ABSTRACT

Title of Dissertation: **CHARACTERIZATION OF THE EUKARYOTIC TRANSLATION INITIATION FACTOR 4E (eIF4E) FAMILY MEMBERS IN THE ZEBRAFISH (*Danio rerio*)**

Kathleen M Gillespie, Doctor of Philosophy 2015

Dissertation directed by: Associate Professor Rosemary Jagus, University of Maryland Center Environmental Sciences

The present study examines the six eIF4E cognates in zebrafish. In addition to the prototypical translation initiation factor eIF4E, eukaryotes have evolved eIF4E variants with distinct characteristics, some of which negatively regulate the recruitment of specific mRNAs. Metazoan eIF4E family members fall into three classes, with Class I containing the canonical translation initiation factor eIF4E-1. eIF4E-1 binds eIF4G to initiate translation, a process inhibited by eIF4E binding proteins such as the 4E-BPs and other eIF4E interactive proteins. Analysis of eIF4E sequences from the twenty fish genomes currently available, as well as those of echinoderm, tunicate and cephalocordate, has allowed a glimpse of the origins and evolution of the eIF4E family in vertebrates. All deuterostomes have one representative from each class of eIF4Es. Early deuterostomes such as sea urchins, tunicates, and lancelets

have only one from each class; eIF4E-1, eIF4E-2 and eIF4E-3. The distribution of the subclasses of eIF4E-1 is consistent with the duplication of Class I prior to the teleost specific whole genome duplication, probably at one of the whole genome duplications at ~550 (1R) and 500 (2R) mya. Evidence of the duplication of Class I eIF4Es can be seen in elephant shark (*Callorhinchus milii*), coelacanth (*Latimeria chalumnae)* and basal ray-finned fish (*Lepisosteus oculatus*), which have eIF4E-1A, -1B, and -1C. eIF4E-1B has neofunctionalized to become a tissue specific regulator of mRNA recruitment. It has been retained in tetrapods, but lost in higher teleosts. eIF4E-1C, appears to have retained function as a prototypical initiation factor. A duplication of Class II eIF4Es occurred prior to the emergence of the tetrapod branch, becoming eIF4E-2A and -2B. The genes proximal to the eIF4E-2A locus appear to be conserved across teleosts and tetrapods, the eIF4E-2B genetic loci are more variable, suggesting that eIF4E-2A is the ancestral form. eIF4E-2B is retained by amphibians and teleosts, but has been lost in coelacanth and amniotes. Although 88 % identical, eIF4E-2B can be distinguished from eIF4E-2A by its ability to bind trimethyl GTP (TMG) and to complement a *S. cerevisiae* strain conditionally deficient in eIF4E. This study has shown that duplication within the different classes of eIF4E occurred early in vertebrate evolution with some neofunctionalization, as well as asymmetric losses in different vertebrate classes

CHARACTERIZATION OF THE EUKARYOTIC TRANSLATION INITIATION FACTOR 4E (eIF4E) FAMILY MEMBERS IN THE ZEBRAFISH (*Danio rerio*)

By

Kathleen M Gillespie

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2015

Advisory Committee:

Associate Professor Rosemary Jagus, Chair
Professor Russell T Hill
Associate Professor Shaojun Du
Associate Professor J. Sook Chung
Associate Professor Joseph Pitula

Dedication

This work is dedicated to my mom, my sisters, my good friends, my peers, my ***self,*** and to my dear, sweet, Robert.

Acknowledgements

I would like to warmly thank those who assisted me in this grand and glorious scientific adventure.

My deep regards go out to all my committee members; Dr Russell Hill, Dr Sook Chung, Dr Shaojun “Jim” Du, and Dr Joe Pitula. All have showed me excellent experimental and intellectual support and provided me the skills to move forward.

Thanks to all the people at IMET; the faculty and admins, facilities, ARC, and BAS lab

I am grateful to the members of the Jagus laboratory; both present (“Jerren”, Grant, Erica), and past (Dr Choi and Dr Joshi) whom have helped in this endeavor.

