dm* Mxt fragments associated with cap-bound eIF4E, but the tripartite fragment exhibited stronger binding (relative to the input) than the bipartite fragment (Fig. 1C, lanes 5,6 vs. lanes 2,3, respectively), suggesting that the auxiliary helix contributes to eIF4E binding.

To further confirm the contribution of Mxt auxiliary sequences to the interaction with eIF4E, we performed in vitro pull-down assays using bacterially expressed versions of the Mxt fragments described above. In contrast to the coimmunoprecipitation assays in cell lysates, the pull-down assays tested the interaction of Mxt with eIF4E in the absence of other 4E-BPs that might compete for binding to eIF4E. Recombinant full-length *Dm* eIF4E (expressed with a hexahistidine [His_6] tag) pulled down similar amounts of the two Mxt fragments (Fig. 1D, lanes 7,10). The binding of the bipartite Mxt fragment was abolished by the II-AA mutations on the lateral surface of eIF4E (Fig. 1D, lane 8). In contrast, the tripartite Mxt fragment retained some binding (Fig. 1D, lane 11), suggesting that the auxiliary sequences interact with a different surface of eIF4E and can partially compensate for the negative effects of the II-AA mutation. Mxt fragments did not interact with His_6-NusA as a negative control (Fig. 1D, lanes 6,9). Thus, *Dm* Mxt contains a canonical and a noncanonical 4E-BM as well as an additional downstream auxiliary sequence that contributes to the interaction with eIF4E.

The auxiliary sequences increase the affinity of Dm *Mxt for eIF4E*

To test the affinities of the bipartite and tripartite *Dm* Mxt fragments for eIF4E, we performed isothermal titration calorimetry (ITC) experiments. The bipartite Mxt fragment has a binding affinity for eIF4E comparable with the affinities observed for other 4E-BPs (Paku et al. 2012; Lukhele et al. 2013; Igreja et al. 2014; Peter et al. 2015), with dissociation constants (K_Ds) in the nanomolar range ($5\times10^{-9} \pm 3\times10^{-9}$ M) (Supplemental Table S2; Supplemental Fig. S3A). Including the auxiliary sequences in the tripartite fragment resulted in a 10-fold increase in affinity ($0.5\times10^{-9} \pm 0.09\times10^{-9}$ M) (Supplemental Table S2; Supplemental Fig. S3B). Notably, the binding of the tripartite Mxt fragment to eIF4E was enthalpically driven, with

a high entropic penalty (Supplemental Table S2). This penalty can be explained by a greater flexibility in the unbound state compared with the bound state, whereas favorable enthalpy indicates an extensive interaction network. Nuclear magnetic resonance (NMR) analysis together with the crystal structures presented below are consistent with these results and indicate that the Mxt peptide is predominantly unfolded in solution but folds into three α helices upon binding to eIF4E (Fig. 2; Supplemental Fig. S3D,E).

The affinity of the tripartite Mxt peptide for the eIF4E II-AA mutant was reduced by three orders of magnitude ($0.3 \times 10^{-6} \pm 0.1 \times 10^{-6}$ M) (Supplemental Table S2; Supplemental Fig. S3C). This reduction supports the results shown in Figure 1D and provides further evidence for the contribution of the lateral surface of eIF4E to complex formation.

Dm *and* Ce *Mxt evolved different binding modes to interact with eIF4E*

A defining feature of the Mxt protein family is an N-terminal MIF4G-like domain followed by a K-homology (KH) domain, both of which are highly conserved (Fig. 1A; Marcotrigiano et al. 2001; Hernández et al. 2013). The sequences C-terminal to the KH domain are variable in length and are less conserved except for the C terminus, which contains the canonical and noncanonical motifs and the connecting linker, which are also conserved among Mxt sequences from different species (Supplemental Fig. S2A,B).

In contrast, the auxiliary linker and helix in *Dm* Mxt are not conserved. Based on the auxiliary sequences, we define three Mxt families. The first family includes *Dm*

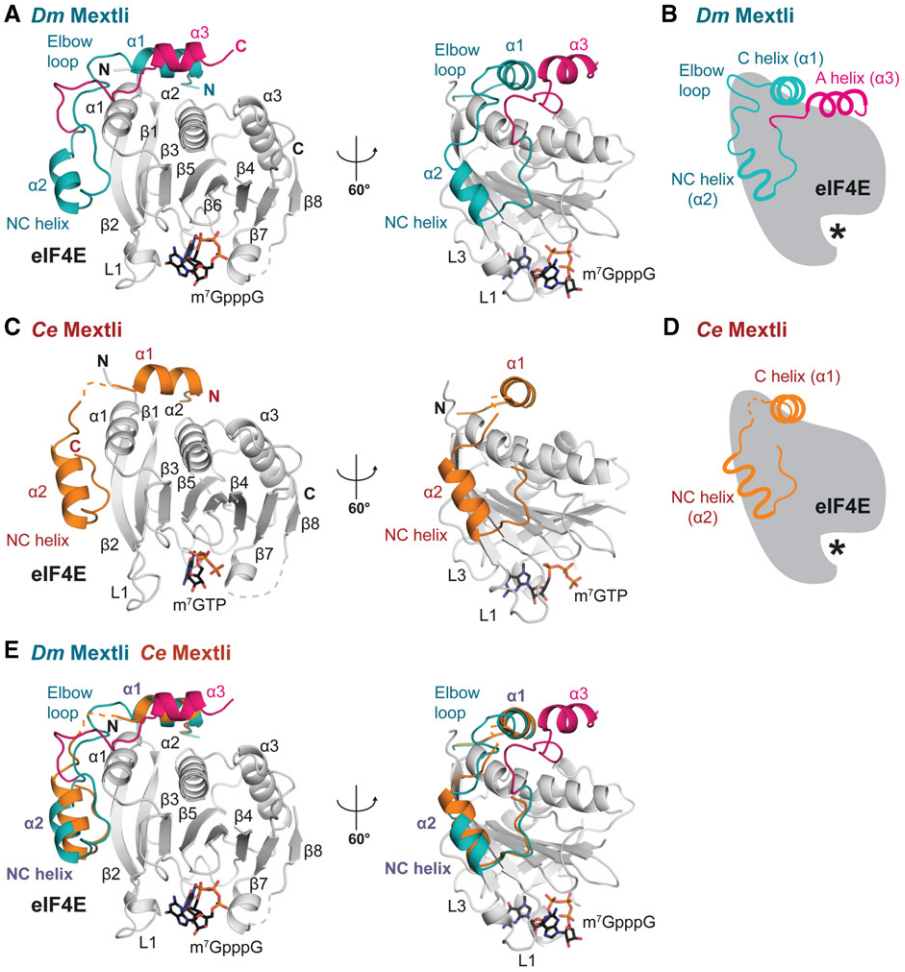


Figure 2. Structures of *Dm* and *Ce* Mxt proteins bound to eIF4E. (*A*) Overview of the structure of the 4E-binding region of *Dm* Mxt bound to eIF4E in two orientations. The region of the Mxt peptide with structural similarity to other 4E-BPs is colored in teal. The auxiliary linker and helix are colored in magenta. Selected secondary structure elements are labeled in black for eIF4E or blue and magenta for Mxt. The bound m^7GpppG cap analog is shown in sticks. (*B*) Schematic representation of Mxt bound to eIF4E, highlighting key structural features. The asterisk represents the mRNA cap structure. (*C*) Overview of the structure of the 4E-binding region of *Ce* Mxt bound to eIF4E in two orientations. The Mxt peptide is colored in orange. Selected secondary structure elements are labeled in black for eIF4E and dark red for Mxt. The bound m^7GTP cap analog is shown in sticks. (*D*) Schematic representation of *Ce* Mxt bound to eIF4E. (*E*) Structural overlay of the *Dm* and *Ce* eIF4E–Mxt complexes. For clarity, the eIF4E molecule from the *Ce* eIF4E–Mxt complex has been removed. Common structural features are labeled in purple.

Mxt and orthologs from dipteran and nondipteran insects; the auxiliary linker and helix are conserved, suggesting that these proteins interact with eIF4E using a mode of interaction similar to that of *Dm* Mxt (Supplemental Fig. S2A,B, family I). The second family includes Mxt proteins from nondipteran insects and arthropods, and their auxiliary sequences are more divergent; the length of the auxiliary linker varies, and a region with helical propensity is present downstream from the noncanonical motif (Supplemental Fig. S2B, family II). However, it is unclear whether these divergent auxiliary sequences contribute to eIF4E binding. Finally, the third family contains Mxt proteins from some arthropods and nematodes that lack the auxiliary linker and helix, as a stop codon is located immediately downstream from the noncanonical motif (Fig. 1A; Supplemental Fig. S2B, family III). These proteins are predicted to interact with eIF4E through the canonical and noncanonical motifs as observed for bipartite 4E-BPs, with no contribution from additional sequences.

To confirm these predictions, we tested the interaction of *Ce* Mxt (fragment 471–507) with *Ce* eIF4E. *Ce* eIF4E interacted with the *Ce* Mxt peptide (Fig. 1E, lane 7). This interaction was abolished by mutations in the canonical motif of *Ce* Mxt as well as mutations on the lateral surface of *Ce* eIF4E (VI-AA mutant, Supplemental Table S1), indicating that *Ce* Mxt binds eIF4E using the canonical and noncanonical motifs.

The structure of Dm *and* Ce *Mxt bound to eIF4E*

To gain insight into the binding modes of *Dm* and *Ce* Mxt to eIF4E, we crystallized the eIF4E-binding region of the proteins in complex with the corresponding eIF4Es and determined the structure of the complexes in the absence or presence of a cap analog (Table 1; Fig. 2A–D; Supplemental Fig. S4). For each organism, the structures in the cap-free and cap-bound states are very similar. They superpose with root mean square deviations (RMSDs) of 0.38 Å over 233 Cα atoms and 207 Cα atoms for the *Dm* and *Ce* structures, respectively (Supplemental Fig. S4). Similarly, no major conformational changes were observed in the eIF4E structures upon Mxt binding, as reported previously for various 4E-BPs (Gross et al. 2003; Volpon et al. 2006; Mizuno et al. 2008; Kinkelin et al. 2012; Paku et al. 2012; Peter et al. 2015).

The most surprising feature was observed in the *Dm* Mxt complex, where the auxiliary linker and helix contact eIF4E using an unprecedented binding mode (Fig. 2A,B). Specifically, the *Dm* Mxt peptide adopts a U-shaped arrangement on the surface of eIF4E, folding into three α helices—a canonical helix (α1), a noncanonical helix (α2), and an auxiliary helix (α3)—connected by linker sequences. The linker sequences consist of a noncanonical linker and auxiliary linker (Fig. 2A,B). This conformation was observed in the two structures of the *Dm* eIF4E–Mxt complex despite different crystal packing (Supplemental Fig. S4), indicating that the observed arrangement reflects a physiologically relevant interaction.

The *Ce* Mxt peptide folds into a canonical (α1) helix and a noncanonical (α2) helix and terminates after seven additional amino acids that form a terminal turn (Fig. 2C,D). Overall, the canonical and noncanonical helices and the connecting noncanonical linker of *Dm* and *Ce* Mxt bind eIF4E using a bipartite binding mode similar to that described for other 4E-BPs (Fig. 2E; Supplemental Fig. S5; Peter et al. 2015).

The presence of Arg/Lys residues at positions 2 and 9 of the canonical motif distinguishes bipartite from tripartite binding modes

The canonical motifs of *Dm* and *Ce* Mxt fold into an α helix that is held in position by interactions analogous to those previously reported for the canonical motifs of eIF4G and 4E-BPs in complex with eIF4E (Fig. 3A,B; Gross et al. 2003; Mizuno et al. 2008; Umenaga et al. 2011; Kinkelin et al. 2012; Paku et al. 2012; Peter et al. 2015). The most conserved interactions are mediated by residues corresponding to LΦ in the $YX_4L\Phi$ consensus sequence (*Dm* Mxt L586 and L587 and *Ce* Mxt L478 and M479). These residues form similar hydrophobic contacts with conserved Val and Trp residues in helix α1 of eIF4E (*Dm* V102 and W106 and *Ce* V64 and W68) (Fig. 3A,B). Furthermore, the hydroxyl group of the Tyr side chain in the canonical motifs (*Dm* Y581 and *Ce* Y473) contacts the backbone of the conserved H-P-L motif at the N terminus of strand β1 of eIF4E, as was observed in previous structures (Fig. 3A,B; Supplemental Fig. S1B).

The most obvious difference between the canonical motifs of *Ce* and *Dm* Mxt is that the *Ce* motif contains Arg residues at positions 2 and 9 (Fig. 3C,D). These Arg (or Lys) residues are also found in the canonical motifs of other 4E-BPs and eIF4G and contribute to the interaction with eIF4E, likely by shielding hydrophobic surface patches of eIF4E from solvent exposure (Kinkelin et al. 2012; Peter et al. 2015). In the canonical motif of *Dm* Mxt, the Arg residues are replaced by Ile583 and Ser590 (Fig. 3A,C,D). The substitution of long, bulky side chains at both positions by residues with shorter side chains is required to accommodate the auxiliary helix (α3), which would otherwise clash with the Arg/Lys residues in the canonical helix (see below). Notably, the canonical motifs of Mxt orthologs from families I and II, which contain auxiliary sequences, lack Arg/Lys residues at positions 2 and 9 (Supplemental Fig. S2A,B). This observation suggests that the substitutions of these residues coevolved with the acquisition of the tripartite binding mode.

The noncanonical linker of Ce *Mxt does not adopt a defined conformation*

Most of the known structures of 4E-BPs contain an elbow loop immediately after the canonical helix that bends the peptide backbone by ~90°, directing the noncanonical linker downward to engage the lateral surface of eIF4E (Peter et al. 2015). Despite the lack of sequence conservation, the elbow loops exhibit common features that are also observed in *Dm* Mxt but are absent in *Ce* Mxt (Fig. 4A–C). Specifically, the *Dm* Mxt elbow loop shows

Table 1. *Data collection and refinement statistics*

	Dm eIF4E–Mxt complex (cap-bound)	*Dm* eIF4E–Mxt complex (cap-free)	*Ce* eIF4E–Mxt complex (cap-bound)	*Ce* eIF4E–Mxt complex (cap-free)
Space group	$P2_12_12_1$	$P2_1$	$P4_12_12$	$P2_12_12$
Unit cell				
Dimensions				
a, b, c	44.9 Å, 56.4 Å, 99.1 Å	65.9 Å, 82.0 Å, 84.7 Å	71.0 Å, 71.0 Å, 85.3 Å	91.2 Å, 158 Å, 55.2 Å
Angles				
α, β, γ	90°, 90°, 90°	90°, 90.1°, 90°	90°, 90°, 90°	90°, 90°, 90°
Data collection				
Wavelength	1.000 Å	1.000 Å	1.000 Å	1.000 Å
Resolution	48.6–2.16 Å (2.22–2.16 Å)	44.0–2.13 Å (2.19–2.13 Å)	43.3–1.66 Å (1.70–1.66 Å)	47.2–1.95 Å (2.0–1.95 Å)
R_{sym}	0.125 (0.559)	0.058 (0.441)	0.060 (1.27)	0.117 (1.17)
Mean $I/\sigma I$	9.0 (2.2)	13.0 (2.0)	24.5 (2.0)	11.8 (2.0)
Completeness	99.3% (99.3%)	98.2% (90.5%)	100% (100%)	100% (100%)
Multiplicity	6.3 (4.5)	4.5 (3.3)	13.1 (12.0)	8.3 (8.6)
Refinement				
Resolution	40.9–2.16 Å	44.0–2.13 Å	43.3–1.66 Å	45.5–1.95 Å
Number of reflections	13,971	49,675	26,447	58,874
R_{work}/R_{free}	0.195/0.238	0.192/0.232	0.173/0.194	0.172/0.209
Number of atoms	2046	7937	1934	5918
Protein	1912	7790	1759	5464
Ligand/ion	53	—	38	20
Water	81	147	137	434
B-factors	41.7 $Å^2$	65.4 $Å^2$	31.6 $Å^2$	34.1 $Å^2$
Protein	41.5 $Å^2$	65.7 $Å^2$	30.8 $Å^2$	33.7 $Å^2$
Ligand/ion	52.6 $Å^2$	—	46.4 $Å^2$	43.6 $Å^2$
Water	39.7 $Å^2$	48.6 $Å^2$	38.0 $Å^2$	37.5 $Å^2$
Ramachandran plot				
Favored	99.1%	97.5%	99.0%	97.5%
Disallowed	0%	0%	0%	0%
RMSD				
Bond lengths	0.002 Å	0.01 Å	0.005 Å	0.008 Å
Bond angles	0.603°	1.05°	0.95°	1.05°

Values in parentheses are for the highest-resolution shell.
Ligands: m^7GpppG (for the cap analog) and one Cl^- ion in the *Dm* eIF4E–Mxt complex (cap-bound); m^7GTP (for the cap analog), four Zn^{2+} ions, and one Cl^- ion in the *Ce* eIF4E–Mxt complex (cap-bound); and two Mg^{2+} ions and three glycerol molecules in the *Ce* eIF4E–Mxt complex (cap-free).

remarkable similarity to the Thor elbow loop (Fig. 4A,B; Supplemental Fig. S5A,B). The Mxt elbow loop starts with residue S590 (structurally equivalent to $R63^{Thor}$) and contains a helical half-turn that begins with a Pro residue located at the tip of the elbow ($P593^{Mxt}$ and $P66^{Thor}$). The elbow loops of *Dm* Mxt and Thor end with similarly arranged Ser residues ($S595^{Mxt}$ and $S68^{Thor}$). The side chains of these Ser residues contact the carbonyl oxygens of the preceding residues ($Y589^{Mxt}$ and $L62^{Thor}$) and fix the backbone, which is further stabilized by internal interactions within the helical half-turn (e.g., between $S592^{Mxt}$ and the amide group of $H594^{Mxt}$) (Fig. 4A,B; Supplemental Fig. S5A,B). The *Dm* Mxt elbow loop is tethered to the lateral surface of eIF4E through interactions of $S595^{Mxt}$ and $L598^{Mxt}$ with the side chains of eIF4E residues S107, N110, and H111 (Fig. 4A).

One important difference between the *Dm* and *Ce* Mxt structures is that the noncanonical linker of *Ce* Mxt adopts a slightly different conformation in each of the three molecules found in the crystal that lacks the cap analog (Fig. 4C; Supplemental Fig. S4F). In the complex with the cap analog, the linker binds in yet another mode and is partially disordered, indicating that this linker is probably flexible and does not adopt a fixed conformation in solution (Fig. 4C).

The Mxt noncanonical motifs adopt an α-helical conformation

The noncanonical motif of *Ce* and *Dm* Mxt engages a lateral pocket of eIF4E that is lined by hydrophobic residues (*Dm* Y80, I96, and I112 and *Ce* Y42, V58, and I74), providing an additional binding site (Fig. 4D–H). A striking feature of Mxt proteins is the formation of an α helix (α2) at this surface, which has been previously observed only in the complex of CUP bound to eIF4E (Supplemental Fig. S5C; Kinkelin et al. 2012). Despite a similar α-helical arrangement, the noncanonical helices of *Dm* and *Ce* Mxt are tilted by ~21°–23° relative to the CUP helix (Supplemental Fig. S5D–F).

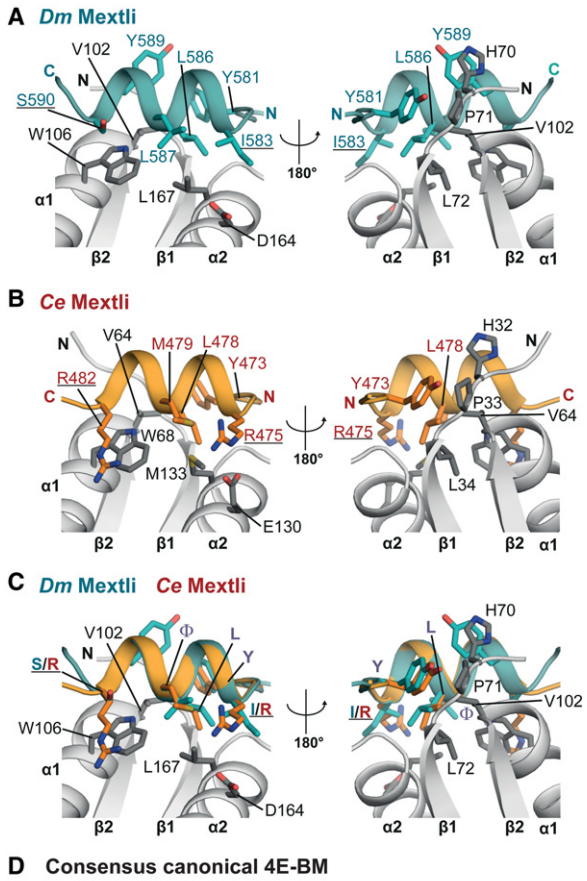


Figure 3. The canonical motifs of *Dm* and *Ce* Mxt. (*A*) Close-up view of the canonical helix of *Dm* Mxt bound to the dorsal surface of eIF4E in two orientations. Selected interface residues are shown as gray sticks for eIF4E and teal sticks for *Dm* Mxt. Residues at positions 2 and 9 of the canonical motif are underlined. (*B*) Close-up view of the canonical helix of *Ce* Mxt bound to the dorsal surface of eIF4E in two orientations. Selected interface residues are shown as gray sticks for *Ce* eIF4E and orange sticks for *Ce* Mxt. The Arg residues at positions 2 and 9 of the canonical motif are underlined. (*C*) Structural overlay of the *Dm* and *Ce* Mxt canonical helices. Selected interface residues of Mxt are shown as colored sticks. The Arg residues flanking the canonical helix of *Ce* Mxt and the corresponding residues in *Dm* Mxt are underlined. For clarity, the molecule of *Ce* eIF4E from the superposition has been removed. (*D*) The extended consensus sequence for the canonical 4E-BM (Peter et al. 2015) and the corresponding sequences for the indicated 4E-BPs and eIF4G.

The noncanonical helix of *Dm* Mxt is shorter and its contribution to the interaction with the lateral pocket of eIF4E is less pronounced than in *Ce* Mxt and *Dm* CUP. The major hydrophobic contacts of *Dm* Mxt with the lateral pocket of eIF4E are mediated by the side chain of M605 in the noncanonical helix and I612 in the linker following the helix (Fig. 4D). The *Ce* Mxt noncanonical helix contains an additional helical turn and uses M493 to establish further contacts with the lateral surface of eIF4E (Fig. 4E,F).

In the turn following the noncanonical helices, the main chain conformation of *Dm* and *Ce* Mxt aligns with CUP; all three proteins have an Ile residue (I612$^{Dm\text{Mxt}}$, I504$^{Ce\text{Mxt}}$, and I373$^{\text{CUP}}$) at a similar position that stacks on the conserved eIF4E Tyr residue (*Dm* Y80 and *Ce* Y42) of the hydrophobic pocket (Fig. 4D,E; Kinkelin et al. 2012). The Mxt peptide backbone is further tethered to eIF4E through interactions between the carbonyl oxygen of *Dm* I612$^{\text{Mxt}}$ and the nitrogen of *Dm* K113^{4E} (*Ce* I504$^{\text{Mxt}}$ to *Ce* Q75^{4E}) and between the amide nitrogen of *Dm* R614$^{\text{Mxt}}$ and the carbonyl oxygen of *Dm* H111$^{\text{Mxt}}$ (*Ce* I506$^{\text{Mxt}}$ to *Ce* H73^{4E}) (Fig. 4G,H).

The auxiliary linker of Dm *Mxt leads back toward the eIF4E dorsal surface*

The distinguishing features of the *Dm* eIF4E–Mxt structure are the auxiliary linker and α helix (α3), which together further anchor Mxt to the eIF4E surface. The side chains of N110^{4E} and H111^{4E} are central to the U-shaped arrangement of the *Dm* Mxt peptide. These residues contact both of the linker regions of Mxt (Fig. 5A–C). Specifically, the nitrogens of the H111^{4E} imidazole ring are in hydrogen-bonding distance to the main chain carbonyl oxygens of W596$^{\text{Mxt}}$ (in the noncanonical linker) and N615$^{\text{Mxt}}$ (in the auxiliary linker), stabilizing the linkers above the lateral surface of eIF4E (Fig. 5C). Furthermore, the side chain of H111^{4E} forms van der Waals contacts with the side chains of L598$^{\text{Mxt}}$ and W602$^{\text{Mxt}}$ (Fig. 5A,C).

Near the dorsal surface of eIF4E, the side chain of N110^{4E} coordinates the backbone of the noncanonical linker and auxiliary linker in a similar fashion. Its amide group contacts the carbonyl oxygen of S595$^{\text{Mxt}}$ in the elbow loop and the backbone nitrogen of F625$^{\text{Mxt}}$ near the auxiliary helix (Fig. 5A,B). Additional stabilization of both of the linkers is probably achieved by van der Waals contacts between W596$^{\text{Mxt}}$ (noncanonical linker) and R624$^{\text{Mxt}}$ (auxiliary linker) (Fig. 5A,B).

The auxiliary helix of Dm *Mxt*

The auxiliary α helix α3 of *Dm* Mxt is stabilized by interactions with eIF4E and runs anti-parallel to the canonical α helix (Fig. 5D,E). The aromatic ring of F625$^{\text{Mxt}}$ (immediately before α helix α3) covers a hydrophobic patch on the eIF4E surface that is formed by W106, L163, and Y109 (Fig. 5D). Residue Y630$^{\text{Mxt}}$ locks the center of the auxiliary α helix to this patch and forms a hydrogen bond with D164^{4E} (Fig. 5D). In addition, Y630$^{\text{Mxt}}$ contacts L587$^{\text{Mxt}}$ on the canonical α helix α1, extending a hydrophobic interface between the two helices that also includes L631$^{\text{Mxt}}$ and forms a small hydrophobic core together with the patch on the eIF4E surface (Fig. 5D).

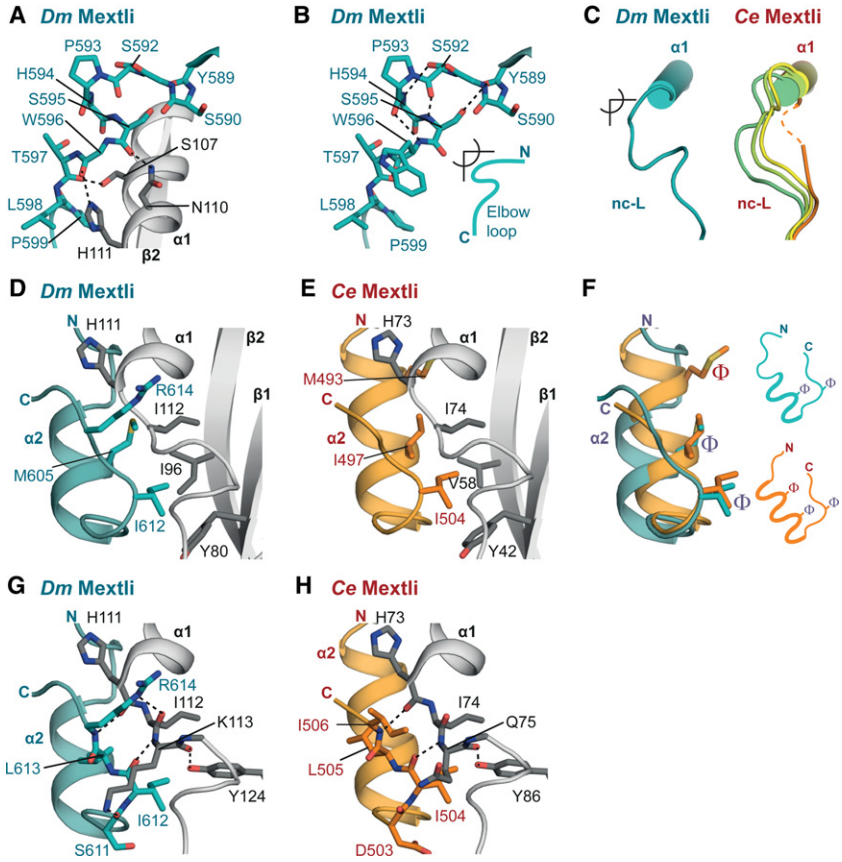


Figure 4. The noncanonical linker and helix. (*A*,*B*) Close-up views of the elbow loop of *Dm* Mxt showing interactions within the elbow loop and eIF4E. Residues within the elbow loop of Mxt and the interface residues of eIF4E are shown as sticks. The side chains of $Y589^{Mxt}$, $H594^{Mxt}$, and $W596^{Mxt}$ and eIF4E secondary structure elements α1 and β2 (in *B*) were omitted for clarity. (*C*) Schematic representations of the noncanonical linker (nc-L) of *Dm* and *Ce* Mxt highlighting the different conformations of the *Ce* noncanonical region. (*D*,*E*) Close-up view of the *Dm* and *Ce* Mxt noncanonical helices bound to the lateral hydrophobic pocket of eIF4E. Selected residues mediating the interactions are shown in gray sticks for eIF4E and teal or orange sticks for *Dm* or *Ce* Mxt, respectively. (*F*) Overlay and schematic representation of the noncanonical helices of *Dm* Mxt (in teal) and *Ce* Mxt (in orange). Common structural features are labeled in purple. The hydrophobic residues interacting with the lateral surface of eIF4E are shown as sticks and are labeled with Φ. (*G*,*H*) Close-up views of interactions between the Mxt noncanonical helices and residues lining the lateral hydrophobic pocket of eIF4E. For clarity, the secondary structure elements β1 and β2 of eIF4E were omitted.

The interactions formed by the *Dm* Mxt auxiliary helix are partially equivalent to the interactions established by the long aliphatic Arg side chains in the canonical α helix of *Ce* Mxt (Fig. 5E). For example, the interaction between *Dm* Mxt residue Y630 and *Dm* eIF4E D164 is equivalent to the interaction between *Ce* Mxt R475 and *Ce* eIF4E E130. Furthermore, *Dm* Mxt F625 shields the surface of eIF4E covered by *Ce* Mxt R482 (Fig. 5E).

Consequently, although the tripartite eIF4E-binding sequence of *Dm* Mxt is longer than the bipartite binding sequence of *Ce* Mxt and other 4E-BPs, the buried surface area on eIF4E is not significantly increased in the *Dm* eIF4E–Mxt complex (1498 $Å^2$) compared with the *Ce* eIF4E–Mxt complex (1370 $Å^2$) or other eIF4E–4E-BP complexes reported previously (Supplemental Table S3; Supplemental Fig. S6; Peter et al. 2015). In particular, the canonical helix of *Dm* Mxt covers a rather small area on the dorsal surface of eIF4E compared with the canonical helices of *Ce* Mxt. However, the auxiliary helix of *Dm* Mxt compensates for the missing binding surface (Supplemental Table S3; Supplemental Fig. S6).

Validation of the eIF4E–Mxt interface

To validate the observed interactions, we substituted key interface residues and tested for complex formation with in vitro pull-down and coimmunoprecipitation assays. In the tripartite *Dm* Mxt peptide (C + NC + α3), substitutions of residues in the canonical (C*), noncanonical (NC*), or auxiliary (α3*) motifs or in the linker regions did not prevent complex formation in vitro, indicating that the remaining sequences were sufficient for binding to eIF4E (Supplemental Fig. S7A, lanes 13–16,18). The association of *Dm* Mxt with eIF4E was prevented only when mutations in the canonical and noncanonical motifs were combined (Supplemental Fig. S7A, lane 17). However, in the bipartite *Dm* Mxt fragment (C + NC), mutations in either the canonical or the noncanonical motif strongly reduced eIF4E-binding (Supplemental Fig. S7A, cf. lanes 20, 21 and 14,16 respectively). Thus, in the tripartite peptide, the auxiliary sequences can compensate for the deleterious effects of the mutations in the canonical or noncanonical motifs. Unlike the results obtained in vitro, results from cell lysates (i.e., in the presence of eIF4G and other 4E-BPs) show that mutations in the canonical or noncanonical motifs abolished the interaction of full-length *Dm* Mxt with endogenous eIF4E; mutations in the linkers and auxiliary helix reduced binding (Supplemental Fig. S7B).

We also analyzed the impact on complex formation of amino acid substitutions in the *Dm* eIF4E residues N110 and H111. As described above, these residues coordinate specific interactions with the *Dm* Mxt linker regions (Fig. 5A–C). Mutations in these residues (NH-EE mutant) abolished binding between full-length eIF4E and full-length Mxt in S2 cells (Supplemental Fig. S7C). In contrast, these mutations reduced but did not eliminate binding to full-length Thor or 4E-T (Supplemental Fig. S7D) and did not affect binding to full-length eIF4G or CUP (Supplemental

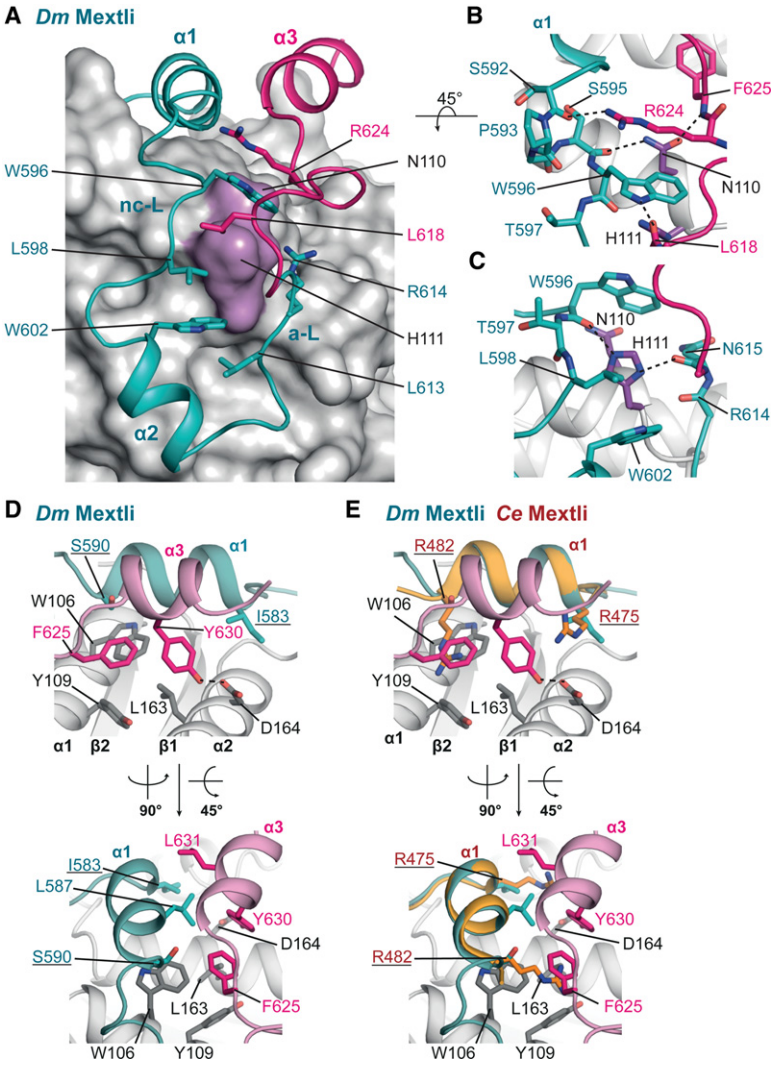


Figure 5. The auxiliary linker and auxiliary helix of *Dm* Mxt. (*A*) Close-up view of the U-shaped arrangement of the *Dm* Mxt peptide and the interactions determining the arrangement of the peptide path. The surface of eIF4E is shown in gray. The positions of the conserved eIF4E residues N110 and H111 are highlighted in purple, and selected residues involved in the arrangement of *Dm* Mxt around the surface of eIF4E are shown as teal and magenta sticks. (*B*,*C*) The eIF4E residues interacting with the Mxt linkers (noncanonical linker [nc-L] and auxiliary linker [a-L]). The side chains of $H594^{Mxt}$ (*B*) and $R614^{Mxt}$ (*C*) were removed for clarity. (*D*) Close-up view of the auxiliary helix of *Dm* Mxt bound to the dorsal surface of eIF4E in two orientations. Selected residues are shown in gray sticks for eIF4E and teal or magenta sticks for *Dm* Mxt. (*E*) Overlay of the canonical helices of *Dm* and *Ce* Mxt proteins reveals equivalent interactions between the Arg residues flanking the canonical helix of *Ce* Mxt and the auxiliary helix of *Dm* Mxt. Selected residues are shown as colored sticks. The Arg residues in the canonical motif of *Ce* Mxt are underlined.

Fig. S7E). These results are consistent with *Dm* Mxt using a different binding mode to interact with eIF4E.

Unlike the results obtained for the bipartite fragment of *Dm* Mxt, mutations in the noncanonical motif of *Ce* Mxt reduced but did not abolish binding to *Ce* eIF4E in vitro (Supplemental Fig. S7F). This suggests a stronger contribution of the canonical helix to the interactions with eIF4E and is likely due to the presence of the Arg residues at positions 2 and 9.

The bipartite and tripartite binding modes confer different abilities to compete with eIF4G

To investigate how the bipartite and tripartite binding modes of Mxt proteins affect their ability to compete with eIF4G, we performed competition assays using preassembled eIF4E–eIF4G complexes. Preassembled *Dm* eIF4E–eIF4G complexes were challenged with a twofold molar excess of bipartite and tripartite *Dm* Mxt fragments. The amount of eIF4G bound to eIF4E was determined over time (Fig. 6A,B). The tripartite *Dm* Mxt fragment (C + NC + α3) displaced eIF4G from preassembled eIF4E–eIF4G complexes faster than the bipartite fragment (C + NC) (Fig. 6A,B). The half-life of the eIF4E–eIF4G complexes was 40 min ± 5 min in the presence of the tripartite fragment compared with >60 min for the bipartite fragment (Fig. 6A,B).

Under the same conditions, a bipartite CUP peptide (C + NC) displaced eIF4G more rapidly, resulting in a half-life of 25 min ± 5 min (Fig. 6A,C). Thus, *Dm* Mxt fragments are less able to displace eIF4G from preassembled eIF4E–eIF4G complexes, although their affinity for eIF4E is either higher than or comparable with that of CUP (Supplemental Table S2; Igreja et al. 2014).

Nevertheless, as shown for CUP and other 4E-BPs, the tripartite *Dm* Mxt peptide requires binding to the lateral surface of eIF4E to displace eIF4G from preassembled eIF4E–eIF4G complexes. The tripartite *Dm* Mxt fragment displaced 80% of eIF4G bound to wild-type eIF4E but failed to displace eIF4G that was prebound to the eIF4E II-AA mutant even after a 180-min incubation (Supplemental Fig. S7G,H).

Next, we analyzed the ability of *Ce* Mxt to compete with preassembled *Ce* eIF4E–eIF4G complexes. Surprisingly,

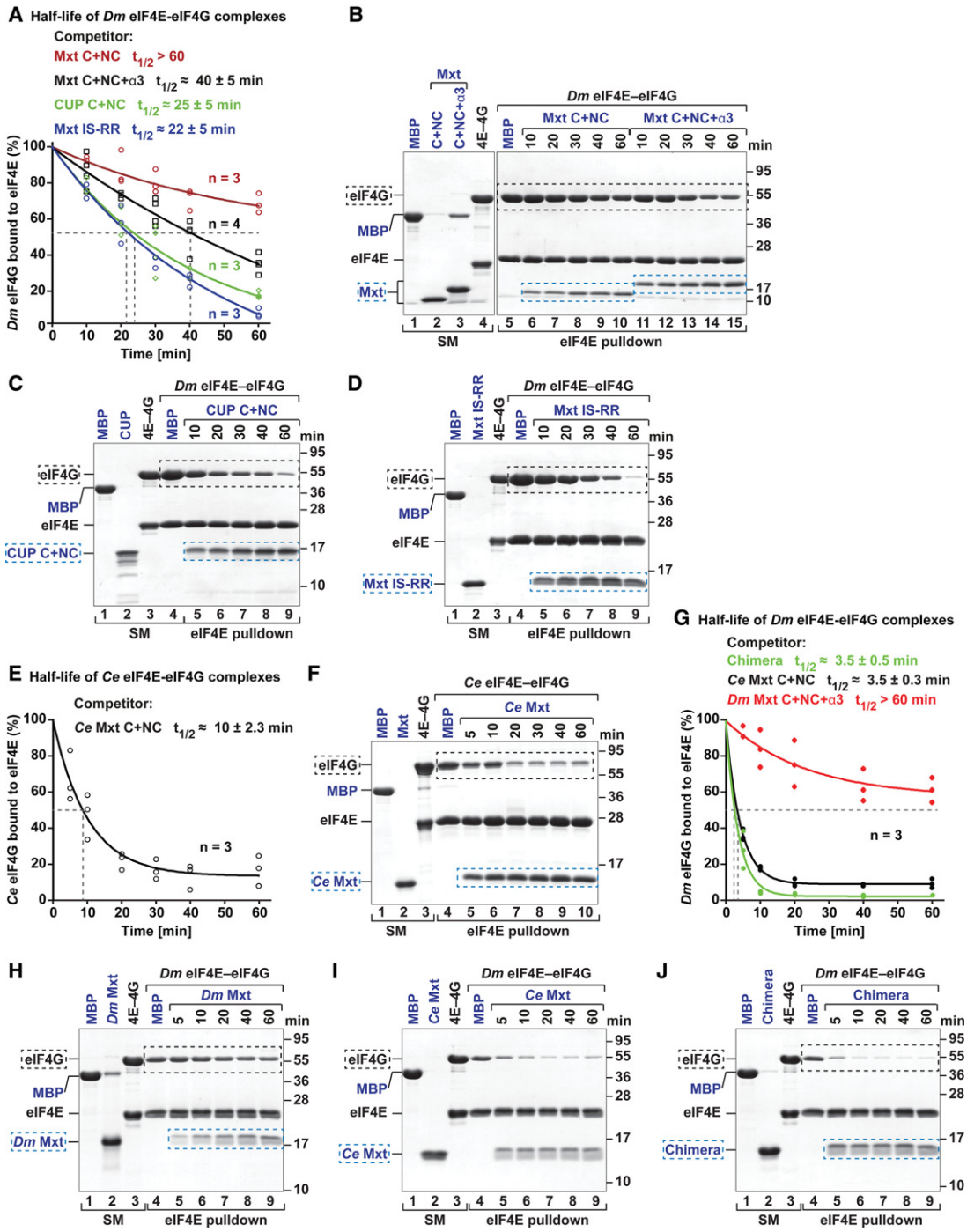


Figure 6. *Ce* Mxt, but not *Dm* Mxt, is a potent eIF4G competitor. (*A–D*) Half-life of *Dm* eIF4E–eIF4G complexes in the presence of a twofold molar excess of the indicated purified and GB1-tagged competitor peptides. MBP served as a negative control. *A* shows data points from three or four independent experiments ($n = 3$ or $n = 4$). *B–D* show representative SDS-PAGE gels for each competition assay. The competitor peptides are labeled in blue, and their positions are highlighted by blue dashed boxes. The black dashed boxes mark the positions of MBP–eIF4G. The lanes labeled SM (starting material) show the purified peptides and complexes used in the competition assay. (*E,F*) Half-life of *Ce* eIF4E–eIF4G complexes in the presence of equimolar amounts of *Ce* Mxt competitor peptide analyzed as described in *A–D*. (*G–J*) Half-life of *Dm* eIF4E–eIF4G complexes in the presence of equimolar amounts of the indicated competitor peptides.

Ce Mxt completely displaced eIF4G after a 5-min incubation when added at twofold molar excess (data not shown). In the presence of equimolar amounts of *Ce* Mxt peptide, the half-life of the *Ce* eIF4E–eIF4G complexes was 10 min (Fig. 6E,F). These results suggest that either *Ce* Mxt is a very efficient eIF4G competitor or the *Ce* eIF4E–

eIF4G complexes are particularly sensitive to competition. We then tested whether *Ce* Mxt also displaced eIF4G from preassembled *Dm* eIF4E–eIF4G complexes, and, indeed, *Ce* Mxt rapidly displaced *Dm* eIF4G (Fig. 6G–J). The half-life of *Dm* eIF4E–eIF4G complexes in the presence of *Ce* Mxt was ~3.5 min. In contrast, under the same conditions, the tripartite *Dm* Mxt peptide slowly competed with *Dm* eIF4G, consistent with Figure 6A,B.

Remarkably, the ability of the *Ce* Mxt peptide to displace *Dm e*IF4G was comparable with that of an engineered chimeric 4E-BP peptide, which contains the canonical motif of *Dm* 4E-T, the linker region of Thor, and the noncanonical helix of CUP (Peter et al. 2015). This peptide is a more efficient eIF4G competitor than the corresponding 4E-BPs (Peter et al. 2015), indicating that *Ce* Mxt is a more potent eIF4G competitor than *Dm* Thor, CUP, or 4E-T.

The bipartite and tripartite binding modes confer differential sensitivity to 4E-BP competition

To gain further insight into the specific properties of the bipartite and tripartite binding modes, we also asked whether eIF4G or the chimeric 4E-BP peptide could dislodge Mxt peptides that were prebound to eIF4E. Complexes containing *Dm* eIF4E bound to bipartite and tripartite *Dm* Mxt fragments were challenged with a fivefold molar excess of eIF4G (residues 578–650) or the chimeric 4E-BP peptide. Proteins bound to eIF4E were monitored by eIF4E pull-down after a 180-min incubation. The eIF4G and chimeric peptides failed to displace the tripartite *Dm* Mxt fragment (Fig. 7A,B, lanes 10–12) but could displace the bipartite *Dm* Mxt fragment (Fig. 7A,B, lanes 7–9). Under the same conditions, the half-life of the *Ce* eIF4E–Mxt complexes in the presence of fivefold molar excess of the chimeric 4E-BP peptide was 25 min (Fig. 7C,D). Thus, the tripartite binding mode of *Dm* Mxt confers resistance to competition by eIF4G and 4E-BPs.

The Arg/Lys residues in the canonical motif are important for competition with eIF4G

Overall, our data indicate that the bipartite *Ce* Mxt peptide is a potent eIF4G competitor, whereas the bipartite *Dm* Mxt peptide only weakly displaces eIF4G. The tripartite *Dm* Mxt peptide shows an intermediate behavior. We hypothesized that the inability of the bipartite *Dm* Mxt peptide to compete with eIF4G may be caused by the absence of Arg residues at positions 2 and 9 in the canonical motif (Fig. 3D); these Arg residues are present in *Ce* Mxt and other 4E-BPs. In the tripartite *Dm* Mxt, the auxiliary helix partially compensates for the lack of Arg residues in the canonical motif, improving the ability of this fragment to compete with eIF4G.

To test this hypothesis, we asked whether the bipartite *Dm* Mxt peptide could be converted into an effective eIF4G competitor if residues $I583^{Mxt}$ and $S590^{Mxt}$ in the canonical motif were replaced by Arg residues (bipartite Mxt IS-RR mutant). Remarkably, the mutations restored the ability of the bipartite *Dm* Mxt peptide to displace *Dm* eIF4G from preassembled eIF4E–eIF4G complexes to the level observed for the bipartite CUP peptide (Fig. 6A–C).

Conversely, eIF4G did not displace the prebound mutated bipartite Mxt peptide from eIF4E but displaced the wild-type bipartite peptide (Fig. 7A,B). In contrast, the chimeric 4E-BP peptide displaced the bipartite Mxt peptide regardless of the mutations, but the mutations extended the half-life of the eIF4E-bipartite Mxt complexes from 14 to 136 min in the presence of the chimeric 4E-BP peptide (Fig. 7A,B,E–G).

In summary, the *Dm* bipartite Mxt peptide behaves as the equivalent peptide from *Ce* Mxt and other 4E-BPs, provided that the extended consensus is restored in the canonical motif with Arg/Lys at positions 2 and 9. In the absence of Arg/Lys residues, the peptide does not efficiently compete with eIF4G for binding to eIF4E and is rapidly displaced by eIF4G and 4E-BPs.

Discussion

Structural evolution of 4E-BPs and the regulation of translation initiation

Unlike many known 4E-BPs, *Dm* Mxt is an unusual 4E-BP in that it is thought to stimulate translation (Hernández et al. 2013). Here we show that Mxt associates with eIF4E using a bipartite binding mode in *C. elegans* and an unprecedented tripartite binding mode in *D. melanogaster*. Our data suggest that Mxt evolved these distinct binding strategies to form complexes with eIF4E that display distinct functional properties. In the case of *Dm* Mxt, the tripartite binding mode compromises the proteins' ability to compete with eIF4G for binding to eIF4E, most likely because *Dm* Mxt has a more complex folding in the bound state. However, once the resulting circularly closed structure is formed, it is more difficult for eIF4G or other 4E-BPs to displace it. As a consequence, the *Dm* eIF4E–Mxt complex is particularly stable and likely helps to maintain *Dm* Mxt function in translation in the presence of other 4E-BPs that would normally displace eIF4G. In contrast, the bipartite binding mode of *Ce* Mxt confers a competitive advantage over eIF4G but is sensitive to competition by other 4E-BPs. Thus, *Ce* Mxt may bind to eIF4E under conditions in which binding of other 4E-BPs is inhibited; for example, by phosphorylation.

It is intriguing that Mxt uses different binding modes in different species, whereas, for example, human 4E-BP1 and its ortholog in *D. melanogaster*, Thor, have almost identical structures (Peter et al. 2015). Our observations suggest that molecular competition for eIF4E binding by the 4E-BPs to regulate translation initiation may represent an important driving force underlying the rapid evolution of the Mxt-binding mode.

The auxiliary sequences coevolved with the substitution of Arg/Lys residues in the canonical helix

The auxiliary sequences of *Dm* Mxt are less conserved than the canonical and noncanonical motifs among Mxt

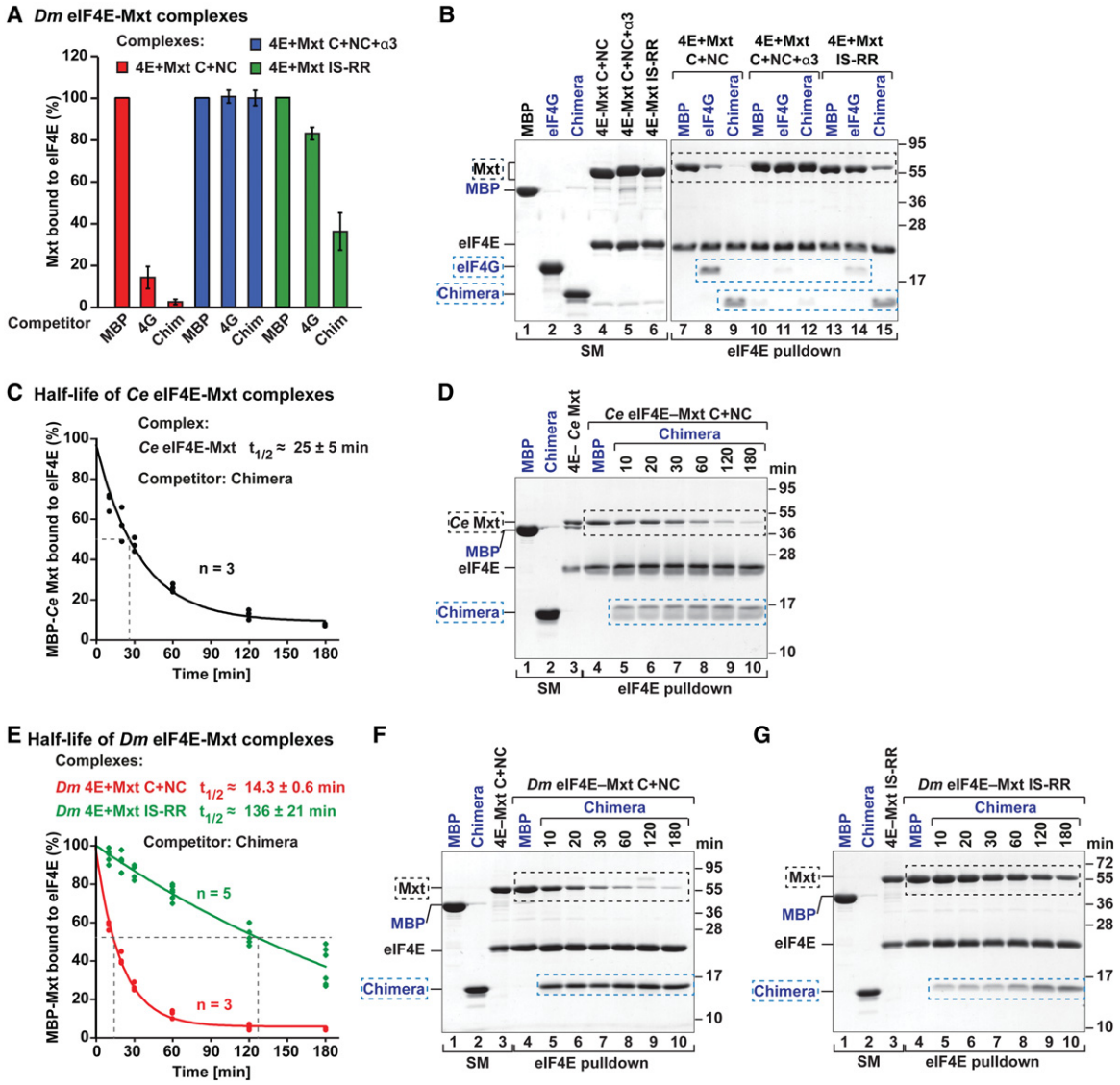


Figure 7. *Dm* eIF4E–Mxt complexes are resistant to 4E-BP competition. (*A*,*B*) The indicated *Dm* eIF4E–Mxt complexes were incubated with a fivefold molar excess of *Dm* eIF4G or chimeric 4E-BP peptides. The eIF4E-bound proteins were pulled down using Ni-NTA beads after a 180-min incubation and analyzed as described in Figure 6, A and B. The competitor proteins are highlighted by blue dashed boxes and are labeled in blue. The positions of the MBP-*Dm* Mxt complexes are marked by black dashed boxes. The lanes labeled SM (starting material) show the purified peptides and complexes used in the competition assay. (*C*,*D*) Half-life of *Ce* eIF4E–Mxt complexes in the presence of a fivefold molar excess of chimeric 4E-BP peptide. (*E*–*G*) Half-life of *Dm* eIF4E–Mxt complexes containing bipartite Mxt peptide (C + NC, wild type, or IS-RR mutant) complexes in the presence of a fivefold molar excess of chimeric 4E-BP peptide. MBP served as a negative control.

orthologs and are absent in *Ce* Mxt as well as in Mxt proteins from other organisms (Supplemental Fig. S2A,B). Sequence analysis indicates that the presence of the auxiliary sequences correlates with the absence of Arg/Lys residues at positions 2 and 9 of the canonical motif (Supplemental Fig. S2A,B; Peter et al. 2015). This may be expected because the auxiliary helix would clash with the side chains of Arg/Lys residues upon binding to the dorsal surface of eIF4E.