My special gratitude is reserved for Dr Rosemary Jagus, my mentor and advisor, and to her work as the IMET director of the Living Marine Resources Science Cooperative Center (LMRCSC) through NOAA-EPP. Without which, none of this would have been possible.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

CRISPR/Cas	Clustered regularly interspaced short palindromic repeats and CRISPR-associated genes
DNA	deoxyribonucleic acid
DTT	dithiothreitol
eIF4E	eukaryotic translation initiation factor 4E
eIF4G	eukaryotic translation initiation factor 4G
4E-BP	eukaryotic translation initiation factor binding protein
4E-T	eukaryotic translation initiation factor transporter protein
4EHP	translation initiation factor 4E homologous protein, aka eIF4E-2
GST	glutathione S- transferase
GTP	guanosine triphosphate
Hpf	hours post fertilization
IACUC	Institutional Animal Care and Use Committee
IFN	interferon
IMET	Institute for Marine and Environmental Technology
ISG15	IFN stimulated gene 15
LCMS	liquid chromatography mass spectrophotometer
LMRCSC	Living Marine Resources Cooperative Science Center
Mya	million years ago
m^7GTP	7-methylguanosine 5'-triphosphate
NOAA-EPP	National Oceanic and Atmospheric Administration Educational Partnership Program
NES	Nuclear export signal
P body	processing body
pGEX	plasmid gene expression system
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
RNA	ribonucleic acid
RNP	ribonucleoprotein
RT-PCR	reverse transcription – PCR
RT-qPCR	reverse transcription – quantitative PCR
S^{35} Met	sulfur 35 isotope radiolabeled methionine
SC	synthetic medium with carbon source
snRNP	small nuclear ribonucleoprotein
SMN	survival of motor neutrons
SPR	surface plasmon resonance
SUMO	small ubiquitin-like modifier

TILLING	targeting induced local lesions in genomes
TMG	trimethylguanosine
TGD	teleost specific whole genome duplication
Ub	ubiquitin
WGD	whole genome duplication
ZFL cell	zebrafish liver cell

CHAPTER 1: INTRODUCTION

1.1. What is prototypical eIF4E and what role does it play in translation?

In eukaryotes, eIF4E-1 is a central and essential component in the initiation and regulation of translation. Through its interaction with the 5'-cap structure of mRNA and its binding partner, eIF4G, eIF4E-1 functions to recruit mRNAs to the ribosome (1-5) Figure 1.1.

48 S preinitiation complex (prior to scanning)

Figure 1.1: Role of eIF4E in mRNA recruitment to the ribosome

eIF4G is a large factor that plays a scaffolding role, coordinating interactions between translation initiation factors (6). eIF4G provides the crucial link to various translation initiation factors associated with the small ribosomal subunit, such as eIF3 (7) (Figure 1.2).

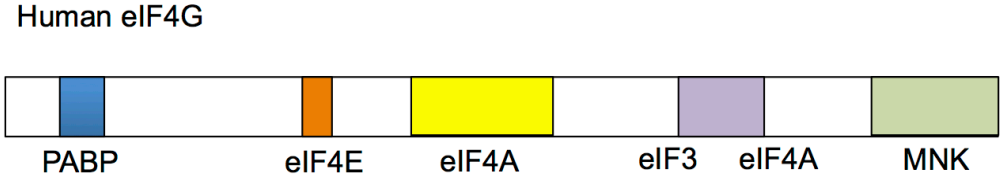


Figure 1.2 Domain structure of eIF4E: middle domain of eukaryotic initiation factor G: PAPB(polyA binding protein) eIF4E, eIF4A (helicase) eIF3, and Mnk domain(MAP kinase interacting kinase 1 domain)

eIF4G also contains an RNA-binding domain, which serves to anchor eIF4E to the mRNA and enhance its interaction with the cap structure (8). Furthermore, eIF4G interacts directly with poly(A) binding protein (PABP) (9). eIF4E interacts with the mRNA cap, PABP interacts with the poly(A) tail and eIF4G bridges the two ends of the mRNA leading to the formation of a closed loop (10). The 'closed-loop' model of translation initiation hypothesizes that mutual interactions of the cap-binding eukaryotic initiation factor eIF4E, eIF4G and PABP hold the 5' and 3' ends of mRNA in close proximity and promote recruitment of the small ribosomal subunit to the mRNA 5' end (4, 11, 12) (Figure 1.3).

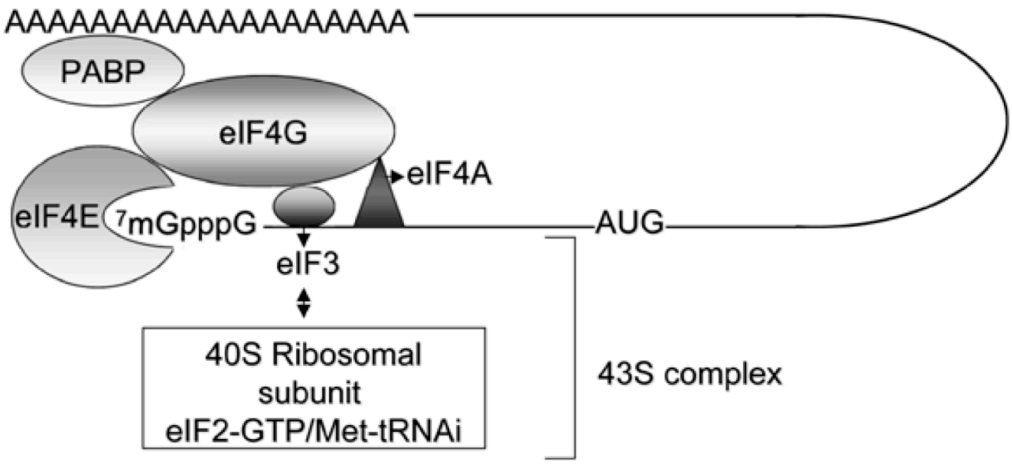


Figure 1.3 Closed loop model of translation Initiation: mechanism of cap-dependent translation initiation: Schematic representation of the closed-loop model of translation initiation. For simplicity, other proteins, as well as a second eIF4A molecule known to interact with eIF4G, have been omitted (from 12)

The anchoring of the eIF4E and eIF4G to the 3'-poly(A) tail ensures that they will remain tethered to the mRNA and gives a competitive advantage in subsequent rounds of initiation.

1.2. Structure of prototypical eIF4E

eIF4E structure and activity is highly conserved across eukaryotic lineages with the core structure representing a novel fold (2, 13). The three-dimensional structure of eIF4E resembles a “cupped-hand” as exemplified by the mouse eIF4E [PDB:1L8B] (14) (Figure 1.4).

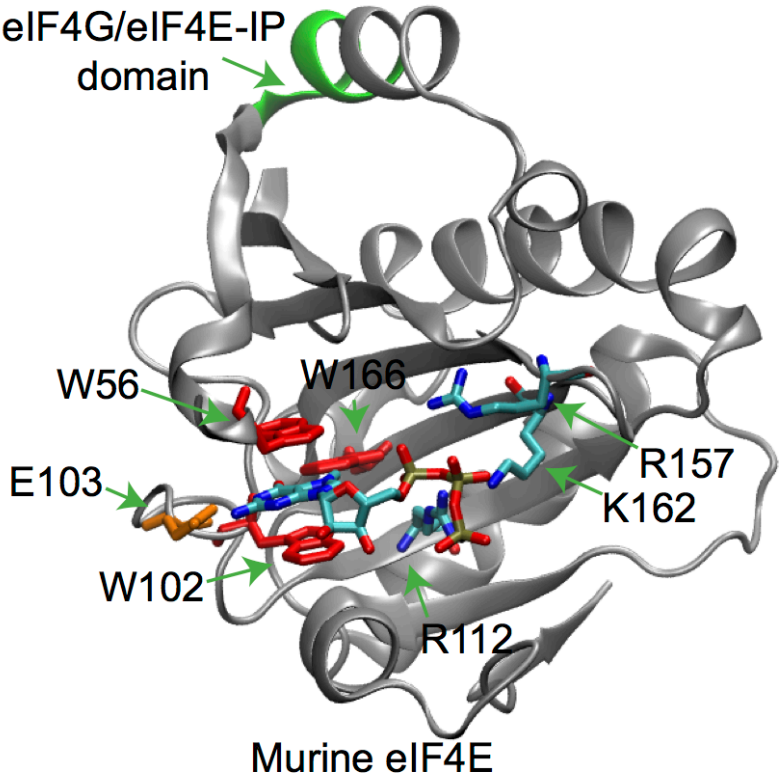


Figure 1.4: Structure of murine translational initiation factor 4E (eIF4E-1) and key binding residues: The crystal structure of Class I murine eIF4E, [PDB:1L8B], was used to show conserved binding domains. Residues W56, W102, and W166 are highlighted in red, as well as E103, which is highlighted in orange, directly interact with the methyl-guanosine moiety. Residues R112, R157 and K162, which are highlighted, contribute charged interactions with the phosphate bridge that links the m^7GTP to the rest of the mRNA chain. A key conserved domain that interacts with eIF4G or eIF4E-interacting proteins is colored green on alpha helix-1 (from 14).