The presence of auxiliary sequences in *Dm* Mxt raises the question of whether equivalent sequences have been overlooked in other 4E-BPs. However, as mentioned earlier, auxiliary helices at a position equivalent to that observed in *Dm* Mxt are incompatible with the presence of Arg/Lys residues at positions 2 and 9 of the canonical motif. This suggests that 4E-BPs containing these residues in their canonical motifs are unlikely to bind eIF4E in a manner similar to that of *Dm* Mxt. (Figs. 3A–C, 5D,E). In addition, mutations in *Dm* eIF4E residues N110 and H111 eliminated binding to *Dm* Mxt but not other 4E-BPs or eIF4G, indicating that the binding modes are different. Nevertheless, it is conceivable that auxiliary sequences in known 4E-BPs may contribute to eIF4E binding by contacting other surfaces of eIF4E.

Implications of distinct eIF4E-binding modes for the regulation of complex assembly

The interaction of vertebrate 4E-BP1–3 and *Dm* Thor with eIF4E is regulated by sequential phosphorylation events at Ser/Thr–Pro sites located upstream of the canonical motif and in the elbow loop (Gingras et al. 1999, 2001). These phosphorylation events regulate the folding of the canonical helix and the conformation of the elbow loop (Bah et al. 2015; Peter et al. 2015). In *Dm* Mxt, the conformation of the elbow loop is strikingly similar to that of *Dm* Thor. In addition to its structural similarity to Thor, the elbow loop of *Dm* Mxt also includes a Ser–Pro phosphorylation site at an equivalent structural position ($S592^{Mxt}$ and $P593^{Mxt}$ correspond to $S65^{Thor}$ and $P66^{Thor}$, respectively). Although no post-translational modifications are currently known for *Dm* Mxt at this position, the conservation of this phosphorylation site suggests that similar mechanisms might regulate the association of *Dm* Mxt with eIF4E. The complex tripartite binding mode of *Dm* Mxt and the stability of the complexes formed with eIF4E also suggest that other mechanisms are likely to regulate the formation of these complexes. Similarly, the flexibility of the noncanonical linker in *Ce* Mxt suggests that binding to eIF4E may involve different mechanisms than for vertebrate 4E-BP1-3.

In summary, the variability and evolution in the binding modes and properties of various 4E-BPs indicate that structural variation among other 4E-BPs might cover additional surfaces on eIF4E, conferring unique properties to the complexes and resulting in different regulatory mechanisms. These differences may have important functional implications; for example by specifying the cell type and conditions in which a specific 4E-BP exerts its regulatory role in translation. More generally, the growing repertoire of 4E-BP-binding mechanisms offers new opportunities for the design of eIF4E inhibitors for therapeutic applications.

Materials and methods

DNA constructs

The DNA constructs used in this study are described in the Supplemental Material and are listed in Supplemental Table S1. All of the constructs and mutations were confirmed by sequencing.

Protein expression and purification

All of the recombinant proteins were expressed in *Escherichia coli* BL21 Star (DE3) cells (Invitrogen) grown in LB medium overnight at 20°C. For ^{15}N-labeling of the GB1-stabilized Mxt peptide (residues 577–640), cells were grown in M9 minimal medium with ammonium–^{15}N chloride (Sigma-Aldrich) as a nitrogen source. The cells were lysed by sonication in lysis buffer containing 50 mM HEPES (pH 7.2), 300 mM NaCl, and 2 mM DTT supplemented with 5 µg/mL DNase I, 1 mg/mL lysozyme, and protease inhibitor cocktail (Roche). To purify the eIF4E–Mxt complexes, His_6-tagged *Dm* eIF4E (residues 69–248) and *Ce* eIF4E3 (residues 30–215) were coexpressed with the maltose-binding protein (MBP)-tagged *Dm* Mxt (residues 577–640) and *Ce* Mxt (residues 471–507), respectively. The complexes were purified from cleared cell lysates using amylose resin (New England Biolabs) followed by removal of the MBP and His_6 tags with HRV3C protease cleavage overnight at 4°C. After cleavage of the tags, the complexes were separated from free MBP and His_6 using a heparin column (5 mL of HiTrap Heparin HP; GE Healthcare) and further purified on a Superdex 75 column (GE Healthcare). Purified proteins were stored at −80°C in a buffer containing 10 mM HEPES (pH 7.2), 200 mM NaCl, 2 mM DTT, and 5% glycerol.

For the ITC measurements and competition assays shown in Figures 6 and 7 and Supplemental Figures S3 and S7, the GB1-stabilized peptides (Mxt, eIF4G, and chimeric 4E-BP) were purified as described previously, with the exception that the GB1-stabilized *Dm* Mxt peptide (residues 577–640) was subjected to an additional round of anion exchange chromatography (5 mL of HiTrap Q HP; GE Healthcare) to remove residual MBP (Igreja et al. 2014; Peter et al. 2015). The ^{15}N-labeled GB1-stabilized Mxt peptide (residues 577–640) used for the NMR experiments was purified using the same procedure. The *Dm* eIF4E construct (residues 69–248) used for ITC measurements and the full-length *Dm* eIF4E construct used for NMR were purified as previously described (Igreja et al. 2014). The complexes of His_6-tagged *Dm* eIF4E (residues 69–248) with MBP-tagged *Dm* eIF4G (residues 578–650) or of His_6-tagged *Ce* eIF4E (full-length) with MBP-tagged *Ce* eIF4G (residues 315–491) that were used for the competition assays were purified as previously described (Igreja et al. 2014).

To obtain the eIF4E–Mxt complexes used in the competition assays shown in Figure 7, His_6-tagged eIF4E (residues 69–248) was coexpressed with MBP-tagged Mxt fragments (residues 577–620, residues 577–640, or the 577–620 IS-RR mutant) that were C-terminally fused to GB1. The complexes were purified from cleared cell lysates using amylose resin (New England Biolabs). For the competition assays shown in Supplemental Figure S7, the complexes containing GST-tagged *Dm* eIF4G (residues 578–650) and SHN-tagged *Dm* eIF4E (full-length; either wild type or II-AA mutant) were expressed and purified as previously described (Igreja et al. 2014). The SHN tag consists of a streptavidin-binding peptide (strep), His_6, and the NusA protein.

Crystallization, data collection, and structure determination

A detailed description of the crystallization conditions and the structure determination process are included in the Supplemental Material. All diffraction data sets were recorded on a Pilatus 6M detector at the PXII beamline of the Swiss Light Source at a temperature of 100 K. The diffraction data and refinement statistics are summarized in Table 1.

Coimmunoprecipitation assays and Western blotting

Coimmunoprecipitation assays in S2 cells and Western blotting were performed as described previously (Igreja et al. 2014). The pull-down assay using m^7GTP beads (Jena Biosciences, AC-155) was performed as previously described (Igreja et al. 2014). All Western blots were developed using the ECL Western blotting detection system (GE Healthcare). The antibodies used in this study are listed in Supplemental Table S4.

Pull-down experiments, ITC, and NMR analysis

The in vitro pull-down assays were performed as previously described (Igreja et al. 2014; Peter et al. 2015). The ITC and NMR measurements are described in the Supplemental Material.

Competition assays

The competition assays were performed as previously described (Igreja et al. 2014; Peter et al. 2015). For the competition assays shown in Figure. 6, A–D and G–J, purified *Dm* eIF4E–eIF4G complexes containing His_6-eIF4E (residues 69–248) and MBP-eIF4G (residues 578–650) were incubated with purified and GB1-tagged 4E-BP peptides, including CUP C + NC, *Dm* Mxt C + NC, *Dm* Mxt C + NC + α3, *Dm* Mxt C + NC IS-RR mutant, *Ce* Mxt C + NC, and 4E-BP chimera. MBP served as a negative control. The eIF4E-bound proteins were pulled down using Ni-NTA beads at the indicated time points and eluted with imidazole for analysis by SDS-PAGE. The amount of eIF4G bound to eIF4E was quantified and normalized to the levels of eIF4E present at each time point. These values were set to 100 in the presence of MBP. Data points from three or four independent experiments were plotted, and the resulting curves were determined using the Levenberg-Marquardt algorithm (exponential decay).

For the experiments shown in Figures 6, E and F, and 7, C and D, purified *Ce* eIF4E–eIF4G complexes containing His_6-*Ce* eIF4E (residues 1–215) and MBP-*Ce* eIF4G (residues 315–491) were incubated with purified GB1-tagged *Ce* Mxt peptide (residues 471–507) or chimeric 4E-BP peptide. MBP served as a negative control. The amount of MBP-*Ce* eIF4G bound to eIF4E at each time point was determined as described above. For the experiment shown in Figure 7, A and B, purified *Dm* eIF4E–Mxt (C + NC, C + NC + α3, or IS-RR) complexes were incubated with a fivefold molar excess of *Dm* eIF4G (residues 578–650) or chimeric peptides that were C-terminally fused to GB1. In the experiment shown in Figure 7, E–G, purified *Dm* eIF4E–Mxt (C + NC wild type or IS-RR mutant) complexes were incubated with a fivefold molar excess of chimeric peptide C-terminally fused to GB1. The eIF4E-bound proteins were pulled down using Ni-NTA beads and analyzed by SDS-PAGE as described above. The amount of competitor used in each experiment is indicated in the figure legends.

Accession numbers

Coordinates for the structures described in this study have been deposited in the Protein Data Bank under accession numbers 5ABU (*Dm* eIF4E–Mxt complex with cap analog), 5ABV (*Dm* eIF4E–Mxt), 5ABX (*Ce* eIF4E–Mxt complex with cap analog), and 5ABY (*Ce* eIF4E–Mxt complex).

Acknowledgments

We are grateful to F. Bono for providing the His-eIF4E and GST-CUP 311–440 constructs, and P. Lasko for kindly providing anti-4E-T antibodies. We thank R. Büttner and T. Raisch for setting up the crystallization screens, E. Valkov for advice in the crystal structure refinement, C. Weiler for technical assistance, and the staff at the PX beamline of the Swiss Light Source for assistance with data collection. This work was supported by the Max Planck Society and the Gottfried Wilhelm Leibniz Program of the Deutsche Forschungsgemeinschaft (DFG; awarded to E.I.).

References

Bah A, Vernon RM, Siddiqui Z, Krzeminski M, Muhandiram R, Zhao C, Sonenberg N, Kay LE, Forman-Kay JD. 2015. Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. *Nature* **519:** 106–109.

Banko JL, Poulin F, Hou L, DeMaria CT, Sonenberg N, Klann E. 2005. The translation repressor 4E-BP2 is critical for eIF4F complex formation, synaptic plasticity, and memory in the hippocampus. *J. Neurosci* **25:** 9581–9590.

Dowling RJ, Topisirovic I, Alain T, Bidinosti M, Fonseca BD, Petroulakis E, Wang X, Larsson O, Selvaraj A, Liu Y, et al. 2010. mTORC1-mediated cell proliferation, but not cell growth, controlled by the 4E-BPs. *Science* **328:** 1172–1176.

Gingras AC, Gygi SP, Raught B, Polakiewicz RD, Abraham RT, Hoekstra MF, Aebersold R, Sonenberg N. 1999. Regulation of 4E-BP1 phosphorylation: a novel two-step mechanism. *Genes Dev* **13:** 1422–1437.

Gingras AC, Raught B, Gygi SP, Niedzwiecka A, Miron M, Burley SK, Polakiewicz RD, Wyslouch-Cieszynska A, Aebersold R, Sonenberg N. 2001. Hierarchical phosphorylation of the translation inhibitor 4E-BP1. *Genes Dev* **15:** 2852–2864.

Gkogkas CG, Khoutorsky A, Ran I, Rampakakis E, Nevarko T, Weatherill DB, Vasuta C, Yee S, Truitt M, Dallaire P, et al. 2013. Autism-related deficits via dysregulated eIF4E-dependent translational control. *Nature* **493:** 371–377.

Gosselin P, Oulhen N, Jam M, Ronzca J, Cormier P, Czjzek M, Cosson B. 2011. The translational repressor 4E-BP called to order by eIF4E: new structural insights by SAXS. *Nucleic Acids Res* **39:** 3496–3503.

Gross JD, Moerke NJ, von der Haar T, Lugovskoy AA, Sachs AB, McCarthy JE, Wagner G. 2003. Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E. *Cell* **115:** 739–750.

Hernández G, Miron M, Han H, Liu N, Magescas J, Tettweiler G, Frank F, Siddiqui N, Sonenberg N, Lasko P. 2013. Mextli is a novel eukaryotic translation initiation factor 4E-binding protein that promotes translation in *Drosophila melanogaster*. *Mol Cell Biol* **33:** 2854–2864.

Igreja C, Peter D, Weiler C, Izaurralde E. 2014. 4E-BPs require non-canonical 4E-binding motifs and a lateral surface of eIF4E to repress translation. *Nat Commun* **5:** 4790.

Jackson RJ, Hellen CU, Pestova TV. 2010. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol* **11:** 113–127.

Kinkelin K, Veith K, Grunwald M, Bono F. 2012. Crystal structure of a minimal eIF4E–Cup complex reveals a general mechanism of eIF4E regulation in translational repression. *RNA* **18:** 1624–1634.

Kong J, Lasko P. 2012. Translational control in cellular and developmental processes. *Nat Rev Genet* **13:** 383–394.

Lukhele S, Bah A, Lin H, Sonenberg N, Forman-Kay JD. 2013. Interaction of the eukaryotic initiation factor 4E with 4E-BP2 at a dynamic bipartite interface. *Structure* **21:** 2186–2196.

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

Marcotrigiano J, Gingras AC, Sonenberg N, Burley SK. 1999. Cap-dependent translation initiation in eukaryotes is regulated by a molecular mimic of eIF4G. *Mol Cell* **3:** 707–716.

Marcotrigiano J, Lomakin IB, Sonenberg N, Pestova TV, Hellen CU, Burley SK. 2001. A conserved HEAT domain within eIF4G directs assembly of the translation initiation machinery. *Mol Cell* **7:** 193–203.

Martineau Y, Azar R, Bousquet C, Pyronnet S. 2013. Anti-oncogenic potential of the eIF4E-binding proteins. *Oncogene* **32:** 671–677.

Matsuo H, Li H, McGuire AM, Fletcher CM, Gingras AC, Sonenberg N, Wagner G. 1997. Structure of translation factor eIF4E

bound to m^7GDP and interaction with 4E-binding protein. *Nat Struct Biol* **4:** 717–724.

Mizuno A, In Y, Fujita Y, Abiko F, Miyagawa H, Kitamura K, Tomoo K, Ishida T. 2008. Importance of C-terminal flexible region of 4E-binding protein in binding with eukaryotic initiation factor 4E. *FEBS Lett* **582:** 3439–3444.

Paku KS, Umenaga Y, Usui T, Fukuyo A, Mizuno A, In Y, Ishida T, Tomoo K. 2012. A conserved motif within the flexible C-terminus of the translational regulator 4E-BP is required for tight binding to the mRNA cap-binding protein eIF4E. *Biochem J* **441:** 237–245.

Peter D, Igreja C, Weber R, Wohlbold L, Weiler C, Ebertsch L, Weichenrieder O, Izaurralde E. 2015. Molecular architecture of 4E-BP translational inhibitors bound to eIF4E. *Mol Cell* **57:** 1074–87.

Umenaga Y, Paku KS, In Y, Ishida T, Tomoo K. 2011. Identification and function of the second eIF4E-binding region in N-terminal domain of eIF4G: comparison with eIF4E-binding protein. *Biochem Biophys Res Commun* **414:** 462–467.

Volpon L, Osborne MJ, Topisirovic I, Siddiqui N, Borden KL. 2006. Cap-free structure of eIF4E suggests a basis for conformational regulation by its ligands. *EMBO J* **25:** 5138–5149.

Mextli proteins use both canonical bipartite and novel tripartite binding modes to form eIF4E complexes that display differential sensitivity to 4E-BP regulation

Daniel Peter, Ramona Weber, Carolin Köne, et al.

Genes Dev. 2015 29: 1835-1849 originally published online August 20, 2015
Access the most recent version at doi:10.1101/gad.269068.115

Supplemental Materials

Mextli proteins use both canonical bipartite and novel tripartite binding modes to form eIF4E complexes that display differential sensitivity to 4E-BP regulation

Daniel Peter, Ramona Weber, Carolin Köne, Min-Yi Chung, Linda Ebertsch, Vincent Truffault, Oliver Weichenrieder, Cátia Igreja and Elisa Izaurralde

DNA constructs

The plasmids used for the expression of *Dm* eIF4E, eIF4G, CUP (full-length or fragments) and chimeric 4E-BP in *Escherichia coli* (*E. coli*) or *Dm* S2 cells have been described previously (Igreja and Izaurralde 2011; Peter et al. 2015). The plasmids for the expression of HA- or GFP-tagged *Dm* Mxt in S2 cells were obtained by inserting the corresponding cDNA into the EcoRI and ApaI sites of the pAc5.1B-λN-HA and pAc5.1B-EGFP vectors. For expression in *E. coli*, DNA fragments encoding *Dm* Mxt residues 577–620 (C+NC, bipartite) and 577–640 (C+NC+α3, tripartite) were inserted into the XhoI and BamHI restriction sites of the pnEA-NpM vector (Diebold et al. 2011), producing an N-terminally fused MBP cleavable by HRV3C protease. A cDNA fragment encoding the *Ce* Mxt eIF4E-binding region [residues 471–507 (C+NC)] was inserted into the NdeI and NheI restriction sites of the pnEA-NpM vector. The cDNA encoding full length *Ce* eIF4E3 and fragments thereof (residues 30–215, trunc 1; and residues 1–215, trunc 2) were inserted into the NdeI and XhoI restriction sites of the pnYC-NpH vector, generating N-terminal fusions with the hexahistidine (His_6) tag cleavable by HRV3C protease. The cDNA fragment coding for the eIF4E-binding region of *Ce* eIF4G (residues 315–491) was inserted into the NdeI and XhoI restriction sites of the pnEA-NpM vector. The cDNA encoding the B1 domain of immunoglobulin-binding protein G (GB1) (Cheng and Patel 2004) was inserted at the C-terminal end of the Mxt fragments by site-directed mutagenesis. All of the mutants used in this study were generated by site-directed mutagenesis using the QuikChange mutagenesis kit (Stratagene) and are listed in Supplemental Table S1.

NMR measurements

All spectra were recorded at 298 K. The ^{15}N-HSQC experiment for the GB1-stabilized Mxt peptide (residues 577–640) was recorded at 600 MHz on a Bruker AVIII-600 spectrometer. The ^{15}N-TROSY experiments for the eIF4E–Mxt-GB1 complex were measured at 800 MHz on a Bruker AVIII-800 spectrometer. The ^{15}N-labeled Mxt-GB1 protein was saturated with a 1.5-fold molar excess of unlabeled full-length *Dm* eIF4E. The molecular weight of the proteins used in the NMR-experiments was determined based on their tumbling behavior in solution. Both diffusion experiments (before and after the addition of full length *Dm* eIF4E to the Mxt-GB1 protein) were recorded at 600 MHz.

Crystallization

Crystals of *Dm* eIF4E (residues 69–248) copurified with a *Dm* Mxt fragment (residues 577–640) and supplemented with a 1.5-fold molar excess of m^{7}GpppG cap analog (New England BioLabs) were obtained by the hanging-drop vapor-diffusion method at 18°C. Crystals appeared one day after mixing the protein solution (12 mg/ml) with a reservoir solution containing 0.1 M MMT buffer and 23% PEG 1500. MMT buffer consists of a 1:2:2 molar ratio of DL-malic acid, 2-(N-morpholino)ethanesulfonic acid (MES), and Tris (pH 4.0), according to the PACT Suite Buffer protocols (Qiagen). Crystals of *Dm* eIF4E (residues 69–248) in complex with *Dm* Mxt (residues 577–640; 20 mg/ml) without the cap analog were obtained by the same method. Crystals appeared in one day over a reservoir of 0.1 M Bis-Tris (pH 6.0) and 23% PEG 3350. The complex of *Ce* eIF4E (residues 30–215), *Ce* Mxt (residues 471–507) and the m^{7}GpppG cap analog (New England BioLabs) was crystallized using the same method. Crystals appeared after two to three days over a

reservoir of 0.1 M sodium acetate (pH 5.0), 0.01 M $ZnCl_2$ and 18% PEG 6000. The same *Ce* complex but without the cap-analog was crystallized over a reservoir of 0.1 M Tris (pH 8.0), 0.2 M $MgCl_2$ and 17% PEG 6000, with crystals appearing after one day. For cryoprotection, crystals were soaked briefly in the reservoir solution supplemented with 10–15% glycerol followed by flash-freezing in liquid nitrogen.

Structure determination

Diffraction data were collected at a wavelength of 1.000 Å on a PILATUS 6M detector at the PXII beamline of the Swiss Light Source. Data were processed with XDS and scaled using XSCALE (Kabsch 2010).

Crystals of the *Dm* eIF4E-Mxt complex and the cap analog diffracted to a resolution of 2.16 Å. The structure was solved by molecular replacement using PHASER (McCoy et al. 2007) with the *Dm* eIF4E structure (PDB-ID 4UE8) (Peter et al. 2015) as the search model. The molecular replacement solution was used to rebuild an initial model of *Dm* eIF4E using the PHENIX AutoBuild wizard (Terwilliger et al. 2008) to prevent model bias. To complete and improve the initial model of eIF4E, iterative cycles of model building and refinement were carried out using COOT (Emsley et al. 2010) and PHENIX (Afonine et al. 2012) against the unmodified native dataset. The Mxt peptide and the m^7GpppG cap analog were then built into the weighted F_o-F_c difference density using COOT, and the structure of the complex was further refined using PHENIX. The stereochemical restraints library for the m^7GpppG cap analog was generated using GRADE (Smart et al. 2011) for subsequent refinement. For the final model, 99.12% of all residues are in the favored regions of the Ramachandran plot and there are no outliers.

Crystals of the *Dm* eIF4E–Mxt complex without the cap analog diffracted to a resolution of 2.13 Å. The crystal belongs to space group $P2_1$ with a β-angle of 90.1° and further data analysis with PHENIX.XTRIAGE (Zwart et al. 2005) indicated the presence of translational pseudosymmetry, generating a strong peak in the Patterson map (37.3% of the origin). Nevertheless, the structure was solved by molecular replacement as described above with the difference being that the asymmetric unit contained four copies of the search model. The structures were rebuilt using the PHENIX AutoBuild wizard (Terwilliger et al. 2008) to prevent model bias. The initial models of eIF4E were then manually completed in COOT (Emsley et al. 2010) and refined using PHENIX (Afonine et al. 2012). The Mxt peptide was built into the weighted F_o-F_c difference density in COOT and the complex was refined using BUSTER with automated NCS restraints and TLS refinement (Bricogne et al. 2011; Smart et al. 2012). For the final model, 97.5% of all residues are in the favored regions of the Ramachandran plot and there are no outliers.

Crystals of the *Ce* eIF4E-Mxt complex without the cap analog diffracted to a resolution of 1.66 Å. Because of bound Zn^{2+} ions from the crystallization condition, there was a strong anomalous signal that was used to derive unbiased experimental phases. Zn^{2+}-sites were identified with SHELX (Sheldrick 2010) and were used to build an initial model of the *Ce* eIF4E-Mxt complex using PHENIX AutoSol (Terwilliger et al. 2009). The model was manually completed and improved in COOT (Emsley et al. 2010) and refined using PHENIX (Afonine et al. 2012). Because the m^7GpppG cap analog was not fully visible, an m^7GTP molecule was placed into the weighted F_o-F_c difference density in COOT, and the Zn^{2+}-sites were confirmed by examining the anomalous difference Fourier map. The final model was refined using

PHENIX and 99% of all residues are in the favored regions of the Ramachandran plot and there are no outliers.

Crystals of the *Ce* eIF4E-Mxt complex with no cap analog diffracted to a resolution of 1.95 Å. The structure was solved by molecular replacement using PHASER (McCoy et al. 2007). The experimentally phased structure of the *Ce* eIF4E-Mxt complex including the cap analog was used as the search model. To prevent model bias, the three complexes found in the asymmetric unit were rebuilt using the PHENIX AutoBuild wizard (Terwilliger et al. 2008). The models were improved in COOT (Emsley et al. 2010) and refined using PHENIX (Afonine et al. 2012). For the final model, 97.5% of all residues are in the favored regions of the Ramachandran plot and there are no outliers.

The stereochemical properties for all of the structures were verified using MOLPROBITY (Chen et al. 2010). Structural images were prepared using PyMOL (http://www.pymol.org). The diffraction data and refinement statistics are summarized in Table 1.

ITC analysis

Dm eIF4E (residues 69–248) used for ITC measurements was purified as described previously (Igreja et al. 2014). The ITC experiments were performed using a VP-ITC microcalorimeter (MicroCal) at 20°C as described previously (Igreja et al. 2014; Peter et al. 2015). The solution of wild-type eIF4E (residues 69–248, 1.0 or 2.0 μM) or eIF4E II-AA mutant (residues 69–248, 9.5 μM) in the calorimetric cell was titrated with tenfold concentrated solutions of GB1-stabilized Mxt peptides (residues 577–620, 20 μM; residues 577–640, 10 μM for eIF4E WT and 95 μM for eIF4E II-AA mutant) that were dissolved in the same buffer (20 mM Na-phosphate pH 7.0 and 150

mM NaCl). The titration experiments consisted of an initial injection of 2 µl followed by 28 injections of 10 µl at 240 s intervals. Each binding experiment was repeated three times. The thermodynamic parameters were calculated using a one-site binding model (Origin version 7.0). The datapoint of the first injection was removed for the analysis (Mizoue and Tellinghuisen 2004).

Supplemental Table S1. Mutants and constructs used in this study.

Protein	Name of the construct	Fragments / mutations	Binding site / motif
***Dm* eIF4E-2** (1–248) (isoform C) P48598-2	4E	Full length	
	II-AA	I96A, I112A	Lateral surface
	W106A	W106A	Dorsal surface
	NH-EE	N110E, H111E	
	trunc	69–248	
***Dm* eIF4G** (1–1666) (isoform A) O61380	4G	578–650	Extended eIF4E-binding region
***Dm* Mxt** (1–653) Q9VR35	Mxt	Full length	
	C+NC+α3	577–640	Tripartite eIF4E-binding region
	C+NC	577–620	Bipartite eIF4E-binding region
	C*	Y581A, L586A, L587A	Canonical motif mutant
	Elbow1*	W596D	Elbow 1oop mutant
	Nc-L*	L598D	Non-canonical linker mutant
	NC1*	W602D	Non-canonical motif mutants
	NC2*	M605D	
	IL-AA	I612A L613A	
	C*+IL-AA	Y581A, L586A, L587A, I612A,L613A	Canonical motif + auxiliary linker mutant
	α3*	F625E	Auxiliary helix mutant
	FY-EE	F625E, Y630E	
	IS-RR	I583R, S590R	Arg mutant
CUP (1–1117) Q9VMA3	C+NC	325–376	eIF4E-binding region
	1–417		CUP N-terminal eIF4E-binding region
***Ce* Mxt** (1–507) Q9XW13	C+NC	471–507	eIF4E-binding region
	C*	Y579A, L584A, M585A	Canonical motif mutant
	NC 1*	I497D, I504D	Non-canonical motif mutants
	NC 2*	M493D, I497D, I504D	
	NC 3*	M493D	
***Ce* eIF4G** (1–1156) Q21531	4G	315–491	Extended eIF4E-binding region
***Ce* eIF4E3** (1–251) O61955	W68A	W68A	Dorsal surface mutant
	VI-AA	V58A, I74A	Lateral surface mutant
	trunc 1	30–215	
	trunc 2	1–215	

Supplemental Table S2. Thermodynamic parameters for the interaction of eIF4E with Mxt peptides.

eIF4E (69–248) wild type					
Mxt peptide	**K_D (M)**	**ΔH (kcal mol^{-1})**	**-TΔS (kcal mol^{-1})**	**ΔG (kcal mol^{-1})**	**Molar ratio**
577–620	$5.0 \pm 3.4 \times 10^{-9}$	-13.1 ± 0.4	1.9	-11.2	0.98 ± 0.02
577–640	$0.52 \pm 0.09 \times 10^{-9}$	-32.4 ± 0.6	19.9	-12.5	1.00 ± 0.03
eIF4E (69–248) II-AA mutant					
577–640	$0.3 \pm 0.12 \times 10^{-6}$	-18.3 ± 1.3	9.5	-8.8	1.00 ± 0.02

Supplemental Table S3. Interface areas of 4E-BP-eIF4E complexes.

The covered surface areas on eIF4E by the indicated 4E-BPs or *Dm* eIF4G and the individual motifs have been calculated using the PISA program from the CCP4 suite.

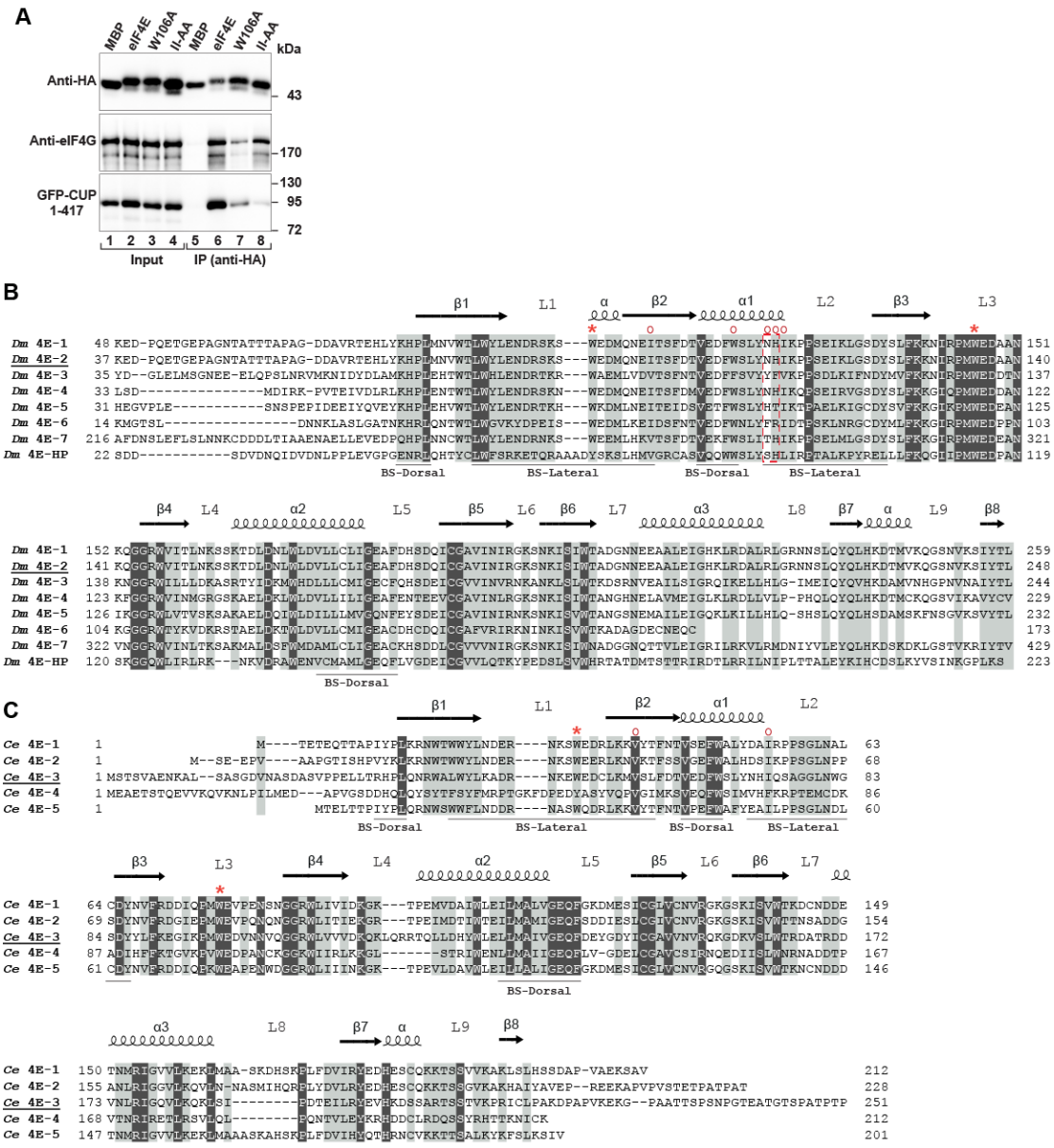
Protein	**motif**	**area (Å^2)**
***Dm* Mxt**	Canonical helix (α1)	471
	Canonical + auxiliary helix (α1+α3)	655
	Complete 4E-binding region	1498
***Ce* Mxt**	Canonical helix (α1)	611
	Complete 4E-binding region	1370
***Dm* Thor**	Canonical helix	632
	Complete 4E-binding region	1450
***Dm* CUP**	Canonical helix	539
***Dm* 4E-T**	Canonical helix	584
	Complete 4E-binding region	1370
***Dm* eIF4G**	Canonical helix	529
***Hs* 4E-BP1**	Canonical helix	550
	Complete 4E-binding region	1235

Supplemental Table S4. Antibodies used in this study.

Antibody	**Source**	**Catalog Number**	**Dilution**	**Monoclonal/ Polyclonal**
Anti-HA-HRP (Western blot)	Roche	12 013 819 001	1:5,000	Monoclonal
Anti-HA (IP)	Covance	MMS-101P	1:1,000	Monoclonal
Anti-*Dm* eIF4E	In house		1:3,000	Rabbit polyclonal
Anti-*Dm* eIF4G	In house		1:3,000	Rabbit polyclonal
Anti-GFP	In house		1:2,000	Rabbit polyclonal
Anti-*Dm* 4E-T	Kindly provided by P. Lasko		1:2,000	Rabbit polyclonal
Anti-rabbit-HRP	GE Healthcare	NA934V	1:10,000	Polyclonal

Supplemental Figure Legends

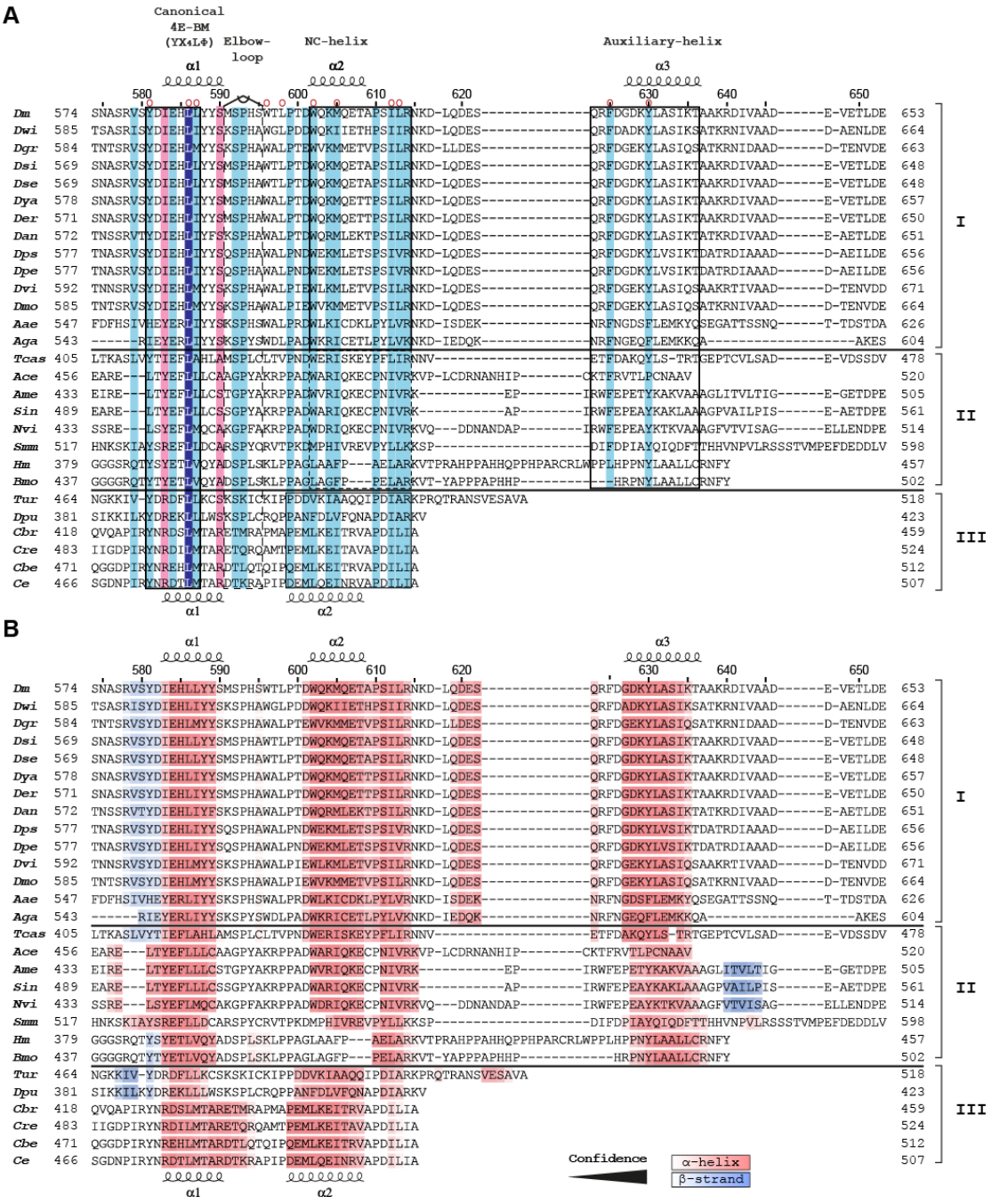
Figure S1



Supplemental Figure S1. Structure-based alignment of eIF4E sequences. (*A*) Western blot showing the interaction of HA-tagged *Dm* eIF4E (either wild-type or mutated) with endogenous eIF4G and GFP-tagged *Dm* CUP (residues 1–417). The inputs (0.6% for eIF4G, 3% for eIF4E and 2% for CUP) and immunoprecipitates (25% for eIF4G and 20% for eIF4E and CUP) were analyzed by western blotting using anti-HA, anti-eIF4G and anti-GFP antibodies. (*B*) Structure-based sequence alignment of eIF4E orthologous proteins from *Drosophila melanogaster* (*Dm*). There

are eight eIF4E isoforms in *D. melanogaster*: 4E-1 (P48598) and 4E-2 (P48598-2; used in this study) are encoded by the same gene and result from alternative splicing, 4E-3 (Q9VSG), 4E-4 (Q9VRY0), 4E-5 (Q9VSB6), 4E-6 (Q9VAR1), 4E-7 (Q9W5B3) and 4E-HP (Q8T3K5) (Hernandez et al. 2005). *Dm* Mxt interacts with eIF4E-1, -2, -3, -4 and -7 in yeast two-hybrid assays (Hernandez et al., 2013). Conserved residues are highlighted with a black background and printed in white. Residues with >70% similarity are shown with a light gray background. Secondary structure elements for *Dm* eIF4E-2 (used in this study) are indicated above the sequences. The Trp residues coordinating the m^7GTP nucleotide are indicated by red asterisks. The lateral and dorsal binding surfaces (BS) are indicated by a line below the sequences. Residues mutated in this study are indicated with red open circles above the sequences. The dashed red box highlights residues N110 and H111 that coordinate interactions with the linker regions in the tripartite binding region of *Dm* Mxt. (C) Structure-based alignment of *Ce* eIF4E sequences. Five eIF4E proteins have been described in *C. elegans:* 4E-1 (IFE-1, O45551), 4E-2 (IFE-2, Q21693), 4E-3 (IFE-3, O61955, used in this study), 4E-4 (4E-HP, Q22888) and 4E-5 (IFE-5, P56570) (Jankowska-Anyszka et al. 1998). Symbols and labels are as in panel (B). Secondary structure elements for *Ce* eIF4E-3 are indicated above the sequences.

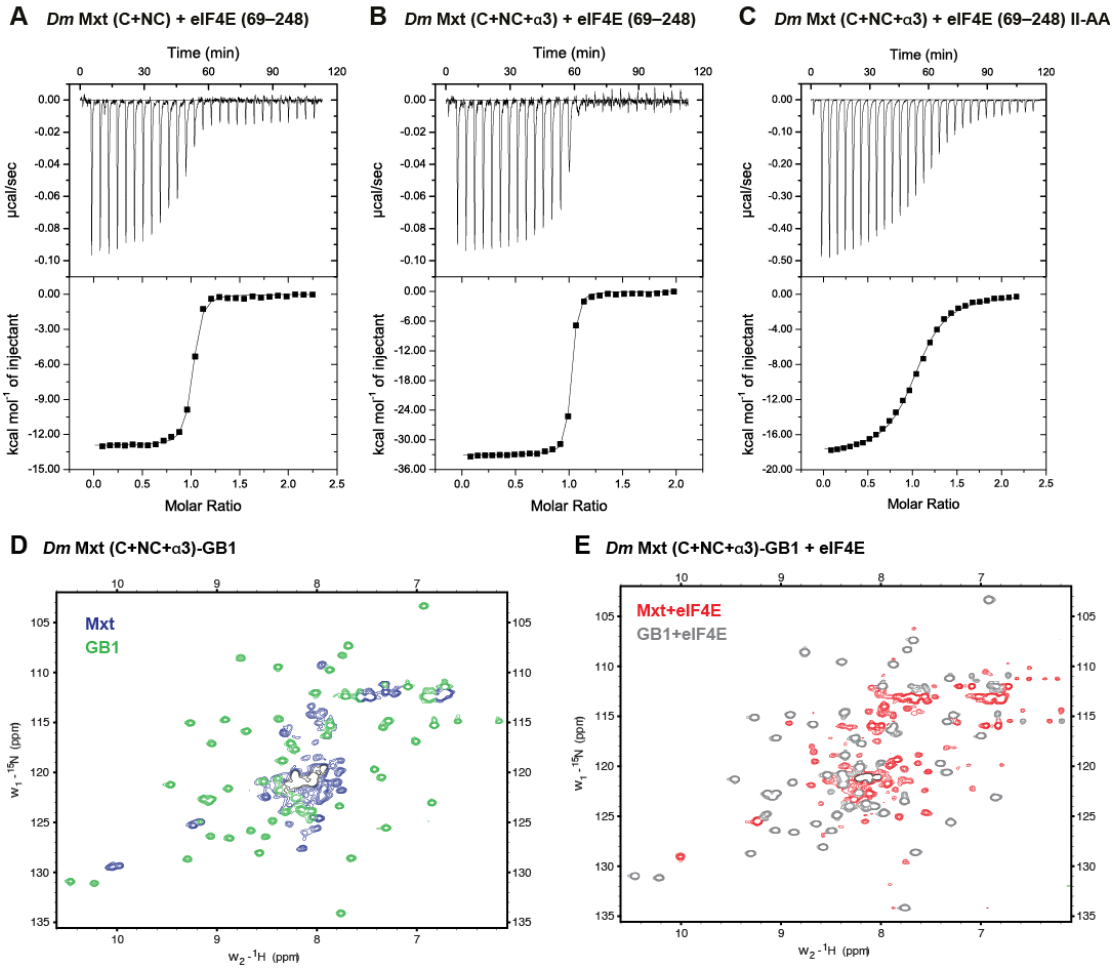
Figure S2



Supplemental Figure S2. Structure-based alignment of the eIF4E-binding region of Mextli proteins. (*A*) Structure-based sequence alignment of Mxt orthologous proteins from *Drosophila* species [*Drosophila melanogaster* (*Dm*), *Drosophila willistoni* (*Dwi*), *Drosophila grimshawi* (*Dgr*), *Drosophila simulans* (*Dsi*), *Drosophila sechellia* (*Dse*), *Drosophila yakuba* (*Dya*), *Drosophila erecta* (*Der*), *Drosophila ananassae* (*Dan*), *Drosophila pseudoobscura* (*Dps*), *Drosophila persimilis* (*Dpe*), *Drosophila*

virilis (*Dvi*), *Drosophila mojavensis* (*Dmo*)], from non-*Drosophila* Dipterans [mosquitos: *Aedes aegypti* (*Aae*), *Anopheles gambiae* (*Aga*)], from non-Dipterans Insects [bee: *Apis mellifera* (*Ame*); wasp: *Nasonia vitripennis* (*Nv*); ants: *Atta cephalotes* (*Ace*), *Solenopsis invicta* (*Sin*); silkmoth: *Bombyx mori* (*Bmo*); human body louse: *Pediculus humanus* (*Hm*); beetle: *Tribolium castaneum* (*Tcas*)], from non-Insects Arthropods [centipede: *Strigamia maritime* (*Smm*); flea: *Daphnia pulex* (*Dpu*); mite: *Tetranychus urticae* (*Tur*)] and from Nematodes [worms: *Caenorhabditis brenneri* (*Cbe*), *Caenorhabditis elegans* (*Ce*), *Caenorhabditis remanei* (*Cre*), *Caenorhabditis briggsae* (*Cbr*)]. Invariant residues are highlighted with a dark-blue background and printed in white. Residues with >70% similarity are shown with a light blue color background. Secondary structure elements are indicated above and below the sequences for the *Dm* and *Ce* proteins, respectively. Residues at positions 2 and 9 of the canonical motifs of Mxt are indicated with a magenta background. The canonical (C), non-canonical (NC) and auxiliary 4E-BMs are boxed in black. Mxt residues in the motifs contacting eIF4E are indicated by red open circles. (*B*) Secondary structure prediction for the eIF4E-binding region of Mxt orthologous proteins using the Ali2D software (http://toolkit.tuebingen.mpg.de/). The degrees of confidence for β-strands (in blue) and α-helices (in red) are indicated by increasing color intensities. Species are the same as in panel (A).

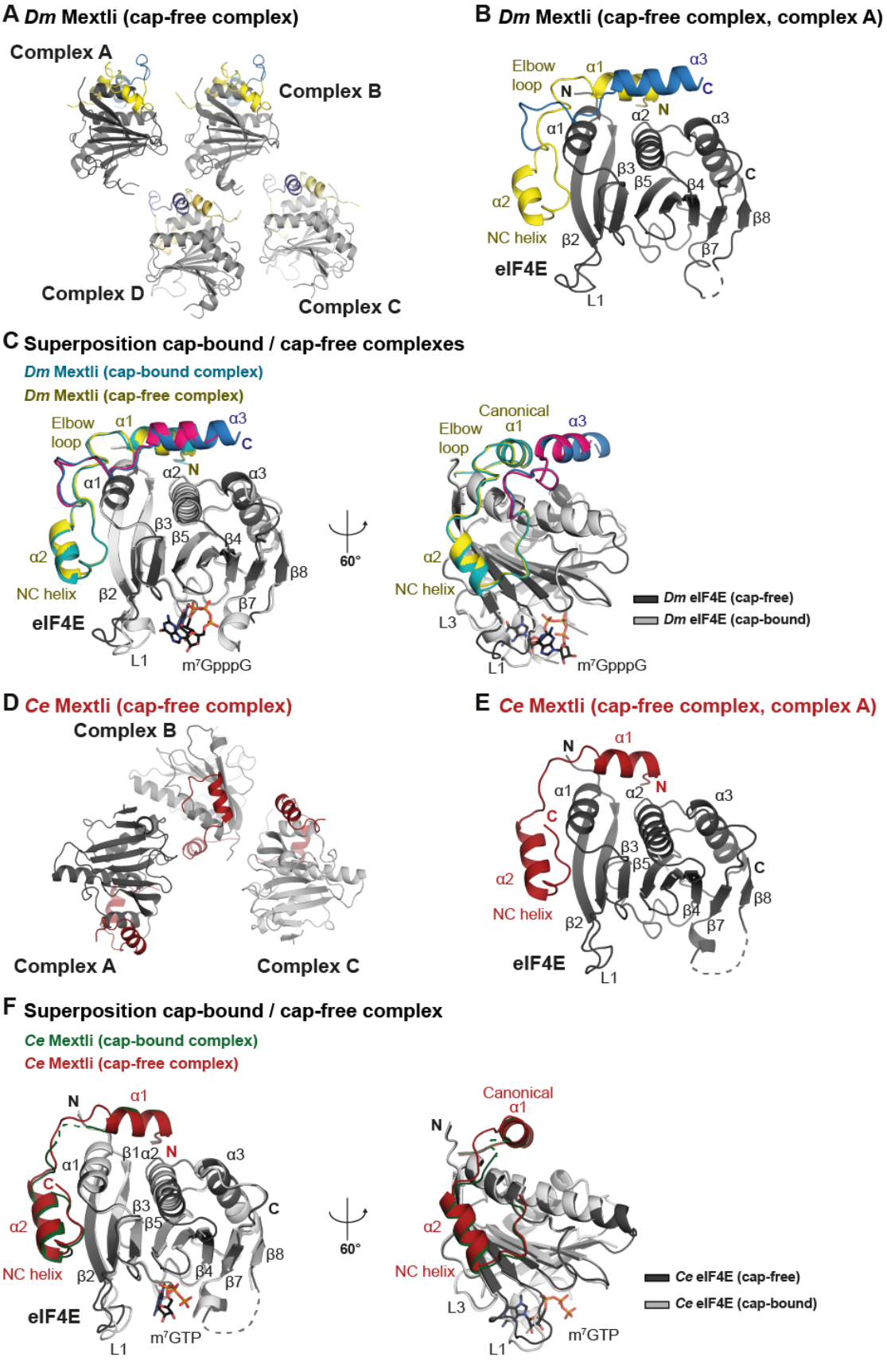
Figure S3



Supplemental Figure S3. The tripartite *Dm* Mxt peptide is unfolded in solution. (*A–C*) Isothermal titration calorimetry data for the interactions of *Dm* eIF4E (residues 69–248, wild type or II-AA mutant) with the indicated *Dm* Mxt peptides. The thermodynamic parameters are shown in Table S2. (*D*) ^{15}N-HSQC spectra of purified ^{15}N-labeled *Dm* Mxt C+NC+α3 peptide C-terminally fused to GB1. The spectrum displays two distinct sets of peaks: a well dispersed spectrum with high intensity peaks corresponding to the folded GB1 protein (green) and a second spectrum with lower intensity and a limited ^{1}H resonance dispersion, which is characteristic of disordered proteins (Mxt; blue). Diffusion experiments indicate that the Mxt-GB1 fusion protein is monomeric in solution (approximately 14 kDa). Black lines indicate overlapping spectra for Mxt and GB1. (*E*) ^{15}N-TROSY spectra of purified ^{15}N-labeled

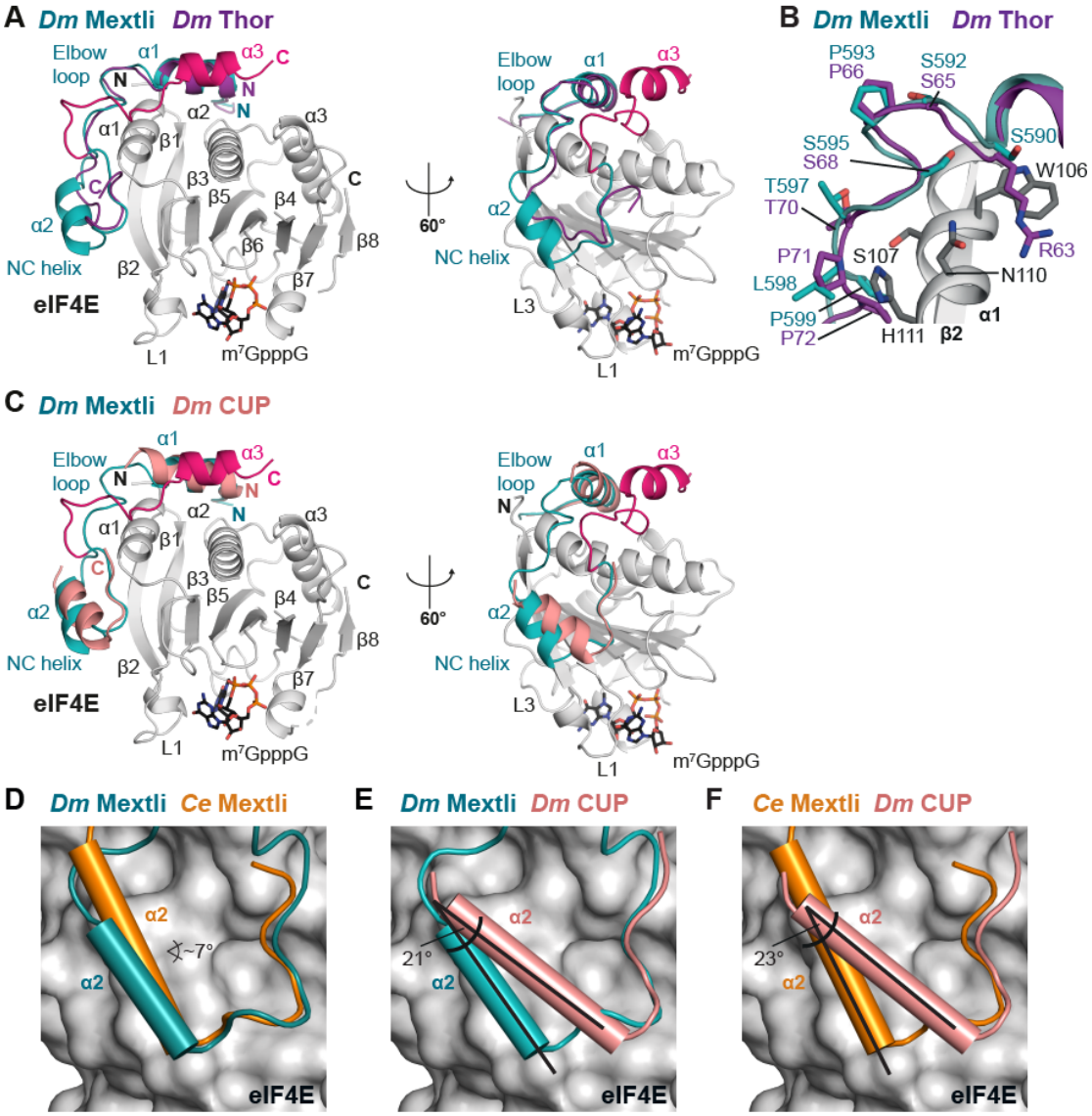
Dm Mxt C+NC+α3 peptide C-terminally fused to GB1 after addition of purified unlabeled *Dm* eIF4E. Two distinct sets of spectra are also identified. In the GB1 protein spectra (gray), the peaks do not show any chemical shift, suggesting that GB1 does not interact with eIF4E. The spectra corresponding to the Mxt peptide (red) gained tremendously in resonance dispersion upon addition of eIF4E. Many peaks left the disordered region of the spectra reflecting a disorder-to-order transition upon binding to eIF4E. The molecular weight of the eIF4E-Mxt-GB1 complex estimated by diffusion experiments corresponds to a 1:1 stoichiometry.