The mRNA cap-binding region is found within a core of 160 to 170 amino acids containing eight aromatic residues with conserved spacing (15). The secondary structure consists of six beta sheets and three major alpha helices (16-18)

(Figure 1.5).

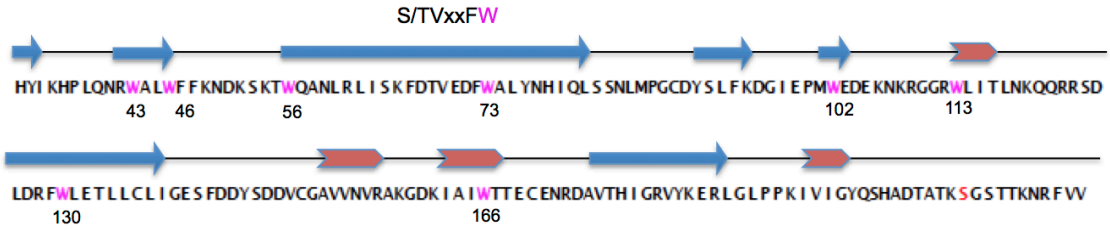


Figure 1.5: Murine eIF4E sequence: mouse *Mus musculus* sequence from Genbank (NP_031943). eight conserved tryptophans (W) are highlighted in fuschia. Alpha helices (blue arrows) and beta sheets (red arrows) are designated. eIF4G and 4E-BP conserved consensus region is indicated above W73.

The beta sheets line the binding pocket, and recognition of the 7-methylguanosine moiety is mediated by cation-π bond stacking between Trp-56 and Trp-102 and H-bonds between Glu-103 and the N-1 and N-2 protons of 7-methylguanine. In addition, W166 interacts with the methyl group on the modified base of the mRNA cap. Furthermore, the triphosphate of the cap forms salt bridges with R112, R157 and K162 (16-19). Aromatic residues Trp, Phe, and Tyr show a distinctive pattern across from N- to C-terminus of the conserved core that contains eight similarly spaced tryptophans (15). The alpha helices form the exterior, solvent accessible side of the protein. Alpha helix one, containing the recognition motif of S/TVEDFW interacts with eukaryotic translation initiation factor 4G (eIF4G) and eIF4E-interacting proteins, the 4E-BPs, and a wealth of other eIF4E-interacting proteins (reviewed, 20, 21).

1.3. Prototypical eIF4E is part of an extended eukaryotic gene family

Multiple eIF4E family members have been identified in a wide range of organisms that include plants, flies, mammals, frogs, birds, nematodes, fish, and

various protists (15). Evolutionarily, it seems that a single early eIF4E gene underwent a series of gene duplications, generating multiple structural classes and in some cases subclasses. eIF4E and its relatives comprise a family of structurally related proteins within a given organism. Through an extensive phylogenetic analysis, it has been shown that eIF4E is part of an extended gene family found exclusively in eukaryotes (15, 22). However, not all eIF4E family members function to promote translation initiation but can be involved in specialized regulatory functions(reviewed, 20,70). Sequence similarity is highest in the core region of 160 to 170 amino acid residues identified by evolutionary conservation and functional analyses (15). Prototypical eIF4E is considered to be eIF4E-1 of mammals, eIF4E and eIF (iso)4E of plants, and eIF4E of *Saccharomyces cerevisiae*.

With the exception of eIF4Es from protists, all eIF4Es can be grouped into one of three classes, Figure 1.6 (15).