Figure S4


A Dm Mextli (cap-free complex)
Complex A
Complex B
Complex D
Complex C
B Dm Mextli (cap-free complex, complex A)
Elbow loop
α1
α3
N
C
α2
N
α1
α3
β3
β5
β4
C
α2
β8
NC helix
β2
β7
eIF4E
L1
C Superposition cap-bound / cap-free complexes
Dm Mextli (cap-bound complex)
Dm Mextli (cap-free complex)
Elbow loop
α1
α3
C
α2
N
α3
α1
β3
β5
β4
α2
β8
NC helix
β2
β7
eIF4E
L1
m7GpppG
60°
Canonical
Elbow loop
α1
α3
α2
NC helix
L3
L1
m7GpppG
Dm eIF4E (cap-free)
Dm eIF4E (cap-bound)
D Ce Mextli (cap-free complex)
Complex B
Complex A
Complex C
E Ce Mextli (cap-free complex, complex A)
N
α1
α1
α2
N
α3
C
β3
β5
C
α2
β8
NC helix
β2
β4
β7
eIF4E
L1
F Superposition cap-bound / cap-free complex
Ce Mextli (cap-bound complex)
Ce Mextli (cap-free complex)
N
α1
α1
β1α2
N
α3
C
β3
β5
C
α2
β8
NC helix
β2
β4
β7
eIF4E
L1
m7GTP
60°
Canonical
α1
N
α2
NC helix
L3
L1
m7GTP
Ce eIF4E (cap-free)
Ce eIF4E (cap-bound)

Supplemental Figure S4. Structure of the *Dm* and *Ce* eIF4E-Mxt complexes in the absence of the cap analog. (*A*) Cartoon representation showing the cap-free crystal form of the *Dm* eIF4E-Mxt complex. The asymmetric unit of the crystal contained four complexes. (*B*) Overview of one *Dm* eIF4E–Mxt complex from the cap-free crystal form. The region of the Mxt peptide with structural similarity to other 4E-BPs is colored in yellow. The auxiliary linker and helix are colored in blue. Selected secondary structure elements are labeled in black for eIF4E and in color for Mxt. (*C*) Structural overlay of the *Dm* eIF4E–Mxt in the cap bound and cap-free state. The two complexes superimpose with an RMSD of 0.38 Å over 233 Cα atoms. (*D*) Cartoon representation showing the crystal form of the *Ce* eIF4E-Mxt complex in the absence of the cap analog. The asymmetric unit of the crystal contained three complexes. (*E*) Overview of one complex from the cap-free crystal form. (*F*) Structural overlay of the *Ce* Mxt-eIF4E complexes in the cap-bound (green) and cap-free (red) form. Selected secondary structure elements are labeled in black for eIF4E and in color for Mxt. The two complexes superimpose with an RMSD of 0.38 Å over 207 Cα atoms.

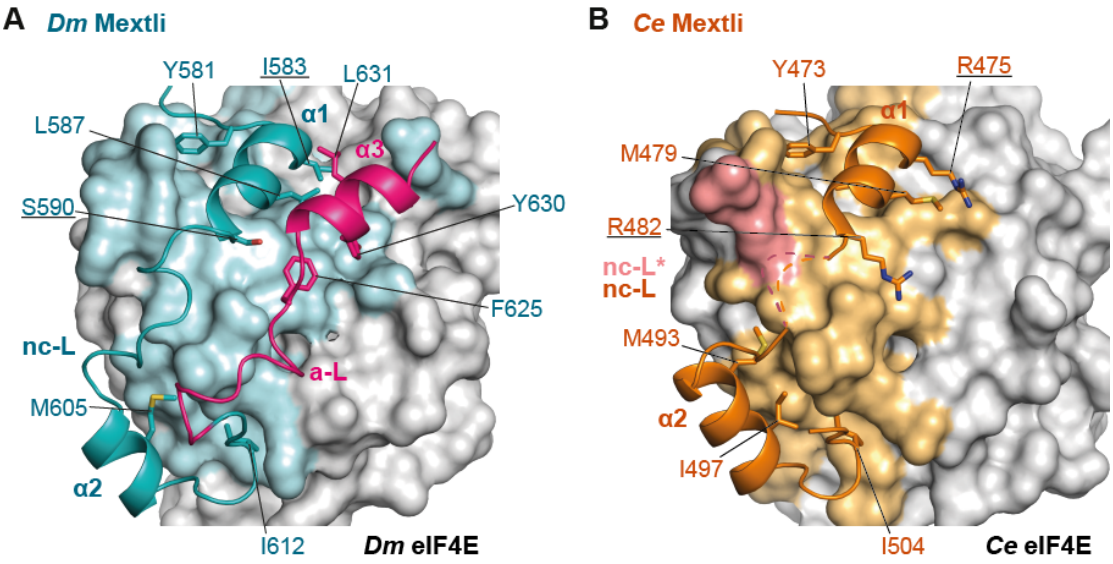
Figure S5



Supplemental Figure S5. Comparison of the *Dm* eIF4E-Mxt complex with the CUP and Thor complexes. (*A*) Overlay of the *Dm* eIF4E-Thor complex (PDB ID: 4UE8; Peter et al., 2015) and the *Dm* eIF4E-Mxt complex. The eIF4E molecule from the eIF4E-Thor complex has been removed for clarity. (*B*) Overlay of the elbow loops of *Dm* Mxt and *Dm* Thor reveals structurally similar residues, shown in colored sticks. (*C*) Overlay of the *Dm* eIF4E-CUP complex (PDB ID: 4AXG; Kinkelin et al., 2012) and the *Dm* eIF4E-Mxt complex bound to the cap analog. The eIF4E molecule from the eIF4E-CUP complex has been removed for clarity. (*D–F*) Close-up views of the

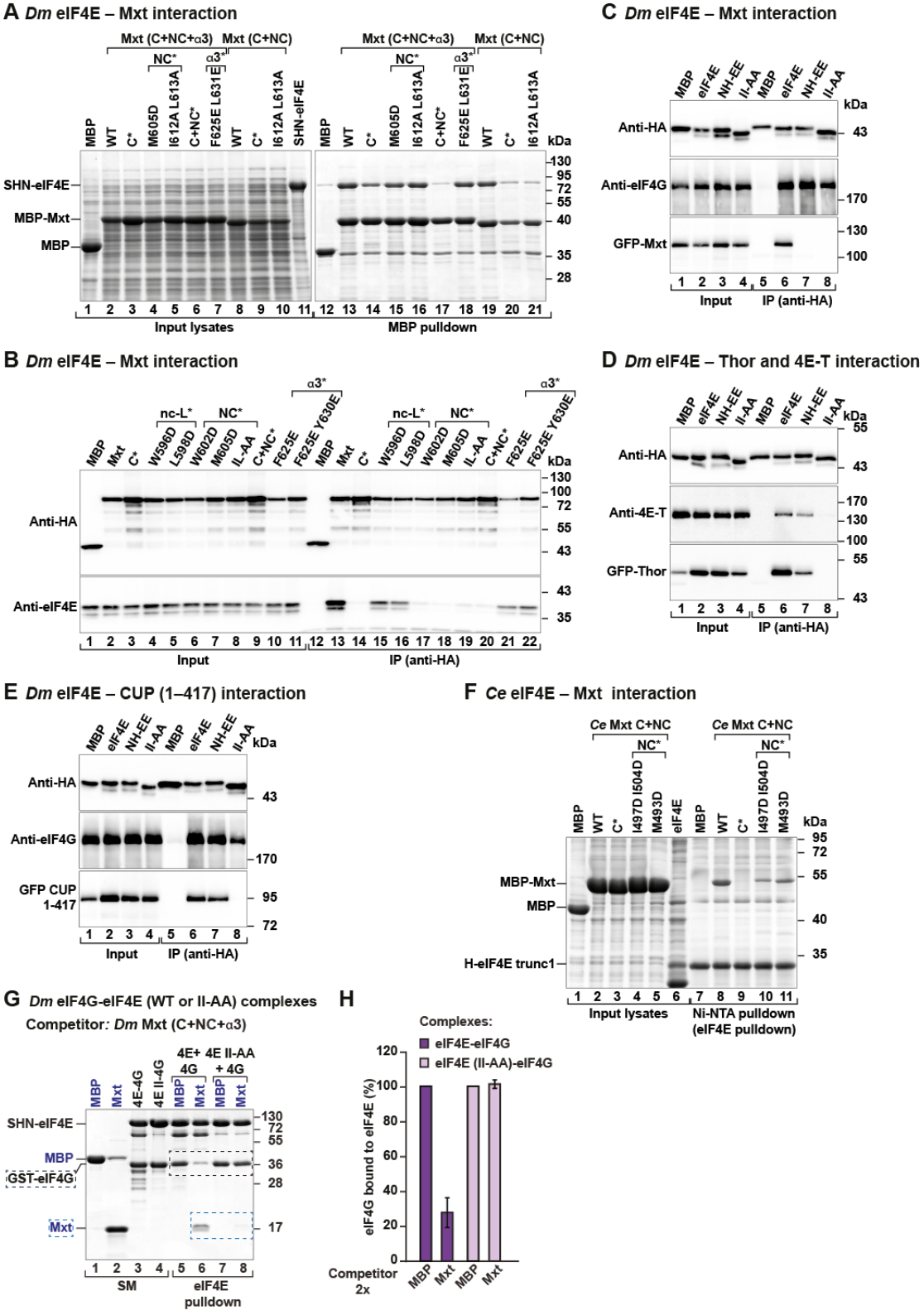
relative orientations of the non-canonical helices of *Dm* Mxt, *Ce* Mxt and *Dm* CUP bound to the lateral hydrophobic pocket of eIF4E. Helices are represented as cylinders.

Figure S6



Supplemental Figure S6. Binding interfaces of *Dm* and *Ce* Mxt on eIF4E. (*A, B*) Surface representation of eIF4E in complex with *Dm* and *Ce* Mxt. Residues of eIF4E within a distance of 4 Å from the bound Mxt peptide are shown in color. Selected Mxt residues that mediate major surface contacts at the dorsal and lateral surfaces of eIF4E are shown in sticks. The residues present at positions 2 and 9 of the *Dm* and *Ce* canonical motifs are underlined. For *Ce* Mxt, the contribution of the nc-L from the cap-free crystal structure is indicated in salmon.

Figure S7



Supplemental Figure S7. Validation of the eIF4E-Mxt interface. (*A*) *E. coli* lysates expressing MBP-tagged *Dm* Mxt fragments (WT or mutated), C-terminally fused to GB1, were incubated with SHN-tagged full length eIF4E. Protein complexes were

pulled down using amylose resin. Inputs (0.5%) and bound fractions (30%) were visualized on a SDS-PAGE gel followed by Coomassie staining. (*B*) The interaction of HA-tagged *Dm* Mxt (full length, either wild-type or mutated) with endogenous eIF4E was tested in *Dm* S2 cells. The proteins were immunoprecipitated using anti-HA antibodies. The inputs (2% for Mxt and 0.5% for eIF4E) and immunoprecipitates (15% for Mxt and 25% for eIF4E) were analyzed by western blotting using anti-HA and anti-eIF4E antibodies. (*C–E*) Western blotting showing the interaction of HA-tagged *Dm* eIF4E (full length, wild-type or mutants) with GFP-tagged *Dm* Mxt (full length, C), Thor (D) or CUP (1–417, E) and endogenous 4E-T and eIF4G (D,E). Input samples (1.5% for eIF4E, Thor and 4E-T; 10% for Mxt; 0.3% for eIF4G and 2% for CUP) and immunoprecipitates (15% for eIF4E; 30% for Mxt and 25% for Thor, 4E-T, eIF4G and CUP) were analyzed as described in (B). (*F*) Ni-NTA pulldown assay showing the association of His_6-tagged *Ce* eIF4E (30–215) with MBP-tagged *Ce* Mxt eIF4E-binding region (471–507, either wild-type or carrying the indicated mutations). Samples were analyzed as described in (A). Protein mutants are described in Table S1. Inputs (0.2%) and bound fractions (12.5%) were analyzed on a SDS-PAGE gel followed by Coomassie staining. (*G*) Purified *Dm* eIF4E-eIF4G complexes (2μM) containing SHN-tagged eIF4E (full length, WT or the II-AA mutant) and GST-eIF4G (residues 578–650) were incubated with two-fold molar excess of the *Dm* Mxt tripartite peptide (C+NC+α3) C-terminally fused to GB1. MBP served as negative control. The eIF4E-interacting proteins were pulled down using Ni-NTA beads. The competitor proteins are labeled in blue and marked by blue dashed boxes. The black dashed box indicates the position of GST-eIF4G. Lanes 1–4 show the purified peptides and complexes (starting material, SM) used in the competition assay. (*H*) The amounts of GST-eIF4G bound to eIF4E in each experimental condition were

determined using the Image J software after Coomassie blue staining. To rule out the possibility that changes in the levels of GST-eIF4G resulted from variations in the loading volume, all values were normalized to the levels of SHN-eIF4E present in each condition. These values were set to 100% in the presence of MBP. The mean values ± SD from three independent experiments are shown.

Supplemental References

Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH, Adams PD. 2012. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* **68**: 352–367.

Bricogne G, Blanc E, Brandl M, Flensburg C, Keller P, Paciorek W, Roversi P, Sharff A, Smart OS, Vonrhein C, Womack TO. 2011. BUSTER version 2.10.2. Cambridge, United Kingdom: Global Phasing Ltd.

Chen VB, Arendall WB 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66**: 12–21.

Cheng Y, Patel DJ. 2004. An efficient system for small protein expression and refolding. *Biochem Biophys Res Commun* **317**: 401–405.

Diebold ML, Fribourg S, Koch M, Metzger T, Romier C. 2011. Deciphering correct strategies for multiprotein complex assembly by co-expression: application to complexes as large as the histone octamer. *J Struct Biol* **175**: 178–188.

Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**: 486–501.

Hernandez G, Altmann M, Sierra JM, Urlaub H, Diez del Corral R, Schwartz P, Rivera-Pomar R. 2005. Functional analysis of seven genes encoding eight translation initiation factor 4E (eIF4E) isoforms in Drosophila. *Mech Dev* **122**: 529–543.

Igreja C, Izaurralde E. 2011. CUP promotes deadenylation and inhibits decapping of mRNA targets. *Genes Dev* **25**: 1955-1967.

Jankowska-Anyszka M, Lamphear BJ, Aamodt EJ, Harrington T, Darzynkiewicz E, Stolarski R, Rhoads RE. 1998. Multiple isoforms of eukaryotic protein synthesis initiation factor 4E in Caenorhabditis elegans can distinguish between mono- and trimethylated mRNA cap structures. *J Biol Chem* **273**: 10538-10542.

Kabsch W. 2010. Xds. *Acta Crystallogr D Biol Crystallogr* **66**: 125–132.

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read R.J. (2007). Phaser crystallographic software. *J Appl Crystallogr* **40**: 658–674.

Mizoue LS, Tellinghuisen J. 2004. The role of backlash in the "first injection anomaly" in isothermal titration calorimetry. *Anal Biochem* **326**: 125–127.

Smart OS, Womack TO, Sharff A, Flensburg C, Keller P, Paciorek W, Vonrhein C, Bricogne G. 2011 GRADE version 1.2.9. Cambridge, United Kingdom, Global Phasing Ltd., http://www.globalphasing.com.

Smart OS, Womack TO, Flensburg C, Keller P, Paciorek W, Sharff A, Vonrhein C, Bricogne G. 2012. Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. *Acta Cryst D Biol Crystallogr* **68**: 368–380.

Sheldrick GM. 2010. Experimental phasing with SHELXC/D/E: combining chaintracing with density modification. *Acta Crystallogr D Biol Crystallogr* **66**: 479–485.

Terwilliger TC, Grosse-Kunstleve RW, Afonine PV, Moriarty NW, Zwart PH, Hung LW, Read RJ, Adams PD. 2008. Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr D Biol Crystallogr* **64**: 61–69.

Terwilliger TC, Adams PD, Read RJ, McCoy AJ, Moriarty NW, Grosse-Kunstleve RW, Afonine PV, Zwart PH, Hung LW. 2009. Decision-making in structure

solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr D Biol Crystallogr* **65**: 582–601.

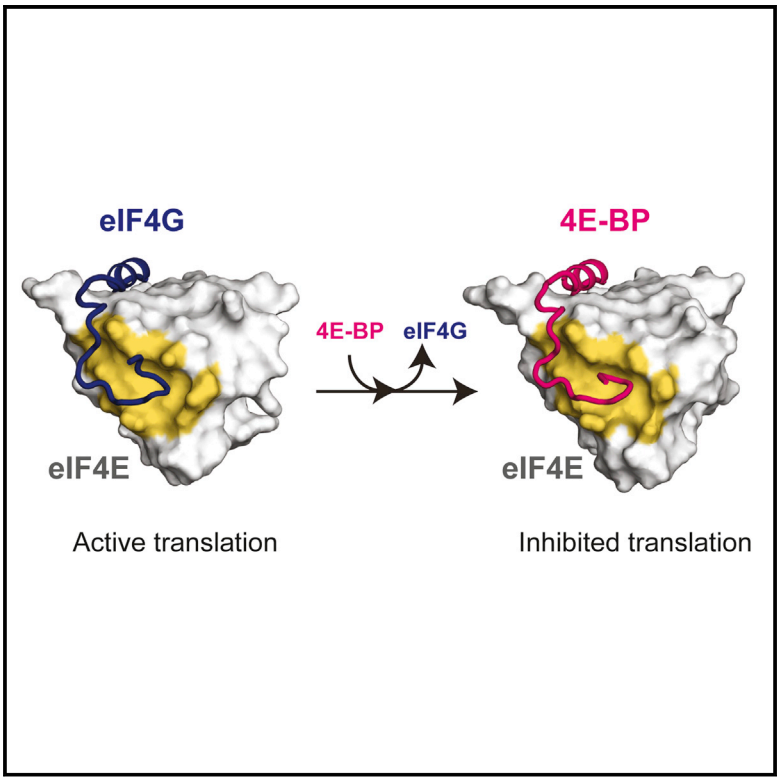
Zwart PH, Grosse-Kunstleve RW, Adams PD. 2005. *CCP4 Newsletter* **42**: contribution 10.

Article

Molecular Cell

The Structures of eIF4E-eIF4G Complexes Reveal an Extended Interface to Regulate Translation Initiation

Graphical Abstract



Authors

Stefan Grüner, Daniel Peter, Ramona Weber, ..., Eugene Valkov, Cátia Igreja, Elisa Izaurralde

Correspondence

oliver.weichenrieder@tuebingen.mpg.de (O.W.), elisa.izaurralde@tuebingen.mpg.de (E.I.)

In Brief

The interaction of eukaryotic translation initiation factors eIF4G and eIF4E is crucial for translation and a key target for synthetic drug design. Grüner et al. determined crystal structures of metazoan eIF4E-eIF4G complexes revealing extended similarity to inhibitory eIF4E-4E-BP complexes and challenging previous models obtained from yeast homologs.

Highlights

- Crystal structures of eIF4E-eIF4G complexes reveal lateral contacts on eIF4E
- Structural similarity of eIF4G and 4E-BPs extends to the lateral contacts with eIF4E
- Stronger lateral binding of inhibitory 4E-BPs helps them to displace eIF4G from eIF4E
- Structural model for the design of translational inhibitors as therapeutic tools

Accession Numbers

5T46
5T47
5T48

Grüner et al., 2016, Molecular Cell *64*, 467–479
November 3, 2016
http://dx.doi.org/10.1016/j.molcel.2016.09.020

Molecular Cell
Article

The Structures of eIF4E-eIF4G Complexes Reveal an Extended Interface to Regulate Translation Initiation

Stefan Grüner,[1] Daniel Peter,[1] Ramona Weber,[1] Lara Wohlbold,[1] Min-Yi Chung,[1] Oliver Weichenrieder,[1,*] Eugene Valkov,[1] Cátia Igreja,[1] and Elisa Izaurralde[1,2,*]
[1]Department of Biochemistry, Max Planck Institute for Developmental Biology, Spemannstrasse 35, 72076 Tübingen, Germany
[2]Lead Contact
*Correspondence: oliver.weichenrieder@tuebingen.mpg.de (O.W.), elisa.izaurralde@tuebingen.mpg.de (E.I.)
http://dx.doi.org/10.1016/j.molcel.2016.09.020

SUMMARY

Eukaryotic initiation factor 4G (eIF4G) plays a central role in translation initiation through its interactions with the cap-binding protein eIF4E. This interaction is a major drug target for repressing translation and is naturally regulated by 4E-binding proteins (4E-BPs). 4E-BPs and eIF4G compete for binding to the eIF4E dorsal surface via a shared canonical 4E-binding motif, but also contain auxiliary eIF4E-binding sequences, which were assumed to contact non-overlapping eIF4E surfaces. However, it is unknown how metazoan eIF4G auxiliary sequences bind eIF4E. Here, we describe crystal structures of human and *Drosophila melanogaster* eIF4E-eIF4G complexes, which unexpectedly reveal that the eIF4G auxiliary sequences bind to the lateral surface of eIF4E, using a similar mode to that of 4E-BPs. Our studies provide a molecular model of the eIF4E-eIF4G complex, shed light on the competition mechanism of 4E-BPs, and enable the rational design of selective eIF4G inhibitors to dampen dysregulated translation in disease.

INTRODUCTION

Eukaryotic initiation factor 4G (eIF4G) plays a central role in translation initiation in eukaryotes (Jackson et al., 2010). Consequently, its function is required for cell growth and differentiation, organism development, and it has been implicated in the pathogenesis of various conditions such as tumor growth and age-related diseases (reviewed in Howard and Rogers, 2014). During initiation, eIF4G mediates protein-protein interactions that are essential for 43S preinitiation complex (PIC) recruitment to capped mRNAs. The 43S PIC consists of a 40S ribosomal subunit, the ternary eIF2-GTP-Met-tRNAi complex, eIF3, and additional initiation factors and interacts with eIF4G through the eIF3 complex (Jackson et al., 2010). In turn, eIF4G associates with the cap-binding protein eIF4E and the RNA helicase eIF4A to assemble into the eIF4F complex, thereby linking translation initiation with mRNA 5′ cap structure recognition (Jackson et al., 2010).

eIF4G interacts with eIF4E via a conserved motif, which is termed the canonical eIF4E-binding motif (canonical 4E-BM), with the consensus sequence $YX_4L\Phi$ (where Y, X, L, and Φ denote Tyr, any amino acid, Leu, and any hydrophobic residue, respectively; Mader et al., 1995; Marcotrigiano et al., 1999). This motif also appears in a diverse class of eIF4E-binding proteins (4E-BPs) that generally inhibit translation initiation by competing with eIF4G for binding to eIF4E (Haghighat et al., 1995; Mader et al., 1995). The canonical 4E-BMs of eIF4G and 4E-BPs adopt an α-helical conformation upon binding to a conserved hydrophobic patch on the dorsal surface of eIF4E, which lies opposite the cap-binding pocket (Marcotrigiano et al., 1999; Matsuo et al., 1997). In addition, sequences that are located C-terminally to the canonical motif also contribute to the interaction between 4E-BPs and eIF4E. These sequences comprise a linker region and a non-canonical motif, which together increase the affinity of 4E-BPs for eIF4E by 2 to 3 orders of magnitude. Although the non-canonical motifs of 4E-BPs are not conserved in sequence, they all bind to a conserved hydrophobic lateral surface of eIF4E and are required for 4E-BPs to compete with eIF4G and repress translation (Igreja et al., 2014; Kinkelin et al., 2012; Lukhele et al., 2013; Paku et al., 2012; Peter et al., 2015a; Sekiyama et al., 2015).

The sequences flanking the canonical motif of eIF4G also contribute to its interaction with eIF4E (Gross et al., 2003; Umenaga et al., 2011). The NMR structure of the *Saccharomyces cerevisiae* (*Sc*) eIF4E-binding region of eIF4G in complex with eIF4E revealed that these flanking auxiliary sequences fold into a bracelet-like structure that wraps around the N terminus of eIF4E and that do not contact the lateral surface of eIF4E (Gross et al., 2003). These observations have led to a model in which 4E-BPs use their non-canonical sequences to dock onto the available lateral surface of eIF4E-eIF4G complexes, enabling them to efficiently compete with eIF4G for binding to the dorsal surface (Igreja et al., 2014; Peter et al., 2015a).

However, it is not known whether a similar bracelet structure is formed in metazoan eIF4E-eIF4G complexes because the N-terminal regions of eIF4E and eIF4G that are involved in bracelet formation are not conserved. Furthermore, no structural information is available for metazoan eIF4E-eIF4G complexes extending beyond the canonical 4E-BM of eIF4G. Importantly, the NMR structure of *Sc* eIF4E-eIF4G was determined in the presence of a solubilizing detergent. The use of detergent-based conditions

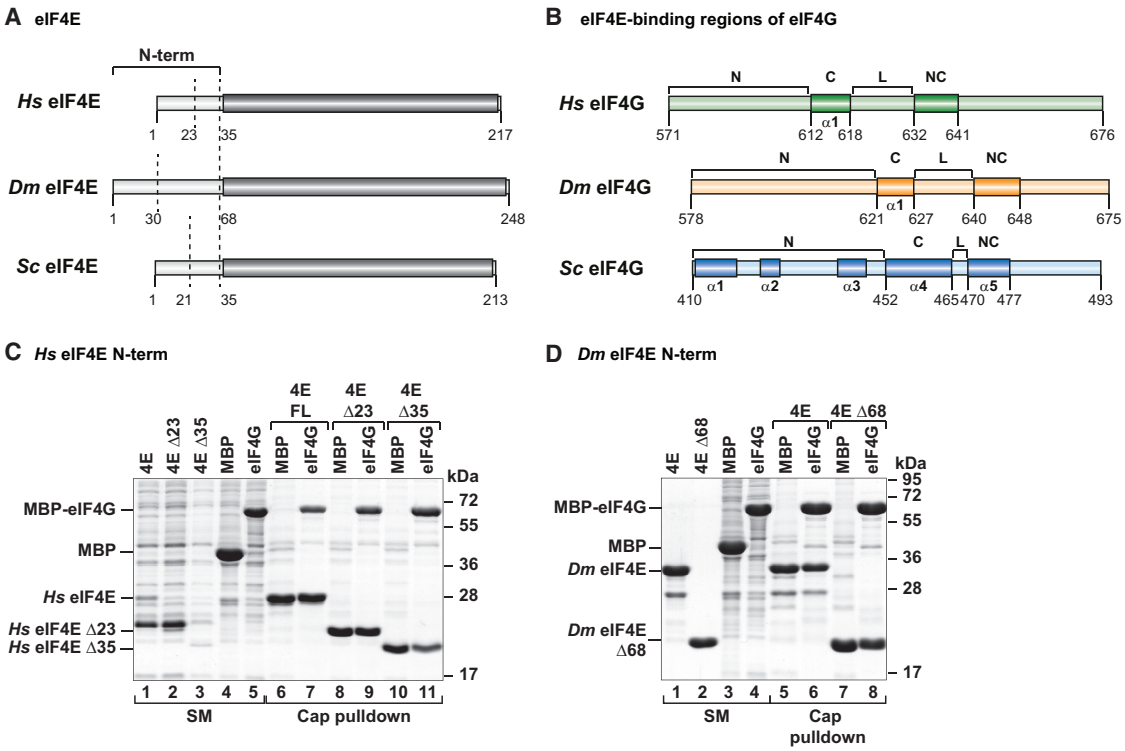


Figure 1. Formation of the Metazoan eIF4E-eIF4G Complex Does Not Require the eIF4E N-Terminal Region

(A) Schematic representation of *Hs*, *Dm*, and *Sc* eIF4E proteins. The N-terminal truncations analyzed in this study are indicated by vertical dashed lines. The conserved folded domain is highlighted in dark gray.

(B) Schematic organization of the eIF4E-binding regions of *Hs*, *Dm*, and *Sc* eIF4G proteins. The N-terminal region (N), the canonical (C) and non-canonical (NC) 4E-BMs, and the connecting linker (L) are indicated.

(C and D) Interaction of N-terminally truncated eIF4E proteins and the eIF4E-binding regions of *Hs* (C) and *Dm* (D) eIF4G. The untagged recombinant *Hs* eIF4E (full-length, Δ23 or Δ35) or purified *Dm* eIF4E (full-length or Δ68) was pulled down using m^7GTP-sepharose beads in the presence of eIF4G fragments. The lanes labeled starting material (SM) show the lysates (2%) and purified proteins (7.5%) used in the pull-down assays. The bound fractions (50% in C and 27% in D) were analyzed by SDS-PAGE and Coomassie blue staining. See also Figure S1A and Table S1.

(which reduces hydrophobic interactions) for the structural determination of the *Sc* complex inevitably raises the question as to whether the observed conformation of *Sc* eIF4G non-canonical sequences is physiologically relevant and representative of the eIF4G-binding mode.

The absence of a complete molecular model for the association of metazoan eIF4G with eIF4E limits our molecular understanding of translation initiation and its inhibition by 4E-BPs. To obtain a comprehensive mechanistic, structural, and evolutionary understanding of eIF4E-eIF4G complex assembly, we determined the structures of human (*Hs*) and *Drosophila melanogaster* (*Dm*) eIF4G proteins bound to the corresponding eIF4E. Unexpectedly, and in contrast to the previous model that was based on the structure of the *Sc* eIF4E-eIF4G complex, our structural data reveal that the eIF4G auxiliary sequences engage the lateral surface of eIF4E in a similar manner to that of 4E-BPs. This similarity explains why 4E-BPs compete efficiently with eIF4G for binding to eIF4E. Furthermore, we show that the amino acid composition of the linker regions and the non-canonical motifs of 4E-BPs are crucial for providing 4E-BPs with a competitive advantage over eIF4G. Together, our studies provide molecular insight into eIF4E-eIF4G complex assembly, a revised model for the action of 4E-BPs, and a structural framework for the rational design of small-molecule translation inhibitors for use as therapeutic agents for diseases that are associated with increased protein synthesis.

RESULTS

The N-terminal Regions of *Hs* and *Dm* eIF4E Are Not Required for Binding to eIF4G

eIF4E proteins comprise a conserved folded domain preceded by an N-terminal unstructured extension of variable length and sequence (Figures 1A and S1A). Despite the lack of conservation, the N-terminal extension of *Sc* eIF4E (residues A21–V35; Figure 1A) is required for high-affinity binding to eIF4G (Gross et al., 2003). To investigate whether the N-terminal unstructured regions of metazoan eIF4Es contribute to the interaction with eIF4G, we tested the ability of N-terminally truncated *Hs* and *Dm* eIF4E proteins to bind fragments of the corresponding eIF4Gs using in vitro pull-down assays. The eIF4G fragments included the canonical (C) 4E-BM and the flanking auxiliary sequences that have previously been shown to be involved in binding to eIF4E in several species. These flanking sequences include

residues that are N-terminal (N) to the canonical motif, as well as C-terminal residues, which we divided into the linker region (L) and the non-canonical motif (NC) based on the structures presented below (Figures 1B and S1B; Table S1; Gross et al., 2003; Hershey et al., 1999; Umenaga et al., 2011). The eIF4E constructs contained progressive deletions of N-terminal residues corresponding to residues that are involved in the eIF4E-eIF4G interaction in *Sc* (Figure 1A; Gross et al., 2003). We observed that the N-terminal deletions did not affect the ability of *Hs* or *Dm* eIF4E to bind to eIF4G in this assay (Figures 1C and 1D).

To validate these observations, we determined the affinity of the eIF4E-binding region of *Dm* eIF4G for *Dm* eIF4E protein variants lacking the first 30 and 68 N-terminal residues (Δ30 and Δ68, corresponding to full-length *Sc* eIF4E and *Sc* eIF4E Δ35, respectively) using isothermal titration calorimetry (ITC). We chose *Dm* proteins for this purpose because they exhibited much greater stability than the human proteins in solution, which is necessary for biophysical analysis. The dissociation constant (K_D) of the eIF4G peptide (N_s+C+L+NC; residues S601–N675) for the eIF4E Δ30 and Δ68 variants was in the low nanomolar range (4.2 ± 0.2 and 5 ± 1.4 nM, respectively), and no differences in binding thermodynamics were observed in the presence or absence of the eIF4E N-terminal extension (Figures S2A and S2B; Table S2). These results contrast with the findings reported in *Sc*, in which the deletion of the first 35 N-terminal residues of eIF4E reduced the affinity for eIF4G by two orders of magnitude (4.2 nM versus 314.4 nM; Gross et al., 2003). Thus, the eIF4E N-terminal extension is not essential for the assembly of the *Dm* eIF4E-eIF4G complex, suggesting that the interaction of eIF4G auxiliary sequences with eIF4E diverged during evolution.

Crystal Structures of Metazoan eIF4E-eIF4G Complexes Reveal a Lateral Binding Site for eIF4G

To gain insight into the assembly of the metazoan eIF4E-eIF4G complex, we determined the structures of *Hs* and *Dm* eIF4E in complex with the eIF4E-binding peptide of the corresponding eIF4G (*Hs* 592–653 and *Dm* 601–660; Figure S1B; Table S1). The *Dm* complex was crystallized both in the absence and presence of the m^7GpppG cap analog, whereas the *Hs* complex was crystallized only in the presence of m^7GpppG (Figures 2A–2G and S3A–S3D; Table 1).

The structures of *Hs* and *Dm* eIF4E in the complexes are similar to the previously determined structures of free and bound eIF4E from diverse organisms, and no major conformational changes occur upon eIF4G binding, as previously described (Figure S3D; Marcotrigiano et al., 1999; Peter et al., 2015a). The overall fold of eIF4E is highly conserved and consists of a strongly curved β sheet of eight antiparallel strands that adopt a horseshoe-like conformation. The cap-binding cavity is located on the concave, ventral surface (Figures 2A, 2B, and 2E). The convex surface is covered by three α helices that form the dorsal binding surface of the protein (Figures 2B, 2D, and 2E).

The structures of *Hs* and *Dm* eIF4E bound to the cap analog superpose with a root-mean-square deviation (RMSD) of 0.46 Å over 137 Cα atoms (Figure S3D). The differences observed in the N-terminal region of eIF4E (Figures 2A versus 2C) are due to a truncation (Δ68) of the *Dm* eIF4E protein for the purpose of crystallization. Although full-length eIF4E was used in the human complex, the first 29 residues (M1–A29) are not visible in the electron density map (Figure S1A).

Unexpectedly, and in contrast to previous observations of the NMR structure of the *Sc* eIF4E-eIF4G complex (Gross et al., 2003; Figures 2H and S4A), the *Hs* and *Dm* eIF4G peptides contact the lateral surface of eIF4E in a mode that is structurally analogous to that observed for diverse 4E-BPs (Figures 2A–2G and S4B–S4D; Peter et al., 2015a, 2015b; Sekiyama et al., 2015). Indeed, the two eIF4G peptides feature three common structural elements: (1) a canonical helix that binds the dorsal surface of eIF4E; (2) a linker region following the canonical helix that forms an elbow loop, which orients the peptide backbone toward the lateral surface of eIF4E; and (3) a non-canonical motif that engages the lateral surface of eIF4E mainly through hydrophobic interactions. Thus, our structures extend the eIF4E-eIF4G interface beyond the previously characterized canonical helix, providing unprecedented information on the binding mode of the metazoan eIF4G. This binding mode is clearly distinct from the bracelet-like structure formed around the eIF4E N terminus on the dorsal surface of eIF4E observed in the NMR structure of the *Sc* eIF4E-eIF4G complex (Gross et al., 2003; Figures 2H and S4A).

Structural Details of the Interaction of eIF4E with eIF4G

The *Hs* and *Dm* eIF4G peptides engage the dorsal and lateral surfaces of eIF4E through interactions that are mediated by the canonical and non-canonical 4E-BMs and the connecting linker (Figures 3A–3F) and bury comparable total surface areas (Table S3).

The canonical helices of *Hs* and *Dm* eIF4G interact with eIF4E through the conserved residues in the consensus sequence in a similar manner to that observed in other eIF4Gs and 4E-BPs in complex with eIF4E (Figures 3A and 3B; Gross et al., 2003; Kinkelin et al., 2012; Marcotrigiano et al., 1999; Peter et al., 2015a, 2015b). The tyrosine residue of the canonical consensus 4E-BM (*Hs* Y612 or *Dm* Y621) forms a hydrogen bond to the backbone of the conserved His-Pro-Leu motif of eIF4E (*Hs* H37-P38-L39 or *Dm* H70-P71-L72; Figures 3A and 3B). The hydrophobic residues at the C terminus of the canonical motif (LΦ; *Hs* L617 and L618 or *Dm* L626 and L627) interact with conserved residues on the dorsal surface of eIF4E (*Hs* V69, W73, and L135 or *Dm* V102, W106, and L167).

The canonical motifs of metazoan eIF4Gs and other 4E-BPs contain Arg/Gln/Lys residues at positions 2 and 9, leading to an extended canonical consensus sequence [YX(R/K) X_2LΦX_2(R/K/Q)]. These residues contribute to the interaction with eIF4E, likely by shielding hydrophobic surface patches of eIF4E from solvent exposure (Kinkelin et al., 2012; Peter et al., 2015a). Furthermore, the Arg residues at position 2 of the metazoan eIF4Gs (*Hs* R614 and *Dm* R623) further stabilize the binding of the canonical helix by forming a salt bridge with an acidic residue in eIF4E (*Hs* E132 or *Dm* D164; Figures 3A and 3B), which is conserved in metazoans (Figure S1A).

Following the canonical helix, the metazoan eIF4G linker forms an elbow loop (Figures 3C and 3D and S5A–S5C). The elbow loops begin immediately after the Arg/Gln residues at position 9 of the canonical motifs (*Hs* Q621 and *Dm* R630) and comprise a half helical turn beginning with a residue located at the tip of the elbow (*Hs* F624 and *Dm* K633) and ending with similarly arranged

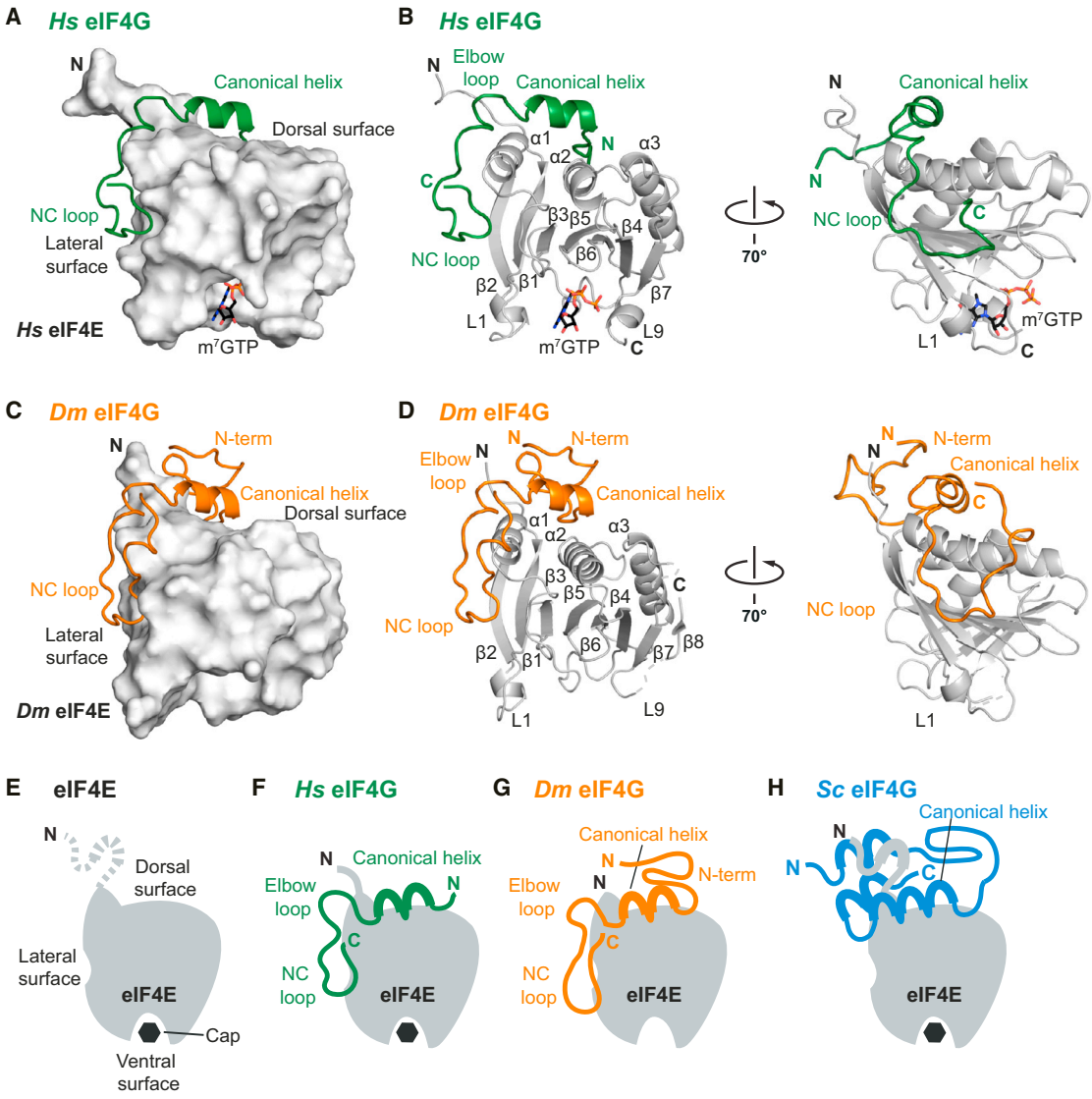


Figure 2. Structures of *Hs* and *Dm* eIF4G Proteins Bound to eIF4E

(A and B) Surface and cartoon representations of the eIF4E-binding region of *Hs* eIF4G (green) in complex with cap-bound eIF4E (gray) in two orientations. The bound m^7GTP cap analog is shown as sticks. Selected secondary structure elements are labeled.

(C and D) Surface and cartoon representations of the structure of *Dm* eIF4G (orange) bound to eIF4E (gray) in two orientations. Selected secondary structure elements are labeled.

(E–H) Schematic representations of eIF4E (E) and of eIF4E bound to *Hs* (F), *Dm* (G), and *Sc* (H) eIF4G. See also Figures S3 and S4; Table S3.

serine residues (*Hs* S626 and *Dm* S635). The side chains of the Ser residues contact the carbonyl oxygens of the preceding residues (*Hs* F620 and *Dm* L629; Figures S5A and S5B), thus fixing the backbone. Further stabilization is provided by internal backbone-mediated hydrogen bonds within the half helical turn (Figures S5A and S5B). The elbow loops are anchored to the surface of eIF4E through interactions with conserved eIF4E Asn (*Hs* N77 and *Dm* N110) and His (*Hs* H78 and *Dm* H111) residues. These interactions orient the linker region toward the lateral surface of eIF4E.

A major difference between the yeast and metazoan complexes is observed in the conformation of the non-canonical motifs. In the metazoan complexes, the non-canonical motifs fold into a loop that engages the lateral surface of eIF4E (Figures 3E and 3F). These non-canonical loops are stabilized on the lateral surface of eIF4E by hydrophobic interactions and bury comparable surface areas in both species (Table S3). Specifically, the hydrophobic residues in the non-canonical motifs (*Hs* L633, I636, V639, and V640 and *Dm* V641, I646, and L647) engage the lateral surface of eIF4E, which is lined with conserved hydrophobic residues (*Hs* F47, I63, L75, and I79 and *Dm* Y80, I96, L108, and I112; Figures 3E and 3F). Thus, binding to the lateral hydrophobic surface of eIF4E is a conserved feature of metazoan eIF4Gs and diverse 4E-BPs. In the yeast complex,

Table 1. Data Collection and Refinement Statistics

	Hs eIF4E-eIF4G Complex (Cap-Bound)	*Dm* eIF4E-eIF4G Complex (Cap-Free)	*Dm* eIF4E-eIF4G Complex (Cap-Bound)
Space Group	$P2_1$	P1	$P2_12_12_1$
Unit Cell			
Dimensions (Å) a, b, and c	45.3, 70.3, and 79.9	46.3, 49.7, and 58.6	45.6, 66.2, and 75.6
Angles (°) α, β, and γ	90.0, 99.4, and 90.0	102.7, 103.9, and 116.1	90.0, 90.0, and 90.0
Data Collection			
Wavelength (Å)	0.9785	0.9999	0.9999
Resolution (Å)[a]	44.65–1.53 (1.56–1.53)	41.42–2.20 (2.27–2.20)	45.62–2.19 (2.26–2.19)
R_{sym}	0.052 (1.013)	0.121 (0.609)	0.081 (1.066)
Mean I/σI	11.6 (1.2)	6.2 (1.7)	17.0 (2.3)
CC (1/2)	0.998 (0.565)	0.992 (0.683)	0.999 (0.850)
Completeness (%)	99.8 (99.9)	96.6 (91.0)	100.0 (99.6)
Multiplicity	3.7 (3.8)	3.0 (2.5)	12.6 (11.4)
Total number of reflections	278,363 (14,011)	62,442 (4,296)	154,909 (11,807)
Number of unique reflections	74,453 (3,718)	20,740 (1,700)	12,336 (1,036)
Refinement			
Resolution (Å)	44.65–1.53	41.42–2.20	39.10–2.19
No. of reflections	74,411	20,729	12,278
R_{work}/R_{free}	0.139/0.162	0.208/0.234	0.209/0.231
Number of Atoms			
Protein	3,574	3,616	1,783
Cap analog	66	–	33
Other ligands	42	12	–
Water	346	157	35
B Factors (Å^2)			
Protein	33.5	43.1	70.6
Cap analog	58.1	–	82.2
Other ligands	53.8	60.2	–
Water	43.1	35.1	56.5
Ramachandran Plot			
Favored (%)	98.1	99.1	98.1
Disallowed (%)	0.0	0.0	0.0
Rmsd			
Bond lengths (Å)	0.008	0.002	0.003
Bond angles (°)	0.918	0.468	0.478

[a]Values shown in parentheses are for the highest resolution shell.

the corresponding non-canonical motif adopts a helical conformation and closes the molecular bracelet embracing the N terminus of eIF4E without contacting the lateral surface of eIF4E (Gross et al., 2003; Figures 2H and S4A).

The eIF4G residues that lie N-terminal to the canonical motifs are arranged differently in the *Dm* and *Hs* complexes. In the *Dm* complex, this N-terminal region adopts a defined loop conformation on the dorsal surface of eIF4E, burying several hydrophobic residues (*Dm* eIF4G residues I603, Y605, and W610; Figures 3G and 3H). This organized loop is less complex than the molecular bracelet that is present in the yeast eIF4E-eIF4G complex (Gross et al., 2003; Figures 2H and S4A) and does not require the N terminus of eIF4E. Instead, the loop is stabilized and anchored to the canonical helix by an extensive intramolecular hydrogen bond network involving the side chains of several residues (N604, N606, Q609, D622, and Q625; Figure 3H). In the *Hs* complex, the equivalent N-terminal eIF4G residues (Q592–L606) were not visible in the electron density map, suggesting either high flexibility or susceptibility to proteolytic degradation during crystallization.

Functional Significance of the eIF4G Linker and Non-Canonical Motif for eIF4E Binding

To determine the contribution of the auxiliary structural elements to the assembly of the eIF4E-eIF4G complexes, we generated eIF4E and eIF4G mutants and tested their binding in pull-down

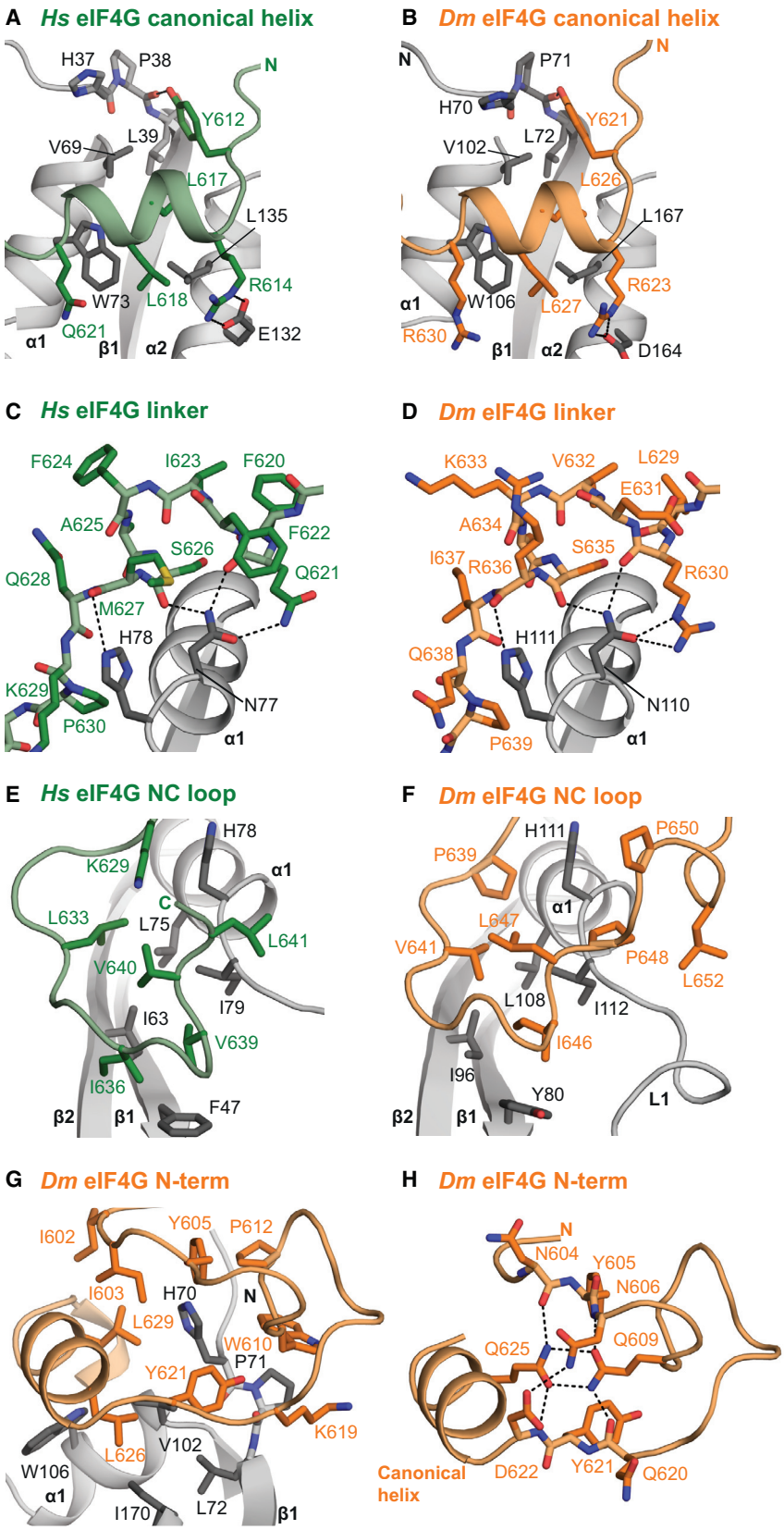


Figure 3. Molecular Details of *Hs* and *Dm* eIF4E-eIF4G Complexes

(A and B) Close-up view of the canonical helix of *Hs* eIF4G (A) and *Dm* eIF4G (B) bound to the dorsal surface of the respective eIF4E. In (B), *Dm* eIF4G residues S601–S616 were omitted for clarity.

(C and D) Close-up views of the linker regions of *Hs* eIF4G (C) and *Dm* eIF4G (D) contacting the respective eIF4Es. The dashed lines indicate hydrogen bonds or salt bridges.

(E and F) Close-up views of the interactions between the non-canonical (NC) loops of *Hs* eIF4G (E) and *Dm* eIF4G (F) and the lateral hydrophobic surface of the corresponding eIF4Es.

(G) Close-up view of the N-terminal loop of *Dm* eIF4G assembled on the dorsal surface of *Dm* eIF4E. Selected residues are shown as sticks.

(H) Intramolecular hydrogen bond (dashed lines) network at the N-terminal region of *Dm* eIF4G. See also Figure S5.

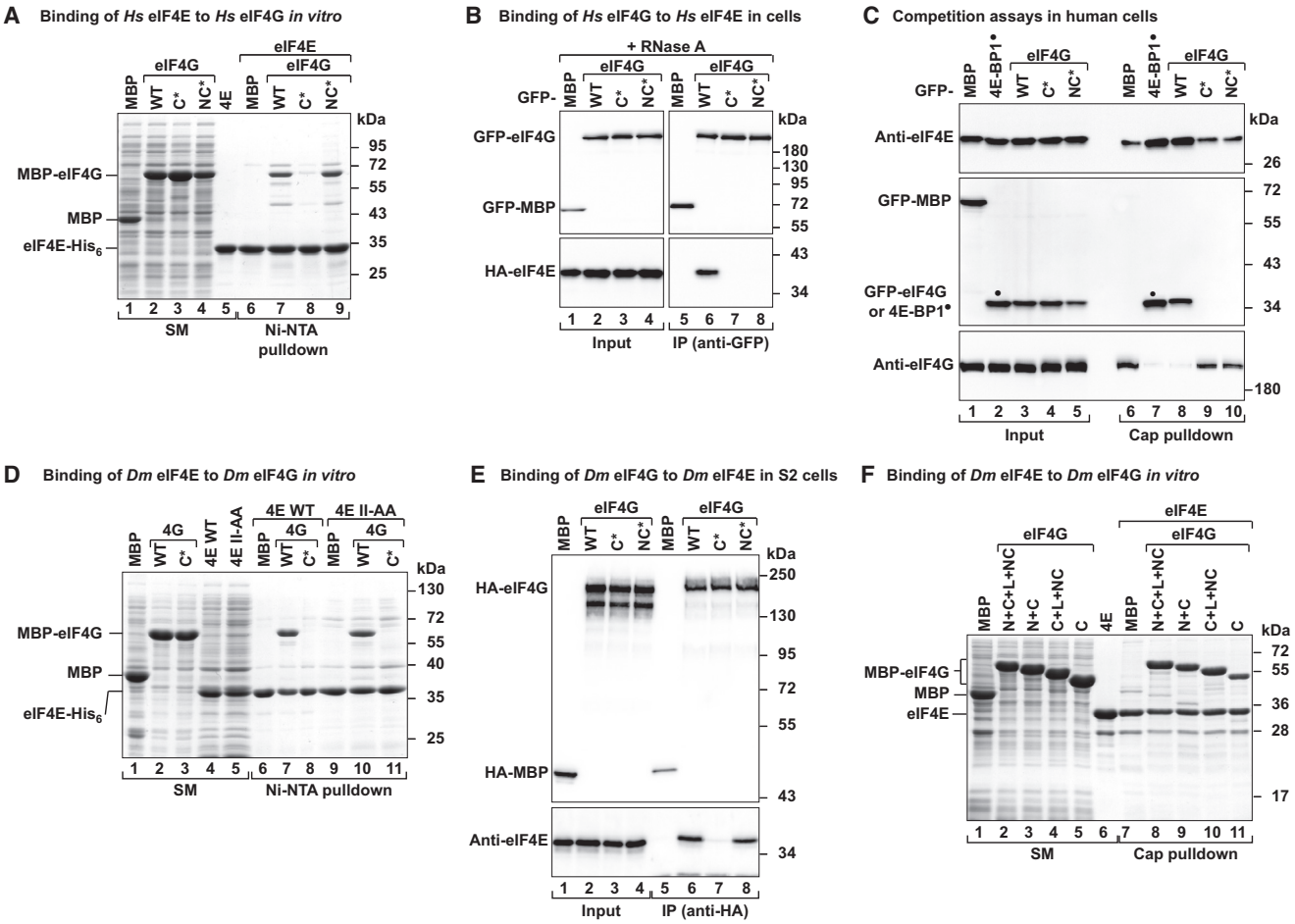


Figure 4. Contribution of eIF4G Structural Elements to eIF4E Binding

(A) The interaction of *Hs* His_6-eIF4E with the eIF4E-binding region of *Hs* eIF4G (N+C+L+NC; wild-type or C* and NC* mutants) was tested using Ni-NTA pull-down assays. The starting material (2% of the lysate or 3% of the purified protein) and bound fractions (25%) were analyzed as described in Figure 1.

(B) The interaction of *Hs* GFP-eIF4G (full-length, wild-type or mutated) with HA-eIF4E in human HEK293T cells was tested by immunoprecipitation using anti-GFP antibodies. The inputs (0.75% for GFP-proteins and 0.5% for HA-eIF4E) and immunoprecipitates (15% for GFP-proteins and 25% for HA-eIF4E) were analyzed by western blotting.

(C) Competition between *Hs* eIF4G and *Hs* 4E-BP1 in vivo. The lysates from human HEK293T cells expressing the indicated GFP-proteins were incubated with m^7GTP-sepharose beads. The input fractions (1% for GFP-proteins and 0.75% for eIF4E and eIF4G) and bound fractions (15% for GFP-proteins, 5% for eIF4E, and 20% for eIF4G) were analyzed by western blotting using anti-eIF4E, anti-eIF4G, and anti-GFP antibodies. The GFP-tagged proteins included MBP, a 4E-BP1 fragment (C+L+NC), and a fragment of *Hs* eIF4G (C+L+NC; wild-type or C* or NC* mutants). The position of the 4E-BP1 fragment is indicated by a black dot.

(D) The interaction of *Dm* His_6-eIF4E (full-length; wild-type or lateral surface mutant [II-AA]) with *Dm* eIF4G (N+C+L+NC; wild-type or C* mutant) was analyzed by Ni-NTA pull-down assay. The starting material (3% for eIF4G and 8% for eIF4E) and bound fractions (13% for eIF4E WT and 40% for eIF4E II-AA) were analyzed by SDS-PAGE.

(E) Interaction of HA-tagged full-length *Dm* eIF4G (wild-type, C* or NC* mutants) with endogenous eIF4E in S2 cell lysates was analyzed by immunoprecipitation using anti-HA antibodies. The inputs (0.75% for eIF4G and 0.15% for eIF4E) and immunoprecipitates (30%) were analyzed by western blotting using anti-HA and anti-eIF4E antibodies.

(F) Cap pull-down assay showing the association of purified, untagged *Dm* eIF4E (full-length) with the indicated *Dm* eIF4G fragments. The starting material (SM, 3% purified eIF4E and 2.5% cell lysates) and bound fractions (25%) were visualized on SDS-PAGE followed by Coomassie staining. See also Tables S1 and S4.

assays in vitro using recombinant proteins as well as in coimmunoprecipitation assays in *Hs* HEK293T cells and in *Dm* Schneider (S2) cells.

The interaction of *Hs* eIF4G with eIF4E was abolished by mutations in the canonical motif (C*; Table S1) both in vitro and in vivo (Figures 4A and 4B). In contrast, amino acid substitutions in the non-canonical motif (NC*; Table S1) did not interfere with eIF4E binding in vitro (Figure 4A, lane 9), but did abolish binding in human cell lysates (Figure 4B, lane 8), probably because the mutant protein cannot compete with other 4E-BPs that are present in cell lysates for binding to eIF4E. Accordingly, the eIF4E-binding peptide of *Hs* eIF4G fused to GFP competed with endogenous full-length eIF4G for binding to eIF4E to a similar extent as 4E-BP1 in human cells (Figure 4C, lanes 7 and 8). However, this

competition was abrogated by mutations in the eIF4G canonical or non-canonical motif (Figure 4C, lanes 9 and 10). Thus, the non-canonical motif of *Hs* eIF4G is not required for binding to eIF4E in the absence of competitor proteins (in vitro), but is necessary in the presence of other 4E-BPs (in human cell lysates), indicating a role in stabilizing the eIF4E-eIF4G interaction.

Similarly, mutations in the canonical motif (C*; Table S1) abolished the interaction of *Dm* eIF4G with eIF4E both in vitro and in S2 cell lysates (Figures 4D, lane 8 and 4E, lane 7). In contrast, the interaction was not detectably impaired by mutations or deletions in the non-canonical motif in either condition (Figures 4E, lane 8 and 4F, lane 9). Accordingly, mutations on the lateral surface of eIF4E (I96A and I112A; II-AA) did not affect the binding of *Dm* eIF4G to eIF4E (Figure 4D, lane 10 versus 7).

Different Contributions of the Linker and Non-Canonical Motifs of eIF4G and 4E-BPs to the Affinity for eIF4E

To determine the contribution of the eIF4G auxiliary sequences to the overall affinity for eIF4E and to gain insight into how these eIF4G sequences compare to the equivalent sequences in 4E-BPs, we performed a detailed ITC analysis of the interactions of *Dm* eIF4E with *Dm* eIF4G and the *Dm* 4E-BP1 ortholog, Thor.

The affinity of the isolated canonical (C) motif of eIF4G for eIF4E (Δ30) was approximately 90-fold lower (K_D = 380 ± 69 nM) than the affinity of the peptide that additionally contained the auxiliary sequences flanking the motif (N_s+C+L+NC; K_D = 4.2 ± 0.2 nM; Figure S2; Table S2). To determine the contribution of the sequences preceding (N_s) or following (L+NC) the canonical motif (C) individually, we measured the affinity of the corresponding eIF4G peptides for eIF4E in independent experiments. The affinity of eIF4G peptides containing the canonical motif was increased 15-fold when the N-terminal sequences were included (N_s+C; K_D = 25 ± 2.9 nM; Table S2) and was increased 10-fold when the linker and non-canonical motif were present (C+L+NC; K_D = 37 ± 6.5 nM). Importantly, addition of the N-terminal sequence of eIF4G (N_s+C peptide) was accompanied by a much favorable binding enthalpy ($\Delta\Delta H = -7.9$ kcal/mol) and an increase in the entropic penalty of binding [$\Delta(-T\Delta S)$ = 6.3 kcal/mol; Table S2] relative to the canonical motif (C). These results are consistent with the observation that the N-terminal region establishes a stabilizing network of intramolecular hydrogen bonds (Figure 3H) and likely undergoes a disorder-to-order transition when binding to eIF4E. In contrast, no such enthalpy-entropy compensation was observed for the addition of the linker and non-canonical motif (C+L+NC peptide), but primarily entropy-driven binding in the absence of an enthalpic contribution [$\Delta(-T\Delta S) = -1.4$ kcal/mol; Table S2], consistent with mainly hydrophobic interactions. Similar enthalpic and entropic differences were obtained when the (N_s) or (L+NC) sequences were deleted from the (N_s+C+L+NC) peptide (Table S2), indicating that overall peptide length has little effect.

In sharp contrast to eIF4G, in which the (L+NC) sequences cause a ~10-fold increase on the affinity for eIF4E, in the case of Thor, this increase was ~1,600-fold (Table S2). Indeed, including the (L+NC) sequences lowers the dissociation constant of the isolated canonical motif (C) of Thor from 2,260 ± 60 nM to 1.4 ± 0.3 nM (C+L+NC; Table S2). These results indicate that the lateral surface of eIF4E is much more important for the interaction with Thor than it is for eIF4G. Accordingly, mutations on the lateral surface of eIF4E reduced the affinity of eIF4G (N_s+C+L+NC) and Thor (C+L+NC) for eIF4E by 3- and 300-fold, respectively (Table S2). The decreased dependence of eIF4G on the lateral surface of eIF4E is also reflected by a different crystal form of the same eIF4E-eIF4G complex preparation, in which eIF4G does not bind to the lateral surface of eIF4E, and most of the linker and non-canonical motif is not visible (Figures S3B and S3C; Table 1).

Taken together, our results indicate that despite having structural similarities, the linker and non-canonical motif of eIF4G contribute less to the overall affinity for eIF4E than the equivalent sequences in 4E-BPs. Nevertheless, these sequences are required to maintain the stability of the eIF4G-eIF4E complex in the presence of competitor proteins.

The Linker and Non-Canonical 4E-BMs Provide a Competitive Advantage to 4E-BPs over eIF4G

The observation that eIF4G and 4E-BPs bind to overlapping eIF4E surfaces raises the question of how, in molecular terms, 4E-BPs compete with eIF4G to repress translation. Given the different contributions of the linker regions and the non-canonical motifs of eIF4G and Thor to eIF4E binding, it is possible that the differences in their ability to compete with each other are determined by these sequences. To evaluate the contribution of the individual sequence elements in conferring a competitive advantage for binding to eIF4E in competition assays in vitro, we designed chimeric peptides in which the canonical motif (C), the linker region (L), or the non-canonical motif (NC) of the *Dm* Thor was replaced by the corresponding sequence from the *Dm* eIF4G.

First, we compared the ability of a Thor peptide (C+L+NC) and the equivalent eIF4G peptide (C+L+NC) to displace the entire eIF4E-binding region of eIF4G (N+C+L+NC) from preassembled eIF4E-eIF4G complexes. The complexes were challenged with a 3-fold molar excess of the competitor peptides, and the proteins associated with eIF4E were recovered by eIF4E pull down at various time points. To compare the different competitor proteins, we estimated the time that each peptide required to displace 50% of the eIF4G molecules from preassembled eIF4E-eIF4G complexes, which we operationally defined as the "half-life" of the complex. In the presence of the Thor peptide, the half-life of the eIF4E-eIF4G complex was 11 ± 3 min. Under the same conditions, the corresponding eIF4G peptide did not efficiently displace eIF4G after a 60 min incubation (Figures 5A–5D), consistent with its 26-fold lower affinity (Table S2).

We next tested a chimeric peptide in which the linker and the non-canonical motif of Thor were replaced by their direct counterparts in eIF4G (chimera-1, C^{Thor}+L^{4G}+NC^{4G}). This peptide exhibited a 78-fold reduction in affinity (K_D = 109 ± 8 nM) compared to the Thor peptide (K_D = 1.4 ± 0.5 nM), thereby confirming that the linker and the non-canonical motif of eIF4G contribute to a lesser extent to the overall binding affinity for eIF4E. Accordingly, the chimera-1 peptide did not efficiently displace eIF4G from preformed complexes in competition assays (Figures 5A and 5E).

In contrast, a chimeric peptide containing the canonical motif of eIF4G fused to the linker and the non-canonical motif of Thor

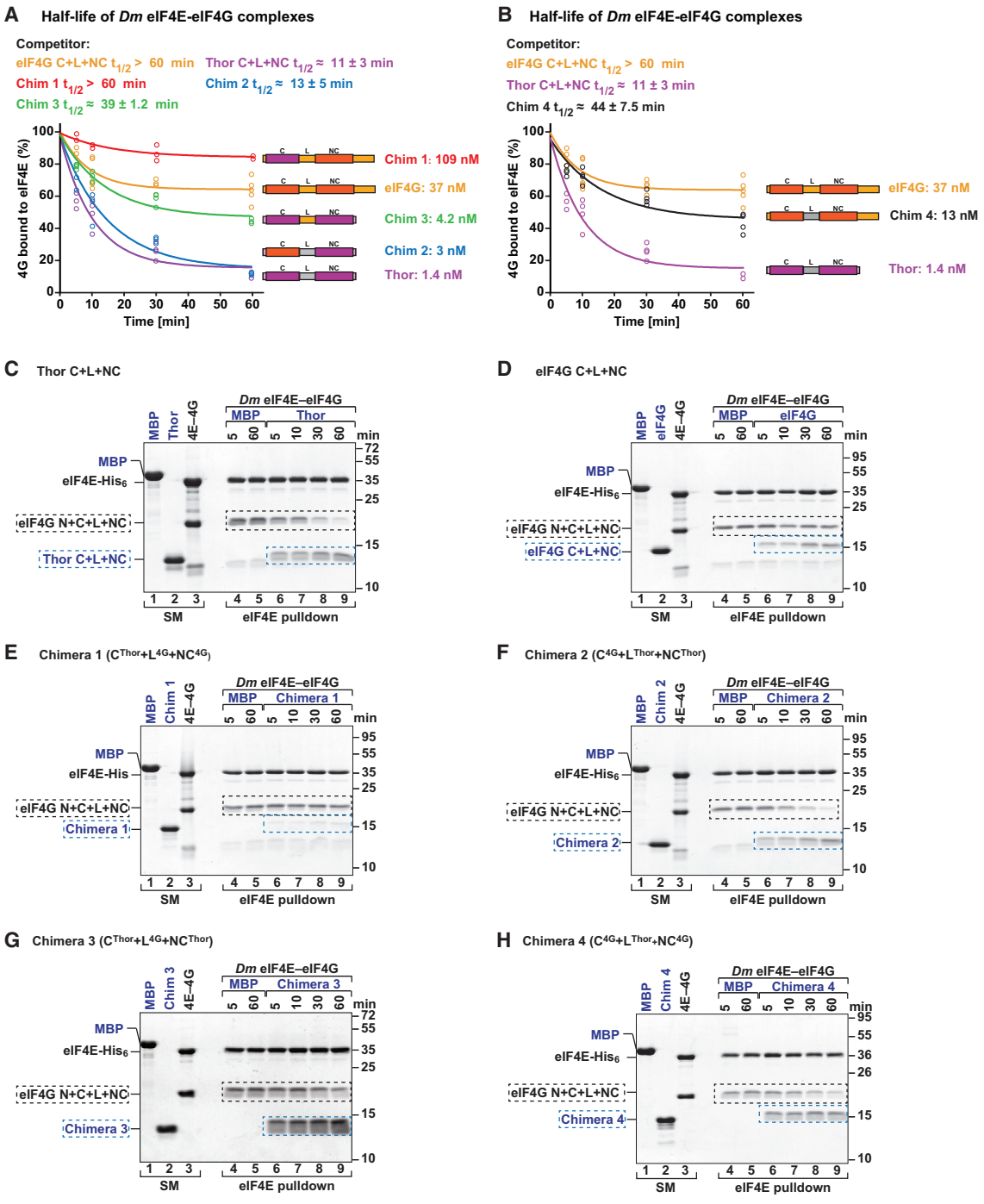


Figure 5. The Linker and Non-Canonical Motif of 4E-BPs Confer a Competitive Advantage over eIF4G

(A–H) In vitro competition assay. *Dm* eIF4E-eIF4G complexes were incubated in the presence of a 3-fold molar excess of the indicated competitor peptides. (A and B) Show the quantification of eIF4G (N+C+L+NC) remaining bound to eIF4E in the presence of the competitor proteins at different time points ($n \geq 3$). The half-life of the eIF4E-eIF4G complex ($t_{1/2}$, mean ± SDs from three independent experiments) in the presence of the competitor protein and the K_D values for the competitor peptides are indicated.

(C–H) Show Representative SDS-PAGE gels for each competition assay. The competitor and eIF4G peptides are boxed in blue and black, respectively. The lanes labeled starting material (SM) show the purified peptides and complexes used in the competition assay. See also Figure S2 and Tables S1 and S2.

(chimera-2, C^{4G}+L^{Thor}+NC^{Thor}) bound to eIF4E with nanomolar affinity (K_D = 3 ± 1.4 nM as compared to K_D = 1.4 ± 0.5 nM for Thor) and was as effective as the Thor peptide in displacing eIF4G from eIF4E in competition assays, thus reducing the half-life of the eIF4E-eIF4G complex to 13 ± 5 min (Figures 5A and 5F). These results indicate that the linker and the non-canonical motif (L+NC) of Thor confer a competitive advantage in displacing eIF4G from eIF4E, whether they are fused to the canonical motif of Thor or to that of eIF4G (Figure 5A).

Next, we investigated the contribution of the linker regions. A peptide in which the linker of Thor was replaced by the linker of eIF4G (chimera-3, C^{Thor}+L^{4G}+NC^{Thor}) bound to eIF4E with similar affinity to Thor and chimera-2 (4.2 ± 0.9 nM; Table S2). In competition assays, however, chimera-3 was less effective at displacing eIF4G than Thor and chimera-2, and the half-life of the eIF4E-eIF4G complex was 39 ± 1.2 min (Figures 5A and 5G). Conversely, when the linker of eIF4G was replaced by the linker of Thor to produce chimera-4 (C^{4G}+L^{Thor}+NC^{4G}), the half-life of the eIF4E-eIF4G complex was 44 ± 7.5 min (Figures 5B and 5H). ITC measurements indicated that the affinity of chimera-4 for eIF4E was increased 3-fold compared to the eIF4G peptide (13 ± 1.6 nM versus 37 ± 6.5 nM, respectively; Table S2). The observation that chimera-4 (C^{4G}+L^{Thor}+NC^{4G}) competes less efficiently than chimera-2 (C^{4G}+L^{Thor}+NC^{Thor}) indicates that the competitive advantage of Thor in binding to eIF4E is conferred by both the linker region and the non-canonical motif.

The Linker Sequence Modulates the Affinity of the eIF4E Interaction

Using the structural data presented in this study, we systematically analyzed the observed atomic interactions of both metazoan eIF4Gs as well as *Dm* Thor and *Hs* 4E-BP1 when bound to eIF4E, in an attempt to identify features that might contribute to the binding advantage of 4E-BPs over eIF4G. A notable feature that distinguishes the linker regions of eIF4Gs from the linkers of *Dm* Thor or *Hs* 4E-BP1–3 is the presence of conserved Pro residues in the latter (Figure 6A).

Because the presence of Pro residues uniquely restricts the conformational flexibility of the peptide backbone (MacArthur and Thornton, 1991; Huang and Nau, 2003), we investigated how the presence of these residues in the linker affects the affinity of 4E-BPs for eIF4E and the competitive ability of 4E-BPs in displacing eIF4G from preformed complexes. We substituted two residues in the linker region of the *Dm* eIF4G peptide (C+L+NC) in positions that are structurally equivalent to those in the Thor peptide with Pro residues (K633P and Q638P; eIF4G-PP mutant; Figures S5B and S5C). Remarkably, this peptide exhibited the same properties as chimera-4, in which the entire linker of eIF4G was replaced by the linker of Thor and displaced 50% of bound eIF4G in 37 ± 10 min compared with >60 min for the wild-type eIF4G (Figures 6B and 6C). These substitutions also increased the affinity of the eIF4G peptide for eIF4E from a dissociation constant of 37 ± 6.5 nM (for eIF4G) to 8 ± 3.5 nM (for eIF4G-PP) to levels that were similar to those observed for chimera-4 (Table S2). The gain in affinity was associated with a more favorable binding entropy compared to the eIF4G C+L+NC peptide [$\Delta(-T\Delta S)$ = − 2.0 kcal/mol; Table S2], possibly indicating that the conformational freedom of the unbound peptide was indeed reduced by the insertion of Pro residues in the linker. Thus, a proline-rigidified linker can apparently improve affinity and competition efficiency. Finally, we generated a Thor peptide in which the two Pro residues were replaced by their structurally equivalent residues from eIF4G (P66K, P71Q, Thor-KQ mutant). Interestingly, these substitutions did not alter the affinity of the peptide for eIF4E, but did slow down the competition kinetics (Figures 6D and 6E; Table S2).

In conclusion, our results indicate that despite having similar binding modes, the sequences in the linker region and the non-canonical motif confer a competitive advantage on 4E-BPs over eIF4G for binding to eIF4E.

DISCUSSION

Recognition of eIF4E via the Dorsal and Lateral Surfaces

In this study, we determined the crystal structures of two metazoan eIF4E-eIF4G complexes, which revealed a much more extensive interface than was previously inferred from metazoan structures in which mainly the canonical motif was visible (Marcotrigiano et al., 1999; Peter et al., 2015a). We observed that eIF4G binds to the lateral side of eIF4E, in addition to the known dorsal interface, using a binding mode similar to that employed by 4E-BPs (Figures 2F and 2G; Kinkelin et al., 2012; Peter et al., 2015a, 2015b; Sekiyama et al., 2015). The two metazoan structures are highly similar to each other (Figure S3D), indicating that the observed mode of interaction is widely conserved across species. However, an NMR structure of the *Sc* eIF4E-eIF4G complex (Gross et al., 2003; Figures 2H and S4B) originally suggested a different interaction mode in which the linker region and the non-canonical motif adopt different conformations and do not contact the lateral surface of eIF4E.

The different conformations of the eIF4G linker and non-canonical motif in the *Sc* and metazoan complexes may reflect the poor conservation of these sequences. However, it is important to note that the NMR structure of the *Sc* complex was determined in the presence of CHAPS (Gross et al., 2003). This zwitterionic detergent has been shown to bind to the lateral surface of *Sc* eIF4E, thus preventing its association with human 4E-BP2 (Matsuo et al., 1997). Consequently, the detergent may in fact interfere with hydrophobically driven contacts such as those of the non-canonical motifs of 4E-BPs and eIF4G with the lateral surface of eIF4E. Thus, the conformation of *Sc* eIF4G bound to eIF4E as observed in the structure of the *Sc* eIF4E-eIF4G complex might not represent the physiologically relevant state required for the initiation of cap-dependent translation.

Together with previous studies (Kinkelin et al., 2012; Peter et al., 2015a, 2015b; Sekiyama et al., 2015), our data reveal a conserved eIF4E recognition mode that is shared by eIF4G and 4E-BPs. This mode comprises a canonical and a non-canonical motifs that are connected by a linker region, which allows eIF4G and 4E-BPs to bind to the dorsal and lateral surfaces of eIF4E. The structural similarity indicates that eIF4G mimicry by 4E-BPs is not restricted to the canonical helix, but that it extends

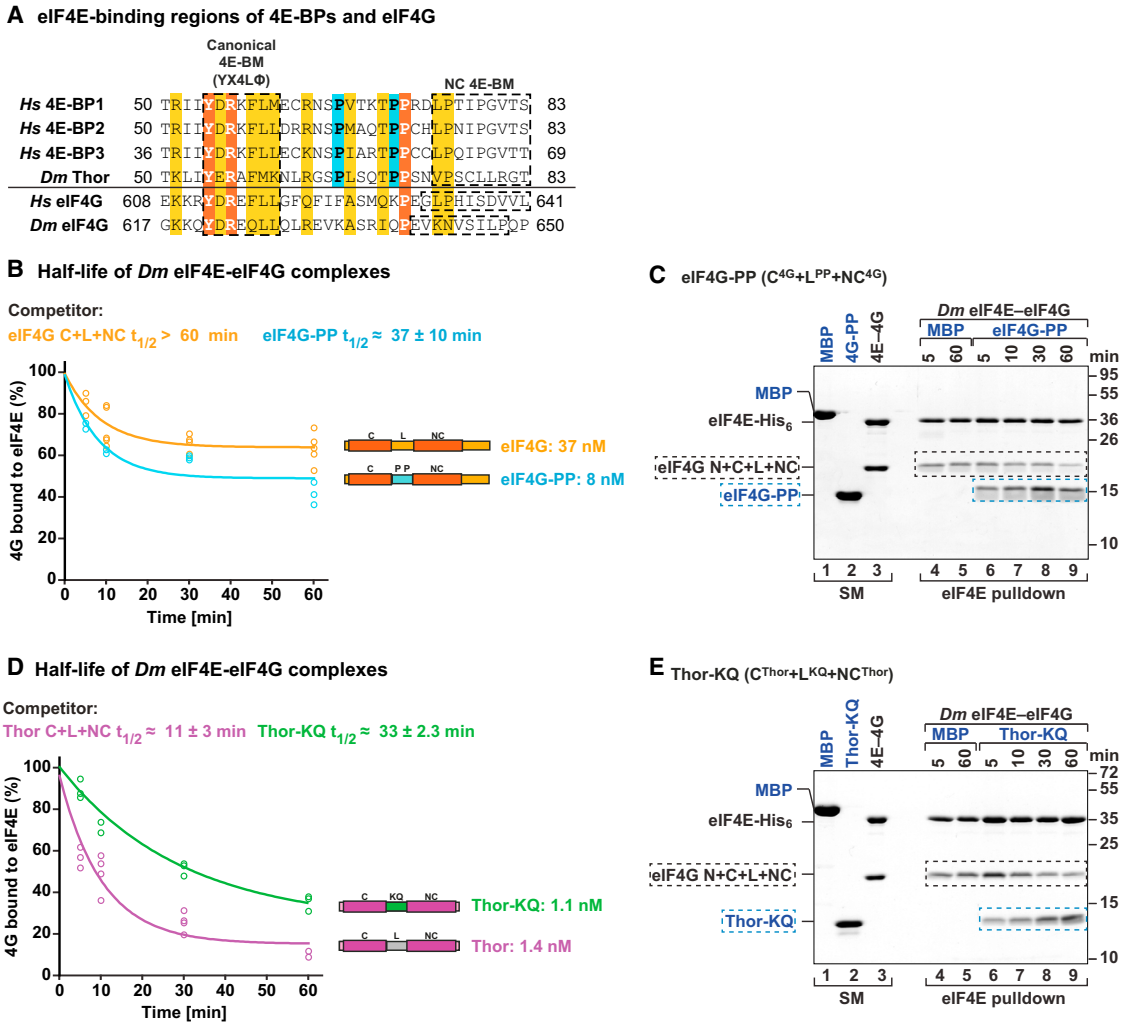


Figure 6. The Amino Acid Composition in the Linker Regions Modulates the Competitive Behavior of the 4E-BP and eIF4G Peptides

(A) Sequence alignment of the eIF4E-interacting regions of human 4E-BP1–3, *Drosophila* Thor, and eIF4G proteins. Canonical (C) and non-canonical (NC) motifs are boxed in black. The Pro residues present in the linker region of 4E-BP1–3 and Thor, but absent in eIF4G, are in bold and highlighted with a cyan background. The conserved residues and residues with >70% similarity are highlighted with an orange and yellow background, respectively.

(B–E) Competition assay. *Dm* eIF4E-eIF4G complexes were incubated in the presence of the indicated competitor peptides as described in Figure 5.

(B and D) Quantification of the amount of eIF4G remaining associated with eIF4E. The K_D values determined by ITC are indicated.

(C and E) Representative SDS-PAGE gels for each competition assay. See also Figure S2 and Table S2.

over the entire interface. This unexpected finding has important implications not only for the development of models to explain eIF4G and 4E-BP function in the cell, but also for the design of synthetic translation inhibitors.

Competition between eIF4G and 4E-BPs

The unexpected observation that the metazoan eIF4Gs and 4E-BPs both contact the dorsal and the lateral surface of eIF4E extends our structural understanding of the mechanism of competition by showing that the lateral surface of eIF4E is no longer exclusive to 4E-BPs and that eIF4G and 4E-BPs also compete on this surface. These interactions need to be considered when developing models of eIF4G-based translation regulation.

Although our results indicate that the linker and non-canonical motif of eIF4G contribute to the stability of the eIF4E-eIF4G complex, there are clear differences in the energetic contributions of individual sequence elements of eIF4G or 4E-BP (Thor)-derived peptides, suggesting that their respective molecular interaction mechanisms may be distinct despite the structural similarity. The linker and the non-canonical motifs of eIF4G both make substantially less significant contributions to the affinity of eIF4G for eIF4E than was observed for Thor. Furthermore, the linker between the canonical and non-canonical motif seems to be more flexible in eIF4G than in Thor, which would allow the two motifs to engage/disengage eIF4E more independently than in the case of Thor and in a consecutive manner. This flexibility has consequences for both the

binding affinity and the temporal stability of the complex that become apparent when the linker sequences are mutated. Additionally, in the bound state, the more flexible eIF4G peptides may sample more conformations, thus providing more opportunities for the 4E-BP competitors to access the transiently available lateral surface on eIF4E; such a mechanism provides a rationale for the competitive advantage of the 4E-BPs for binding to eIF4E.

The Lateral Surface of eIF4E Is a Key Target for Translational Inhibitors

The eIF4E-eIF4G interface is a key target for the development of small-molecule inhibitors that selectively disrupt this interaction, thereby leading to the arrest of cap-dependent translation initiation. These inhibitors can be used as therapeutic tools to repress translation in diseases that are associated with increased protein synthesis (Jia et al., 2012). The most extensively characterized inhibitor is 4EGI-1, which has been shown to displace eIF4G from eIF4E in a dose-dependent manner, while demonstrating a high degree of selectivity and no effect on the eIF4E interactions with 4E-BP1 (Moerke et al., 2007).

Surprisingly, however, 4EGI-1 was found to bind to the hydrophobic lateral surface of eIF4E, which made it difficult to explain how it would displace eIF4G from the dorsal surface. An allosteric mechanism has been proposed and is supported by NMR experiments and molecular dynamics simulations (Papadopoulos et al., 2014; Salvi et al., 2016; Sekiyama et al., 2015). Our finding that the eIF4G interaction extends to the lateral surface of eIF4E now provides an additional explanation for the inhibitory mechanism of 4EGI-1, wherein it directly competes with eIF4G for binding to the lateral surface (Figures S5D and S5E). It is conceivable that the affinity of 4EGI-1 is sufficient to displace the weakly binding non-canonical motif of eIF4G, thus selectively destabilizing the eIF4E-eIF4G complex and/or accelerating its dissociation, whereas the strongly binding non-canonical motif of 4E-BP1 with its rigidifying linker is substantially more resistant to such action.

Regarding the future design of small-molecule translational inhibitors, our studies indicate that a highly selective translational inhibitor that would target the interaction of eIF4G with eIF4E without affecting the interactions with 4E-BPs may be challenging to attain. For therapeutic applications, however, a highly potent small-molecule inhibitor that also displaces 4E-BPs should prove equally useful since this would still result in the reduction of eIF4E-dependent translation. Furthermore, 4E-BP activity has been shown to be frequently downregulated under pathological conditions (e.g., via hyperphosphorylation; Jia et al., 2012; Martineau et al., 2013); in this case, the synthetic inhibitor would effectively compete only with eIF4G for the lateral surface of eIF4E.

In conclusion, we have identified and characterized the key structural principles that govern the assembly and stability of the metazoan eIF4E-eIF4G complex. The identification of an extended and conserved eIF4E-eIF4G interface not only advances our mechanistic understanding of cap-dependent translation initiation, but also offers vantage points to guide the development of translational inhibitors with greater potency, with the prospect of using such molecules as therapeutic tools to target exacerbated eIF4E activity in diseases such as cancer and autism spectrum disorders (Gkogkas et al., 2013; Martineau et al., 2013).

EXPERIMENTAL PROCEDURES

DNA Constructs

All the mutants used in this study were generated by site-directed mutagenesis using the QuikChange mutagenesis kit (Stratagene). All mutants were confirmed by DNA sequencing and are listed in Table S1.

Protein Expression and Purification

All proteins were expressed in *E. coli* BL21 Star (DE3) cells (Invitrogen) grown in LB medium overnight at 20°C. A detailed description of the purification procedures is provided in the Supplemental Experimental Procedures.

Pull-Downs, ITC Analysis, and Competition Assays

The in vitro pull-down and competition assays and the ITC analysis were performed as described previously (Igreja et al., 2014; Peter et al., 2015a) and are described in the Supplemental Information.

Crystallization, Data Collection, and Structure Determination

The purified eIF4E-eIF4G complexes were subjected to crystallization trials directly after purification. When indicated, cap analog (m^7GpppG; NEB) was added to the purified eIF4E-eIF4G complexes prior to crystallization, at a 1.1- to 1.2-fold molar excess. Crystals of the *Hs* eIF4E (residues M1–V217)-*Hs* eIF4G (residues Q592–D653) complex were obtained after mixing 0.2 μL of the purified protein (18 mg/mL [540 μM] in 10 mM HEPES-NaOH [pH 7.5], 200 mM NaCl, and 2 mM DTT supplemented with m^7GpppG [648 μM]) with 0.2 μL of reservoir solution containing 0.1 M HEPES/NaOH (pH 7.25) and 25% (w/v) PEG6000. Crystal grew at 20°C over a 7-day period using the sitting-drop vapor diffusion method. Crystals of the *Dm* eIF4E (residues K69–L248)-*Dm* eIF4G (residues S601–N660) complex (18 mg/mL in 10 mM HEPES-NaOH [pH 7.5], 200 mM NaCl, and 2 mM DTT) were obtained by the hanging-drop vapor-diffusion method at 18°C. These crystals grew within 7 days after mixing 0.8 μL of protein solution with 0.8 μL of reservoir solution containing 0.1 M Tris-HCl (pH 8.25), 0.2 M $MgCl_2$, and 32% (w/v) PEG4000. The crystals containing the cap analog were obtained after mixing 0.2 μL of the purified protein (21 mg/mL [734 μM] in HEPES-NaOH [pH 7.5], 200 mM NaCl, and 2 mM DTT) in the presence of m^7GpppG (807 μM) with 0.2 μL of reservoir solution containing 0.1 M Tris-HCl (pH 8.0), 0.01 M $ZnCl_2$, and 20% (w/v) PEG6000. Crystals grew at 20°C in 3 to 5 days using the sitting-drop vapor diffusion method. All crystals were briefly transferred into mother liquor supplemented with 17%–20% (v/v) glycerol for cryoprotection and flash-frozen in liquid nitrogen.

Diffraction data for the *Hs* eIF4E-eIF4G complex bound to cap analog were recorded at a wavelength of 0.97857 Å on a DECTRIS PILATUS 6M detector at the PROXIMA 1 beamline of the SOLEIL synchrotron (France). The diffraction data for the *Dm* eIF4E-eIF4G complexes were collected at a wavelength of 0.9999 Å at 100 K and on a PILATUS 6M detector at the PXII beamline of the Swiss Light Source. A detailed description of the structure determination process can be found in the Supplemental Experimental Procedures. The stereochemical properties for all structures were verified using MOLPROBITY (Chen et al., 2010), and the structural images were prepared using PyMOL (http://www.pymol.org). Data collection and refinement statistics are summarized in Table 1.

ACCESSION NUMBERS

The accession numbers for the coordinates for the structures reported in this paper are PDB: 5T46 (*Hs* eIF4E–eIF4G–cap analog), 5T47 (*Dm* eIF4E–eIF4G), and 5T48 (*Dm* eIF4E–eIF4G–cap analog).

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, five figures, and four tables and can be found with this article online at http://dx.doi.org/10.1016/j.molcel.2016.09.020.

AUTHOR CONTRIBUTIONS

S.G. purified most of the proteins used in this study, performed and analyzed the ITC experiments, collected crystal diffraction data, and solved the structures under the supervision of E.V. and O.W. S.G., D.P., E.V., and O.W. analyzed structural data. R.W. and D.P. performed competition assays. S.G., R.W., D.P., L.W., and M.-Y.C. performed pull-down and coimmunoprecipitation assays under the supervision of C.I. E.I. oversaw the project as principal investigator. D.P. and C.I. initiated the project. C.I. coordinated the study. S.G., E.V., C.I., O.W., and E.I. wrote the manuscript with contributions from all authors.

ACKNOWLEDGMENTS

We thank R. Büttner and T. Raisch for setting up crystallization screens, F. Sandmeir for help in protein purification, A. Bläßle for help in data analysis, and C. Weiler for technical assistance. We acknowledge the staff at the PXII beamline of the Swiss Light Source for excellent support with data collection and also the beamline staff at Proxima 1 of SOLEIL Synchrotron for granting beam time and providing assistance with data collection. This work was supported by the Max Planck Society.

Received: June 10, 2016
Revised: August 22, 2016
Accepted: September 14, 2016
Published: October 20, 2016

REFERENCES

Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. *66*, 12–21.

Gkogkas, C.G., Khoutorsky, A., Ran, I., Rampakakis, E., Nevarko, T., Weatherill, D.B., Vasuta, C., Yee, S., Truitt, M., Dallaire, P., et al. (2013). Autism-related deficits via dysregulated eIF4E-dependent translational control. Nature *493*, 371–377.

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

Haghighat, A., Mader, S., Pause, A., and Sonenberg, N. (1995). Repression of cap-dependent translation by 4E-binding protein 1: competition with p220 for binding to eukaryotic initiation factor-4E. EMBO J. *14*, 5701–5709.

Hershey, P.E., McWhirter, S.M., Gross, J.D., Wagner, G., Alber, T., and Sachs, A.B. (1999). The Cap-binding protein eIF4E promotes folding of a functional domain of yeast translation initiation factor eIF4G1. J. Biol. Chem. *274*, 21297–21304.

Howard, A., and Rogers, A.N. (2014). Role of translation initiation factor 4G in lifespan regulation and age-related health. Ageing Res. Rev. *13*, 115–124.

Huang, F., and Nau, W.M. (2003). A conformational flexibility scale for amino acids in peptides. Angew. Chem. Int. Ed. Engl. *42*, 2269–2272.

Igreja, C., Peter, D., Weiler, C., and Izaurralde, E. (2014). 4E-BPs require non-canonical 4E-binding motifs and a lateral surface of eIF4E to repress translation. Nat. Commun. *5*, 4790.

Jackson, R.J., Hellen, C.U., and Pestova, T.V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. Nat. Rev. Mol. Cell Biol. *11*, 113–127.

Jia, Y., Polunovsky, V., Bitterman, P.B., and Wagner, C.R. (2012). Cap-dependent translation initiation factor eIF4E: an emerging anticancer drug target. Med. Res. Rev. *32*, 786–814.

Kinkelin, K., Veith, K., Grünwald, M., and Bono, F. (2012). Crystal structure of a minimal eIF4E-Cup complex reveals a general mechanism of eIF4E regulation in translational repression. RNA *18*, 1624–1634.

Lukhele, S., Bah, A., Lin, H., Sonenberg, N., and Forman-Kay, J.D. (2013). Interaction of the eukaryotic initiation factor 4E with 4E-BP2 at a dynamic bipartite interface. Structure *21*, 2186–2196.

MacArthur, M.W., and Thornton, J.M. (1991). Influence of proline residues on protein conformation. J. Mol. Biol. *218*, 397–412.

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.

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.

Martineau, Y., Azar, R., Bousquet, C., and Pyronnet, S. (2013). Anti-oncogenic potential of the eIF4E-binding proteins. Oncogene *32*, 671–677.

Matsuo, H., Li, H., McGuire, A.M., Fletcher, C.M., Gingras, A.C., Sonenberg, N., and Wagner, G. (1997). Structure of translation factor eIF4E bound to m7GDP and interaction with 4E-binding protein. Nat. Struct. Biol. *4*, 717–724.

Moerke, N.J., Aktas, H., Chen, H., Cantel, S., Reibarkh, M.Y., Fahmy, A., Gross, J.D., Degterev, A., Yuan, J., Chorev, M., et al. (2007). Small-molecule inhibition of the interaction between the translation initiation factors eIF4E and eIF4G. Cell *128*, 257–267.

Paku, K.S., Umenaga, Y., Usui, T., Fukuyo, A., Mizuno, A., In, Y., Ishida, T., and Tomoo, K. (2012). A conserved motif within the flexible C-terminus of the translational regulator 4E-BP is required for tight binding to the mRNA cap-binding protein eIF4E. Biochem. J. *441*, 237–245.

Papadopoulos, E., Jenni, S., Kabha, E., Takrouri, K.J., Yi, T., Salvi, N., Luna, R.E., Gavathiotis, E., Mahalingam, P., Arthanari, H., et al. (2014). Structure of the eukaryotic translation initiation factor eIF4E in complex with 4EGI-1 reveals an allosteric mechanism for dissociating eIF4G. Proc. Natl. Acad. Sci. USA *111*, E3187–E3195.

Peter, D., Igreja, C., Weber, R., Wohlbold, L., Weiler, C., Ebertsch, L., Weichenrieder, O., and Izaurralde, E. (2015a). Molecular architecture of 4E-BP translational inhibitors bound to eIF4E. Mol. Cell *57*, 1074–1087.

Peter, D., Weber, R., Köne, C., Chung, M.Y., Ebertsch, L., Truffault, V., Weichenrieder, O., Igreja, C., and Izaurralde, E. (2015b). Mextli proteins use both canonical bipartite and novel tripartite binding modes to form eIF4E complexes that display differential sensitivity to 4E-BP regulation. Genes Dev. *29*, 1835–1849.

Salvi, N., Papadopoulos, E., Blackledge, M., and Wagner, G. (2016). The role of dynamics and allostery in the inhibition of the eIF4E/eIF4G translation initiation factor complex. Angew. Chem. Int. Ed. Engl. *55*, 7176–7179.

Sekiyama, N., Arthanari, H., Papadopoulos, E., Rodriguez-Mias, R.A., Wagner, G., and Léger-Abraham, M. (2015). Molecular mechanism of the dual activity of 4EGI-1: Dissociating eIF4G from eIF4E but stabilizing the binding of unphosphorylated 4E-BP1. Proc. Natl. Acad. Sci. USA *112*, E4036–E4045.

Umenaga, Y., Paku, K.S., In, Y., Ishida, T., and Tomoo, K. (2011). Identification and function of the second eIF4E-binding region in N-terminal domain of eIF4G: comparison with eIF4E-binding protein. Biochem. Biophys. Res. Commun. *414*, 462–467.

Molecular Cell, Volume *64*

Supplemental Information

The Structures of eIF4E-eIF4G Complexes Reveal an Extended Interface to Regulate Translation Initiation

Stefan Grüner, Daniel Peter, Ramona Weber, Lara Wohlbold, Min-Yi Chung, Oliver Weichenrieder, Eugene Valkov, Cátia Igreja, and Elisa Izaurralde

SUPPLEMENTAL FIGURES

Figure S1

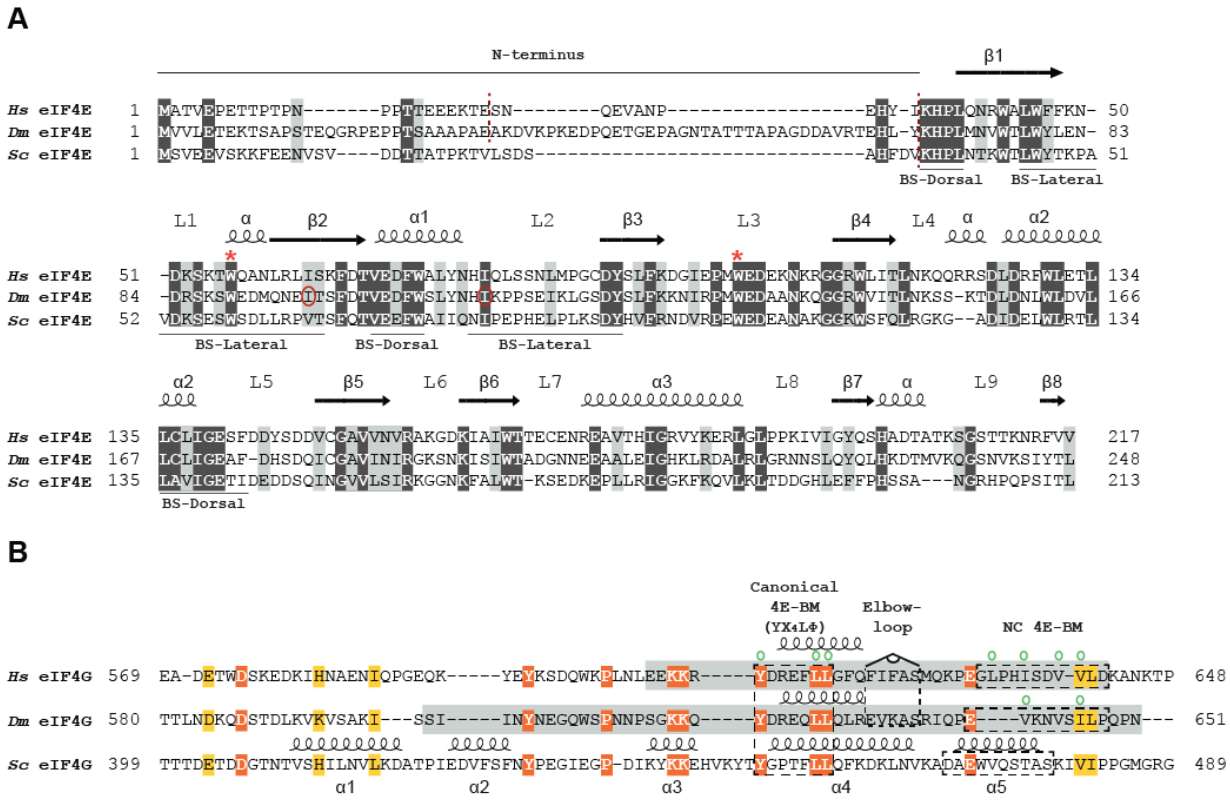


Figure S1, related to Figure 1. Sequence alignments

(A) Structure-based sequence alignment of eIF4E orthologous proteins from *Homo sapiens* (*Hs*), *Drosophila melanogaster* (*Dm*) and *Saccharomyces cerevisiae* (*Sc*). Conserved residues are highlighted with a black background and printed in white. Residues with >70% similarity are shown with a gray background. The Trp residues coordinating the mRNA cap structure are indicated by red asterisks. The residues mutated in this study are indicated by red circles. The lateral and dorsal binding surfaces (BS) are indicated by a line below the sequences. The positions of the N-terminal deletions on the different eIF4E proteins are indicated by vertical red dashed lines. Secondary structure elements of *Hs* eIF4E are indicated above the sequences with the exception of L9 and β8, which are indicated for *Dm* eIF4E because the corresponding residues are not defined in our human eIF4E structure.

(B) Sequence alignment of the eIF4E-interacting regions of eIF4G orthologous proteins. Species abbreviations are as in panel (A). The eIF4G sequences visible in the crystal structures are highlighted by a light gray background. Conserved residues are highlighted with an orange background and printed in white. Residues with >70% similarity are shown with a yellow background. The canonical (C) and non-canonical (NC) motifs are boxed in black. Key eIF4G residues within the motifs contacting eIF4E and mutated in this study are indicated by green open circles. The residues forming the elbow loops in the *Hs* and *Dm* proteins are indicated by a triangle above the alignment. The secondary structure elements indicated above the *Sc* sequences are based on the structure of the *Sc* eIF4E–eIF4G complex [PDB ID: 1RF8; Gross et al., 2003).