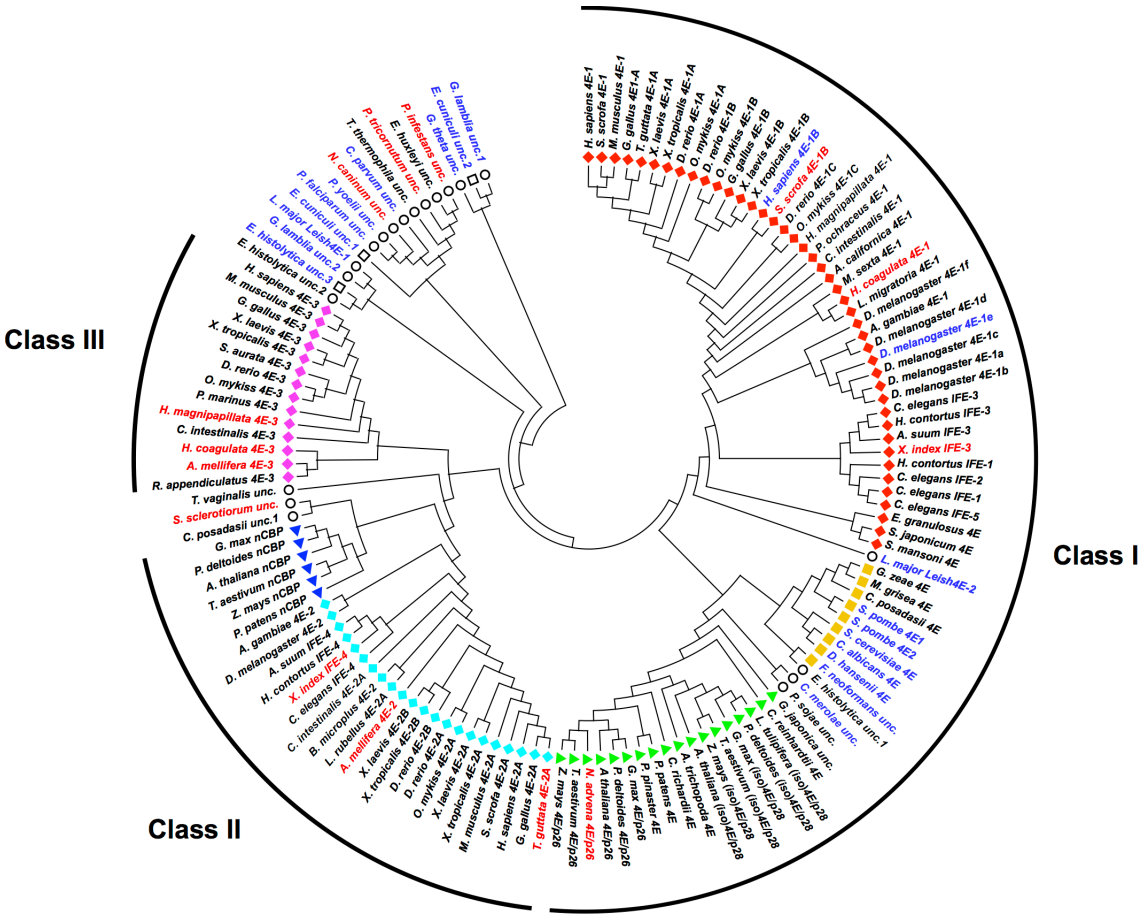


Figure 1.6: Classification of eIF4E family members: eIF4E-family member names in black or red indicate whether or not the complete sequence of the conserved core region of the member could be predicted from consensus cDNA sequence data, respectively. eIF4E-family member names in blue indicate that genomic sequence data was used to either verify or determine the nucleotide sequence representing the core region of the member. The shape of a 'leaf' indicates the taxonomic kingdom from which the species containing the eIF4E-family member derives: Metazoa (diamonds); Fungi (squares); Viridiplantae (triangles); and Protista (circles); respectively. The color of a 'leaf' indicates the sub-group of the eIF4E-family member: metazoan eIF4E-1 and IFE-3-like (red); fungal eIF4E-like (gold); plant eIF4E and eIF(iso)4E-like (green); metazoan eIF4E-2-like (cyan); plant nCBP-like (blue); fungal nCBP/eIF4E-2-like (purple); metazoan eIF4E-3-like (pink); atypical eIF4E-family members from some protists(white). eIF4E-family members within structural classes Class I, Class II, and Class III are indicated (from 15).

Class I members from Viridiplantae, Metazoa, and Fungi carry Trp residues equivalent to W43, W46, W56, W73, W102, W113, W130, and W166 of *Homo sapiens* eIF4E-1**.** Prototypical eIF4Es bind eIF4G through the motif S/TVE/DE/DFW in which the Trp is W73. Substitution of a nonaromatic amino acid for W73 has been shown to disrupt the ability of eIF4E to interact with eIF4G and the regulatory eIF4E binding proteins, the 4E-BPs (23). Substitution of a Gly residue in place of V69 creates an eIF4E variant that still binds mammalian 4E-BP1 but has a reduced capacity to interact with both eIF4G and 4E-BP2 (23). Only Class I eIF4Es are known to function as translation factors. Class I members include the prototypical initiation factor but may also include eIF4Es that recognize alternative cap structures such as IFE-1, -2,and -5 of *Caenorhabditis elegans (24, 25)***,** or eIF4Es that fulfill regulatory functions such as the vertebrate eIF4E-1Bs (26-28) and the Class I deIF4E3 of *Drosophila* (29).