Figure S2

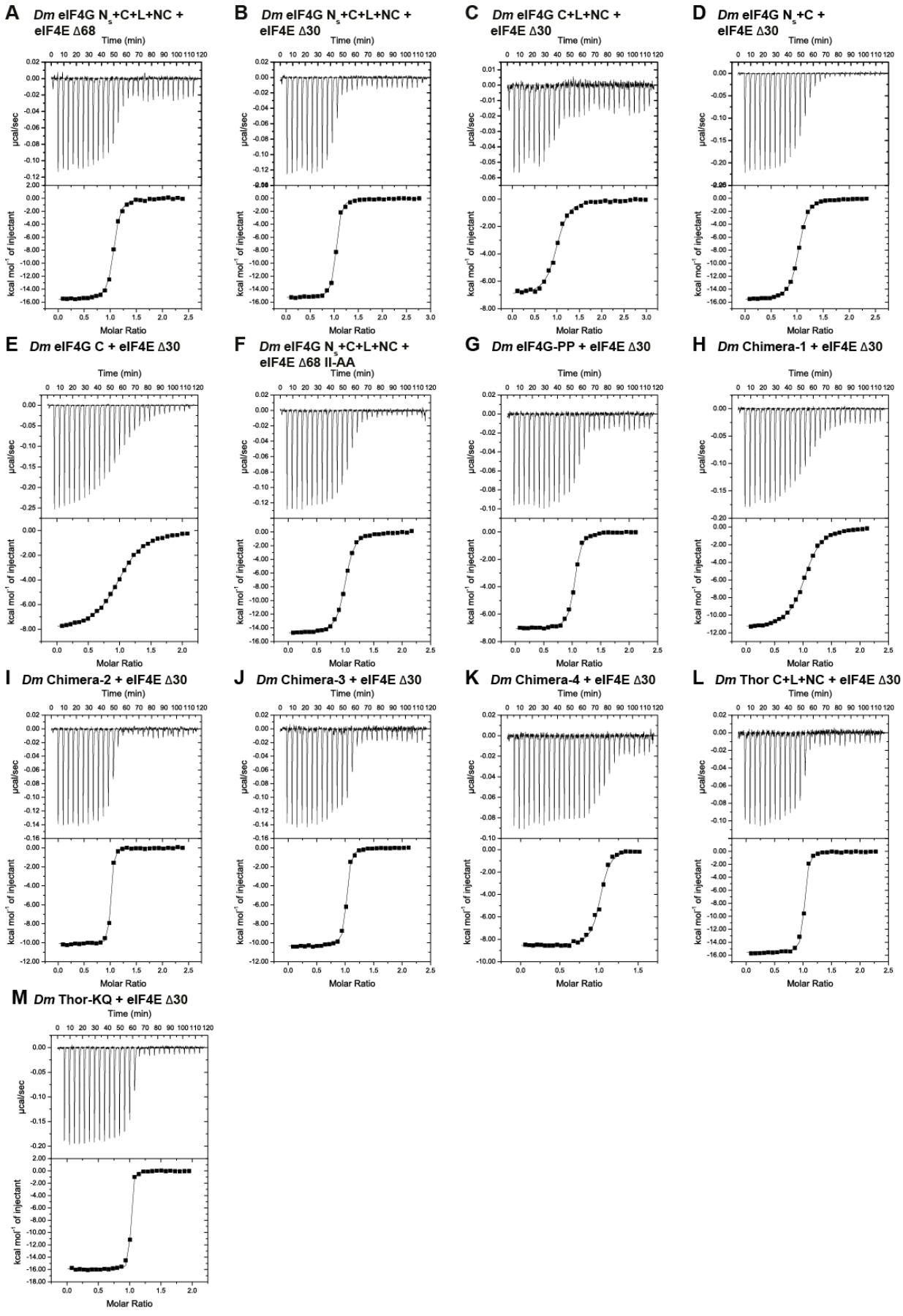

A Dm eIF4G Ns+C+L+NC + eIF4E Δ68
B Dm eIF4G Ns+C+L+NC + eIF4E Δ30
C Dm eIF4G C+L+NC + eIF4E Δ30
D Dm eIF4G Ns+C + eIF4E Δ30
E Dm eIF4G C + eIF4E Δ30
F Dm eIF4G Ns+C+L+NC + eIF4E Δ68 II-AA
G Dm eIF4G-PP + eIF4E Δ30
H Dm Chimera-1 + eIF4E Δ30
I Dm Chimera-2 + eIF4E Δ30
J Dm Chimera-3 + eIF4E Δ30
K Dm Chimera-4 + eIF4E Δ30
L Dm Thor C+L+NC + eIF4E Δ30
M Dm Thor-KQ + eIF4E Δ30
Time (min)
µcal/sec
kcal mol-1 of injectant
Molar Ratio

Figure S2, related to Figure 1, Figure 5, Figure 6 and Table S2. Isothermal titration calorimetry (ITC) data for the interaction of *Dm* eIF4E with the indicated peptides

(A to M) Representative isothermal titration calorimetry (ITC) thermograms for the interactions of *Dm* eIF4E proteins (Δ30, Δ68, wild-type or II-AA mutant) with the indicated *Dm* eIF4G, *Dm* Thor and Chimera peptides C-terminally fused to GB1. The fitted thermodynamic parameters are shown in Table S2. Upper panels show raw data in (μcal sec^{-1}), and lower panels represent the integration of heat changes associated with each injection (kcal mol^{-1} of injectant). Data were fitted using a one-site binding model.

Figure S3

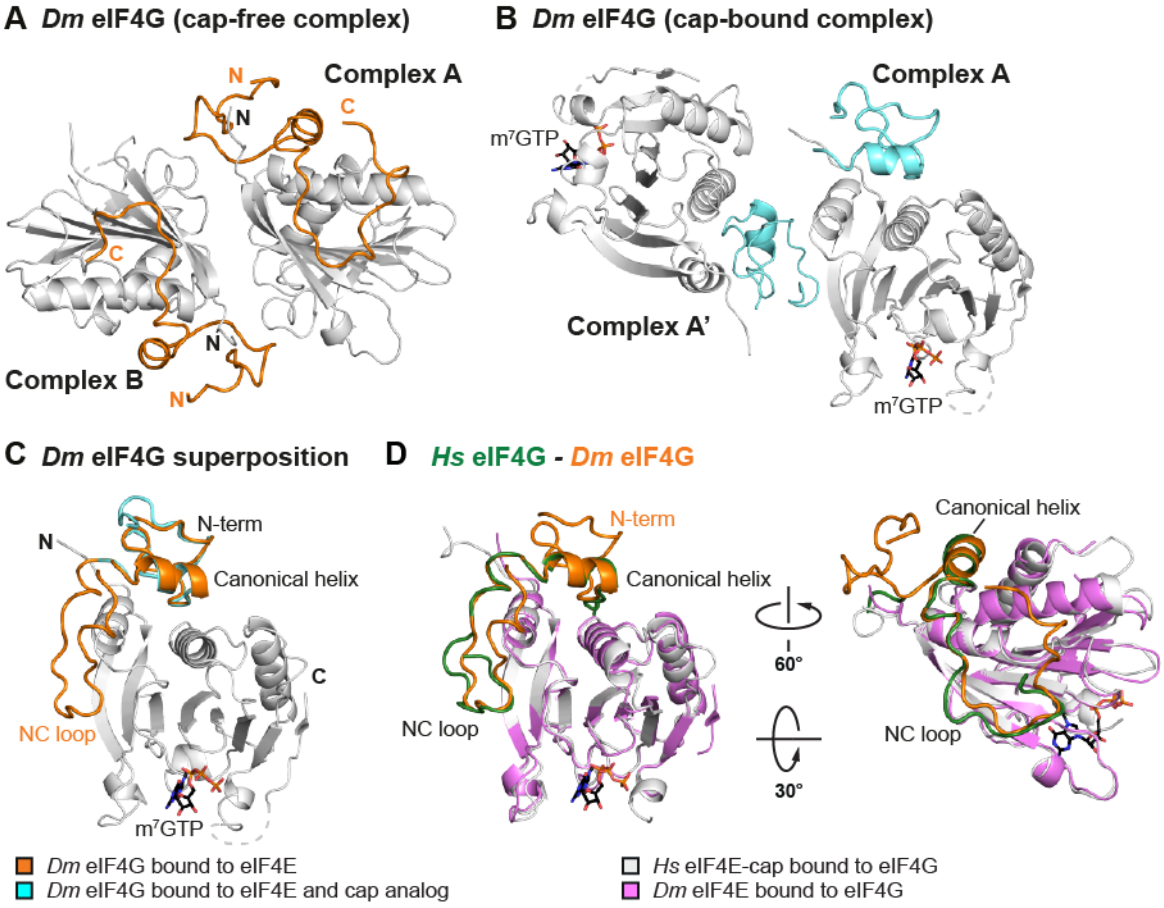


Figure S3, related to Figure 2. Crystal structures and packing of *Dm* eIF4E–eIF4G complexes

(A) Cartoon representation of the *Dm* eIF4E–eIF4G complex in the absence of cap analog. The asymmetric unit of the crystal contains two copies of the complex (complex I, chains A and B and complex II chains C and D). The non-canonical motif was better defined in complex II. eIF4E is shown in gray and eIF4G in orange.

(B) Schematic representation of the crystal form of the *Dm* eIF4E–eIF4G complex in the presence of cap analog m^7GpppG. Only the electron density for the m^7GTP-moiety was visible. eIF4E is shown in gray and eIF4G in cyan. The asymmetric unit contained one complex (complex I). In this crystal form, the non-canonical motif of eIF4G is not visible in the electron density. Instead, the expression tag and the N-terminal region of the eIF4G molecule from a symmetry-related complex (complex I') occupy the lateral surface of eIF4E, with residues

H599-N604 from the symmetry mate taking the position of residues N643-P648 as observed in PDB-ID 5T47.

(C) Structural superposition of the *Dm* eIF4E–eIF4G complexes with and without cap analog. The RMSD is 0.33 Å over 127 Cα atoms of eIF4E. Cap-free eIF4E is omitted for clarity. The N-terminal region and the canonical motif of eIF4G are highly similar between the two crystal forms. Selected secondary structure elements are labeled in black for eIF4E and in color for eIF4G.

(D) Overlay of *Dm* and *Hs* eIF4E–eIF4G complexes. The RMSD is 0.39 Å over 123 Cα atoms of eIF4E. *Hs* and *Dm* eIF4G are colored in green and orange, respectively. *Hs* and *Dm* eIF4E are colored in gray and purple, respectively.

Figure S4

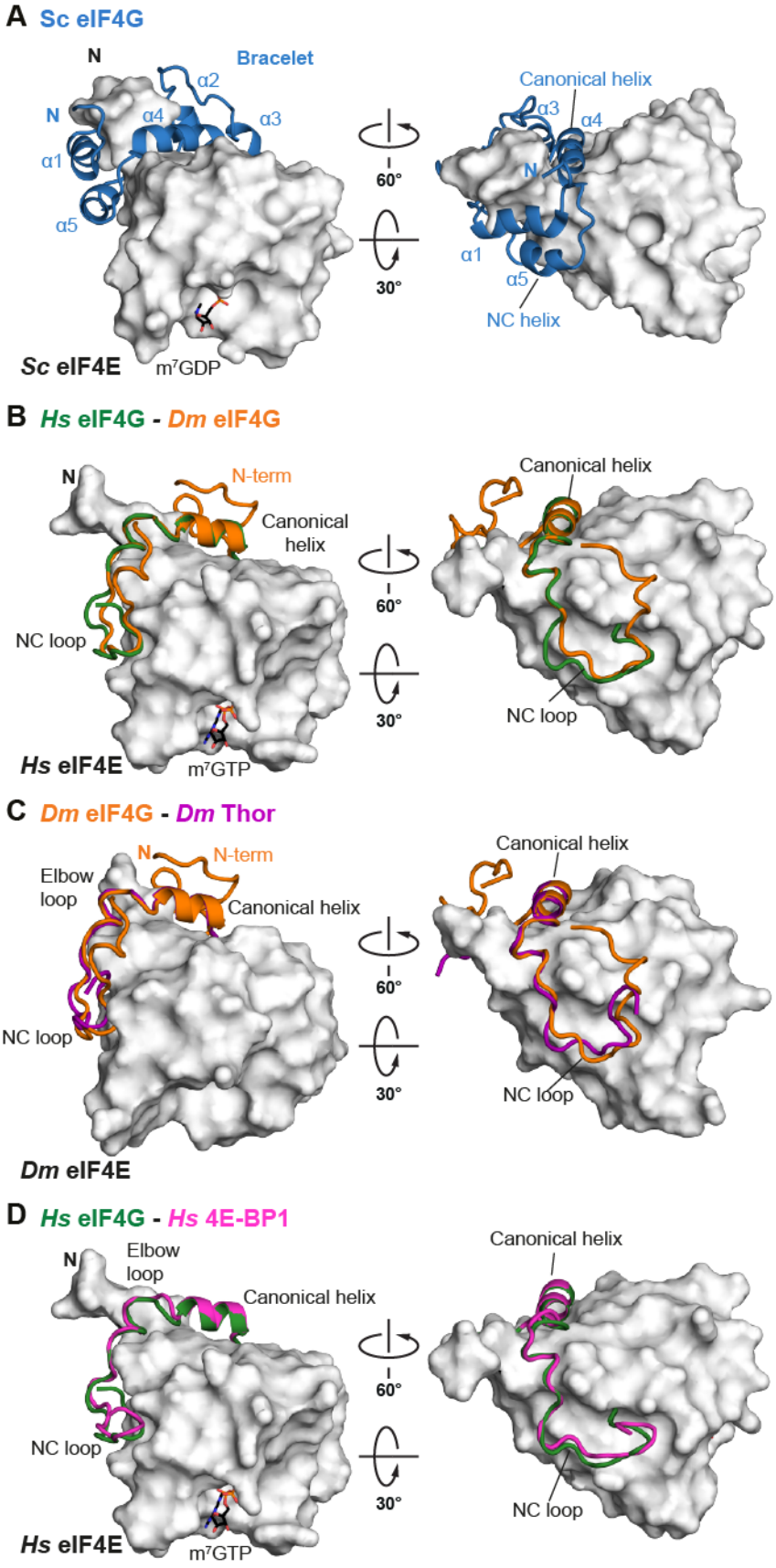

A Sc eIF4G
N
Bracelet
α2
N
α4
α3
α1
α5
Sc eIF4E
m7GDP
60°
30°
Canonical helix
α3
α4
N
α1
α5
NC helix
B Hs eIF4G - Dm eIF4G
N
N-term
Canonical helix
NC loop
Hs eIF4E
m7GTP
60°
30°
Canonical helix
NC loop
C Dm eIF4G - Dm Thor
N
N-term
Elbow loop
Canonical helix
NC loop
Dm eIF4E
60°
30°
Canonical helix
NC loop
D Hs eIF4G - Hs 4E-BP1
N
Elbow loop
Canonical helix
NC loop
Hs eIF4E
m7GTP
60°
30°
Canonical helix
NC loop

Figure S4, related to Figure 3. Structural comparisons of the fungal and metazoan eIF4E–eIF4G and eIF4E–4E-BP complexes

(A) Surface representations of *Sc* eIF4E in complex with *Sc* eIF4G (PDB ID: 1RF8; Gross et al., 2003). eIF4G is shown in blue. The m^7GDP ligand is shown as sticks. Residues M1–A21 of *Sc* eIF4E and I391–T404 of *Sc* eIF4G are omitted for clarity.

(B) Superposition of eIF4E in complex with *Hs* and *Dm* eIF4G. The eIF4E molecule from the *Dm* complex is omitted for clarity. The eIF4G molecules are shown as cartoon representation with *Hs* eIF4G colored in green and *Dm* eIF4G in orange.

(C) Comparison of the *Dm* eIF4E-eIF4G complex (this study) with the *Dm* eIF4E–Thor complex (PDB ID: 4UE8; Peter et al., 2015a). *Dm* eIF4G and Thor are colored in orange and purple, respectively.

(D) Superposition of the *Hs* eIF4E–eIF4G complex (this study) with the *Hs* eIF4E–4E-BP1 complex (PDB ID: 4UED; Peter et al., 2015a). *Hs* eIF4G and *Hs* 4E-BP1 are colored in green and magenta, respectively.

Figure S5

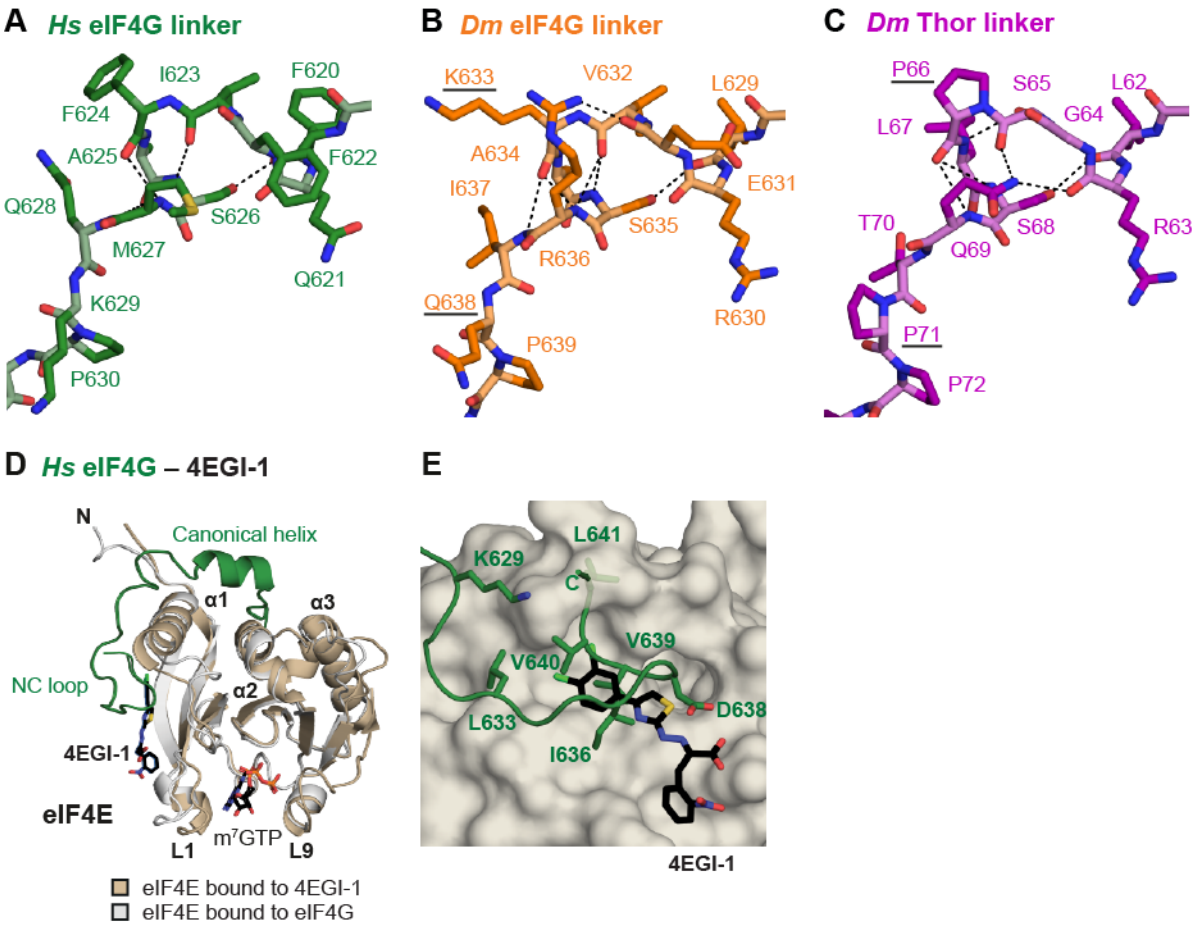


Figure S5, related to Figure 4. Metazoan eIF4G linker interactions and modulation of *Hs* eIF4E–eIF4G interactions by 4EGI-1

(A–C) Close-up views of the *Hs* (A) and *Dm* (B) eIF4G elbow loops compared to the *Dm* Thor elbow loop (C) (PDB ID: 4UE8; Peter et al., 2015a). Selected residues are shown as sticks. Hydrogen bonds are indicated by dashed lines. Pro residues in the Thor linker region and the corresponding residues in the *Dm* eIF4G linker mutated in this study are underlined.

(D) Superposition of the *Hs* eIF4E-eIF4G structure with the structure of eIF4E bound to the 4EGI-1 inhibitor (PDB ID: 4TPW; Papadopoulos et al., 2014). The eIF4E and eIF4G molecules are shown in cartoon representation, with selected secondary structure elements labeled in black for eIF4E and green for eIF4G. The 4EGI-1 inhibitor is shown as black sticks.

(E) Close-up view of superposed eIF4E molecules bound to either eIF4G or 4EGI-1, with eIF4E shown in surface representation. Selected residues from the non-canonical motif of eIF4G and 4EGI-1 are shown as sticks.

SUPPLEMENTAL TABLES

Table S1 related to Figure 1 and 4–6. Mutants and constructs used in this study

Protein	Name of the construct	Fragments / mutations	Comments
***Dm* eIF4E** (1–248) (isoform C) P48598-2	4E	Full-length (1–248)	
	II-AA	I96A, I112A	Lateral surface mutant
	Δ30	31–248	
	Δ68	69–248	eIF4E used for crystallization
***Dm* eIF4G** (1–1666) (isoform A) O61380	N+C+L+NC	578–675	Extended eIF4E-binding region
	N_s+C	601–634	N-term short + canonical
	N_L+C	578–634	N-term long + canonical (pulldown)
	C+L+NC	617–675	Canonical +linker + non-canonical
	C	617–634	Canonical
	N_S+C+L+NC	601–675	eIF4E-binding region
	N_S+C+L+NC	601–660	eIF4E-binding region used for crystallization
	C*	Y621A, L626A, L627A	Canonical mutant
	NC*	V641A, I646A	Non-canonical mutant
	eIF4G-PP	617–675 (K633P, Q638P)	Linker mutant
***Hs* eIF4E** (1–217) P06730	4E	Full–length (1–127)	eIF4E used for crystallization
	Δ23	24–217	
	Δ35	36–217	
***Hs* eIF4G Isoform A** (1–1599) Q04637	N+C+L+NC	592–653	eIF4E-binding region used for crystallization
	N_L+C+L+NC	571–676	eIF4E-binding region (pulldown)
	C+L+NC	608–647	Canonical +linker + non-canonical
	C*	Y612A, L617A, L618A	Canonical
	NC*	L633A, V639A	Non-canonical
Thor (1–117) Q9XZ56	C+L+NC	50–83	eIF4E-binding region
	Thor-KQ	50–83 P66K, P71Q	
C^{Thor}+L^{4G}+NC^{4G}	Chimera 1	*Dm* Thor 50–63 + *Dm* eIF4G 631–675	
C^{4G}+L^{Thor}+NC^{Thor}	Chimera 2	*Dm* eIF4G 617–630 + *Dm* Thor 64–83	
C^{Thor}+L^{4G}+NC^{Thor}	Chimera 3	*Dm* Thor 50–63 + *Dm* eIF4G 631–638 + *Dm* Thor NC 72–83	
C^{4G}+L^{Thor}+NC^{4G}	Chimera 4	*Dm* eIF4G 617–630 + *Dm* Thor 64–71+*Dm* eIF4G 639–675	

For the *in vitro* experiments, *Dm* eIF4G, *Dm* Thor and the chimeric peptides were tagged C-terminally with GB1 (see Supplemental Experimental Procedures).

Table S2, related to Figure 5, 6 and S2. Thermodynamic parameters for the interaction of *Dm* eIF4E with the indicated eIF4G, Thor and Chimera peptides

Peptide	K_D (nM)	ΔH (kcal mol^{-1})	-TΔS (kcal mol^{-1})	ΔG (kcal mol^{-1})	Molar ratio
eIF4G peptide + eIF4E Δ68					
eIF4G N_s+C+L+NC	5 ± 1.4	-15.0 ± 0.5	3.9 ± 0.6	-11.1 ± 0.2	1.02 ± 0.02
eIF4G or Thor peptide + eIF4E Δ68 II-AA					
eIF4G N_s+C+L+NC	16 ± 2.6	-15.1 ± 0.2	4.6 ± 0.3	-10.4 ± 0.1	0.99 ± 0.02
*Thor C+L+NC	470 ± 30	-7.6 ± 0.3	-0.9	-8.5	0.97 ± 0.01
eIF4G peptide + eIF4E Δ30					
eIF4G N_s+C+L+NC	4.2 ± 0.2	-15.2 ± 0.2	4.0 ± 0.2	-11.2 ± 0.1	1.01 ± 0.01
eIF4G C+L+NC	37 ± 6.5	-7.8 ± 0.6	-2.1 ± 0.6	-10.0 ± 0.1	1.00 ± 0.04
eIF4G N_s+C	25 ± 2.9	-15.8 ± 0.3	5.6 ± 0.3	-10.2 ± 0.1	1.00 ± 0.01
eIF4G C	380 ± 69	-7.9 ± 0.1	-0.7 ± 0.2	-8.6 ± 0.1	1.00 ± 0.01
eIF4G-PP	8 ± 3.5	-6.8 ± 0.2	-4.1 ± 0.3	-10.8 ± 0.2	1.00 ± 0.01
Chimeric peptides + eIF4E Δ30					
Chimera 1 C^{Thor}+L^{4G}+NC^{4G}	109 ± 8	-11.5 ± 0.1	2.2 ± 0.1	-9.3 ± 0.1	1.01 ± 0.01
Chimera 2 C^{4G}+L^{Thor}+NC^{Thor}	3 ± 1.4	-10.0 ± 0.9	-1.4 ± 0.9	-11.4 ± 0.3	0.99 ± 0.01
Chimera 3 C^{Thor}+L^{4G}+NC^{Thor}	4.2 ± 0.9	-10.1 ± 0.3	-1.1 ± 0.3	-11.2 ± 0.1	1.00 ± 0.02
Chimera 4 C^{4G}+L^{Thor}+NC^{4G}	13 ± 1.6	-7.4 ± 1.2	-3.2 ± 1.2	-10.6 ± 0.1	1.00 ± 0.01
Thor peptide + eIF4E Δ68					
*Thor C+L+NC	1.4 ± 0.3	-16.8 ± 1	4.9	-11.9	0.95 ± 0.02
*Thor C	2260 ± 60	-12.1 ± 2.8	2.2	-9.8	1.06 ± 0.01
Thor peptide + eIF4E Δ30					
Thor C+L+NC	1.4 ± 0.5	-15.7 ± 0.5	3.9 ± 0.5	-11.9 ± 0.2	0.99 ± 0.01
Thor-KQ	1.1 ± 0.5	-15.5 ± 0.4	3.5 ± 0.4	-12.0 ± 0.2	1.00 ± 0.01

Values are averages of at least three experiments and errors represent standard deviations. All fragments contained a C-terminal GB1-tag connected via a GSASG linker to the respective protein fragment (see Supplemental Experimental Procedures). * Taken from Igreja et al., (2014).

Table S3, related to Figure 3. Binding surfaces of eIF4E–eIF4G and eIF4E–4E-BP1 complexes

Species	Ligand	eIF4G or 4E-BP peptides	Total buried surface area in Å^2
Hs	eIF4G	All (607–642)	2873
		Canonical (607–621)	1398
		Linker (622–629)	707
		Non-canonical (630–642)	995
	4E-BP1 (4UED.pdb)	All (49–83)	2441
		Canonical (50–63)	1366
		Linker (64–71)	452
		Non-canonical (72–83)	780
Dm	eIF4G	All (601–656)	2988
		N-terminus (601–616)	394
		Canonical (617–630)	1111
		Linker (631–638)	477
		Non-canonical (639–656)	1227
	Thor (4UE8.pdb)	All (47–83)	2827
		Canonical (50–63)	1434
		Linker (64–71)	659
		Non-canonical (72–83)	964

Buried surface areas were calculated using the PISA server at the European Bioinformatic Institute (http://www.ebi.ac.uk/msd-srv/prot_int/cgi-bin/piserver). For calculation of the buried surface area of *Dm* and *Hs* eIF4E–eIF4G chains C and D were used in both cases.

Table S4, related to Figures 4. Antibodies used in this study

Antibody	Source	Catalog Number	Dilution	Monoclonal/ Polyclonal
Anti-HA-HRP (Western blot)	Roche	12 013 819 001	1:5,000	Monoclonal
Anti-HA (Immunoprecipitation)	Covance	MMS-101P	1:1,000	Monoclonal
Anti-*Dm* eIF4E	In house		1:3,000	Rabbit polyclonal
Anti-*Hs* eIF4E	BETHYL laboratories	A301-154A	1:2,000	Rabbit polyclonal
Anti-*Hs* eIF4G	BETHYL laboratories	A301-776A	1:2,000	Rabbit polyclonal
Anti-GFP	In house		IP	Rabbit polyclonal
Anti-GFP	Roche	11814460001	1 :2,000	Monoclonal
Anti-rabbit-HRP	GE Healthcare	NA934V	1:10,000	Polyclonal
Anti-mouse-HRP	GE Healthcare	RPN4201	1:10,000	Polyclonal

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

DNA constructs

For expression in *E. coli*, a DNA fragment encoding *Hs* eIF4E (residues M1–V217) was inserted into the XhoI and NheI restriction sites of the pnYC vector, to express untagged eIF4E, and of the pnYC-NpH vector (Diebold et al., 2011), to express *Hs* eIF4E with an N-terminal His_6-tag cleavable by a HRV3C protease. The same DNA fragment was also inserted into the XhoI and BamHI sites of the pnYC-CvH vector to express *Hs* eIF4E with a C-terminal His_6-tag cleavable by TEV protease. The DNA fragments encoding *Hs* eIF4E N-terminal deletions (residues S24–V217 and K36–V217) were inserted into the NdeI and NheI restriction sites of the pnYC vector. DNA fragments encoding *Dm* eIF4E (residues M1–L248, A31–L248 and K69–L248) were inserted into pnYC-NpH and pnYC-CvH as described above. *Dm* eIF4G DNA fragments were inserted into the NdeI and XhoI restriction sites of the pnEA-NpM vector for expression with an N-terminal maltose binding protein (MBP)-tag cleavable by HRV3C protease. DNA fragments encoding *Hs* eIF4G and chimeric peptides were also inserted into the pnEA-NpM vector using the BamHI and NheI restriction sites (Table S1). The *Dm* and *Hs* eIF4G DNA fragments were also inserted into a modified pnEA-NpM vector additionally encoding a C-terminal fusion of the B1 domain of immunoglobulin-binding protein G (GB1; Cheng and Patel, 2004), connected by a GSASG linker sequence. The plasmid used for the expression of *Dm* Thor 50–83-GB1 has been described previously (Igreja et al., 2014). The plasmids used for the expression of HA-tagged *Dm* eIF4G in S2 cells and of GFP-tagged *Hs* eIF4E-BP1 or HA-tagged *Hs* eIF4E in human cells, have also been described previously (Peter et al., 2015a). The plasmids for the expression of GFP-tagged full-length (FL) *Hs* eIF4G or the corresponding eIF4E-binding region (residues E608–T647) were obtained by inserting the corresponding cDNAs between the XhoI and HindIII (FL) and XhoI and BamHI sites of the

pT7-EGFP-C1 vector, respectively. All constructs used in this study are listed in Supplementary Table S1.

Protein expression and purification

All proteins were expressed in *E. coli* BL21 Star (DE3) cells (Invitrogen) grown in LB medium overnight at 20°C. To purify the eIF4E–eIF4G complexes used in crystallization, N-terminally His_6-tagged *Dm* eIF4E (residues K69–L248) or *Hs* eIF4E (residues M1–V217) was co-expressed with MBP-tagged *Dm* eIF4G (residues S601–N660) or *Hs* eIF4G (residues Q592–D653), respectively. For competition assays, C-terminally His_6-tagged full length *Dm* eIF4E was coexpressed with MBP-tagged *Dm* eIF4G (residues P578–N675) fused C-terminally to the GB1 domain.

All protein purification steps were performed at 4°C. The cells were lysed by sonication or using the Avestin Emulsiflex-C3 homogenizer in lysis buffer containing 20 mM Na-phosphate (pH 7.5), 300 mM NaCl and 2 mM dithiothreitol (DTT) supplemented with lysozyme (1 mg/mL), DNase I (5 μg/mL) and protease inhibitor cocktail (Roche). All complexes were purified from the cleared lysates using amylose resin (New England Biolabs) followed by overnight cleavage of tags with HRV3C protease and dialysis into 20 mM HEPES-NaOH (pH 7.5), 100 mM NaCl and 2 mM DTT. The complexes were further purified using a HiTrap Heparin column (GE Healthcare) and eluted using a gradient from 100 to 500 mM NaCl. In the case of the *Hs* eIF4E–eIF4G complex, an additional purification on amylose resin and SP cation exchange chromatography on a HiTrap SP column (GE Healthcare) was performed. As a final step, all complexes were subjected to size exclusion chromatography on a Superdex 75 column (GE Healthcare) equilibrated in 10 mM HEPES-NaOH (pH 7.5), 200 mM NaCl and 2 mM DTT (for complexes used in crystallization) or 10 mM Na-phosphate (pH 7.5), 200 mM NaCl and 2 mM DTT (for complexes used in the competition assays). The complexes used in

the competition assays were additionally supplemented with 5% (v/v) glycerol and all proteins were stored at -80°C after flash-freezing in liquid nitrogen.

The C-terminally GB1-tagged peptides used in the ITC measurements and competition assays were purified as described above with the following changes. After removal of the MBP tag, the proteins were diluted to 20 mM HEPES-NaOH (pH 7.5), 75 mM NaCl and 2 m DTT and subjected to anion exchange chromatography on a HiTrap Q column (GE Healthcare) to remove the cleaved MBP. The proteins were further purified by size exclusion chromatography on a Superdex 75 column equilibrated in 10 mM Na-phosphate (pH 7.5), 150 mM NaCl and 2 mM DTT, followed by purification over amylose resin to remove the remaining MBP. The Thor-GB1 peptide used in the competition assays was purified as described previously (Igreja et al., 2014).

For the ITC measurements, eIF4E proteins were expressed with an N-terminal His_6-tag and the cells were lysed in lysis buffer containing 25 mM imidazole. The cell lysates were applied to a Ni^{2+}-charged IMAC column (GE Healthcare) and eluted with a gradient of 25-500 mM imidazole. After the cleavage of the His_6-tag with HRV3C protease, the proteins were dialyzed in 20 mM HEPES-NaOH (pH 7.5), 100 mM NaCl and 2 mM DTT and further purified using a HiTrap Heparin column (GE Healthcare). The *Dm* eIF4E proteins were subjected to a final purification on a Superdex 75 (GE Healthcare) in 10 mM Na-phosphate (pH 7.5), 150 mM NaCl and 2 mM dithiothreitol (DTT). Note that the precision protease treatment to remove the N-terminal His_6- and MBP-tags leaves behind four extra amino acids (GPHM) from the expression vector at the N-terminus of the proteins.

Data collection and structure determination

Diffraction data were processed using XDS (Kabsch, 2010) and scaled using AIMLESS as part of the CCP4 package (Evans, 2006; Evans, 2011; Winn et al., 2011). All structures were solved

by molecular replacement using PHASER (McCoy et al., 2007). In the case of the *Hs* eIF4E–eIF4G complex bound to cap analog, the human eIF4E structure (PDB 4TPW, Papadopoulos et al., 2014) was used as a search model. To solve the structures of the *Dm* eIF4E–eIF4G complexes, molecular replacement was performed using the *Dm* eIF4E structure (PDB 4UE8; Peter et al., 2015a) as a search model. The molecular replacement solutions were used to rebuild the initial models of all complexes using ARP/wARP (Langer et al., 2008). To complete and improve the initial models, iterative cycles of model building and refinement were performed using COOT (Emsley et al., 2010) and PHENIX (Afonine et al., 2012), respectively. Final refinement rounds were carried out with non-crystallographic symmetry (NCS) torsional restraints for both the *Hs* eIF4E–eIF4G complex bound to cap analog (PDB ID: 5T46) and the *Dm* eIF4E–eIF4G complex (PDB ID: 5T47). For the *Hs* eIF4E–eIF4G complex, atoms other than water were refined with anisotropic individual B-factors. For the *Dm* eIF4E–eIF4G complexes (PDB IDs: 5T47, 5T48), translation/libration/screw (TLS) parameters were refined for the peptide chains in addition to individual B-factors. Ligand restraints for refinement were generated using GRADE (Smart et al., 2012). The unmodified guanosine base of the m^7GpppG cap analog used in the crystallization condition was not visible in electron density map. Therefore the cap analog was modeled as m^7GTP.

The final model of the *Hs* eIF4E–eIF4G complex with to cap analog comprises two eIF4E–eIF4G complexes, two m^7GTP, seven glycerol and 346 water molecules in the asymmetric unit. The following residues are missing in the model: eIF4E M1–N30 and S207–V217 (chain A), eIF4E M1–A29 and S207–V217 (chain C), eIF4G Q592–E607 and K643–D653 (chain B) and eIF4G Q592–L606 and K643–D653 (chain D). The final model of the *Dm* eIF4E–eIF4G complex comprises two eIF4E–eIF4G complexes, two glycerol and 157 water molecules in the asymmetric unit. The following residues are missing from the model: eIF4E S86–W89 and T234–N241 (chain A), eIF4E R222–N223 and T234–K243 (chain C), eIF4G S601–I602 and

Q649–N660 (chain B) and eIF4G I657–N660 (chain D). The final model of the *Dm* eIF4E–eIF4G complex with cap analog comprises one eIF4E–eIF4G complex, one m^7GTP and 35 water molecules in the asymmetric unit. The following residues are missing from the model: eIF4E G239–S240 (chain A) and eIF4G K632–N660 (chain B).

The diffraction data and refinement statistics are summarized in Table 1.

Isothermal titration calorimetry

The ITC experiments were performed using a VP-ITC microcalorimeter (Microcal) at 20°C as previously described (Igreja et al., 2014; Peter et al., 2015a). Briefly, a 2 to 10 μM solution of *Dm* eIF4E (residues A31–L248 or K69–L248; WT or II-AA mutant) in the calorimetric cell was titrated with tenfold concentrated solutions of GB1-stabilized eIF4G fragments (20 to 100 μM). All proteins were dissolved in 10 mM Na-phosphate (pH 7.0) and 150 mM NaCl. Each titration experiment consisted of a first injection of 2 μL followed by 28 injections of 10 μL at 240 s intervals. At least three independent experiments were performed. The collected data were analyzed using ORIGIN software 7.0 (Microcal). Correction for the heat of dilution and mixing was achieved by subtracting the final baseline, which consisted of small peaks of similar size. The first data point was removed from the analysis (Mizoue and Tellinghuisen, 2004), and the thermodynamic parameters were estimated by fitting a single-site binding isotherm yielding the equilibrium association constant (K_a) and enthalpy of binding (ΔH). Representative thermograms are shown in Figure S2.

Coimmunoprecipitation assays and western blotting

The coimmunoprecipitation assays from S2 cell or human HEK293T cell lysates and western blotting were described previously (Igreja et al., 2014; Peter et al., 2015a). Western blots were

developed using the ECL Western blotting detection system (GE Healthcare), following the manufacturer's recommendations. The antibodies used in this study can be found in Table S4.

In vitro competition and pulldown assays

The competition and pulldown assays were performed as described previously (Igreja et al., 2014; Peter et al., 2015a). In the pulldown assays shown in Figures 1 and 4, bacterial lysates expressing recombinant eIF4E or purified eIF4E (2 μM; 50 μg) were incubated with m^7GTP-sepharose beads (Jena Bioscience) or Ni-NTA beads for 30 min. The immobilized eIF4E was then incubated for 30 min with recombinant eIF4G fragments tagged N-terminally with MBP and C-terminally with GB1. Proteins associated with eIF4E were eluted with imidazole (in the case of the Ni-NTA pulldown) or with SDS-PAGE sample buffer and analyzed by SDS-PAGE followed by Coomassie staining. For the competition assays shown in Figures 5 and 6, purified *Dm* eIF4E–eIF4G complexes containing eIF4E-His_6 and eIF4G-GB1 (residues P578–N675) were incubated with Ni-NTA beads for 30 min in 50 mM Na-phosphate (pH 7.0) and 200 mM NaCl. The immobilized complexes were then incubated with a three-fold molar excess of purified and GB1-tagged competitor peptides as indicated or with MBP as a negative control. After the indicated time points, the supernatant was discarded. The beads were washed three times in the same buffer, and the eIF4E-associated proteins were eluted with imidazole for analysis by SDS-PAGE, followed by Coomassie staining. The amount of eIF4G bound to eIF4E was quantified using the ImageJ software and normalized to the levels of eIF4E present at each time point. These values were set to 100 in the presence of MBP. Data points from at least three independent experiments were plotted together and the resulting fitting curves were determined using the Levenberg-Marquardt algorithm for a single exponential decay function. The R^2 values associated with the fitting of the exponential decay curves were between 0.88 and 0.97.

Sequence analysis

eIF4E and eIF4G sequences were aligned using MAFFT as accessible in Jalview (Waterhouse et al., 2009).

SUPPLEMENTAL REFERENCES

Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., and Adams, P.D. (2012). Towards automated crystallographic structure refinement with phenix.refine. Acta Crystallogr D Biol Crystallogr *68*, 352-367.

Cheng, Y., and Patel, D.J. (2004). An efficient system for small protein expression and refolding. Biochem. Biophys. Res. Commun. *317*, 401-405.

Diebold, M.L., Fribourg, S., Koch, M., Metzger, T., and Romier, C. (2011). Deciphering correct strategies for multiprotein complex assembly by co-expression: application to complexes as large as the histone octamer. J. Struct. Biol. *175*, 178-188.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr D Biol Crystallogr *66*, 486-501.

Evans, P. (2006). Scaling and assessment of data quality. Acta Crystallogr D Biol Crystallogr *62*, 72-82.

Evans, P.R. (2011). An introduction to data reduction: space-group determination, scaling and intensity statistics. Acta Crystallogr D Biol Crystallogr *67*, 282-292.

Kabsch, W. (2010). Xds. Acta Crystallogr D Biol Crystallogr *66*, 125-132.

Langer, G., Cohen, S.X., Lamzin, V.S., and Perrakis, A. (2008). Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. Nat Protoc *3*, 1171-1179.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J Appl Crystallogr *40*, 658-674.

Mizoue, L.S., and Tellinghuisen, J. (2004). The role of backlash in the "first injection anomaly" in isothermal titration calorimetry. Anal Biochem. *326*, 125-127.

Smart, O.S., Womack, T.O., Flensburg, C., Keller, P., Paciorek, W., Sharff, A., Vonrhein, C., and Bricogne, G. (2012). Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. Acta Crystallogr D Biol Crystallogr *68*, 368-380.

Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M., and Barton, G.J. (2009). Jalview Version 2-a multiple sequence alignment editor and analysis workbench. Bioinformatics *25*, 1189-1191.

Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G., McCoy, A.*, et al.* (2011). Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr *67*, 235-242.

GIGYF1/2 proteins use auxiliary sequences to selectively bind to 4EHP and repress target mRNA expression

Daniel Peter, Ramona Weber, Felix Sandmeir, Lara Wohlbold, Sigrun Helms, Praveen Bawankar, Eugene Valkov, Cátia Igreja and Elisa Izaurralde

Department of Biochemistry, Max Planck Institute for Developmental Biology, Spemannstrasse 35, 72076 Tübingen, Germany

Corresponding author: Elisa Izaurralde (elisa.izaurralde@tuebingen.mpg.de)

Running title: Structures of 4EHP–GIGYF complexes

Keywords: eIF4E, translational regulation, translational repression

Abstract

The eIF4E-homologous protein (4EHP) is thought to repress translation by competing with eIF4E for binding to the 5'-cap structure of specific mRNAs, to which it is recruited through interactions with various proteins, including the GRB10-interacting GYF proteins 1 and 2 (GIGYF1/2). Despite its similarity to eIF4E, 4EHP does not interact with eIF4G and therefore fails to initiate translation. In contrast to eIF4G, GIGYF1/2 selectively bind to 4EHP but not to eIF4E. Here, we present crystal structures of the 4EHP-binding regions of GIGYF1 and GIGYF2 in complex with 4EHP, which reveal the molecular basis for the selectivity of the GIGYF1/2 proteins for 4EHP. Complementation assays in a GIGYF1/2-null cell line using structure-based mutants indicate that 4EHP requires interactions with GIGYF1/2 to downregulate target mRNA expression. Our studies provide structural insights into the assembly of 4EHP–GIGYF1/2 repressor complexes and reveal that rather than merely facilitating 4EHP recruitment to transcripts, GIGYF1/2 proteins are required for repressive activity.

Introduction

The initiation of cap-dependent translation involves a series of sequential steps that start with the assembly of the eukaryotic initiation factor 4F (eIF4F) on the mRNA 5'-cap structure (Jackson et al. 2010). The eIF4F complex consists of the cap-binding protein eIF4E, the RNA helicase eIF4A, and the scaffold protein eIF4G, which bridges the interaction between the other two subunits in the complex. eIF4G also interacts with eIF3 and mediates the recruitment of the pre-initiation complex (PIC, comprising a 40S ribosomal subunit and associated factors) to the mRNA to initiate translation (Jackson et al. 2010).

The assembly of the eIF4F complex is regulated by multiple mechanisms. One major mechanism involves a broad class of eIF4E-binding proteins (4E-BPs) that compete with eIF4G for binding to eIF4E, thereby inhibiting translation initiation (Mader et al. 1995; Marcotrigiano et al. 1999). eIF4G and the 4E-BPs share a conserved, canonical 4E-binding motif (C) with the sequence $YX_4L\Phi$ (where Y, X, L and Φ represent Tyr, any amino acid, Leu, and a hydrophobic residue, respectively), which binds to the dorsal surface of eIF4E opposite to the cap-binding pocket (Matsuo et al. 1997; Marcotrigiano et al. 1999; Gross et al. 2003). Both eIF4G and the 4E-BPs also contain variable non-canonical 4E-binding motifs (NC) that bind to an eIF4E hydrophobic lateral surface increasing the affinity of the interaction (Kinkelin et al. 2012; Paku et al. 2012; Lukhele et al. 2013; Igreja et al. 2014; Peter et al. 2015a,b; Sekiyama et al. 2015; Grüner et al. 2016). Because eIF4G and 4E-BPs bind to the same surfaces on eIF4E their binding is mutually exclusive, resulting in translation activation and inhibition, respectively.

An alternative mechanism that inhibits the recruitment of the eIF4F complex involves the recognition of the mRNA 5'-cap by another member of the eIF4E family, the 4E-homologous protein (4EHP, also known as eIF4E2; Joshi et al. 2004; Rom et al. 1998). Despite its sequence and structural similarity to eIF4E (Supplemental Fig. S1A, Rosettani et

al. 2007), 4EHP does not interact with eIF4G and thus fails to initiate translation (Rom et al. 1998; Joshi et al. 2004; Hernandez et al. 2005).

4EHP is recruited to specific mRNAs by RNA-binding proteins and thus acts as a sequence-specific rather than a general translational repressor. For example, *Drosophila melanogaster* (*Dm*) 4EHP is specifically recruited to and represses translation of *caudal* and *hunchback* mRNAs through interactions with the RNA-binding proteins Bicoid and Brain tumor, respectively (Cho et al. 2005; Cho et al. 2006). Mammalian 4EHP has been implicated in posttranscriptional mRNA regulation through its interaction with the nucleocytoplasmic shuttling protein 4E-T (eIF4E-transporter), which is a component of P-bodies (Kubacka et al. 2013). In mouse oocytes, the homeobox protein Prep1 recruits 4EHP to inhibit the translation of *Hoxb4* mRNA (Villaescusa et al. 2009). 4EHP also forms a translational repressor complex with GIGYF2 (GRB10-interacting GYF protein 2, hereafter GYF2), a protein involved in the insulin signaling pathway (Giovannone et al. 2009; Morita et al. 2012). This repressor complex is recruited to specific mRNAs by the Zn-finger protein ZNF598 (Morita et al. 2012). Alternatively, the 4EHP–GYF2 complex is recruited to mRNAs containing AU-rich elements in their 3'-UTRs by tristetraprolin (TTP; Tao and Gao 2015; Fu et al. 2016). Thus, through its association with diverse binding partners, 4EHP regulates the translation of mRNAs involved in a broad range of biological process, and disruption of its expression results in perinatal lethality in mice (Morita et al. 2012).

Current models suggest that 4EHP-binding proteins (4EHP-BPs) interact with 4EHP through a canonical 4EHP-binding motif with the sequence $YXYX_4L\Phi$ that is present in GYF1/2 proteins (Fig. 1A; Supplemental Fig. S1B). Although this motif consists of a canonical 4E-binding motif extended by only two N-terminal residues (YX; Cho et al. 2005; Morita et al. 2012), GYF2 does not bind to eIF4E *in vivo* (Morita et al. 2012). Conversely, eIF4G, which contains a canonical motif, binds to eIF4E but not to 4EHP (Rom et al. 1998;

Joshi et al. 2004; Hernandez et al. 2005). In contrast, some 4E-BPs, such as *Hs* 4E-BP1–3 and 4E-T, which lack the additional YX residues, interact with both eIF4E and 4EHP (Rom et al. 1998; Rosettani et al. 2007; Kubacka et al. 2013). This suggests that the canonical motif is unlikely to be the sole specificity determinant for 4EHP or eIF4E and that the structural basis for this molecular discrimination is unknown. Additionally, structural insights into 4EHP complexes are limited to a complex with the 4E-BP1 canonical motif, which binds to 4EHP *in vitro* but not *in vivo* (Rom et al. 1998; Rosettani et al. 2007).

To obtain molecular insights into the assembly of 4EHP repressor complexes, we have determined the crystal structures of 4EHP in complex with the binding regions of human 4E-BP1, GYF1 and GYF2. The structures reveal that in addition to the known canonical motifs that bind to the dorsal surface of 4EHP, 4E-BP1 and GYF1/2 also make contacts with the lateral surface of 4EHP using non-canonical motifs, indicating that lateral binding is a common feature observed in both 4EHP and eIF4E complexes. Remarkably, GYF1/2 proteins, but not 4E-BP1, contain C-terminal auxiliary sequences that extend the interface, contacting 4EHP residues that are not conserved in eIF4E. Our studies reveal the molecular basis for the selectivity of GYF1/2 proteins for 4EHP over eIF4E and provide mechanistic insights into the regulation of cap-dependent translation initiation by 4EHP repressive complexes.

Results

GYF1/2 proteins bind to the dorsal and lateral surfaces of 4EHP

Given that the canonical motif in 4E-BP1 binds to the dorsal surface of 4EHP in a similar way as it binds to eIF4E *in vitro* (Rosettani et al. 2007), we initially investigated whether the non-canonical sequences in 4E-BP1 could also bind to the lateral surface of 4EHP as observed in eIF4E–4E-BP1 complexes (Igreja et al. 2014; Peter et al. 2015a). We substituted residues I85 and M101 on the lateral surface of 4EHP with alanine (IM-AA mutant, Supplemental Fig. S1A; Supplemental Table S1). These residues are structurally equivalent to eIF4E residues I63 and I79, which are required for the non-canonical motifs in 4E-BPs to bind to the lateral surface of eIF4E (Igreja et al. 2014; Peter et al. 2015a). As a control, a 4EHP mutant carrying the W95A substitution on the dorsal surface (mutant W-A, Supplemental Table S1, corresponding to eIF4E W73A mutant) was designed to disrupt canonical motif binding.

The substitutions in either the dorsal or the lateral surface of V5-SBP-tagged 4EHP disrupted interactions with GFP-tagged 4E-BP1 in human HEK293T cells (Fig. 1B, lanes 7 and 8). Conversely, substitutions in either the canonical (C* mutant) or non-canonical motif (NC* mutant) of 4E-BP1 abolished its interaction with 4EHP (Supplemental Fig. S2A). These results indicate that the interactions between the non-canonical 4E-BP1 sequences and the 4EHP lateral surface are also critical for complex stability.

The interaction between GYF2 and 4EHP requires a canonical 4EHP-binding motif at the N-terminus of the protein ($YXYX_4L\Phi$; Morita et al. 2012; Fig. 1A; Supplemental Fig. S1B). Accordingly, alanine substitutions of the four conserved residues (Y, Y, L and Φ) in this motif abolished full-length GYF1/2 binding to V5-SBP-4EHP in human cells (Supplemental Fig. S2B,C). However, it is not known whether GYF proteins contain non-canonical sequences. We therefore examined the effects of substitutions at the dorsal and

lateral surfaces of 4EHP on interactions with GYF1/2 proteins. The substitutions at either surface reduced but did not abolish 4EHP binding to endogenous GYF2 in human cells (Fig. 1C, lanes 8 and 9 vs. 7). The interaction was abolished only when the substitutions on both surfaces were combined (Fig. 1C, lane 10). In contrast, the dorsal and lateral substitutions disrupted binding with endogenous GYF1 (Supplemental Fig. S2D). Thus, 4EHP uses its dorsal and lateral surfaces to interact with the GYF1/2 proteins, suggesting that these proteins also contain non-canonical motifs.

GYF1/2 proteins contain non-canonical and auxiliary 4EHP-binding sequences

The non-canonical motifs in 4E-BPs are typically located 12–30 residues C-terminal to the canonical motifs and contain hydrophobic residues (Igreja et al. 2014; Peter et al. 2015a,b). During a search for potential non-canonical motifs in GYF1/2 proteins, we identified a hydrophobic motif, 12 residues C-terminal to the canonical motif containing a conserved Phe that we termed the non-canonical motif (NC; Fig. 1A; Supplemental Fig. S1B). The following 30 residues (which we termed auxiliary sequences, A) are also well conserved and contain several short motifs that may potentially interact with 4EHP, as was previously observed in the *Dm* protein Mextli in complex with eIF4E (Peter et al. 2015b).

To more precisely define the GYF1/2 sequences that interact with 4EHP, we performed *in vitro* pulldown assays using recombinant proteins expressed in *E. coli*. In particular, we tested the binding of 4EHP to GYF1/2 fragments comprising the canonical motif and the non-canonical sequences (i.e., non-canonical linker and motif; L+NC) with and without the auxiliary sequences (fragments C+L+NC and C+L+NC+A; Supplemental Table S1), as well as a fragment comprising only the canonical motif (C). 4EHP expressed with a hexa-histidine (His_6) tag pulled down all three GYF1/2 fragments expressed with an N-terminal maltose binding protein (MBP) tag (Fig. 1D; Supplemental Fig. S2E), indicating that the canonical

motifs are sufficient for 4EHP binding *in vitro*. The eIF4E-binding region of human 4E-BP1 interacted with 4EHP to a similar extent (Fig. 1D; Supplemental Fig. S2E, lane 11). However, the GYF1/2 fragments bound to eIF4E much less efficiently than 4E-BP1 (Fig. 1E; Supplemental S2F), indicating that GYF1/2 proteins exhibit selectivity for 4EHP over eIF4E *in vitro* in the absence of cellular factors.

In contrast to the results obtained *in vitro*, endogenous GYF1/2 proteins interacted with 4EHP but not with eIF4E in cell lysates, as previously reported (Fig. 1F; Supplemental Fig. S2G; Rom et al. 1998; Morita et al. 2012). Similar results were obtained with overexpressed GYF1/2 (Supplemental Fig. S2H,I), suggesting that although GYF1/2 proteins can bind to eIF4E *in vitro*, their binding affinity may be too low to compete with other 4E-BPs present in cell lysates for binding to eIF4E. As expected, the 4E-T protein interacted with both eIF4E and 4EHP (Fig. 1F; Kubacka et al. 2013). In contrast, although endogenous 4E-BP1 did not bind to 4EHP in cell lysates (Fig. 1F; Supplemental Fig. S2G; Rom et al. 1998), it did bind when it was overexpressed (Fig. 1B). Thus, in cell lysates, the selectivity of GYF1/2 proteins and 4E-BP1 for 4EHP and eIF4E, respectively, is likely to be determined by their affinities and concentrations relative to those of other competing proteins.

GYF1/2 auxiliary sequences increase affinity for 4EHP

Although the GYF1/2 fragments with and without auxiliary sequences associated with wild-type 4EHP in *in vitro* pulldown assays (Fig. 1G; Supplemental Fig. S2J, lanes 10 and 13), they were differentially affected by mutations on the 4EHP dorsal and lateral surfaces. The GYF1/2 fragments including the auxiliary sequences were insensitive to the mutations (Fig. 1G; Supplemental Fig. S2J, lanes 10 to 12). In contrast, the binding of the fragments lacking the auxiliary sequences was reduced or abolished by the mutations (Fig. 1G; Supplemental Fig. S2J, lanes 14,15). Thus, the GYF1/2 auxiliary sequences contribute to the stability of the

complexes with 4EHP and compensate for the destabilizing effects of the mutations in the dorsal and lateral 4EHP surfaces.

To evaluate the thermodynamic contribution of the GYF1/2 canonical, non-canonical and auxiliary sequences to the affinity for 4EHP, we performed isothermal titration calorimetry (ITC) experiments. GYF1/2 peptides containing only the canonical motif (C) exhibited dissociation constants (K_Ds) for 4EHP in the high nanomolar range (360 ± 120 nM^{GYF1} and 290 ± 160 nM^{GYF2}; Supplemental Fig. S3A,B; Table S2). Addition of the non-canonical linker and motif (C+L+NC peptides) increased the affinity for 4EHP by 20–30-fold (K_Ds of 12 ± 2 nM^{GYF1} and 14 ± 1 nM^{GYF2}, Supplemental Fig. S3C,D; Table S2), confirming the importance of the non-canonical sequences for complex formation. Importantly, addition of the auxiliary sequences increased the affinity even further by 30-40-fold (C+L+NC+A peptides; K_Ds of 0.4 ± 0.2 nM^{GYF1} and 0.3 ± 0.1 nM^{GYF2}; Supplemental Fig. S3E,F; Table S2) relative to that of the peptides lacking the auxiliary sequences.

Collectively, the affinity measurements indicate that the GYF1/2 auxiliary sequences contribute substantially to the affinity for 4EHP by further stabilizing the interactions mediated by the canonical motif and the non-canonical sequences.

The overall architecture of the 4EHP–GYF1/2 and 4EHP–4E-BP1 complexes

To understand the structural basis of complex formation and selectivity, we determined the crystal structures of human 4EHP bound to GYF1 and GYF2 fragments (C+L+NC+A peptides; residues 33–103^{GYF1} and 35–105^{GYF2}) at 2.9 Å and 2.3 Å resolution, respectively (Fig. 2A–E; Supplemental Fig. S4A,B, Table 1). We also determined the crystal structure of 4EHP bound to a peptide comprising the 4E-BP1 canonical motif and non-canonical sequences (C+L+NC) and the corresponding peptide in GYF2 at 1.9 Å and 2.0 Å resolution, respectively (Fig. 2F–H; Supplemental Fig. S4C,D; Table 1).

The cap-binding protein 4EHP adopts a eIF4E-like fold (Fig. 2C–G; Rosettani et al. 2007) and its dorsal and lateral surfaces are very similar to those of eIF4E at the structural and sequence levels (Supplemental Fig. S1A and Fig. S4E,F). Upon 4E-BP1 or GYF1/2 binding no major conformational changes are observed between the 4EHP structures presented in this study compared with a previously determined structure of 4EHP bound to the 4E-BP1 canonical motif (Rosettani et al. 2007).

The 4E-BP1 and GYF1/2 canonical motifs adopt a helical conformation on the dorsal surface of 4EHP, whereas the non-canonical sequences bind to the lateral surface of 4EHP using a similar binding mode as described for 4E-BP1 in complex with eIF4E (Fig. 2C–J; Supplemental Fig. S4E,F; Peter et al. 2015a; Sekiyama et al. 2015). The binding mode and conformation for the GYF2 fragment comprising the canonical motif and the non-canonical sequences was not influenced by the auxiliary region, as the two GYF2 structures in complex with 4EHP are very similar across common elements irrespective of whether the auxiliary region was present (Fig. S4G).

The distinguishing structural feature observed in the 4EHP–GYF1/2 complexes containing the auxiliary sequences is the unprecedented binding mode between these sequences and 4EHP (Fig. 2A–E). The auxiliary sequences extend the binding interface to 2190 $Å^2$ compared to 1380 $Å^2$ in their absence. The auxiliary sequences in GYF1/2 can be delineated into three short sequence motifs, termed auxiliary motifs 1–3 (A1, A2 and A3), and connect to the non-canonical motif by the auxiliary linker 1 (a-L1) and are interconnected by auxiliary linkers 2 and 3 (a-L2 and a-L3; Fig. 1A, Fig. 2A,B; Supplemental Fig. S1B).

The GYF1/2 canonical helix stabilizes the interaction with 4EHP and the auxiliary sequences

The canonical helices in GYF1/2 and 4E-BP1 bind to the 4EHP dorsal surface through interactions analogous to those observed for the eIF4G and 4E-BPs canonical motifs in

complex with eIF4E (Fig. 3A,B; Supplemental Fig. S5A–D; Marcotrigiano et al. 1999; Gross et al. 2003; Kinkelin et al. 2012; Peter et al. 2015a,b; Grüner et al. 2016). The most conserved interactions are mediated by residues corresponding to LΦ in the $YXYX_4L\Phi$ consensus sequence (M46, $L47^{GYF1}$; M48, $L49^{GYF2}$; L59, $M60^{BP1}$) and the second Tyr side chain in the canonical 4EHP-binding motif ($Y41^{GYF1}$, $Y43^{GYF2}$ corresponding to $Y54^{BP1}$; Fig. 3A,B; Supplemental Fig. S5A–D).

The first Tyr in the canonical 4EHP-binding motif ($\underline{Y}XYX_4L\Phi$) was suggested to contribute to the binding specificity of 4EHP-binding proteins (Cho et al. 2005; Villaescusa et al. 2009). Our structural analysis does not support such a role for this tyrosine ($Y39^{GYF1}$, $Y41^{GYF2}$). Although its aromatic ring is in contact with $P55^{4EHP}$ (Fig. 3A; Supplemental Fig. S5C,D), this interaction is not unique to 4EHP-binding proteins, as $P55^{4EHP}$ forms a similar contact with $I52^{BP1}$ at an equivalent position in the motif in the 4E-BP1 complex ($\underline{I}XYX_4L\Phi^{BP1}$; Fig. 3B).

An important difference between 4E-BP1 and GYF1/2 canonical motifs is that the latter do not possess an Arg/Lys/Gln residue at position 9 in the extended canonical motif [extended motif: $YX(R/K)X_2L\Phi X_2$(**R/K/Q**)]. This residue typically contributes to the interaction with eIF4E (Marcotrigiano et al. 1999; Peter et al. 2015a,b). Instead, GYF1/2 proteins contain an aromatic residue at this position ($Y50^{GYF1}$, $F52^{GYF2}$; Fig. 3A; Supplemental Fig. S1B and S5C,D), which establishes hydrophobic contacts with $W95^{4EHP}$ and stabilizes the GYF1/2 auxiliary motifs through intramolecular interactions with the invariant Pro residue in the PLAL motif ($P76^{GYF1}$, $P78^{GYF2}$, see below), thus rationalizing the conservation of both residues.

The non-canonical linkers contribute to structural stability in the 4EHP complexes

Following the canonical helix, the non-canonical linkers orient the peptide chains to engage with the 4EHP lateral surface (Fig. 2A–G). In the structures of 4E-BPs bound to eIF4E, these linkers adopt a specific 'elbow loop' conformation (Peter et al. 2015a,b). An analogous structural feature is observed in the 4EHP–4E-BP1 complex, where $P66^{BP1}$, $P71^{BP1}$ and $P72^{BP1}$ restrict the flexibility of the backbone conformation in the elbow loop (Fig. 3D; Supplemental Fig. S5E,F). Stabilizing hydrophobic contacts, such as between the invariant $H100^{4EHP}$ (corresponding to $H78^{4E}$) and $P71^{BP1}$, ensure that the overall elbow conformation is almost identical in the eIF4E- and 4EHP-bound complexes (Fig. 3D; Supplemental Fig. S5E,F).

One important difference, however, is that the $N77^{4E}$ side chain is in hydrogen bonding distance to $R63^{BP1}$ guanidium group in position 9 of the extended canonical motif and to the $T68^{BP1}$ carbonyl oxygen in the linker region in the eIF4E–4E-BP1 complex (Supplemental Fig. S5F; Peter et al. 2015a). In the 4EHP complex, these contacts cannot be maintained by S99, and consequently, $R63^{BP1}$ does not contribute to complex stability (Fig. 3D; Supplemental Fig. S5E,F).

Comparison of the GYF1/2 conformations in complex with 4EHP reveals that the linker region is also arranged in a single preferred conformation (Fig. S4H) that is distinct from the elbow loop conformation observed in the complexes with 4E-BP1 (Fig. 3C vs. Fig. 3D). The invariant $H100^{4EHP}$ plays a crucial role in anchoring the GYF1/2 linker to the 4EHP surface through van der Waals contacts with $I58^{GYF2}$ ($V56^{GYF1}$), while its imidazole ring is also in hydrogen bonding distance to the $D55^{GYF2}$ ($E53^{GYF1}$) carbonyl oxygen (Fig. 3C; Supplemental Fig. S5G,H). The principal stabilizing hydrophobic interaction in the GYF1/2 linker region is between the invariant $P59^{GYF2}$ ($P57^{GYF1}$) and $F97^{4EHP}$, which is structurally equivalent to the interaction between $P72^{BP1}$ and $F97^{4EHP}$ (Fig. 3C,D; Supplemental Fig. S5E–H).

Non-canonical loops mediate conserved contacts at the 4EHP lateral surface

The non-canonical loops in 4E-BP1 and the GYF1/2 proteins engage a hydrophobic pocket on the lateral surface of 4EHP, which is lined by residues Y64, I85 and M101 (corresponding to eIF4E residues F47, I63 and I79, respectively; Fig. 3E,F; Supplemental Fig. S5I–L). The conformation of the GYF1/2 non-canonical loops is stabilized by an extensive and conserved network of contacts across the 4EHP lateral surface. The 4E-BP1 and GYF1/2 non-canonical loops differ in conformation and only align at major contact points (Fig. 3E vs. Fig. 3F).

Strikingly, a carbon-π interaction through which the conserved aromatic residue ($Y64^{4EHP}$ equivalent to $F47^{4E}$) makes contacts with $I70^{GYF2}$ ($V68^{GYF1}$) to anchor this loop at the lateral surface of 4EHP (Fig. 3E; Supplemental Fig. S5K,L) is conserved in the 4EHP–4E-BP1 complex ($Y64^{4EHP}$–$V81^{BP1}$) as well as in all eIF4E–4E-BP and eIF4E–eIF4G complex structures (Peter et al. 2015a; Grüner et al. 2016), underscoring the role of this aromatic residue ($Y64^{4EHP}$, $F47^{4E}$) in positioning the non-canonical loops at the lateral surface of eIF4E proteins. The invariant $F67^{GYF2}$ ($F65^{GYF1}$) is critically positioned at the sharp turn of the peptide and stabilizes this conformation via hydrophobic contacts with $Y64^{4EHP}$, $K83^{4EHP}$ and $I85^{4EHP}$. The C-terminal residues in the GYF1/2 and 4E-BP1 non-canonical loops (V68, $Q70^{GYF1}$; I70, $Q72^{GYF2}$; and V81, $S83^{BP1}$) mediate similar backbone interactions with the 4EHP residues H100 and V102 (Fig. 3E,F; Supplemental Fig. S5I-L).

4EHP-specific interactions with GYF1/2 auxiliary motifs close to the cap-binding site

The auxiliary linker-1 (a-L1) and the invariant PLAL motif at the start of the auxiliary sequences adapt to a composite surface formed between the GYF1/2 canonical helix and the 4EHP surface (Fig. 4A–D). The linker (a-L1) is fixed in position via $H100^{4EHP}$, which can participate in polar contacts with the carbonyl oxygens of $Q70^{GYF1}$ or $D71^{GYF1}$ ($Q72^{GYF2}$ or

E73^{GYF2}). The R103^{4EHP} guanidinium group is in hydrogen bonding distance to the Q75^{GYF1} (P77^{GYF2}) carbonyl oxygen, which rationalizes the conservation of the R103 residue in 4EHP but not in eIF4E. The invariant P76^{GYF1} (P78^{GYF2}) in the PLAL motif coordinates an intramolecular carbon-π interaction with Y50^{GYF1} (F52^{GYF2}). Importantly, key interactions of 4EHP with the auxiliary linker-1 (a-L1) and PLAL motif are mediated by 4EHP-specific residues, e.g., E149, which fixes the orientation of the GYF1/2 chain via hydrogen bonds to L77^{GYF1} (L79^{GYF2}) and A78^{GYF1} (A80^{GYF2}) as well as R146^{4EHP}, which contacts E80^{GYF1} (V82^{GYF2}) and L77^{GYF1} (L79^{GYF2}; Fig. 4C,D). Therefore, the auxiliary linker-1 and the PLAL motif interactions are highly specific and involve residues present only in 4EHP (R103 and E149) and thus would not be possible with eIF4E.

The auxiliary sequences A2 and A3 arrange into two helical elements (auxiliary helices α2 and α3, respectively; Fig. 2A–E), which are connected by a conserved VNS linker (linker 3; a-L3). Helix α2 (A2) shows some conformational heterogeneity across all complex structures (Supplemental Fig. S4H), most likely due to weak contacts at the interface (Fig. 4E–H). In contrast, auxiliary linker-3 (a-L3) aligns well between the six complex structures. The linker VNS sequence enters a surface groove on 4EHP and interacts closely with 4EHP, e.g., through invariant S96^{GYF1} (S98^{GYF2}), which maintains a backbone hydrogen bond to R138^{4EHP}, while its hydroxyl group is in polar contact with the 4EHP-specific residue E177 (Fig. 4G,H). As a consequence of these interactions, helix α3 (A3) is positioned in close proximity to the 4EHP cap-binding pocket (Fig. 2C,D) and is stabilized in this orientation through hydrophobic contacts between V99^{GYF1} (V101^{GYF2}) and L100^{GYF1} (L102^{GYF2}) with the aliphatic side chain of R138 and I211 in 4EHP (Fig. 4G,H). However, the GYF1/2 peptides containing all the 4EHP-binding elements did not contribute to the 4EHP affinity for the m^{7}GpppG cap analog as observed by ITC (Supplemental Table S2 and Fig. S3G,H),

suggesting that additional GYF1/2 sequences may contribute to enhance 4EHP binding to capped mRNAs.

The auxiliary sequences contribute to complex stability in vivo

To assess the biological significance of the interactions mediated by the auxiliary sequences, we substituted 4EHP residues R103 and E149, which interact with the GYF1/2 PLAL motif, with leucine residues, as is observed in eIF4E (4EHP RE-LL mutant). The RE-LL substitutions strongly reduced binding to endogenous GYF2 compared to wild-type 4EHP or the 4EHP dorsal and lateral mutants in human cells (Fig. 5A, lanes 7-10). All of the mutations disrupted binding to endogenous GYF1 (Supplemental Fig. S6A, lanes 7-10). As a control, binding of 4E-BP1 was not affected by the RE-LL substitutions (Fig. 5A, lane 10), indicating that the mutations do not disrupt the 4EHP fold.

We also analyzed the impact of amino acid substitutions in the GYF1/2 proteins on complex formation. Substitutions in the GYF1/2 non-canonical (NC*) or auxiliary motifs (A1* and A2+3*) did not affect binding to overexpressed 4EHP in human cells. This is consistent with structural data showing that R103 and E149 in 4EHP interact via their side chains with the backbone atoms of the GYF1/2 auxiliary sequences (Fig. 5B; Supplemental Fig. S6B, compare lanes 12-14 and 10). However, binding was disrupted when the mutations in the non-canonical and auxiliary motifs were combined (Fig. 5B; Supplemental S6B, lanes 15 and 16 *vs* 10). Together with the data showing that the mutations in the canonical motif prevent GYF1/2 binding to 4EHP (Fig. 5B; Supplemental Fig. S6B) this indicates that the canonical motif is necessary but not sufficient for 4EHP-binding *in vivo*.

We further assessed the relevance of R103 and E149 toward complex stability in competition assays using pre-assembled 4EHP–4E-BP1 complexes containing either the wild-type or the RE–LL 4EHP mutant. These complexes were challenged with an equimolar

amount of the GYF2 fragment comprising all 4EHP-interacting elements. The amount of 4E-BP1 that remained bound to 4EHP was determined over time (Fig. 5C–E). The GYF2 fragment displaced 50% of 4E-BP1 from the pre-assembled 4EHP–4E-BP1 complexes in 5 ± 1.2 min. Under the same conditions, the GYF2 fragment displaced only 40% of 4E-BP1 bound to the 4EHP RE-LL mutant after 60 min incubation (Fig. 5C–E). Collectively, the competition experiments together with the observation that the GYF1/2 proteins do not associate with the 4EHP RE-LL mutant in cell lysates (Fig. 5A; Supplemental Fig. S6A) indicate that the auxiliary sequences afford GYF2 a competitive advantage over 4E-BP1 for binding to 4EHP.

The canonical and auxiliary regions promote complex self-association in solution

In the asymmetric unit in the 4EHP–GYF1/2 crystals, the GYF1/2 canonical and auxiliary motifs are part of a large interface (1008 $Å^2$) connecting two neighboring complexes (Supplemental Fig. S7A–C), suggesting dimerization. We analyzed the solution properties of these complexes by small-angle X-ray scattering (SAXS). The SAXS parameters for the 4EHP–GYF2 complex were indeed consistent with those of a dimer (Supplemental Fig. S7A,D,E). We also observed that the self-association of the complexes was concentration-dependent (Supplemental Table S3). However, the SAXS measurements of complexes lacking the GYF2 auxiliary sequences, were consistent with a monomeric state (Fig. S7F,G and Supplemental Table S3).

The putative dimer interface is stabilized by residues that are 4EHP- and GYF-specific, including R202, M161 and Q159 in 4EHP and $E46^{GYF1}$ ($E47^{GYF2}$) in GYF1/2 (Supplemental Fig. S1A,B and S7A,B). We designed mutations in GYF2 and 4EHP to disrupt dimerization (dimerization mutant, D*; Supplemental Table S1). These mutations did not affect complex assembly, as the mutated proteins still co-purified as a complex (Supplemental Fig. S7H) and

retained the same affinity for their partner as the wild-type proteins (Supplemental Fig. S3I,J; and Table S2). Importantly, however, the SAXS scattering profile of the mutated complex demonstrated the best fit to a monomeric state (Supplemental Fig. S7I and Table S3), indicating that the mutations effectively disrupt dimerization and validate the observed interface.

A 4EHP-specific residue reduces 4E-BP1 binding

To probe the molecular basis for the binding preference of 4E-BP1 for eIF4E over 4EHP observed in cell lysates, we measured the binding affinity of the 4E-BP1 peptide for 4EHP and eIF4E using ITC. The affinity of the 4E-BP1 peptide for 4EHP was 10-fold lower compared with eIF4E (K_D = 55 ± 14 nM and K_D = 5 ± 2 nM, respectively; Supplemental Fig. S3K,L and Table S2) and 100-fold lower compared to the GYF1/2 peptides for 4EHP (Supplemental Table S2).

A possible explanation for the lower affinity of 4E-BP1 for 4EHP compared to eIF4E is that $R63^{BP1}$ directly interacts with N77 in the eIF4E complex, but this residue is replaced by a Ser (S99) in 4EHP, which breaks this critical contact (Supplemental Fig. S5E,F). Interestingly, a 4EHP mutant in which Ser99 was substituted with Asn (4EHP S99N) showed a 10-fold gain in affinity for 4E-BP1 to a level comparable with eIF4E (4 ± 1 nM and 5 ± 2 nM, respectively; Supplemental Fig. S3M and Table S2). The S99N 4EHP mutant also bound to endogenous 4E-BP1 (Supplemental Fig. S7J, lanes 8 vs 7), although not to the same level as observed for eIF4E. Binding of GYF2 was not affected by the S99N mutation because GYF proteins have an aromatic residue at the equivalent $R63^{BP1}$ position ($Y50^{GYF1}$ and $F52^{GYF2}$), which mediates hydrophobic contacts with 4EHP.

The affinity measurements indicate that endogenous 4E-BP1 is unlikely to effectively compete with GYF1/2 proteins for 4EHP binding under equilibrium conditions in cell lysates.

This is consistent with the *in vivo* data that shows that 4E-BP1 bound to 4EHP only when overexpressed (Fig. 1B). Furthermore, a single amino acid substitution is responsible for the different affinities of eIF4E and 4EHP for 4E-BP1.

4EHP requires interaction with GYF1/2 proteins to downregulate mRNA expression

To assess the functional relevance of the 4EHP–GYF1/2 complex in repressing mRNA targets, we tethered λN-HA-tagged 4EHP to an R-Luc reporter containing five binding sites for the λN tag (BoxB hairpins) in the 3'-UTR. To uncouple the effects on translation from the effects on mRNA stability, the reporter contained an internal polyadenosine stretch of 95 residues followed by the 3'-end of the non-coding RNA MALAT1, which is generated through endonucleolytic cleavage by RNase P and is thus not polyadenylated (Wilusz et al. 2012). An F-Luc-GFP reporter served as a transfection control. The λN-HA-4EHP protein repressed the expression of the R-Luc-5BoxB-A_{95}-MALAT1 reporter relative to the λN-HA peptide (Fig. 6A), without causing corresponding changes in mRNA levels (Fig. S8A,B). The levels of a reporter lacking the BoxB hairpins were not affected (Supplemental Fig. S8C–E), indicating that 4EHP recruitment is prerequisite for repression.

We used CRISPR-Cas9 gene editing to generate a GYF1/2-null HEK293T human cell line in which the GYF1/2 levels were reduced below 10% of their control levels, whereas the expression of endogenous 4EHP was not affected (Fig 6B, lane 4 *vs* 1). In this cell line, the repression of the R-Luc mRNA reporter by tethered 4EHP was impaired (Fig. 6A), even though 4EHP was expressed at levels comparable to those observed in control cells (Fig. 6C, lanes 2 *vs* 4). The 4EHP-mediated repression was restored by transient expression of wild-type GFP-tagged GYF2 but not by the GYF2 canonical mutant (C*) that does not interact with 4EHP (Fig. 6A), despite comparable expression levels (Fig. 6C, lanes 5 and 6). Thus, 4EHP requires interactions with GYF1/2 proteins for full repressive activity.

In agreement with this conclusion, 4EHP activity in tethering assays correlated with GYF1/2 binding and was independent of cap binding. Indeed, 4EHP mutants with impaired GYF1/2 binding (+/-) exhibited reduced repressive activity, and repression was abolished by combined mutations that disrupt binding to GYF1/2 (Fig. 6D,E; Supplemental Fig. S8F,G). Unexpectedly, however, a 4EHP mutant that does not bind to the cap (cap*; Fig. S8H), but still binds to GYF1/2 (Supplemental Fig. S8I), repressed the expression of the reporter mRNA in a GYF1/2-dependent manner (Fig. 6D; Supplemental Fig. S9A,B). The 4EHP mutants did not repress the expression of a control mRNA lacking the BoxB binding sites (Supplemental Fig. S9C–E).

Our results suggest that the GYF1/2 proteins confer repressive activity to the 4EHP–GYF1/2 complexes (Fig. 6F,G). Accordingly, GYF2 repressed mRNA reporter expression in tethering assays. GYF2 activity was independent of 4EHP because the GYF2 canonical mutant, still repressed the expression of the R-Luc-6xMS2-A_{95}-MALAT1 reporter as efficiently as wild-type GYF2 (Fig. 6F,G). GYF2 did not repress a reporter lacking the MS2 binding sites (Supplemental Fig. S9F,G).

Finally, to investigate 4EHP activity without artificial tethering we used an R-Luc reporter that included two copies of the AU-rich element present in the 3' UTR of the TNF-α mRNA (R-Luc-ARE-A_{90}-MALAT1). This reporter was repressed in control cells expressing TTP (Fig. 6H). TTP-induced repression was relieved in GYF1/2-null cells (Fig 6H), although TTP was expressed at levels comparable to those observed in control cells (Fig. 6I, lene 4 vs 2). In the GYF1/2-null cell line repression was restored only when GYF2 and 4EHP were coexpressed but not when each protein was expressed individually (Fig. 6H,I). No restoration was observed when 4EHP was coexpressed with a GYF2 canonical mutant (C*). In contrast to the observations in tethering assays, a 4EHP mutant that does not bind the cap (cap*) was impaired in restoring TTP-mediated repression, although it was expressed at levels

comparable to wild-type (Fig. 6H,I). Collectively, these results indicate that the assembly of the 4EHP-GYF2 complex is required for full repression of target mRNA expression.

Discussion

GYF1/2 proteins are able to discriminate between 4EHP and eIF4E, but the molecular basis for this discrimination remained unknown (Morita et al. 2012). Here, we show that the 4EHP-binding region of GYF1/2 proteins comprises canonical and non-canonical motifs connected by a linker, which recognize the dorsal and lateral surfaces of 4EHP, respectively, in a similar manner to that observed for the diverse 4E-BPs and eIF4G bound to eIF4E (Kinkelin et al. 2012; Peter et al. 2015a,b; Sekiyama et al. 2015; Grüner et al. 2016). Thus, dorsal and lateral binding is conserved and widespread among eIF4E family proteins. Given this common binding interface, GYF1/2 proteins achieve their remarkable selectivity for 4EHP by virtue of unique auxiliary sequences (C-terminal to the non-canonical motif) that contact a surface on 4EHP, which is more divergent among eIF4E paralogs. In particular, the 4EHP-specific residues R103, R140 and E149 interact with GYF1/2 auxiliary sequences and are important for complex formation *in vivo*. These interactions stabilize the complex assembly by increasing the affinity of the interaction and may have evolved to ensure that *in vivo* GYF1/2 proteins efficiently compete for 4EHP binding with other potential binding partners such as 4E-BP1, which is more abundant than GYF1/2 but has lower affinity for 4EHP (this study; Hein et al. 2015).

Intriguingly, as a consequence of the 4EHP-specific interactions by the auxiliary sequences, the helix α3 (A3) of GYF1/2 is oriented in close structural proximity to the cap-binding pocket of 4EHP. However, the GYF2 peptide containing all the 4EHP-binding elements did not increase the affinity of 4EHP for an m^7GpppG cap analog (Table S2). In the GYF1/2 proteins, there is a long stretch of Gly/Arg-rich sequence immediately following the auxiliary motifs. Given that such low-complexity Gly/Arg-rich regions often confer non-specific RNA binding properties to the proteins that contain them (Thandapani et al. 2013), it is tempting to speculate that the GYF1/2 proteins may play a role in stabilizing 4EHP bound

to capped transcripts. Furthermore, because 4EHP has a reduced affinity for the cap structure compared to eIF4E (Rom et al. 1998; Zuberek et al. 2007), it is possible that the auxiliary sequence-mediated dimerization observed in this study may have some as yet undefined functions, e.g., increasing local concentration of repressor complexes on the mRNA but this hypothesis needs to be tested in future studies. The affinity of 4EHP for capped mRNAs may also be stimulated by post-translational modifications such as ISG15 modification (Okumura et al. 2007) and mono/di-ubiquitinylation (von Stechow et al. 2015), but whether these are synergistic with the GYF1/2 proteins is currently not known.

Mechanism of repressor complex assembly is likely to be divergent among the 4EHP interactors

Our study provides mechanistic insights into the assembly of 4EHP repressor complexes and raises the question of whether the binding mode is conserved among other 4EHP-binding proteins. However, it is important to note that although the *Dm* Brat protein sequence contains a canonical 4EHP-binding motif, the motif is buried within the hydrophobic core of a folded domain (the NHL domain) and therefore unlikely to participate in interaction with 4EHP (Cho et al. 2006). Furthermore, the canonical motifs in *Dm* Bcd and the mammalian Prep1 proteins contain internal proline residues and are unlikely to adopt helical conformations, which are crucial for stable binding to 4EHP (Cho et al. 2006; Villaescusa et al. 2009). Thus, the interaction with Bcd and Prep1 is either indirect or the mode of binding has diverged. Only 4E-T features canonical and non-canonical motifs that bind directly to eIF4E and are likely to bind 4EHP in a similar manner (Kubacka et al. 2013; Peter et al. 2015a). Although the 4E-T orthologs do not contain motifs with similarity to the GYF1/2 auxiliary motifs, our structural data indicate that the precise sequence composition of these motifs may not be critical for interactions because the 4EHP-specific residues principally

stabilize the complex via contacts with the auxiliary sequences backbone. However, it remains to be seen whether 4E-T may indeed contain auxiliary sequences and what is their mode of binding to 4EHP.

4EHP–GYF1/2 complex assembly is required for post-transcriptional mRNA regulation

Because 4EHP has a reduced affinity for the cap structure compared to eIF4E (Rom et al. 1998; Zuberek et al. 2007), it has been proposed that it is recruited to specific mRNAs through interactions with proteins that are bound (directly or indirectly) to the mRNA, thus increasing its local concentration and competing with eIF4E *in cis* for binding to the 5'-cap (Cho et al. 2005; Cho et al. 2006; Villaescusa et al. 2009; Morita et al. 2012). According to this model, 4EHP should repress translation independently of GYF1/2 proteins when directly tethered to the 3'-UTR of an mRNA reporter. Unexpectedly, however, we observed that 4EHP not only loses its repressive activity in GYF1/2-null cells, but also that its interaction with GYF1/2 proteins is in fact required for full repression. Thus, rather than merely facilitating 4EHP recruitment to an mRNA (e.g., by bridging the interaction between 4EHP and the Zn-finger proteins ZNF598 or TTP), the GYF1/2 proteins act directly in the repression. In agreement with this conclusion, GYF1/2 repressed target transcripts in tethering assays independently of 4EHP binding. However, it is also evident that regulation of endogenous transcripts is dependent on the 4EHP–GYF1/2 complex assembly. Indeed, the TTP-mediated repression of an ARE-containing reporter in the GYF1/2-null cell line was restored only when 4EHP was coexpressed with GYF2 that was competent for binding to the 4EHP. Intriguingly, the cap binding by 4EHP was necessary for full repression in this context.

In summary, our studies reveal the structural basis for the assembly of a translational repressor complex consisting of 4EHP and its specific binding partners, the GYF1/2 proteins.

We show that the GYF1/2 proteins directly contribute to the repressive activity of 4EHP, thus uncovering an unexpected facet of a mechanism that regulates mRNA expression.

Materials and Methods

DNA constructs

The DNA constructs used in this study are described in the Supplemental Material and are listed in Supplemental Table S1. All of the constructs and mutations were confirmed by sequencing.

Protein expression and purification

All of the recombinant proteins were expressed in *E. coli* BL21 Star (DE3) cells (Thermo Fisher Scientific) grown in LB medium overnight at 20°C. The cells were lysed by sonication in lysis buffer containing 50 mM HEPES (pH 7.2), 200 mM (4EHP–4E-BP1) or 300 mM (4EHP–GYF1/2) NaCl and 2 mM DTT supplemented with DNase I (5 µg/ml), lysozyme (1 mg/ml) and protease inhibitor cocktail (Roche). To purify the complexes containing 4EHP bound to GYF1, GYF2 or 4E-BP1 for crystallization and SAXS, His_6-tagged 4EHP (residues A52–F234) was coexpressed with MBP-tagged GYF1 (residues K33–M103), GYF2 (residues A35–T105) or 4E-BP1 (residues T50–S83). The complexes were purified from cleared cell lysates using an amylose resin (New England Biolabs), followed by cleavage of the MBP and His_6 tags with HRV3C protease overnight at 4°C. After cleavage of the tags, the complexes were separated from the MBP and His_6 tags using a heparin column (HiTrap Heparin HP 5 ml, GE Healthcare) and further purified on a Superdex 75 column (GE Healthcare) in a buffer consisting of 10 mM HEPES (pH 7.2), 200 mM NaCl, and 2 mM DTT. The complexes were stored at -80°C or used directly for crystallization and SAXS. The 4EHP-complexes used in the competition assays were expressed and purified as described above with the difference that the C-terminal His_6 tag was not removed from the 4EHP and that the copurified 4E-BP1 peptide contained a C-terminal GB1-tag. The complexes were stored in a buffer containing 20 mM Na-phosphate (pH 7.0), 200 mM NaCl and 5% (w/v) glycerol.

For the pulldown assays, eIF4E (full-length) was expressed with a C-terminal His_6 tag, purified from cleared cell lysates using a nickel column (HisTrap HP 5 ml, GE Healthcare) and further purified on a heparin column (HiTrap Heparin HP 5 ml, GE Healthcare) followed by size exclusion chromatography (Superdex 75 column, GE Healthcare) without removing the C-terminal His_6 tag. The purified eIF4E-His_6 was stored at -80°C in a buffer consisting of 20 mM Na-phosphate (pH 7.0) and 200 mM NaCl.

Pulldowns,,competition assays, coimmunoprecipitation and western blotting

The *in vitro* pulldown and competition assays were performed as previously described (Igreja et al. 2014; Peter et al. 2015a,b). All coimmunoprecipitation and pulldown assays in HEK293T cell lysates were performed in the presence of RNase A as described previously (Peter et al. 2015a). All of the western blots were developed using the ECL western blotting detection system (GE Healthcare). The antibodies used in this study are listed in Supplemental Table S4. A detailed description of these assays is included in the Supplemental Material.

ITC measurements and small-angle X-ray scattering (SAXS)

The ITC measurements and small-angle X-ray scattering (SAXS) experiments are described in the Supplemental Material.

Crystallization, data collection and structure determination

A detailed description of the crystallization conditions and the structure determination process are included in the Supplemental Material. All diffraction data sets were recorded on a PILATUS 6M detector at the PXII beamline of the Swiss Light Source at a temperature of 100 K. The diffraction data and refinement statistics are summarized in Table 1.

Tethering and complementation assays

A detailed description of the procedure to generate the GYF1/2-null cell line is included in the Supplemental Material For the complementation assays, HEK293T cells (wild-type or GYF1/2-null cells) were seeded in six-well plates ($0.6x10^6$ cells per well) and transfected using Lipofectamine 2000 (Invitrogen). The tethering reporters have been previously described (Kuzuoğlu-Öztürk et al. 2016). The transfection mixtures contained 0.25 μg of pEGFP-N3-F-Luc transfection control reporter, 0.5 μg of pCIneo-R-Luc-5BoxB-A_{95}-MALAT1 (or pCIneo-R-Luc-A_{95}-MALAT1 without BoxB) and 0.3 and 0.7 μg of the plasmids expressing the λN-HA and λN-HA-tagged 4EHP proteins, respectively. Cells were also co-transfected with plasmids expressing GFP-tagged proteins (0.25 μg GFP-MBP, 1.8 μg GFP-GYF2 wild-type and 1.2 μg GFP-GYF2 C* mutant).

For the 4EHP tethering assays in HEK293T cells, wild-type or mutant λN-HA-4EHP proteins (0.3 μg for WT and cap* mutant; 0.8 μg for W-A; 1 μg for IM-AA; 1.2 μg for WIM-AAA and 0.4 μg RE-LL proteins) were co-transfected with the same amounts of reporter plasmids as described for the complementation assay. In the tethering assay with the GYF2 protein, the transfection mixture contained 0.25 μg of pEGFP-N3-F-Luc transfection control reporter, 0.5 μg of pCIneo-R-Luc-6xMS2-A_{95}-MALAT1 or pCIneo-R-Luc-A_{95}-MALAT1, 0.3 μg and 1 μg of the plasmids expressing the MS2-HA and MS2-HA-tagged GYF2 proteins, respectively, and 0.25 μg of GFP-MBP.

For the assay with the pCIneo-R-Luc-ARE-A_{90}-MALAT1 reporter, wild-type and GYF1/2-null HEK293T cells were transfected with 1 μg of the ARE reporter and 0.25 μg of pEGFP-N3-F-Luc transfection control reporter in the presence or absence of plasmids expressing λN-HA-TTPΔNIM (50 ng), 0.2 μg of GFP-MBP (0.2 μg), GFP-GYF2 (1 μg, wild-type or canonical mutant C*) and λN-HA-4EHP (0.5 μg).

Firefly and *Renilla* luciferase activities were measured two days after transfection using the Dual-Luciferase Reporter Assay System (Promega).

Accession numbers

Atomic coordinates and structure factors have been deposited in the Protein Data Bank (PDB) under accession codes 5NVK (4EHP–GYF1 C+L+NC+A), 5NVL (4EHP–GYF2 C+L+NC+A), 5NVM (4EHP–GYF2 C+L+NC) and 5NVN (4EHP–4E-BP1 C+L+NC).

Author contributions

D.P. purified, crystallized, collected and solved the structures of all 4EHP complexes presented in this study together with F.S. ITC measurements were performed by D.P. and F.S. S.H. performed pulldowns and competition assays. R.W. generated the GYF1/2-null cell line, performed tethering, complementation and pulldown assays, purified peptides/proteins, and generated many of the constructs used in pulldown assays. L.W. generated several constructs for expression in human cells and conducted immunoprecipitation experiments in HEK293T cells. D.P., F.S. and E.V. performed the SAXS experiments and the structural analysis. P. B. performed immunoprecipitations an generated the ARE reporter and the TTP expression vector. C.I. coordinated the project. E.I. was the principal investigator who supervised the project. D.P., C.I., E.V. and E.I. wrote the manuscript. All authors corrected the manuscript.

Acknowledgments

We are grateful to M.Y. Chung and C. Weiler for excellent technical assistance and H. Budde for help with cloning. We thank L. Langer for help with cloning, purification of some of the 4EHP-complexes and the initial crystallisation trials. We thank R. Büttner and T. Raisch for

the setup of crystallization screens and the staff at the PX beamline of the Swiss Light Source, B21 beamline at the Diamond Light Source (Didcot, U.K.) and the SWING beamline at the SOLEIL synchrotron (Saint-Aubin, France) for assistance with data collection. We also thank R. Rambo for helpful comments on the SAXS data analysis and O. Weichenrieder for insightful comments on the manuscript. This work was supported by Biostruct-X (ID 9926) and the Max Planck Society.

References

Cho PF, Gamberi C, Cho-Park YA, Cho-Park IB, Lasko P, Sonenberg N. 2006. Cap-dependent translational inhibition establishes two opposing morphogen gradients in Drosophila embryos. *Curr Biol* **16:** 2035–2041.

Cho PF, Poulin F, Cho-Park YA, Cho-Park IB, Chicoine JD, Lasko P, Sonenberg N. 2005. A new paradigm for translational control: inhibition via 5'-3' mRNA tethering by Bicoid and the eIF4E cognate 4EHP. *Cell* **121:** 411–423.

Fu R, Olsen MT, Webb K, Bennett EJ, Lykke-Andersen J. 2016. Recruitment of the 4EHP-GYF2 cap-binding complex to tetraproline motifs of tristetraprolin promotes repression and degradation of mRNAs with AU-rich elements. *RNA* **22:** 373–382.

Giovannone B, Tsiaras WG, de la Monte S, Klysik J, Lautier C, Karashchuk G, Goldwurm S, Smith RJ. 2009. GIGYF2 gene disruption in mice results in neurodegeneration and altered insulin-like growth factor signaling. *Hum Mol Genet* **18:** 4629–4639.

Gross JD, Moerke NJ, von der Haar T, Lugovskoy AA, Sachs AB, McCarthy JE, Wagner G. 2003. 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 MY, Weichenrieder O, Valkov E, Igreja C, Izaurralde E. 2016. The Structures of eIF4E-eIF4G Complexes Reveal an Extended Interface to Regulate Translation Initiation. *Mol Cell* **64:** 467–479.

Hein MY, Hubner NC, Poser I, Cox J, Nagaraj N, Toyoda Y, Gak IA, Weisswange I, Mansfeld J, Buchholz F et al. 2015. A human interactome in three quantitative dimensions organized by stoichiometries and abundances. *Cell* **163:** 712–723.

Hernandez G, Altmann M, Sierra JM, Urlaub H, Diez del Corral R, Schwartz P, Rivera-Pomar R. 2005. Functional analysis of seven genes encoding eight translation initiation factor 4E (eIF4E) isoforms in Drosophila. *Mech Dev* **122:** 529–543.

Igreja C, Peter D, Weiler C, Izaurralde E. 2014. 4E-BPs require non-canonical 4E-binding motifs and a lateral surface of eIF4E to repress translation. *Nat Commun* **5:** 4790.

Jackson RJ, Hellen CU, Pestova TV. 2010. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol* **11:** 113–127.

Joshi B, Cameron A, Jagus R. 2004. Characterization of mammalian eIF4E-family members. *Eur J Biochem* **271**: 2189–2203.

Kinkelin K, Veith K, Grunwald M, Bono F. 2012. Crystal structure of a minimal eIF4E-Cup complex reveals a general mechanism of eIF4E regulation in translational repression. *RNA* **18:** 1624–1634.

Kubacka D, Kamenska A, Broomhead H, Minshall N, Darzynkiewicz E, Standart N. 2013. Investigating the consequences of eIF4E2 (4EHP) interaction with 4E-transporter on its cellular distribution in HeLa cells. *PLoS One* **8:** e72761.

Kuzuoglu-Ozturk D, Bhandari D, Huntzinger E, Fauser M, Helms S, Izaurralde E. 2016. miRISC and the CCR4-NOT complex silence mRNA targets independently of 43S ribosomal scanning. *EMBO J* **35:** 1186–1203.

Lukhele S, Bah A, Lin H, Sonenberg N, Forman-Kay JD. 2013. Interaction of the Eukaryotic Initiation Factor 4E with 4E-BP2 at a Dynamic Bipartite Interface. *Structure* **21:** 2186–2196.

Mader S, Lee H, Pause A, 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.

Marcotrigiano J, Gingras AC, Sonenberg N, Burley SK. 1999. Cap-dependent translation initiation in eukaryotes is regulated by a molecular mimic of eIF4G. *Mol Cell* **3:** 707–716.

Matsuo H, Li H, McGuire AM, Fletcher CM, Gingras AC, Sonenberg N, Wagner G. 1997. Structure of translation factor eIF4E bound to m7GDP and interaction with 4E-binding protein. *Nat Struct Biol* **4:** 717–724.

Morita M, Ler LW, Fabian MR, Siddiqui N, Mullin M, Henderson VC, Alain T, Fonseca BD, Karashchuk G, Bennett CF et al. 2012. A novel 4EHP-GIGYF2 translational repressor complex is essential for mammalian development. *Mol Cell Biol* **32:** 3585–3593.

Okumura F, Zou W, Zhang DE. 2007. ISG15 modification of the eIF4E cognate 4EHP enhances cap structure-binding activity of 4EHP. *Genes Dev* **21**: 255-260.

Paku KS, Umenaga Y, Usui T, Fukuyo A, Mizuno A, In Y, Ishida T, Tomoo K. 2012. A conserved motif within the flexible C-terminus of the translational regulator 4E-BP is required for tight binding to the mRNA cap-binding protein eIF4E. *Biochem J* **441:** 237–245.

Peter D, Igreja C, Weber R, Wohlbold L, Weiler C, Ebertsch L, Weichenrieder O, Izaurralde E. 2015a. Molecular Architecture of 4E-BP Translational Inhibitors Bound to eIF4E. *Mol Cell* **64:** 467–479.

Peter D, Weber R, Kone C, Chung MY, Ebertsch L, Truffault V, Weichenrieder O, Igreja C, Izaurralde E. 2015b. Mextli proteins use both canonical bipartite and novel tripartite binding modes to form eIF4E complexes that display differential sensitivity to 4E-BP regulation. *Genes Dev* **29:** 1835–1849.

Rom E, Kim HC, Gingras AC, Marcotrigiano J, Favre D, Olsen H, Burley SK, Sonenberg N. 1998. Cloning and characterization of 4EHP, a novel mammalian eIF4E-related cap-binding protein. *J Biol Chem* **273:** 13104–13109.

Rosettani P, Knapp S, Vismara MG, Rusconi L, Cameron AD. 2007. Structures of the human eIF4E homologous protein, h4EHP, in its m7GTP-bound and unliganded forms. *J Mol Biol* **368:** 691–705.

Sekiyama N, Arthanari H, Papadopoulos E, Rodriguez-Mias RA, Wagner G, Leger-Abraham M. 2015. Molecular mechanism of the dual activity of 4EGI-1: Dissociating eIF4G from eIF4E but stabilizing the binding of unphosphorylated 4E-BP1. *Proc Natl Acad Sci USA* **112:** E4036–4045.

Tao X, Gao G. 2015. Tristetraprolin Recruits Eukaryotic Initiation Factor 4E2 To Repress Translation of AU-Rich Element-Containing mRNAs. *Mol Cell Biol* **35:** 3921–3932.

Tee AR, Tee JA, Blenis J. 2004. Characterizing the interaction of the mammalian eIF4E-related protein 4EHP with 4E-BP1. FEBS Lett. **564:** 58–62.

Thandapani P, O'Connor TR, Bailey TL, Richard S. 2013. Defining the RGG/RG motif. *Mol Cell* **50:** 613–623.

Villaescusa JC, Buratti C, Penkov D, Mathiasen L, Planaguma J, Ferretti E, Blasi F. 2009. Cytoplasmic Prep1 interacts with 4EHP inhibiting Hoxb4 translation. *PLoS One* **4:** e5213.

von Stechow L, Typas D, Carreras Puigvert J, Oort L, Siddappa R, Pines A, Vrieling H, van de Water B, Mullenders LH, Danen EH. 2015. The E3 ubiquitin ligase ARIH1 protects against genotoxic stress by initiating a 4EHP-mediated mRNA translation arrest. *Mol Cell Biol* **35**: 1254-1268.

Wilusz JE, JnBaptiste CK, Lu LY, Kuhn CD, Joshua-Tor L, Sharp PA. 2012. A triple helix stabilizes the 3' ends of long noncoding RNAs that lack poly(A) tails. *Genes Dev* **26:** 2392–2407.

Zuberek J, Kubacka D, Jablonowska A, Jemielity J, Stepinski J, Sonenberg N, Darzynkiewicz E. 2007. Weak binding affinity of human 4EHP for mRNA cap analogs. *RNA* **13:** 691–697.

Figure Legends

Figure 1. GYF1/2 proteins use canonical, non-canonical and auxiliary sequences to bind to 4EHP. (*A*) GYF1/2 proteins contain a central glycine-tyrosine-phenylalanine (GYF) domain and an N-terminal 4EHP-binding region (4EHP-BR). The 4EHP-binding region includes canonical (C), non-canonical (NC) and auxiliary motifs (A1–3) connected by linker sequences (nc-L and a-L1–3). The 4E-binding region (4E-BR) of 4E-BP1 contains canonical and non-canonical motifs. (*B, C*) Western blots showing the interaction between V5-SBP-4EHP [wild-type (WT) or the indicated mutants] and GFP-4EBP1 (full-length) or endogenous GYF2. The proteins were pulled down using streptavidin-coated beads. V5-SBP-MBP served as negative control. The inputs (1.5% for the V5-tagged proteins and 1% for the GFP-tagged proteins) and bound fractions (3–5% for the V5-tagged proteins and 20% for GYF2 and GFP-4E-BP1) were analyzed by western blotting using anti-V5, anti-GFP and anti-GYF2 antibodies. (*D, E*) Ni-NTA pulldown assays showing the interactions between GYF2 fragments (C+L+NC+A, C+L+NC and C) and 4EHP-His_6 (M1–F234; D) or eIF4E-His_6 (E). 4E-BP1 and MBP served as positive and negative controls, respectively. The GYF2 and 4E-BP1 peptides contain an N-terminal MBP tag and a C-terminal GB1 tag. The starting material (SM; 1.3% for MBP-proteins, 6% for 4EHP and purified eIF4E) and bound fractions (7–10%) were analyzed by SDS-PAGE followed by Coomassie blue staining. (*F*) The interaction between V5-SBP-eIF4E or 4EHP proteins with endogenous GYF2, 4E-T and 4E-BP1 was analyzed in HEK293T cell lysates using streptavidin pulldowns. The input (1% for 4E-BP1 and 4E-T and 1.5% for V5-SBP-tagged proteins and GYF2) and bound fractions (20% for 4E-BP1 and 4ET, 30% for GYF2 and 5% for the V5-SBP-tagged proteins) were analyzed by western blotting using the indicated antibodies. (*G*) Ni-NTA pulldown assay showing the interaction between 4EHP (M1–F234, WT or the indicated mutants) and GYF2 fragments. MBP served as a negative control. Samples were analyzed as described in (D).

The starting material (2% for the MBP-tagged proteins and 4–12% for the 4EHP proteins) and bound fractions (10%) were analyzed by SDS-PAGE followed by Coomassie blue staining.

Figure 2. Overall structures of GYF1, GYF2 and 4E-BP1 bound to 4EHP. (*A, B*) Overview of the structures of 4EHP bound to GYF1/2 (C+L+NC+A) fragments. The 4EHP surface is shown in yellow and surface residues within a radius of 4 Å of the bound GYF1 or GYF2 peptides are colored in orange. The GYF1 and GYF2 peptides are colored in purple and blue, respectively. Selected secondary structure elements in the GYF1/2 peptides are indicated. The invariant PLAL motif of GYF1/2 is circled with a dashed line. (*C, D*) Cartoon representation of the structures of 4EHP bound to GYF1/2. Selected secondary structure elements are labeled in black for 4EHP and in color for GYF1/2. (*E*) Superposition of the structures of 4EHP bound to GYF1 and GYF2. For clarity, the 4EHP molecule from the 4EHP–GYF1 complex was omitted. The structures of the complexes are very similar and overall RMSDs do not exceed 0.32 Å over 227 Cα atoms. (*F*) Structure of 4EHP bound to 4E-BP1. Selected secondary structure elements are labeled in black for 4EHP and in cyan for 4E-BP1. (*G*) Structure of 4EHP bound to the GYF2 C+L+NC fragment. Selected secondary structure elements are labeled in black and red for 4EHP and GYF2, respectively. (*H*) Superposition of the structures of 4EHP bound to the 4EBP1 and GYF2 C+L+NC peptides. For clarity, the 4EHP molecule from the 4EHP–4E-BP1 complex was omitted. (*I, J*) Schematic representations of 4EHP bound to GYF1/2 and 4E-BP1.

Figure 3. The interactions between the canonical and non-canonical sequences of GYF2 and 4E-BP1 with 4EHP. (A, B) Close-up views of the interactions between the 4EHP dorsal surface and the GYF2 and 4E-BP1 canonical helices. 4E-BP1 residue R63 is colored in dark

blue after its Cγ atom and highlighted with a black dashed box. The corresponding residues in GYF2 (F52) are also highlighted by a black dashed box. (C, D) Close-up views of the interaction between the 4EHP lateral surface and the GYF2 and 4E-BP1 non-canonical linkers. (E, F) Close-up views of the interactions between the 4EHP lateral surface and the GYF2 and 4E-BP1 non-canonical loops. Selected interface residues are shown as sticks. For clarity reasons, all residues labeled with an asterisk are shown without their side chain.

Figure 4. Interaction between the GYF1/2 auxiliary sequences and 4EHP. (*A, B*) Close-up view of the arrangement of the linker a-L1 and the PLAL motif (A1) in GYF1/2 proteins at the 4EHP dorsal surface. The surface of 4EHP is shown in yellow and the surface of the GYF1/2 canonical helices is shown in gray. The positions of the 4EHP-unique residues R103 and E149 are highlighted in orange and selected GYF1/2 residues are shown as either purple (GYF1) or blue (GYF2) sticks. (*C–H*) Close-up views of the interactions between 4EHP and the GYF1/2 auxiliary sequences (A1, A2 and A3). Selected GYF1/2 residues and 4EHP interface residues are shown as sticks. The GYF1/2 canonical helices are colored in gray.

Figure 5. The auxiliary interactions are crucial for the formation of the 4EHP-GYF complex. (*A*) Western blot showing the interaction of endogenous GYF2 or GFP-4E-BP1 with V5-SBP-4EHP (WT or the indicated mutants). The proteins were pulled down using streptavidin-coated beads. The inputs (1.5% for the V5-tagged proteins and 1% for GYF2 and GFP-4E-BP1) and bound fractions (3% for the V5-tagged proteins and 20% for GYF2 and GFP-4E-BP1) were analyzed by western blot using the indicated antibodies. (*B*) Interaction of V5-SBP-4EHP with GFP-GYF2 (residues 1-180; either wild-type or the indicated mutants). The proteins were immunoprecipitated using anti-GFP antibodies. GFP-MBP served as negative control. The inputs (1.5% for the GFP-tagged proteins and 0.5% for

V5-SBP-4EHP) and immunoprecipitates (7.5% for the GFP-tagged proteins and 30% for V5-SBP-4EHP) were analyzed by western blot using anti-GFP and anti-V5 antibodies. (*C–E*) Purified 4EHP–4E-BP1 complexes containing 4EHP-His_6 (wild-type or the RE-LL mutant) were incubated in the presence of equimolar amounts of the GYF2 C+L+NC+A peptide C-terminally tagged with GB1 or MBP as a negative control. The proteins bound to 4EHP were pulled down using Ni-NTA beads at the indicated time points and analyzed by SDS-PAGE and Coomassie blue staining. Panel (C) shows the quantification of the amount of 4E-BP1 still associated with 4EHP (n=3). The half-life of the 4EHP-4E-BP1 complex ($t_{1/2}$) in the presence of the competitor protein is represented as the mean +/- SD. Panels (D) and (E) show representative SDS-PAGE gels. The positions of the GYF2 and 4E-BP1 peptides are marked by blue and black dashed boxes, respectively. The lanes labeled SM (starting material) show the purified complexes and peptides used in the assay.