1.4. Roles for eIF4E family members that do not function as translation factors

1.4.1. Vertebrate eIF4E-1B

Unlike the prototypical eIF4E-1A, eIF4E-1B does not bind to eIF4G, but instead binds directly to the so-called eIF4E transporter protein, 4E-T (27). eIF4E-1B can be found complexed with CPEB, RNA helicase Xp54, P100 (Pat1) and the eIF4E transporter (4E-T) and is responsible for suppressing the

translation of CPE-containing mRNAs (27, 28) Figure 1.7.

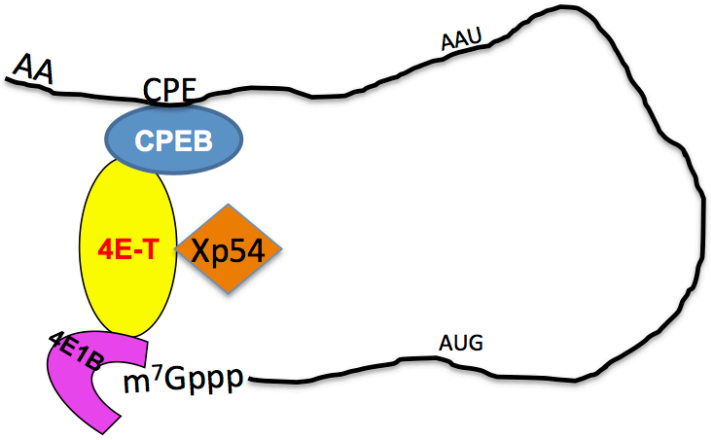


Figure 1.7: Model for translational repression by eIF4E-1B: In *Xenopus* oocytes, CPE (cytoplasmic polyadenylation element) mRNAs will be repressed by the binding of eIF4E-1B. This eIF4E acts as a co-repressor of the CEP binding complex(4E-transporter and Xp54) when tethered to the 3' UTR (from 27).

The target mRNAs are specifically inhibited by a weak binding of eIF4E-1B to the cap structure from its tether at the 3'-UTR. Antisense morpholino nucleotides to eIF4E-1B injected into *Xenopus tropicalis* fully-grown stage VI oocytes down-regulate eIF4E-1B and cause a significant acceleration of oocyte maturation due to increased translation of these mRNA targets (30). eIF4E-1B does not show the characteristics of the prototypical initiation factor. eIF4E-1Bs from zebrafish, *Xenopus laevis* and mouse show weak binding to the 5'-cap structure (26, 31). Although eIF4E-1B does not interact with eIF4G or the 4E-binding proteins (4E-BPs) all the residues critical for 5'-cap mRNA binding and interactions with eIF4Gs or eIF4E-BPs are absolutely conserved among eIF4E-1Bs. eIF4E-1B is a tissue specific translational regulatory factor expressed primarily in ovary and testis (26, 27, 30, 32). In the amphibian *X. laevis,* eIF4E-1B is involved in translational repression in early oogenesis (27, 30, 32). Mouse eIF4E-1B is

distinguishable from eIF4E1A by a set of 15 dispersed amino acid changes (33). The residues are proximal to the region of amino acids that bind the cap. It was determined that Ser105 and Arg106, the amino acids at positions equivalent to Glu103 and Lys104 in *H. sapiens* eIF4E-1A may directly influence the position of Trp102 in the structure and hence modify the stacking interaction with the cap. After mutagenesis of select residues to match those found in eIF4E-1B, the cap-binding ability of eIF4E-1A decreases to a level approximating that of eIF4E-1B (31). The N-terminal region of the eIF4E-1B is enriched in basic residues such as lysine and arginine, suggestive of the ability to induce nuclear import, although the location of eIF4E-1B appears to be entirely in the cytoplasm of the oocyte. Furthermore, domain swap experiments in zebrafish eIF4E-1B have shown that the inability of eIF4E-1B to support protein synthesis is a characteristic of the conserved core (26).