Figure 6. 4EHP requires interaction with GYF1/2 proteins to repress translation.

(*A*) A complementation assay using the R-Luc-5BoxB-A_{95}-MALAT1 reporter and λN-HA-4EHP (either wild-type or the indicated mutants) was performed in control and GYF1/2-null HEK293T cells expressing GFP-MBP or GFP-GYF2 (wild-type or canonical mutant). A plasmid expressing F-Luc-GFP served as the transfection control. R-Luc activity was normalized to that of the F-Luc transfection control and set to 100% in cells expressing the λN-HA peptide. Bars represent the mean values and error bars represent standard deviations from three independent experiments. (*B*) Western blot analysis showing that full-length GYF1/2 levels were strongly reduced relative to control levels in the GYF1/2-null cell line. (*C*) Western blot analysis showing the expression of the λN-HA-4EHP and GFP-GYF2 proteins used in the assay shown in A. (*D*) Tethering assay using the R-Luc-5BoxB-A_{95}-MALAT1 reporter and λN-HA-4EHP (wild-type or mutants) in HEK293T cells. Samples

were analyzed as described in A. The symbols (+), (+/-) and (-) indicate binding, reduced binding and no binding to the GYF1/2 proteins, respectively. (*E*) Western blot showing the equivalent expression of the λN-HA-4EHP proteins used in the assay shown in D. (*F*) Tethering assay using the R-Luc-6xMS2-A_{95}-MALAT1 reporter and MS2-HA-GYF2 (wild-type or canonical mutant) in HEK293T cells. The cells were also co-transfected with GFP-MBP and F-Luc-GFP as transfection controls. R-Luc activity was normalized to that of the F-Luc transfection control and set to 100% in cells expressing the MS2-HA Samples were analyzed as described in A. (*G*) Western blot analysis showing the equivalent expression of the MS2-HA-GYF2 proteins. (*H*) Control HEK293T cells or cells depleted of GYF1/2 (KO) were transfected with the R-Luc-ARE-A_{90}-MALAT1 reporter and plasmids expressing the indicated proteins. F-Luc-GFP reporter served as a transfection control. R-Luc activity was normalized to that of the F-Luc transfection control and set to 100% in the absence of TTP for each cell line. (*I*) Western blot showing the expression of the proteins in the experiment shown in H. Note that TTP is stabilized in GYF1/2-null cells expressing GYF2. However, repression did not correlate with TTP levels but with the coexpression of wild-type GYF2 and 4EHP.

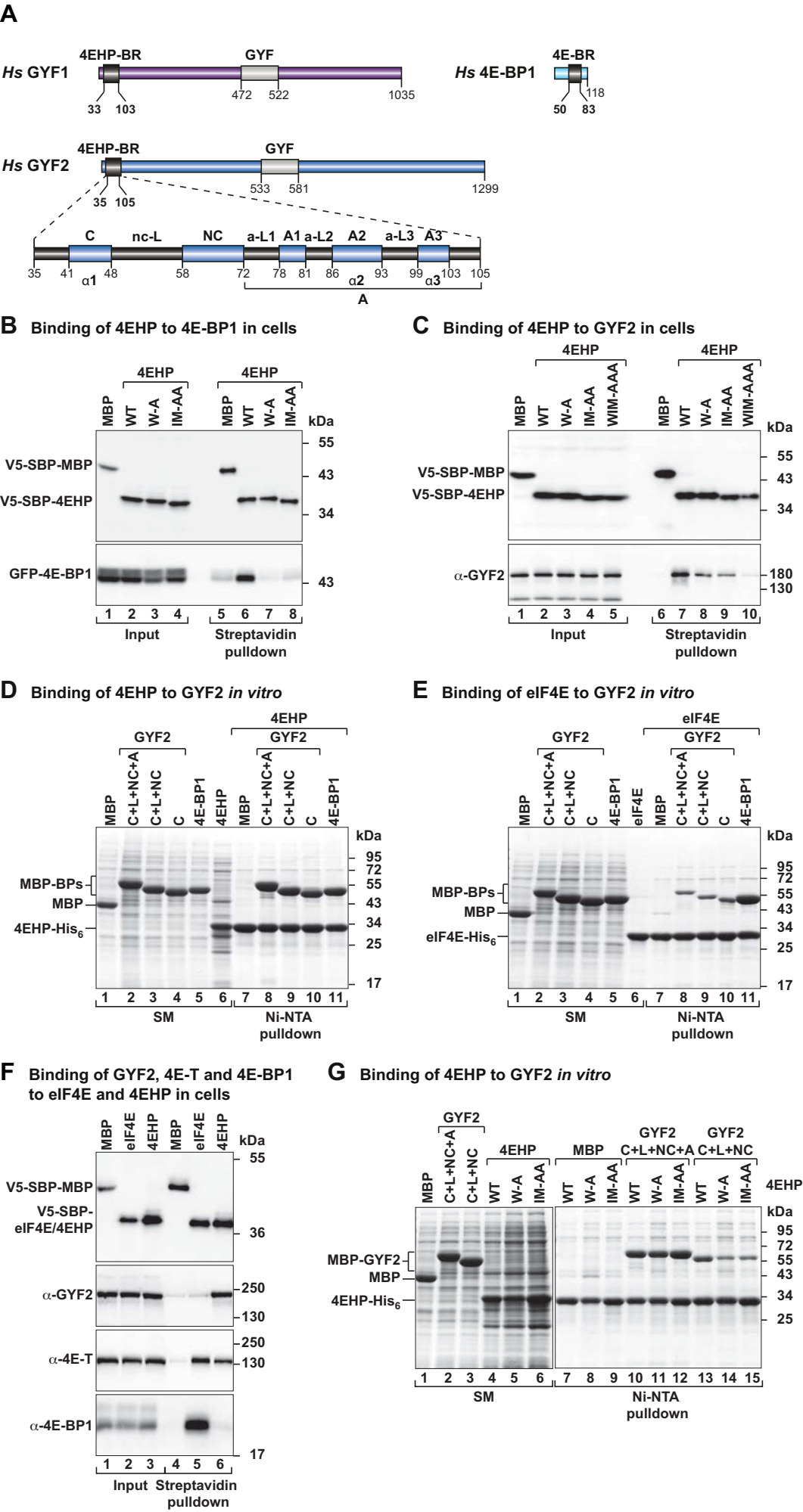
Table 1. Data collection and refinement statistics

	Hs 4EHP–GIGYF1 complex (GYF1 C+L+NC+A)	*Hs* 4EHP–GIGYF2 complex (GYF2 C+L+NC+A)	*Hs* 4EHP–GIGYF2 complex (GYF2 C+L+NC)	*Hs* 4EHP-4E-BP1 complex (4E-BP1 C+L+NC)
Space group	$P4_2$	$P4_12_12$	C2	$P2_1$
Unit Cell				
Dimensions (Å)				
a, b, c	135.3, 135.3, 60.9	82.6, 82.6, 148.5	152.2, 98.6, 39.3	38.4, 83.4, 70.5
Angles (°)				
α, β, γ	90, 90, 90	90, 90, 90	90, 99.6, 90	90, 104.3, 90
Data collection				
Wavelength (Å)	1.000	1.000	1.000	0.999
Resolution (Å)	47.8-2.9	45.9-2.3	44.6-2.0	41.7-1.9
R_{sym}	0.111 (0.647)	0.125 (1.14)	0.066 (0.611)	0.145 (1.14)
Mean $I/\sigma I$	10.9 (1.95)	13.2 (2.06)	12.0 (2.09)	8.5 (2.07)
Completeness (%)	99.6 (99.9)	99.8 (98.3)	99.8 (99.9)	99.7 (99.6)
Multiplicity	3.4 (3.4)	11.2 (10.6)	5.1 (4.8)	6.6 (6.7)
Refinement				
Resolution (Å)	47.8-2.9	45.9-2.3	44.6-2.0	41.7-1.9
No. reflections	24694	23497	38550	33898
R_{work}/ R_{free}	0.204/0.254	0.205/0.242	0.195/0.231	0.226/0.251
No. atoms	7833	3811	3536	3547
Protein	7833	3741	3386	3305
Ligand/ion	/	/	20	21
Water	/	70	130	221
B-factors (Å^2)	49.4	62.0	63.9	27.8
Protein	49.4	62.3	64.2	27.3
Ligand/ion	/	/	88.6	33.5
Water	/	46.3	52.7	28.7
Ramachandran Plot				
Favored (%)	95.9	97.3	95.5	97.7
Disallowed (%)	0	0	0	0
Root-Mean-Square Deviation				
Bond lengths (Å)	0.003	0.004	0.011	0.003
Bond angles (°)	0.494	0.540	1.002	0.529

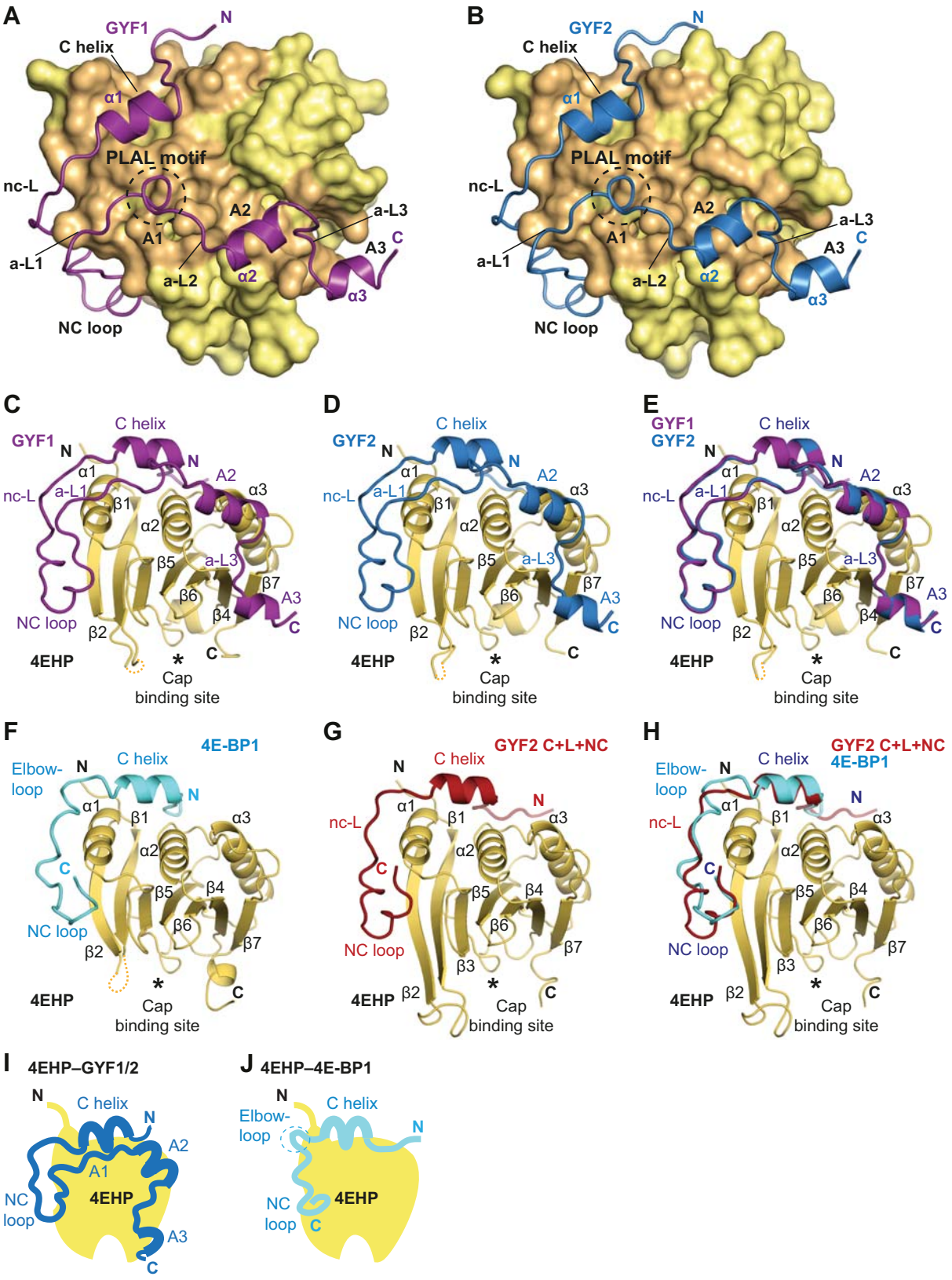
Values in parentheses are for highest-resolution shell.

Ligands: four $PO4^{3-}$ ions in the *Hs* 4EHP-GIGYF2 (C+NC) complex, seven formic acid molecules in the *Hs* 4EHP-4E-BP1 (C+NC) complex

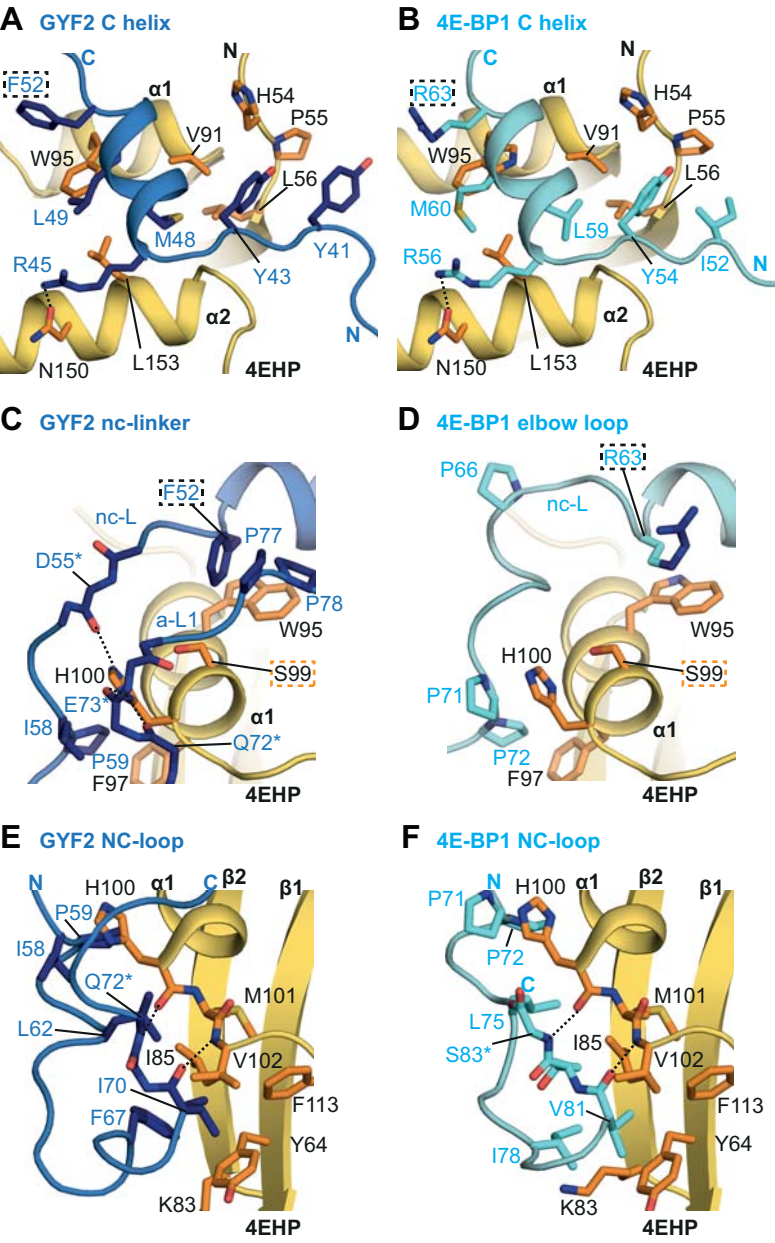
Peter *et al.* Fig. 1



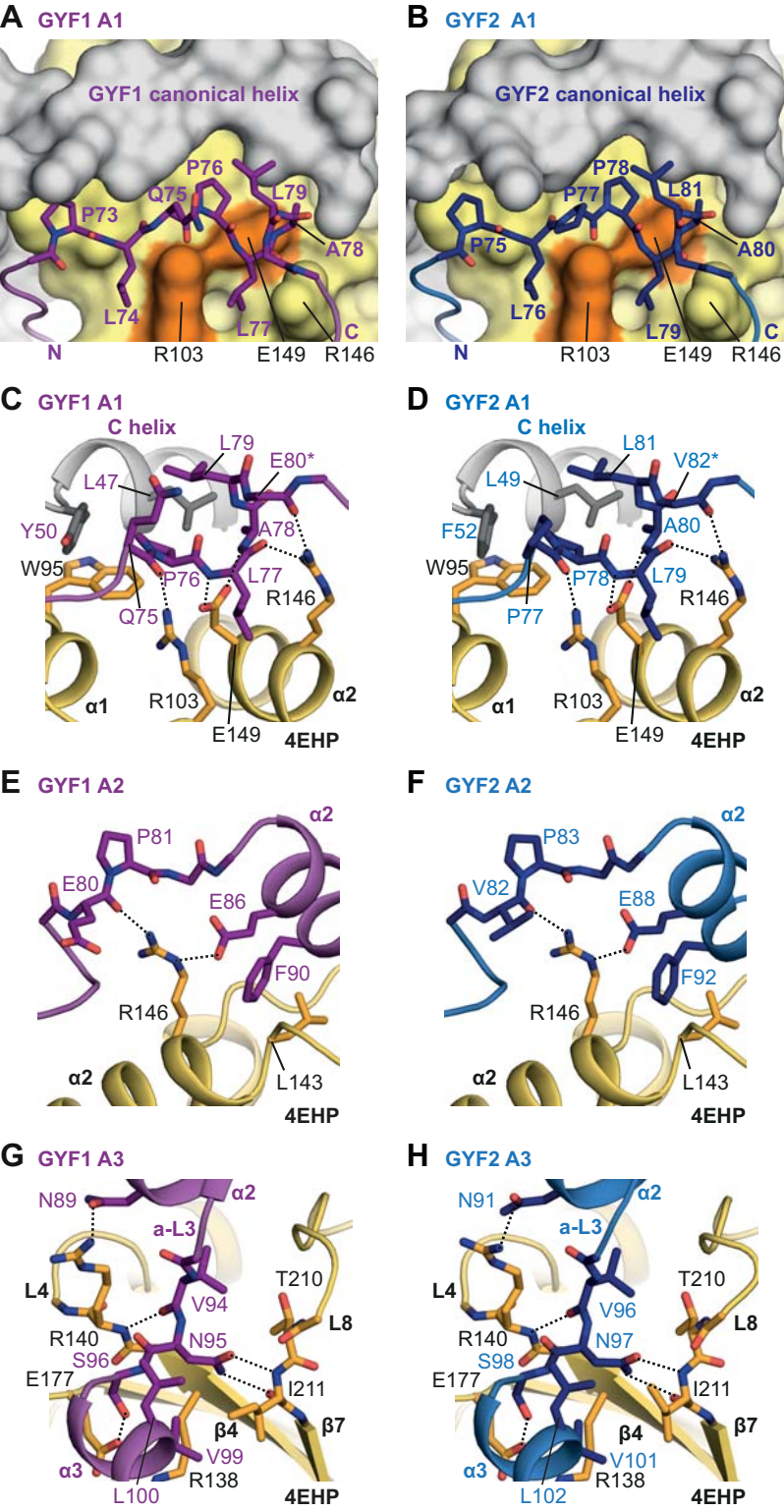
Peter *et al.* Fig. 2



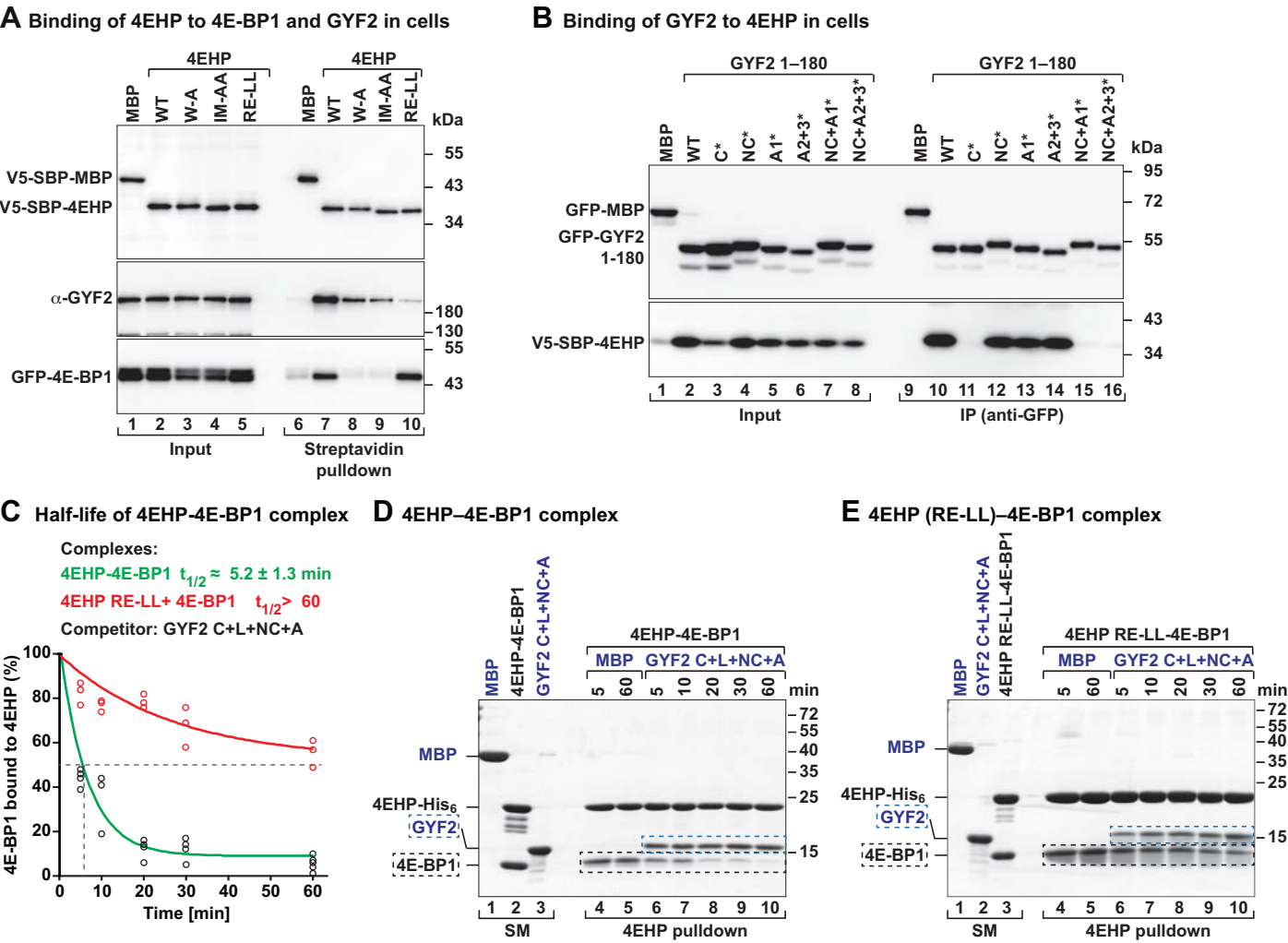
Peter *et al.* Fig. 3



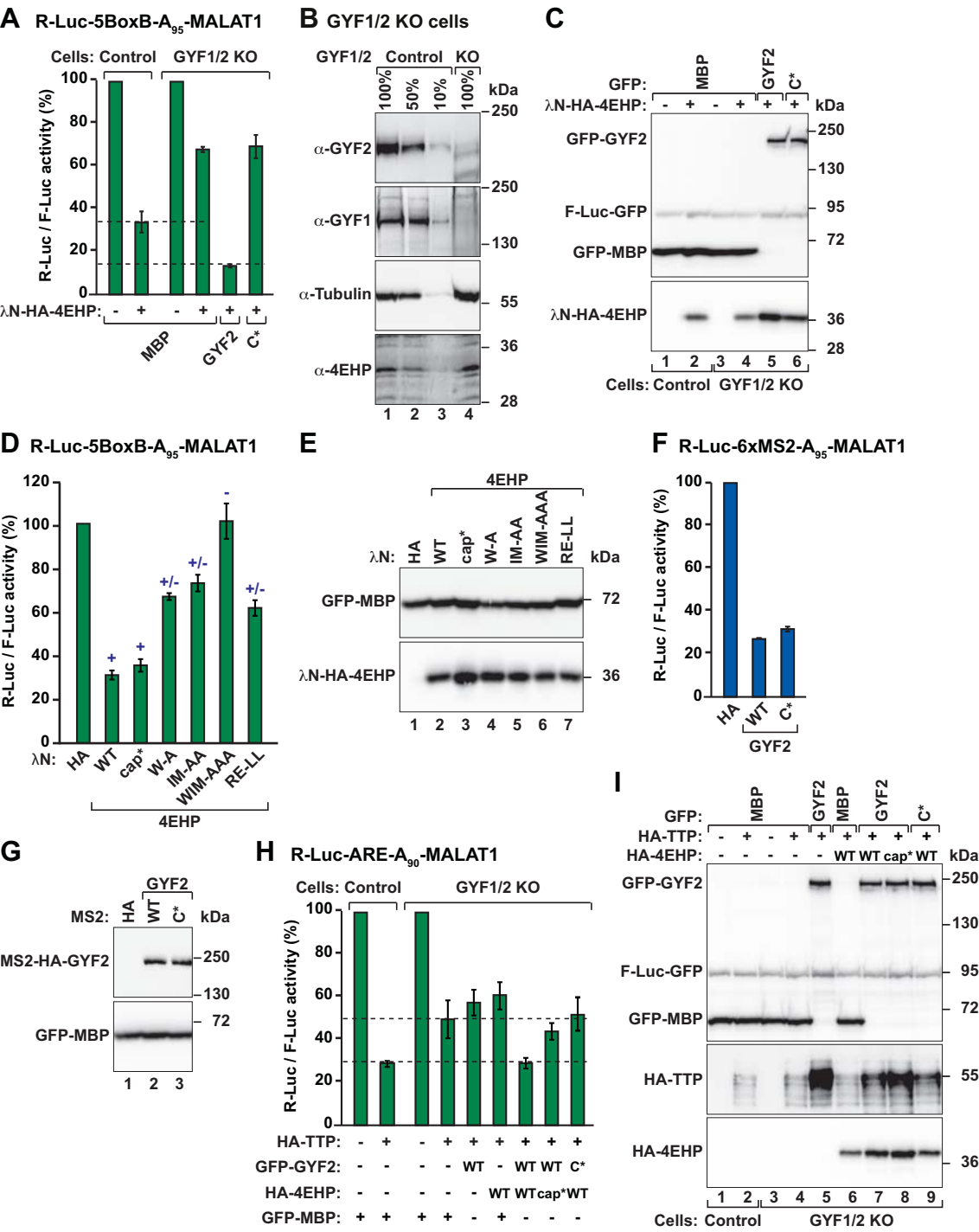
Peter *et al.* Fig. 4



Peter *et al.* Fig. 5



Peter *et al.* Fig. 6



Supplemental Materials

GIGYF1/2 proteins use auxiliary sequences to selectively bind to 4EHP and repress target mRNA expression

Daniel Peter, Ramona Weber, Felix Sandmeir, Lara Wohlbold, Sigrun Helms, Praveen Bawankar, Eugene Valkov, Cátia Igreja and Elisa Izaurralde

DNA constructs

The plasmids used for the expression of human eIF4E, 4E-BP1 and eIF4G1 (full-length or fragments) in *Escherichia coli* or in human cells have been previously described (Peter et al. 2015a; Grüner et al. 2016). The plasmids for the expression of 4EHP fragments (M1–F234) and (A52–F234) in *E. coli* were obtained by inserting the corresponding cDNA fragments either into the pnYC-NpH (between the XhoI and NheI restriction sites) or the pnYC-CvH (between the XhoI and BamHI restriction sites) vectors that include N- and C-terminal His_6 tags (Diebold et al. 2011), respectively. DNA fragments encoding for GYF1 [residues K33–K52 (C), K33–D71 (C+L+NC) and K33–M103 C+L+NC+A)] and GYF2 [residues K35–K54 (C), K35–Q72 (C+L+NC) and K35–T105 (C+L+NC+A)] were inserted into the NdeI-NheI and NdeI-XbaI restriction sites in the pnEA-NpM vector (Diebold et al. 2011), respectively. These constructs express GYF fragments that are N-terminally fused to an MBP-tag, which is cleavable by the HRV3C protease. A DNA fragment encoding the B1 domain of immunoglobulin-binding protein G (GB1; Cheng and Patel 2004) was inserted C-terminally to the GYF fragments by site-directed mutagenesis using the QuikChange mutagenesis kit (Stratagene).

The plasmids for the expression of V5-streptavidin binding protein (SBP)-tagged and λN-hemagglutinin (HA)-tagged 4EHP in human cells were obtained by inserting the full-length 4EHP cDNA into the XhoI and BamHI sites in the pT7-V5-SBP and pλN-HA-C1 vectors (Kuzuoglu-Ozturk et al. 2016), respectively. The plasmids for the expression of full-length GYF1 and GYF2, which are N-terminally fused to GFP or MS2-HA were obtained by inserting the GYF1 cDNA (XhoI-EcoRI, obtained from the Kazusa DNA Research Institute; sj03926) or the GYF2 cDNA (XhoI-BamHI) into the corresponding sites of the pT7-EGFP-C1 and pT7-MS2-HA-C1 vectors. cDNA fragments encoding for GYF1 (residues M1–C177) and GYF2 (residues M1–P180) were introduced into the XhoI and BamHI restriction sites in

the pT7-EGFP-C1 vector. The cDNA encoding 4E-T (eIF4E-Transporter protein; EIF4ENIF1) was inserted into the HindIII and BamHI restriction sites in the pT7-EGFP-C1 vector. The cDNA encoding TTP (Tristetraprolin, residues M1–P313; Fabian et al. 2013) was inserted between the XhoI and EcoRI restriction sites of the pλN-HA-C1 vector. To generate a reporter containing the ARE-element (pCIneo-R-Luc-ARE-A_{90}-MALAT1), the sequence of the ARE element present in the 3' UTR of the TNF (Tumor Necrosis Factor)-α mRNA was inserted twice into the 3' UTR of the pCIneo-R-Luc parental plasmid by site-directed mutagenesis. A cDNA containing a stretch of 90 A and the MALAT1 sequence was then inserted into the XhoI and NotI restriction sites of the R-Luc-ARE vector. The DNA sequence of the TNF-α ARE is as follows: TTATTTATTATTTATTTATTATTTATTTATTT. All of the mutants used in this study were generated by site-directed mutagenesis using the QuikChange mutagenesis kit (Stratagene). All of the constructs and mutations were confirmed by sequencing and are listed in Supplemental Table S1.

Pulldown and competition assays

In the pulldown assays shown in Fig. 1 and Supplemental Fig. S2 and S7, bacterial lysates expressing recombinant 4EHP-His_6 (residues M1–F234, wild-type and mutants) or purified eIF4E-His_6 (2 μM; 50 μg) were incubated with Ni-NTA beads for 30 min. The immobilized 4EHP and eIF4E proteins were then incubated for 30 min with bacterial lysates expressing GYF1, GYF2 or 4E-BP1 fragments (wild-type and mutants) that were N-terminally tagged with MBP and C-terminally tagged with GB1. Proteins associated with 4EHP or eIF4E were eluted with imidazole and analyzed by SDS-PAGE followed by Coomassie staining.

For the competition assays, purified 4EHP–4EBP1 complexes (2 μM) containing 4EHP (residues A52–F234)-His_6 and 4E-BP1 C+L+NC (residues R50–S83; with a C-terminal GB1

tag) were incubated with Ni-NTA beads for 30 min in 50 mM Na-phosphate (pH 7.0) and 200 mM NaCl. The immobilized complexes were then incubated with equimolar amount of purified, GB1-tagged competitor peptides or with MBP as a negative control. After the specified time points, the beads were pelleted and washed three times in the same buffer. Proteins bound to the Ni-NTA beads were eluted with the same buffer containing 500 mM imidazole and analyzed by SDS-PAGE followed by Coomassie staining. The amount of 4E-BP1 bound to 4EHP was quantified using the ImageJ software and normalized to 4EHP levels present at each time point. These values were set to 100 in the presence of MBP. Data points from at least three independent experiments were plotted and the resulting fitting curves were determined using the Levenberg-Marquardt algorithm for single exponential decay functions. The R^2 values associated with the fitting of the exponential decay curves were between 0.82 and 0.96.

ITC analysis

For the ITC measurements, the GB1-stabilized GYF1/2 peptides (wild-type and mutants) and 4E-BP1 peptides were purified as previously described for the other 4E-BPs (Igreja et al. 2014; Peter et al. 2015a,b). The 4EHP protein (residues A52–F234; wild-type and mutants) used in the ITC measurements was expressed with an N-terminal His_6 tag and purified from cleared cell lysates using a nickel column (HisTrap HP 5 ml, GE Healthcare). The His_6 tag was cleaved by HRV3C protease overnight at 4°C. The protein was further purified using a heparin column (HiTrap Heparin HP 5 ml, GE Healthcare) and a final purification on a Superdex 75 column (GE Healthcare). The 4EHP-GYF2 complex (GYF2 residues K35–T105) used for measuring the affinity for m^7GpppG cap analog was purified as described for the 4EHP–GYF2 complex used for crystallization. All of the proteins used in the ITC

measurements were stored at -80°C in a buffer consisting of 20 mM Na-phosphate (pH 7.0) and 200 mM NaCl.

The ITC experiments were performed using a VP-ITC microcalorimeter (MicroCal) at 20°C as described previously (Igreja et al. 2014; Peter et al. 2015a,b). A solution containing either 4EHP (residues A52–F234, wild-type, S99N mutant and dimerization mutant, 1-5 μM) or eIF4E (residues K36–V217, 5 μM) in a calorimetric cell was titrated with tenfold concentrated solutions of GB1-stabilized peptides that were dissolved in the same buffer (20 mM Na-phosphate (pH 7.0) and 150 mM NaCl). The following peptides were used: GYF1 (C, residues K33–K52, 50 μM; C+L+NC, residues K33–D71, 20 μM; C+L+NC+A wild type or dimerization mutant, residues K33–M103, 10 μM), GYF2 (C, residues K35–K54, 50 μM; C+L+NC, residues K35–Q72, 20 μM; C+L+NC+A wild type or dimerization mutant, residues K35–T105, 10 μM) and 4E-BP1 (C+L+NC, residues T50–S83, 50 μM). The affinity for the m^7GpppG cap analog was measured in a buffer containing 20 mM HEPES (pH 7.5) and 200 mM NaCl by titrating a solution of m^7GpppG (400 μM; New England Biolabs) into a solution of 4EHP (residues A52–F234, 40 μM) or 4EHP in complex with GIGYF2 (residues K35–T105, 40 μM) diluted in the same buffer.

The titration experiments consisted of an initial injection of 2 μl followed by 28 injections of 10 μl at 240 s intervals. Each binding experiment was repeated three times. Correction for dilution heating and mixing was achieved by subtracting the final baseline, which consisted of small peaks of similar size. The thermodynamic parameters were estimated using a one-site binding model (Origin version 7.0), whereby the data points for the first injection were removed from the analysis (Mizoue and Tellinghuisen 2004). Because the protein concentration used in these measurements is low (1 μM for 4EHP in the calorimetric cell), dimerization of the 4EHP–GYF1/2 complexes is unlikely to occur and thus it does not contribute to the measured binding constants. Accordingly, GYF2 and 4EHP dimerization

mutants still display a binding affinity similar to that observed for the complexes containing the wild type proteins (Supplemental Table S2 and Fig. S7).

Crystallization

Crystals of 4EHP (residues A52–F234) in complex with GYF1 (residues K33–M103) were obtained at 18°C using the hanging-drop vapor diffusion method two days after mixing the protein solution (16 mg/ml; 1 μl) with the crystallization solution (1 μl) containing 20% PEG 3350 in 0.2 M potassium nitrate. Crystals of 4EHP (residues A52–F234) bound to GYF2 (residues A35–T105) were obtained at 18°C using the hanging-drop vapor diffusion method one day after mixing the protein solution (16 mg/ml, 1 μl) with the crystallization solution (1 μl) containing 0.1 M sodium citrate (pH 5.0), 0.1 M magnesium chloride and 12% PEG 4000.. Crystals of 4EHP (residues A52–F234) in complex with GYF2 (residues A35–Q72) were obtained at 18°C using the hanging-drop vapor diffusion method. Crystals grew in one day after mixing the protein solution (16 mg/ml, 1 μl) with the crystallization solution (1 μl) containing in 0.1 M sodium acetate (pH 4.6) and 0.6 M diammonium phosphate. All of the crystals containing GYF peptides were soaked in mother liquor supplemented with 10–15% glycerol for cryoprotection before flash-freezing in liquid nitrogen.

Crystals of 4EHP (residues A52–F234) in complex with 4E-BP1 (residues T50–S83) were obtained at 18°C using the sitting-drop vapor diffusion method. The crystals grew three days after mixing the protein solution (16.5 mg/ml; 0.1 μl) with the crystallization solution (0.1 μl) containing 0.1 M sodium acetate (pH 4.6) and 1.7 M sodium formate. The crystals were cryoprotected in mother liquor supplemented with 3.5 M sodium formate and flash-frozen in liquid nitrogen.

Data collection and structure determination

The data for all the crystals were collected at 100K on a PILATUS 6M detector at the PXII beamline at the Swiss Light Source. Diffraction data were processed with XDS and scaled using XSCALE (Kabsch 2010). The phases were obtained by molecular replacement using PHASER (McCoy et al., 2007). For the 4EHP–GYF2 (residues K35–T105; C+L+NC+A) complex, the structure of human 4EHP (PDB 2JGB; Rosettani et al. 2007) was used as a search model with an asymmetric unit containing two copies of the model. To solve the structures of the 4EHP–GYF2 (residues K35–Q72; C+L+NC) and 4EHP–4E-BP1 (residues T50–S83; C+L+NC) complexes, two copies of 4EHP from the 4EHP–GYF2 (C+L+NC+A) complex were used as a search model. In the case of the 4EHP–GYF1 (residues K33–M103; C+L+NC+A) complex, four copies of the 4EHP–GYF2 (C+L+NC+A) complex were used as a search model. To minimize model bias, the molecular replacement solutions were used to rebuild the initial models using the PHENIX AutoBuild wizard (Terwilliger et al., 2008). To complete the structure, iterative cycles of model building and refinement were performed with COOT (Emsley et al. 2010) and PHENIX (Afonine et al. 2012), respectively. The GYF2 (C+L+NC+A and C+L+NC) and 4E-BP1 (C+L+NC) peptide chains were manually built into the difference density in COOT and further refined with PHENIX. In the final refinement rounds for the 4EHP–GYF1 (C+L+NC+A) and 4EHP–GYF2 (C+L+NC+A) complexes, translation/libration/screw (TLS) parameters were refined for the peptide chains in addition to the individual B-factors; in the case of the GYF1 complex, non-crystallographic symmetry (NCS) torsional restraints were also used in refinement.

The stereochemical properties for all of the structures were verified with MOLPROBITY (Chen et al. 2010), and structural images were prepared with PyMOL (http://www.pymol.org). The diffraction data and refinement statistics are summarized in Table 1.

Small-angle X-ray scattering (SAXS)

SAXS experiments were conducted at the SWING beamline at the SOLEIL synchrotron. Data collection for the 4EHP–GYF2 complexes was performed in-line with size exclusion chromatography (Superdex 200 Increase 5/150 GL, GE Healthcare) using an Agilent HPLC system in a buffer containing 20 mM HEPES pH 7.5, 200 mM NaCl, and 1 mM TCEP [Tris(2-carboxyethyl)phosphine]. The scattering data were collected at 1 s exposures using an Aviex charge-coupled device detector at a sample-detector distance of 1798 mm and a wavelength of 1.033 Å. Data reduction to absolute units, frame averaging and buffer subtraction were performed using the FOXTROT software (Xenocs, France). Theoretical scattering curves and fitting to the experimental SAXS data was performed using the FoXS software (Schneidman-Duhovny et al. 2013). To ensure protein stability during SAXS data collection, all the 4EHP–GYF2 complexes were measured with a 1.5x molar excess of m^7GpppG cap analog (New England Biolabs) in the protein samples. Therefore, the coordinates of the structures used during the fitting procedures were adjusted such that the 4EHP cap-binding loops were fixed in the bound conformation including the cap analog, which was based on the structure of the m^7GTP-bound 4EHP (PDB 2JGB; Rosettani et al. 2007).

Generation of GYF1/2-null cell line

Two sgRNAs targeting GYF1 and two sgRNAs targeting GYF2 were designed using the DNA 2.0 (ATUM) or CHOPCHOP (http://chopchop.cbu.uib.no) online tools and cloned into the pSpCas9(BB)-2A-Puro (PX459) vector [a gift from F. Zhang, Addgene plasmid 48139; (Ran et al., 2013)]. HEK293T cells were transfected with the sgRNA-Cas9 vectors and selected with puromycin (3 μg ml^{-1}) to obtain stable GYF1/2 knockout cells. To obtain clonal cell lines, single cells were distributed in 96-well plates using serial dilutions. Genomic

DNAs from single clones were isolated using the Wizard SV Genomic DNA Purification System (Promega) and the targeted GYF1 and GYF2 loci were PCR amplified and sequenced to confirm gene editing. For GYF1 we observed two frameshift mutations in exon 7 (4 bp deletion and a 8 bp deletion together with a C->T mutation) targeted by sgGYF1-a and two indels in exon 16 (insertion of 43 bp or 93 bp) produced by sgGYF1-b. These mutations changed the GYF1 reading frame after the respective targeted site and introduced premature STOP codons. One frameshift mutation (16 bp deletion in the first exon removing the start codon) was detected for the GYF2 locus (targeted by sgGYF2-a and sgGYF2-b). This deletion was caused by sgGYF2-b. In contrast, sgGYF2-a did not target the genomic locus as the sequence around this target site is wild-type. The knockouts of GYF1/2 were further confirmed by western blotting. For the GYF2 gene we observe low levels of truncated protein fragments that are consistent with translation initiation at internal AUGs (Figure 6B, lane 4). Taking the GYF2 sequence and the position of the mutations into account, these truncated forms lack the 4EHP-binding region and the expression levels are approximately 10% of wild-type levels. The following guide sequences were used: sgGYF1-a: 5'-GCCAGCGGTCGCCGTCTCGC-3'; sgGYF1-b: GACAAGGACCGGCTCATCGT-3'; sgGYF2-a: 5'- ATTTTGAAAACTCACCATTC-3'; sgGYF2-b: 5'-AATACGGAAAAGAATGGCAG

Supplemental Table S1. Mutants and constructs used in this study.

Protein	Name of the construct	Fragments / mutations	Binding site / motif
***Hs* 4EHP** (isoform 1) O60573-1	4EHP	Full-length (1–245)	
	4EHP ΔC-term	1–234	Δ235–245
	4EHP truncated	52–234	Δ1–51 & 235–245
	W-A	W95A	Dorsal surface
	IM-AA	I85A, M101A	Lateral surface
	WIM-AAA	W95A, I85A, M101A	Dorsal + lateral surface
	RE-LL	R103L, E149L	Auxiliary surface
	WRE-ALL	W95A, R103L, E149L	Dorsal + auxiliary surface
	S99N	S99N	Dorsal surface
	Cap mutant (cap*)	W124A	Cap-binding pocket
	D* (dimer mutant)	Q159S, M161D, R202E	Dimer interface
***Hs* eIF4E** (isoform 1) P06730-1	4E	Full-length (1–217)	
	4E trunc	36–217	
	Cap mutant (cap*)	W102A	Cap-binding pocket
***Hs* GIGYF1** O75420	GYF1	Full-length (1–1035)	
	C+L+NC+A	33–103	Complete 4EHP-binding region
	C+L+NC	33–71	Bipartite 4EHP-binding region
	C	33–52	Canonical 4EHP-binding region
	C*	Y39A, Y41A, M46A, L47A	Canonical
	NC*	L60D, F65D, V68D	Non-canonical
	A1*	P76D, L77A	Auxiliary site 1
	A2*	E86A, N95F	Auxiliary site 2
	NC+A1*	L60D, F65D, V68D, P76D, L77A	Non-canonical + auxiliary site 1
	NC+A2+3*	L60D, F65D, V68D, E86A, N95F	Non-canonical + auxiliary site 2
	1-177		N-terminus
	D* (dimer mutant)	E44A, E45F, Q87A	Dimer interface
***Hs* GIGYF2** (isoform 1) Q6Y7W6-1	GYF2	Full-length (1–1299)	
	C+L+NC+A	35–105	Complete 4EHP-binding region
	C+L+NC	35–72	Bipartite 4EHP-binding region
	C	35–54	Canonical 4EHP-binding region
	C*	Y41A, Y43A, M48A, L49A	Canonical motif
	NC*	L62D, F67D, I70D	Non-canonical
	A1*	P78D, L79A	Auxiliary site 1
	A2*	E88A, N96F	Auxiliary site 2
	NC+A1*	L62D, F67D, I70D, P78D, L79A	Non-canonical + auxiliary site 1
	NC+A2+3*	L62D, F67D, I70D, E88A, N97F	Non-canonical + auxiliary site 2
	1-180		N-terminus
	D* (dimer mutant)	E46A, E47F, Q89A	Dimer interface
***Hs* 4E-BP1** Q13541	4E-BP1	Full-length (1–118)	
	4E-BP1 C+L+NC	50–83	eIF4E-binding region
	C*	Y54A, L59A	Canonical motif
	NC*	L75A, V81A	Non-canonical
	C+NC*	Y54A, L59A, L75A, V81A	Canonical+ non-canonical
***Hs* 4E-T** Q9NRA8	4E-T	Full-length (1–985)	
***Hs* TTP** (1–326) P26651	TTP ΔNIM	1–313	Δ314-326, deletion of the NOT1 interacting motif (NIM)

Supplemental Table S2. Thermodynamic parameters for the interaction of 4EHP and eIF4E with the indicated peptides.

GYF peptides *vs* 4EHP					
GYF protein	**K_D (nM)**	**ΔH (kcal mol^{-1})**	**-TΔS (kcal mol^{-1})**	**ΔG (kcal mol^{-1})**	**Molar ratio**
GYF1 C (33-52)	360 ± 120	-23 ± 3	14.6	-8.7	1.00 ± 0.01
GYF1 C+L+NC (33-71)	12 ± 2	-30 ± 1	19.3	-10.6	1.00 ± 0.01
GYF1 C+L+NC+A (33-103)	0.4 ± 0.2	-37 ± 4	24.4	-12.6	1.00 ± 0.02
GYF2 C (35-54)	290 ± 160	-22 ± 2	13.5	-8.8	1.00 ± 0.01
GYF2 C+L+NC (35-72)	14 ± 1	-23 ± 2	12.3	-10.6	1.00 ± 0.01
GYF2 C+L+NC+A (35-105)	0.3 ± 0.1	-32 ± 1	19.1	-12.8	1.00 ± 0.01
GYF peptides *vs* 4EHP dimerization mutants					
GYF protein	**K_D (nM)**	**ΔH (kcal mol^{-1})**	**-TΔS (kcal mol^{-1})**	**ΔG (kcal mol^{-1})**	**Molar ratio**
GYF1 C+L+NC+A (33-103) D*	0.4 ± 0.3	-34 ± 1	21.6	-12.7	1.01 ± 0.01
GYF2 C+L+NC+A (35-105) D*	0.5 ± 0.3	-30.8 ± 0.5	18.1	-12.7	1.01 ± 0.02
4EBP1 C+L+NC *vs* eIF4E or 4EHP					
4E molecule	**K_D (nM)**	**ΔH (kcal mol^{-1})**	**-TΔS (kcal mol^{-1})**	**ΔG (kcal mol^{-1})**	**Molar ratio**
eIF4E	5 ± 2	-18 ± 1	6.4	-11.2	1.00 ± 0.01
4EHP	55 ± 14	-16.4 ± 0.8	6.6	-9.8	1.01 ± 0.01
4EHP S99N	4 ± 1	-21 ± 2	9.7	-11.3	1.01 ± 0.01
m^{7}GpppG cap analog *vs* 4EHP or 4EHP-GIGYF2 complex					
Protein	**K_D (μM)**	**ΔH (kcal mol^{-1})**	**-TΔS (kcal mol^{-1})**	**ΔG (kcal mol^{-1})**	**Molar ratio**
4EHP	4 ± 1	-7.3 ± 0.4	0.1 ± 0.5	-7.2	1.01 ± 0.01
4EHP–GYF2 (35–105) complex	6 ± 3	-9 ± 2	1 ± 3	-7.2	1.01 ± 0.01

Note that the presence of the auxiliary sequences increases the entropic penalty (-TΔS) of the interaction between GYF1/2 and 4EHP compared to that of the peptides lacking these sequences [$\Delta(-T\Delta S)^{GYF1}$ = 5.1 kcal/mol^{-1}, $\Delta(-T\Delta S)^{GYF2}$ = 6.8 kcal/mol^{-1}]. One explanation for the increase in the entropic penalty is a higher disorder-to-order transition for the binding of the GYF1/2 C+L+NC+A peptides compared to the C+L+NC peptides. This is supported by the crystal structures in which the auxiliary sequences fold into two α-helices in complex with 4EHP.

Supplemental Table S3. Experimental and theoretical SAXS parameters for different 4EHP–GYF complexes.

Experimental parameters					
4EHP bound to:	**R_g (Guinier) [Å]**	**R_g (real space) [Å]**	**D_{max} [Å]**	**Exp. I(0) [x10^{-2}]**	**Concentration [mg/ml]**
GYF2 C+L+NC	21.1	21.2	71.1	4	10
GYF2 C+L+NC+A	26.1	26.2	90.3	7.9	10
GYF2 C+L+NC+A	25.8	25.8	89.3	3.6	5
GYF2 C+L+NC+A	24.7	24.8	83.7	1.7	2.5
GYF2 C+L+NC+A	23.7	23.7	80.2	0.7	1.25
4EHP D mutant bound to:	**R_g (Guinier) [Å]**	**R_g (real space) [Å]**	**D_{max} [Å]**	**Exp. I(0) [x10^{-2}]**	**Concentration [mg/ml]**
GYF2 C+L+NC+A D* (dim. mutant)	20.6	20.6	73.7	1.3	5

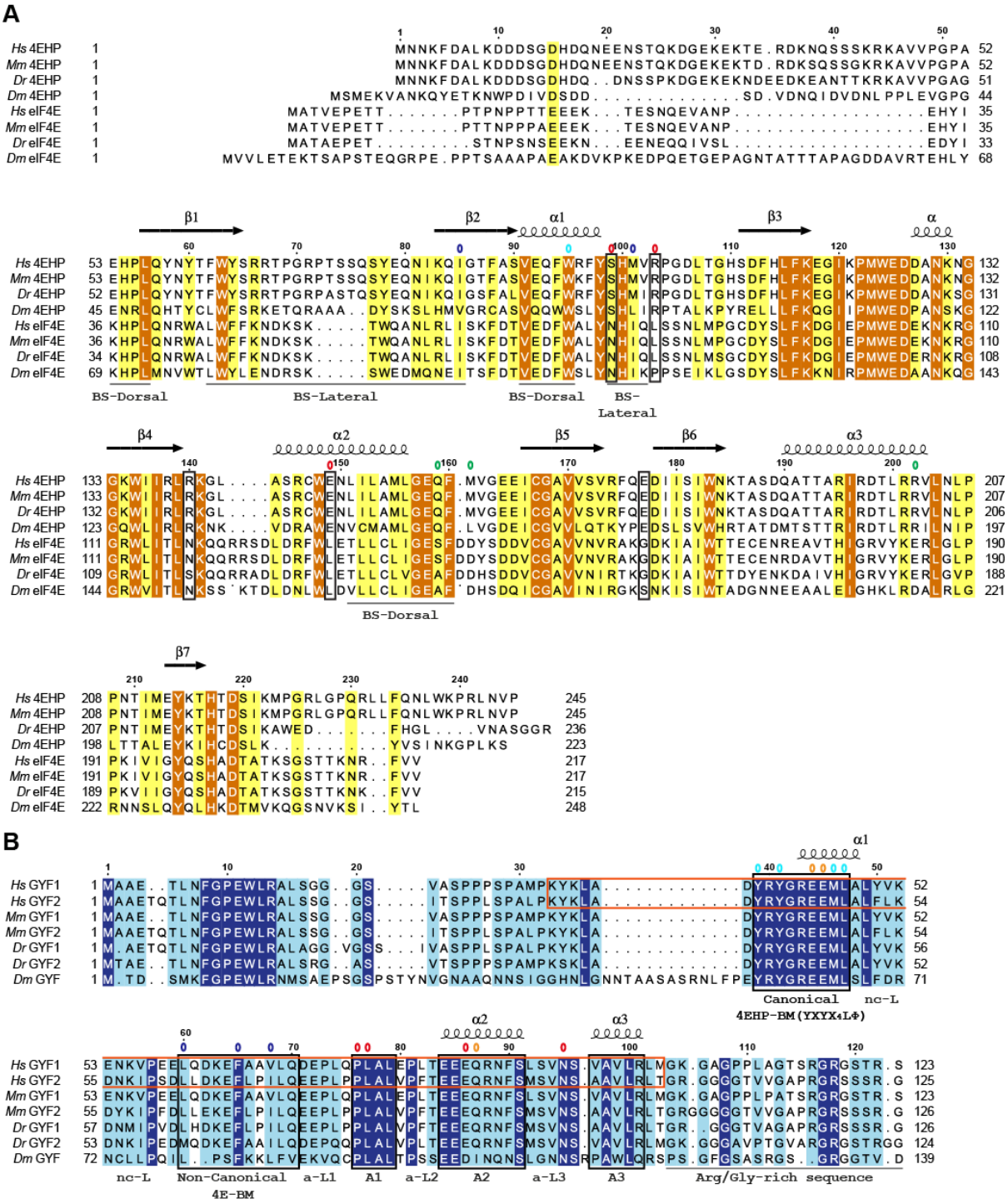
Theoretical parameters			
4EHP–GYF2 C+L+NC+A			
Single complex		**Symmetric dimer**	
R_g (Guinier) [Å]	**D_{max} [Å]**	**R_g (Guinier) [Å]**	**D_{max} [Å]**
17.7	64	24.7	85
4EHP–GYF2 C+L+NC			
Single complex		**Symmetric dimer**	
R_g (Guinier) [Å]	**D_{max} [Å]**	**R_g (Guinier) [Å]**	**D_{max} [Å]**
17.1	64	26.1	90

Supplemental Table S4. Antibodies used in this study.