1.4.2. Class II eIF4E family members

Class II eIF4E family members, the eIF4E-2s (also called 4EHP, 34) have been shown to regulate specific mRNA recruitment in *Drosophila* (35), *C. elegans* (36) and mouse embryos (37, 38). Translational inhibition of a specific mRNA by Class II eIF4Es involves tethering of the mRNA 5’ and 3’ ends, giving a configuration much like the classic closed-loop configuration of mRNAs being actively translated, but preventing the interaction of eIF4E and eIF4G. Class II members possess W→Y/F/L and W→Y/F substitutions relative to W43 and W56 of *H. sapiens* eIF4E, respectively (39). Although Class II eIF4Es are found throughout plants, fungi and metazoa, they are absent from the model

ascomycetes, *S. cerevisiae* and *Schizosaccharomyces pombe*. In the *D. melanogaster* embryo the Class II eIF4E, termed d4EHP (deIF4E-8) regulates the synthesis of caudal protein (35, 40). Caudal protein is synthesized asymmetrically in the *D. melanogaster* embryo because translation of its mRNA is inhibited in the anterior region by bicoid (35) Figure 1.8.

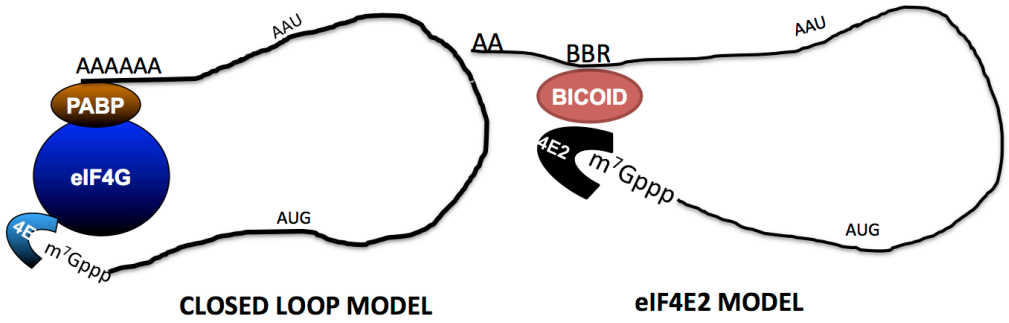


Figure 1.8: Model for translational repression by eIF4E-2: In *Xenopus* eIF4E2 protein will directly interact with bicoid by recognition of the 3' UTR element bicoid-binding region (BBR) (from 27)

The Class II eIF4E family member d4EHP, which binds the cap but not eIF4G, specifically interacts with bicoid to suppress *caudal* mRNA translation. The inhibition is dependent on the bicoid-binding region present in the 3'-UTR of *caudal* mRNA. In another study, it was found that translation of *hunchback* mRNA is regulated by the same Class II eIF4E family member, d4EHP, but in this case, the eIF4E-binding partner is Brat (41).

In mouse oocytes, eIF4E-2 co-localizes with prep1, a homeodomain transcription factor, which contains an eIF4E-binding motif (37). The Prep1/eIF4E-2 interaction seems to bridge the 3'-UTR of *Hoxb4* mRNA to the 5'-cap structure suppressing its translation. This has been the first demonstration that a mammalian homeodomain transcription factor regulates translation, raising

the possibility that this function could be involved in mammalian zygote development. There are over 200 homeodomain proteins that are predicted to contain an eIF4E binding motif, which may all interact with eIF4E-2. In order to determine the role of eIF4E-2 in mouse, an eIF4E-2 knockout was created (38). In the absence of eIF4E-2, translation rates were increased in the eIF4E$^{-/-}$ mouse. However, the deletion was an embryonic lethal, confirming that eIF4E-2 is essential for embryonic development. Conversely, over-expression of Hoxb4 in mouse zygotes *in vitro* resulted in the slowing of development.

1.4. 3. Class III eIF4Es

Class III eIF4Es have been identified primarily in chordates with rare examples in other Coelomata and in Cnidaria (15). Class III members, like eIF4E-3 of mouse possess a Trp residue equivalent to W43 of *H. sapiens* eIF4E but carry a W→C/Y substitution relative to *H. sapiens* W56 (15, 39). Their biological function has not yet been determined, although mouse eIF4E-3 has been shown to bind both cap and eIF4G (39). Mammalian eIF4E-3 binds the m^{7}G cap in the absence of an aromatic sandwich, using instead a cluster of hydrophobic and charged residues in the C-terminus to make extensive contact with the cap to increase affinity (42). Only one form of eIF4E-3 has been found across chordates and it appears to have a limited tissue distribution. Its role in the regulation of gene expression is not well established, with its role as a tumor suppressor appearing at odds with its ability to prevent muscle atrophy (42, 43).