Antibody	**Source**	**Catalog Number**	**Dilution**	**Monoclonal/ Polyclonal**
Anti-HA-HRP (Western blot)	Roche	12 013 819 001	1:5,000	Monoclonal
Anti-HA (Immunoprecipitation)	Biolegend	MMS-101P	1:1,000	Monoclonal
Anti-*Hs* GYF2	Bethyl laboratories	A303-731A	1:1,000	Rabbit polyclonal
Anti-*Hs* GYF1	Bethyl laboratories	A304-132A-M	1:1,000	Rabbit polyclonal
Anti-*Hs* 4E-T	Abcam	ab95030	1:2,000	Rabbit polyclonal
Anti-*Hs* 4EHP	In house		1:200	Rabbit polyclonal
Anti-*Hs* eIF4E	Bethyl laboratories	A301-154A	1:2,000	Rabbit polyclonal
Anti-*Hs* 4E-BP1	Cell Signaling Technology	9452	1:1,000	Rabbit polyclonal
Anti-GFP	In house		IP	Rabbit polyclonal
Anti-GFP	Roche	11814460001	1:2,000	Monoclonal
Anti-rabbit-HRP	GE Healthcare	NA934V	1:10,000	Polyclonal
Anti-mouse-HRP	GE Healthcare	RPN4201	1:10,000	Polyclonal
Anti-V5	QED Bioscience Inc.	18870	1:5,000	Rabbit polyclonal
Anti-V5	LSBio LifeSpan BioSciences, Inc.	LS-C57305	1:5,000	Monoclonal
Anti-tubulin	Sigma Aldrich	T6199	1:10,000	Monoclonal

Supplemental Figures

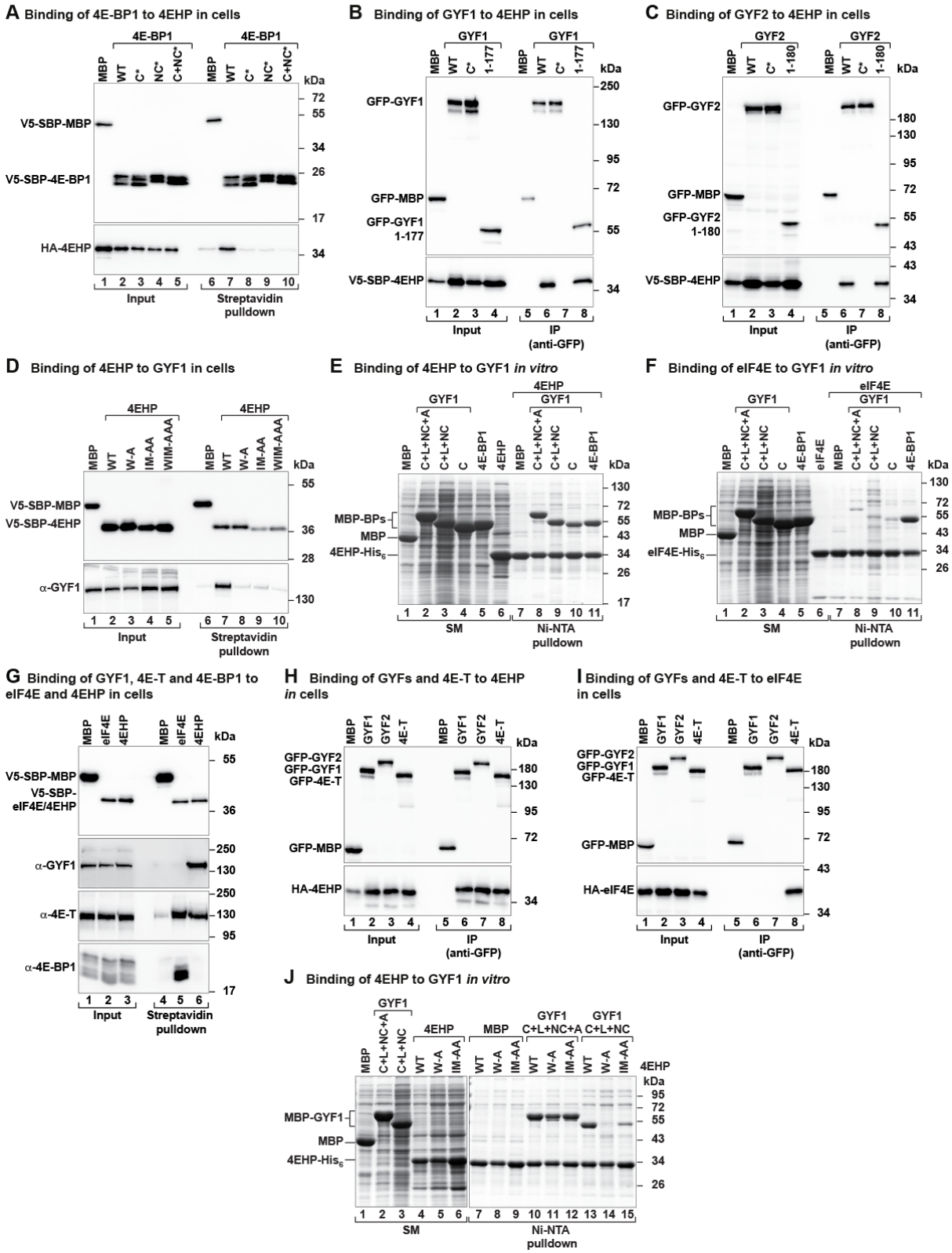
Figure S1



Supplemental Figure S1. Sequence alignments. In all aligned sequences, residues with >70% similarity are shown with a light color background and conserved residues are highlighted with a darker background and printed in white. Secondary structure elements are indicated above the sequences for 4EHP and GYF1 and are based on the structures presented

in this study. (*A*) Sequence alignment of 4EHP and eIF4E orthologous proteins from *Homo sapiens* (*Hs*), *Mus musculus* (*Mm*), *Danio rerio* (*Dr*) and *Drosophila melanogaster* (*Dm*). Residues highlighted in black boxes are specific for 4EHP and are relevant for the interactions described in this study. The dorsal and lateral binding surfaces (BS) of 4EHP are indicated by a line below the sequences. Residues that were mutated in this study are indicated by open circles colored as follows: cyan (dorsal surface), blue (lateral surface), red (4EHP specific residues) and green (dimerization). (*B*) Sequence alignment of GYF proteins. The canonical (C), non-canonical (NC) and auxiliary (A1, A2, A3) sequences are boxed in black. The GYF1/2 sequences visible in the crystal structures are indicated with a red box. Only a short stretch of the Arg/Gly-rich sequence adjacent to the auxiliary motif is shown and underlined. The species are as in A. Open circles above the alignment indicate the residues mutated in this study and are colored as follows: cyan (canonical), blue (non-canonical), red (auxiliary) and orange (dimerization).

Figure S2

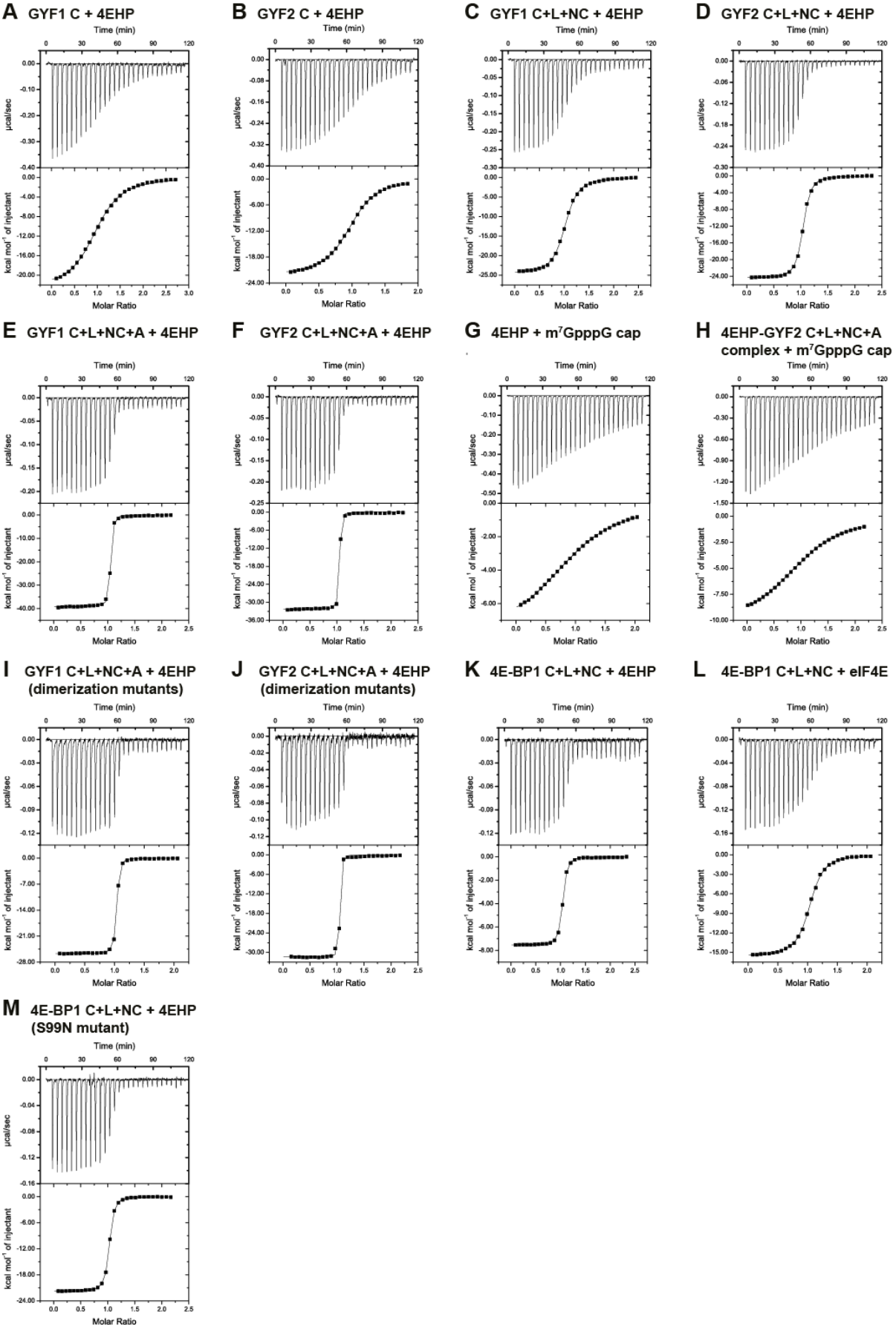


Supplemental Figure S2. Interaction of GYF1, GYF2 and 4E-BP1 with 4EHP. (*A*) The interaction of HA-4EHP with V5-SBP-4E-BP1 (wild-type or the indicated mutants) was tested in HEK293T cell lysates. The proteins were pulled down using streptavidin-coated

beads. V5-SBP-MBP served as negative control. The inputs (1.5%) and bound fractions (3% for the V5-proteins and 5% for HA-4EHP) were analyzed by western blotting using anti-HA and anti-V5 antibodies. (*B*) The interaction of GFP-GYF1 [either full-length, canonical mutant (C*) or N-terminal fragment (residues 1–177)] with V5-SBP-4EHP was analyzed by immunoprecipitation assay in HEK293T cells using anti-GFP antibodies. GFP-MBP served as negative control. The input samples (1.5%) and the immunoprecipitates (10%) were analyzed by western blotting using anti-V5 and anti-GFP antibodies. GYF1 residues 1–177 bound to 4EHP to a similar extent as the full-length protein, indicating that this protein fragment contains the principal 4EHP-binding region of the protein. (*C*) The interaction of GFP-GYF2 [either full-length, canonical mutant (C*) or N-terminal fragment (residues 1–180)] with V5-SBP-4EHP was analyzed as described in B. GYF2 residues 1–180 bound to 4EHP to a similar extent as the full-length protein. (*D*) Western blot showing the interaction of V5-SBP-4EHP (wild-type or the indicated mutants) with endogenous GYF1. The proteins were pulled down using streptavidin-coated beads. V5-SBP-MBP served as negative control. The inputs (1.5% for the V5-tagged proteins and 3% for GYF1) and bound fractions (3% for V5-tagged proteins and 35% for GYF1) were analyzed by western blotting using anti-V5 and anti-GYF1 antibodies. (*E, F*) Ni-NTA pulldown assays showing the interactions of GYF1 fragments (C+L+NC+A, C+L+NC and C) with 4EHP-His_6 (E) or eIF4E-His_6 (F). The eIF4E-binding region of 4E-BP1 (C+L+NC) binds similarly to both 4EHP and eIF4E, whereas GYF1 associates preferentially with 4EHP. The GYF1 and 4E-BP1 peptides contain an N-terminal MBP-tag and a C-terminal GB1 tag. The starting material (4% for the GYF1 fragments, 6% for 4EHP and recombinant eIF4E) and bound fractions (10% and 15% in panels E and F, respectively) were analyzed by SDS-PAGE followed by Coomassie blue staining. MBP served as a negative control. (*G*) The interaction of V5-SBP-tagged eIF4E or 4EHP proteins with endogenous GYF1, 4E-T and 4E-BP1 was analyzed in HEK293T cell

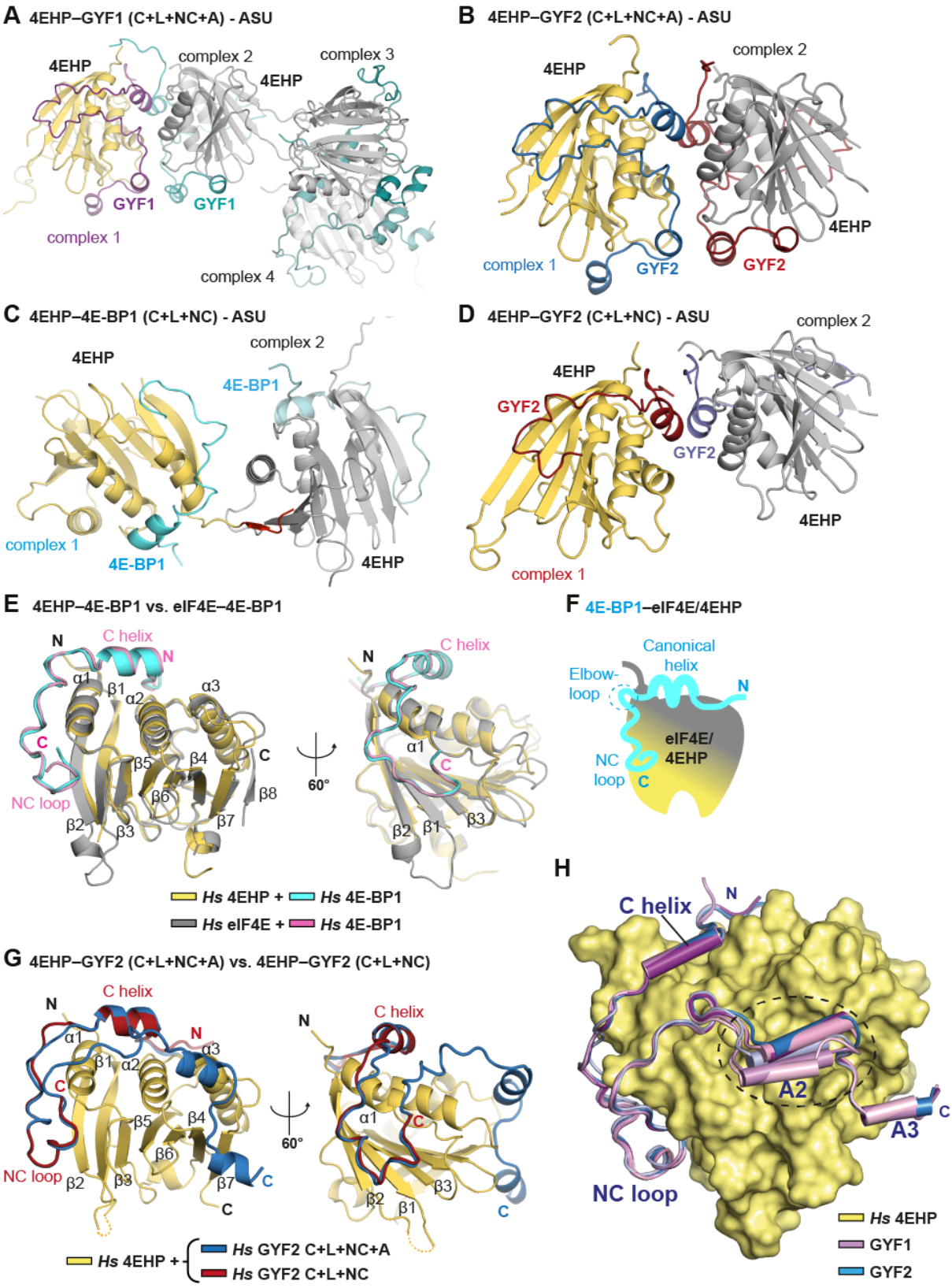
lysates. The proteins were pulled down using streptavidin-coated beads. Inputs (1.5%) and bound fractions (30% for 4E-BP1, GYF1 and 4E-T and 5% for the V5-SBP-tagged proteins) were analyzed by western blotting using anti-V5, anti-4E-BP1, anti-4E-T and anti-GYF1 antibodies. (*H, I*) Western blot analysis showing the interaction of GFP-tagged GYF1, GYF2 and 4E-T with HA-4EHP (H) or HA-eIF4E (I) in HEK293T cells. The proteins were immunoprecipitated using anti-GFP antibodies. The inputs (0.75% for GFP-tagged proteins and 0.5% for the HA-tagged proteins) and immunoprecipitates (15% for GFP-tagged proteins and 25% for HA-tagged proteins) were analyzed using anti-GFP and anti-HA antibodies, respectively. (*J*) Ni-NTA pulldown assay showing the interaction of 4EHP-His_6 (wild-type, W-A and IM-AA mutants) with GYF1 fragments with or without the auxiliary region (C+L+NC+A vs. C+L+NC). MBP served as a negative control. The starting material (4%) and bound fractions (9%) were analyzed by SDS-PAGE followed by Coomassie blue staining.

Figure S3


A GYF1 C + 4EHP
B GYF2 C + 4EHP
C GYF1 C+L+NC + 4EHP
D GYF2 C+L+NC + 4EHP
E GYF1 C+L+NC+A + 4EHP
F GYF2 C+L+NC+A + 4EHP
G 4EHP + m7GpppG cap
H 4EHP-GYF2 C+L+NC+A complex + m7GpppG cap
I GYF1 C+L+NC+A + 4EHP (dimerization mutants)
J GYF2 C+L+NC+A + 4EHP (dimerization mutants)
K 4E-BP1 C+L+NC + 4EHP
L 4E-BP1 C+L+NC + eIF4E
M 4E-BP1 C+L+NC + 4EHP (S99N mutant)
Time (min)
µcal/sec
kcal mol-1 of injectant
Molar Ratio

Supplemental Figure S3. Calorimetric titration data for the interaction of 4EHP with peptides derived from GYF1, GYF2 and 4E-BP1 or with m^7GpppG cap analog. (*A–F*) Isothermal titration calorimetry (ITC) profiles for the interaction of 4EHP (residues 52–234) with the following peptides: (A) GYF1 C; (B) GYF2 C; (C) GYF1 C+L+NC; (D) GYF2 C+L+NC; (E) GYF1 C+L+NC+A; (F) GYF2 C+L+NC+A. (*G*) ITC profile for the binding of 4EHP (residues 52-234) to m^7GpppG cap analog. (*H*) ITC profile for the binding of 4EHP-GYF2 C+L+NC+A complex to m^7GpppG cap analog. (*I, J*) ITC profiles for the interaction of 4EHP dimerization mutant with GYF1 and GYF2 (C+L+NC+A) dimerization mutant peptides. (*K - M*) ITC profiles for the interaction of 4E-BP1 (residues 50–83, C+L+NC) with the following proteins: (K) wild-type 4EHP (residues 52–234); (L) wild-type eIF4E (residues 36–217); (M) 4EHP (residues 52–234) S99N mutant. The thermodynamic parameters are shown in Table S2. Upper panels show raw data in (μcal sec^{-1}) and lower panels represent the integration of heat changes associated with each injection (kcal mol^{-1} of injectant). Data was fit using a one-site binding model.

Figure S4

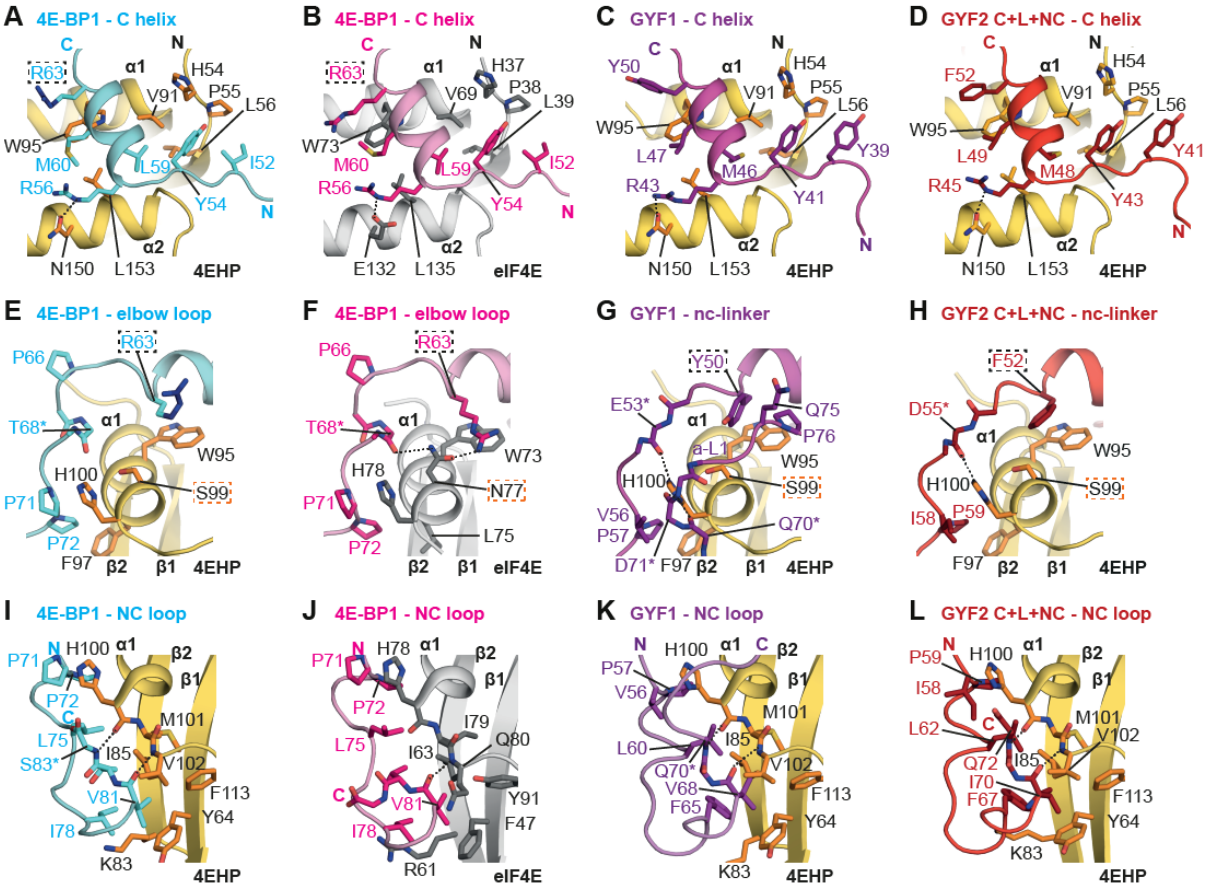


Supplemental Figure S4. Structures of 4EHP bound to GYF1, GYF2 and 4EBP1. (*A*) Cartoon representation showing the asymmetric unit (ASU) of the 4EHP–GYF1 crystal

form. The ASU contains four 4EHP–GYF1 complexes. In complex 1, which was used for representation, GYF1 is colored in purple and 4EHP is colored in yellow; in the other complexes of the ASU GYF1 is colored in cyan and 4EHP in grey. (*B*) Cartoon representation showing the ASU of the 4EHP–GYF2 crystal form. The ASU contains two 4EHP–GYF2 complexes. In complex 1, GYF2 is colored in blue and 4EHP in yellow. In complex 2, GYF2 is colored in red and 4EHP in grey. (*C*) ASU of the 4EHP–4E-BP1 (C+L+NC) crystal form. There are two 4EHP–4E-BP1complexes in the ASU. In complex 1, 4E-BP1 is colored in cyan and 4EHP is colored in yellow. In complex 2, 4EHP is colored in grey. The N-terminal portion of the 4EHP molecule from complex 1 is colored in red and contains residues from the expression tag, which mediate contacts to complex 2. (*D*) Cartoon representation showing the ASU of the 4EHP–GYF2 (C+L+NC) crystal structure. The ASU contains two 4EHP–GYF2 complexes. In complex 1, GYF2 is colored in red and 4EHP yellow. In complex 2, GYF2 is colored in blue and 4EHP in grey. The structural arrangement of the two complexes that lack the GYF2 auxiliary sequences appears similar to the dimeric arrangement of the complexes containing the auxiliary sequences (panel B). (*E*) Superposition of the structure of 4E-BP1 (cyan) bound to 4EHP (yellow) to the structure of 4E-BP1 (magenta) bound to eIF4E (grey; PDB: 4UED, Peter et al. 2015a). Selected secondary structural elements in 4EHP are label in black. The structures superpose with an RMSD of 0.41 Å over 194 Cα atoms. (*F*) Schematic representation of eIF4E and 4EHP bound to 4E-BP1. (*G*) Superposition of the structure of 4EHP bound to GYF2 C+L+NC+A (blue) with the structure of 4EHP bound to GYF2 C+L+NC (red) peptides. Selected secondary structural elements in 4EHP are label in black. The structures superpose with an RMSD of 0.38 Å over 207 Cα atoms. (*H*) Overlay of all complex structures of 4EHP bound to GYF1 and GYF2 peptides to illustrate the conformational flexibility of helix α2 (A2), which is circled with a black dashed line. The surface of 4EHP is shown in pale yellow and

the GYF peptides are colored in purple and blue for GYF1 and GYF2, respectively. Helical secondary elements (canonical helix, helices α2 and α3) are represented as cylinders.

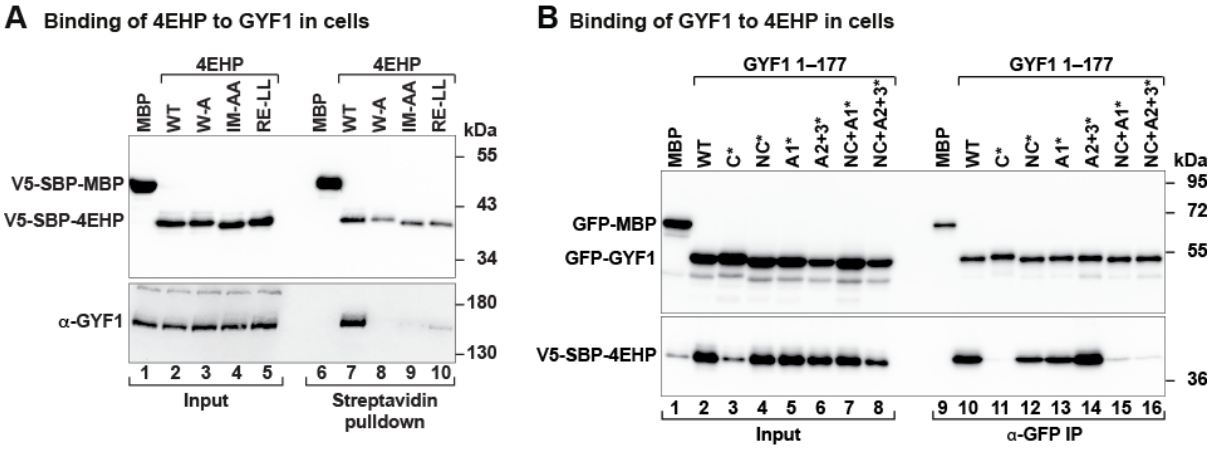
Figure S5



Supplemental Figure S5. Interactions of GYF1, GYF2 and 4E-BP1 with 4EHP and eIF4E. (*A–D*) Close-up views of the interaction between the dorsal surface of 4EHP (A,C,D) or eIF4E (B) and the canonical helices of 4E-BP1 (A, B; Peter et al. 2015a), GYF1 C+L+NC+A (C) and GYF2 C+L+NC (D). Selected residues are shown as sticks. Selected secondary structure elements are labeled in black for 4EHP or eIF4E and in color for the interacting partners. Residue $R63^{4E\text{-}BP1}$ is colored in dark blue following the Cγ atom in A and highlighted by a black dashed box in A and B. (*E–H*) Close-up views of the interaction between 4EHP (E,G,H) or eIF4E (F) and the non-canonical linkers of 4E-BP1 (E, F; Peter et al. 2015a), GYF1 C+L+NC+A (G) and GYF2 C+L+NC (H). Selected residues are shown as

sticks. Residue R63$^{4E\text{-}BP1}$ is colored in dark blue following the Cγ atom in E and highlighted by a black dashed box in E and F. The corresponding residues in GYF1 (Y50) or GYF2 (F52) are also highlighted by a black dashed box. For visual clarity, only backbone atoms are shown for the residues labeled with an asterisk. The residues N77 in eIF4E and S99 in 4EHP are highlighted with orange dashed boxes. (*I–L*) Close-up views of the interaction between the lateral surface of 4EHP (I,K,L) or eIF4E (J) and the non-canonical loops of 4E-BP1 (I, J; Peter et al. 2015a), GYF1 C+L+NC+A (K) and GYF2 C+L+NC (L).

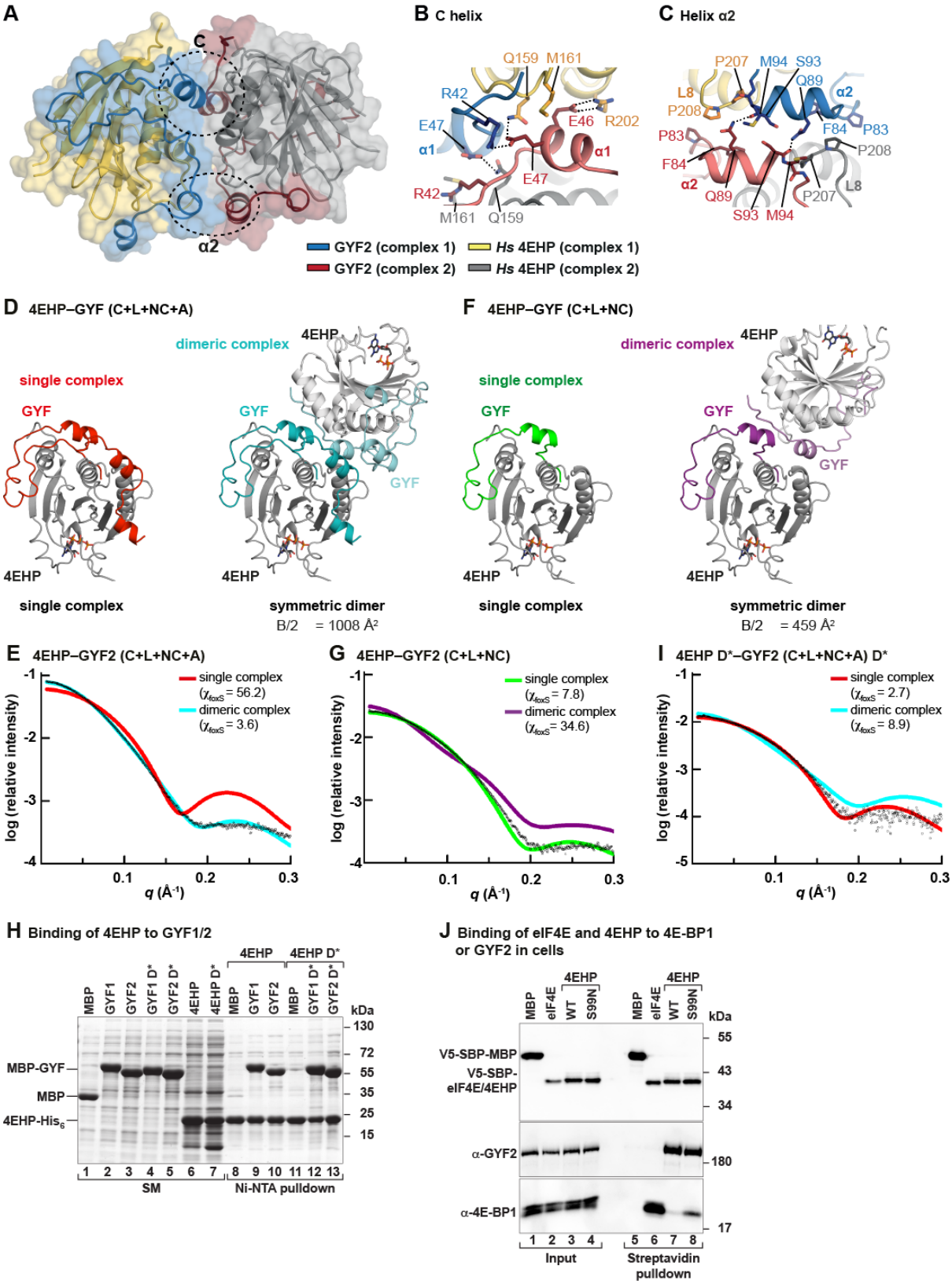
Figure S6



Supplemental Figure S6. Validation of the interfaces observed in the 4EHP–GYF1/2 and 4EHP–4E-BP1 complex structures. (*A*) Western blot analysis showing the interaction of endogenous GYF1 with V5-SBP-tagged 4EHP (WT or the indicated mutants). The proteins were pulled down using streptavidin-coated beads. V5-SBP-MBP served as negative control. The inputs (2.5%) and immunoprecipitates (3% for the V5-tagged proteins and 20% for GYF1) were analyzed by western blotting using anti-V5 and anti-GYF1 antibodies. (*B*) Interaction of GFP-GYF1 N-terminal fragment (residues 1–177; either wild-type or the indicated mutants) with V5-SBP-tagged full-length 4EHP. The proteins were immunoprecipitated from HEK293T cell lysates using anti-GFP antibodies. GFP-MBP served as negative control. The inputs (1.5% for the GFP-tagged proteins and 0.5% for V5-

SBP-4EHP) and immunoprecipitates (7.5% for the GFP-tagged proteins and 30% for V5-SBP-4EHP) were analyzed by western blotting using anti-GFP and anti-V5 antibodies.

Figure S7

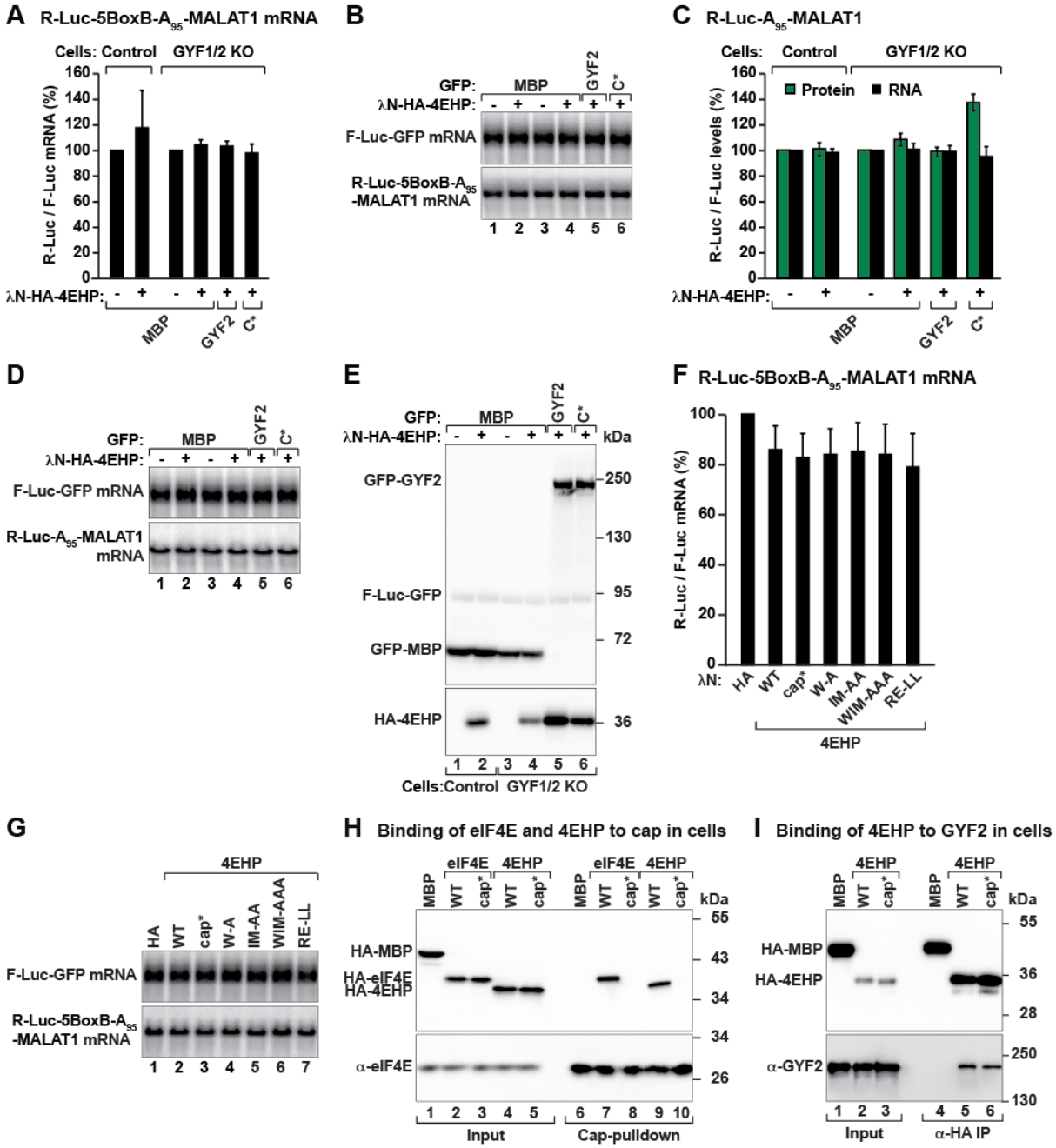


Supplemental Figure S7. The 4EHP–GYF1/2 complexes form dimers in solution. (*A*) Representation of the dimeric arrangement of the 4EHP–GYF2 complexes in the asymmetric

unit of the crystal. In this dimeric arrangement, the two 4EHP–GYF2 complexes adopt a two-fold rotational symmetry with their dorsal surfaces facing each other. The interface of this arrangement involves: I) the canonical helix of GYF2, which is in contact with the neighboring 4EHP molecule and is extended by the canonical helix of the GYF2 molecule of the neighboring complex, and II) the auxiliary region of GYF2, where helix α2 of one complex faces, in an antiparallel fashion, helix α2 of the GYF2 and loop 8 of the 4EHP present in the neighboring complex. The contacts between the two complexes are highlighted with dashed black circles. In complex 1, GYF2 is colored in blue and 4EHP in yellow. In complex 2, GYF2 is colored in red and 4EHP in grey. (*B, C*) Close-up views on the dimeric interface involving the canonical helices (C helix, panel B) and the auxiliary helix 2 (helix α2, panel C). Selected residues are shown as sticks and colored as in panel A. GYF2 residues E46 and E47 within the canonical motifs of interacting GYF2 molecules contact the guanidinium group of $R202^{4EHP}$ and the side chain of $Q159^{4EHP}$ on neighboring 4EHP molecules. $M161^{4EHP}$ contacts the aliphatic portion of $E46^{GYF2}$ in the canonical helix of the neighboring complex. Pro residues in 4EHP loop L8 (P207 and P208) are facing residues proximal to the helix α2 of GYF2 from the neighboring complex (P83, F84, $Q89^{GYF2}$). (*D, F*) Crystallographic models of the 4EHP–GYF2 complexes. The radius of gyration (R_G) and the maximum particle size (D_{max}) were calculated using Scatter and are summarized in Table S3. In the case of the dimeric assemblies, the dimer interface (B/2) was calculated using PISA from the CCP4 package and is indicated below the structures. 4EHP is shown in grey. The GYF2 (C+L+NC+A) peptide is colored in red and cyan in the single and dimeric arrangements, respectively. The GYF2 (C+L+NC) peptide is colored in green and purple in the single and dimeric arrangements, respectively. (*E, G, I*) Small-angle X-ray scattering profiles comparing single and dimeric arrangements of the 4EHP–GYF2 complexes with the experimental data. The data are plotted with the logarithmic scattering intensity on the y-axis

and the scattering angle *q* on the x-axis. Experimental scattering data of the complexes in solution are shown as open black circles and the fits for the single and dimeric arrangements are shown as a line colored as indicated on the right. The goodness-of-fit χ values, calculated using FoXS, are indicated for each fit. (*H*) Ni-NTA pulldown assay showing the interaction of 4EHP-His_6 (M1–F234, wild-type and dimerization mutant) with MBP-tagged GYF1 and GYF2 proteins [(wild-type and dimerization mutant (D*)]. MBP served as a negative control. The input (07% for MBP-tagged proteins and 2% for 4EHP) and bound fractions (9%) samples were analyzed by SDS-PAGE followed by Coomassie blue staining. (*J*) Analysis of the interaction of V5-SBP-tagged eIF4E, 4EHP and the 4EHP S99N mutant with endogenous GYF2 and 4E-BP1 proteins in HEK293T cell lysates. The proteins were pulled down using streptavidin-coated beads. V5-SBP-MBP served as negative control. Input samples (1% for 4E-BP1 and GYF2 and 1.5% for the V5-SBP-tagged proteins) and bound fractions (20% for 4E-BP1 and GYF2 and 5% for V5-SBP-tagged proteins) were analyzed by western blotting using anti-V5, anti-GYF2 and anti-4E-BP1 antibodies.

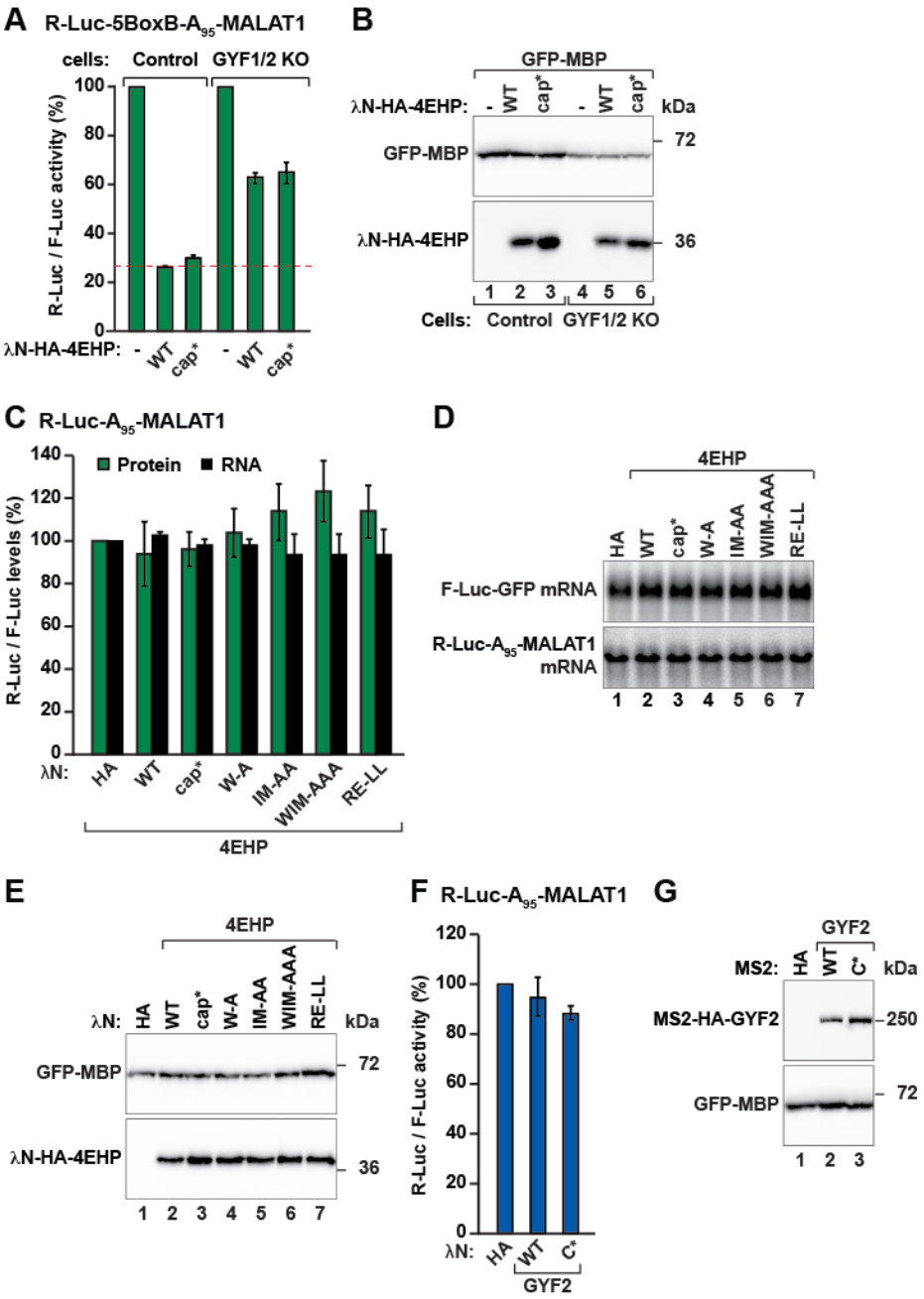
Figure S8



Supplemental Figure S8. 4EHP requires GYF1/2 proteins to repress the expression of bound mRNAs (*A*) A complementation assay using the R-Luc-5BoxB-A$_{95}$-MALAT1 reporter and λN-HA-4EHP (either wild-type or the indicated mutants) was performed in control and GYF1/2-null HEK293T cells expressing GFP-MBP or GFP-GYF2 (wild-type or canonical mutant, C*). A plasmid expressing F-Luc-GFP served as the transfection control. R-Luc activity and mRNA levels were normalized to those of the F-Luc transfection control

and set to 100% in cells expressing the λN-HA peptide. Normalized R-Luc activities are shown in Figure 6A. The panel shows the corresponding normalize R-Luc-5BoxB-A_{95}-MALAT1 mRNA levels. Bars represent the mean values and error bars represent standard deviations from three independent experiments. (*B*) Northern blot of representative RNA samples corresponding to the experiment shown in A and Fig. 6A. (*C*) Complementation assay in WT and GYF1/2-null cells using the R-Luc-A_{95}-MALAT1 reporter lacking the BoxB hairpins. A plasmid expressing F-Luc-GFP was used as a transfection control. R-Luc activity and mRNA levels were normalized to those of the F-Luc transfection control and set to 100% in cells expressing the λN-HA peptide. (*D*) Northern blot of representative RNA samples corresponding to the experiment shown in C. (*E*) Western blot analysis showing the equivalent expression of the λN-HA-4EHP and GYF2 proteins used in the complementation shown in C and D. (*F*) Normalized R-Luc-5BoxB-A_{95}-MALAT1 mRNA levels corresponding to the experiment shown in Fig. 6D,E. (*G*) Northern blot of representative RNA samples corresponding to the experiment shown in F. (*H*) Lysates from HEK293T cells expressing HA-tagged eIF4E or 4EHP (wild-type or cap mutant, cap*) were pulled down with m^7GTP-sepharose beads. Endogenous eIF4E served as positive control. Inputs (0.75% for the HA-tagged proteins and 1% for endogenous eIF4E) and bound fractions (15% for the HA-tagged proteins and 5% for endogenous eIF4E) were analyzed by Western blot using anti-HA and anti-eIF4E antibodies. (*I*) Interaction of HA-tagged 4EHP (wild-type or cap mutant, cap*) with endogenous GYF2 in HEK293T cells. HA-tagged MBP served as a negative control. Inputs (0.37% for the HA-tagged proteins and 0.75% for endogenous GYF2) and immunoprecipitates (15% for the HA-tagged proteins and 20% for endogenous GYF2) were analyzed by Western blot using anti-HA and anti-GYF2 antibodies.

Figure S9



Supplemental Figure S9. 4EHP requires GYF1/2 proteins to repress translation of bound mRNAs (*A*) A complementation assay using the R-Luc-5BoxB-A_{95}-MALAT1 reporter and λN-HA-4EHP [either wild-type or cap mutant (cap*, W124A)] was performed in control and GYF1/2-null HEK293T cells expressing GFP-MBP or GFP-GYF2 (wild-type or canonical mutant). A plasmid expressing F-Luc-GFP served as the transfection control. For each cell type, R-Luc activity was normalized to that of the F-Luc transfection control and set

to 100% in cells expressing the λN-HA peptide. Samples were analyzed as described in Fig. 6A. (*B*) Western blot showing similar expression of the proteins used in A. (*C*) Tethering assay using the R-Luc-A_{95}-MALAT1 reporter and λN-HA-4EHP (wild-type or mutants) in HEK293T cells. Samples were analyzed as described in Fig. 6A. (*D*) Northern blot of representative RNA samples corresponding to the experiment shown in C. (*E*) Western blot showing the equivalent expression of the λN-HA-4EHP proteins used in the tethering assay shown in D.

(*E*) Tethering assay using the R-Luc-A_{95}-MALAT1 reporter lacking MS2 binding sites and MS2-HA-GYF2 (wild-type or canonical mutant). Samples were analyzed as described in Fig. 6F. The corresponding assay with the reporter containing the MS2 binding sites is shown in Fig. 6F. (*G*) Western blot analysis showing the expression of the GFP-GYF2 proteins used in the tethering assays shown in F.

Supplemental references

Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH, Adams PD. 2012. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* **68:** 352–367.

Chen VB, Arendall WB, 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. 2010. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **66:** 12–21.

Cheng Y, Patel DJ. 2004. An efficient system for small protein expression and refolding. *Biochem Biophys Res Commun* **317:** 401–405.

Diebold ML, Fribourg S, Koch M, Metzger T, Romier C. 2011. Deciphering correct strategies for multiprotein complex assembly by co-expression: application to complexes as large as the histone octamer. *J Struct Biol* **175:** 178–188.

Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66:** 486–501.

Fabian MR, Frank F, Rouya C, Siddiqui N, Lai WS, Karetnikov A, Blackshear PJ, Nagar B, Sonenberg N. 2013. Structural basis for the recruitment of the human CCR4-NOT deadenylase complex by tristetraprolin. *Nat Struct Mol Biol* **20**: 735-739.

Kabsch W. 2010. Xds. *Acta Crystallogr D Biol Crystallogr* **66:** 125–132.

Kuzuoglu-Ozturk D, Bhandari D, Huntzinger E, Fauser M, Helms S, Izaurralde E. 2016. miRISC and the CCR4-NOT complex silence mRNA targets independently of 43S ribosomal scanning. *EMBO J* **35:** 1186–1203.

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. 2007. Phaser crystallographic software. *J Appl Crystallogr* **40:** 658–674.

Mizoue LS, Tellinghuisen J. 2004. The role of backlash in the "first injection anomaly" in isothermal titration calorimetry. *Anal Biochem* **326:** 125–127.

Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. 2013. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* **8:** 2281–2308.

Schneidman-Duhovny D, Hammel M, Tainer JA, Sali A. 2013. Accurate SAXS profile computation and its assessment by contrast variation experiments. *Biophys J* **105:** 962–974.

Terwilliger TC, Grosse-Kunstleve RW, Afonine PV, Moriarty NW, Zwart PH, Hung LW, Read RJ, Adams PD. 2008. Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr D Biol Crystallogr* **64:** 61–69.