1.5. Modulation of eIF4E-1 activity by covalent modification

1.5.1. Phosphorylation of eIF4E-1

Phosphorylation occurs at Ser209 in the human and mouse proteins eIF4E-1 (44). The primary signal transduction pathway leading to eIF4E phosphorylation is that involving the *ras* gene; RAS activation leads to the phosphorylation and activation of MAP-interacting kinase-1 (Mnk1) that in turn phosphorylates eIF4E. Although it has long been known that eIF4E-1 can be phosphorylated at Ser209, the functional consequences are still unclear. The effect of eIF4E phosphorylation appears to be a reduction of binding affinity to 5' cap structures (45, 46) although increased phosphorylation invariably accompanies increased protein synthetic rates. Mouse mutants that cannot phosphorylate eIF4E have been shown to be less susceptible to viral infection (46). Knock-in mice expressing a nonphosphorylatable form of eIF4E-1 are resistant to tumorigenesis in a prostate cancer model (47). *Drosophila* eIF4E-1 Ser209 mutants show arrested larval development (48).

1.5.2. Sumoylation of eIF4E

Small ubiquitin-like modifier (SUMO) proteins are a family of small proteins that are covalently attached to and detached from other proteins to modify their function. Sumoylation of eIF4E-1A has been shown to activate mRNA translation (49, 50). Sumoylation is involved in various cellular processes, such as nuclear-cytoplasmic transport, transcriptional regulation, apoptosis, protein stability, response to stress, and progression through the cell cycle. Phosphorylation has been shown to enhance SUMO modification for several SUMO substrates and

phosphorylation-dependent SUMO conjugation motifs are a key example of this (51). Sumoylation of eIF4E-1A has been shown to activate mRNA translation by enhancing the interaction between eIF4E and eIF4G SUMO-2. Knockdown of SUMO-2 via shRNA partially impaired cap-dependent translation and cell proliferation (52). It may be that phosphorylation at eIF4E-1 facilitates changes in sumoylation that in turn modulates eIF4E-1 activity/localization (49, 50).

1.5.3. Modification of eIF4E activity by ubiquitin like molecules

Under stress conditions, modifications involving eIF4E proteins may occur by interaction with ubiquitin (Ub). When degradation of mis-folded protein or regulation of a cell signaling pathway (ie PCNA) is necessitated, the ubiquitin system steps in to process the targeted protein via a series of conjugation/de-conjugation steps (53) The presence of a lysine residue is a key point of recognition for attachment of these modification molecules, such as the interferon stimulated gene 15 (ISG15). This small 15 kDa sized protein is induced by type I interferon (54, 55) and consists of two tandem domains, both of which have high identity to an ubiquitin conjugation motif. The functional region of binding for ISG15 was first identified in yeast as having a unique binding motif LRLRGG (56). Induction of ISG15 expression is triggered by cell stressors, such as infections from bacterial or viral agents (57, 58) ,and cellular insults such as radiation and aging (59). ISG15 has been identified in fish kidney and spleen in an antiviral immune response (60, 61) and has the same conserved motif as the mammalian homologue (62). ISG15 may play a role in potential interaction with human eIF4E-2 (4EHP). ISGylated eIF4E-2 displays enhancement of cap-

binding, evident from the m^7GTP pull downs (63). The mechanism is speculated to be a conformational change in eIF4E-2. This reaction does not interfere with general translation.

1.6. Regulation of activity of prototypical eIF4E by 4E-BPs in deuterostomes

A family of eIF4E binding proteins can prevent the interaction between eIF4E-1 and eIF4G. These are known as the 4E-BPs, which are capable of suppressing translation (1)**.** Binding of eIF4E to the 4E-BPs or eIF4G is mutually exclusive. In high affinity binding of eIF4G or 4E-BP, the signature binding motif is YXXXXLϕ (in which X represents any amino acid and ϕ is a hydrophobic residue). In addition, a conserved PGVTS/T motif within the C-terminal region of 4E-BP has been discovered that plays a role in strengthening the binding of eIF4E to the core motif YXXXXLφ and accounts for the higher affinity of 4E-BPs to eIF4E-1 compared to the binding of eIF4G (64, 65). This association is reversible and is regulated by phosphorylation (66, 67). Hypo-phosphorylated 4E-BPs will bind strongly to eIF4E and phosphorylated forms will not. Phosphorylation of 4E-BP is regulated by the mTOR signal transduction pathway (1, 48), which is activated by hormones, growth factors, and amino acids and by cellular energy status (68). 4E-BPs act as global regulators of protein synthesis, with more pronounced effects on mRNAs with high secondary structure content in the 5’ untranslated

regions (69). Figure 1.9

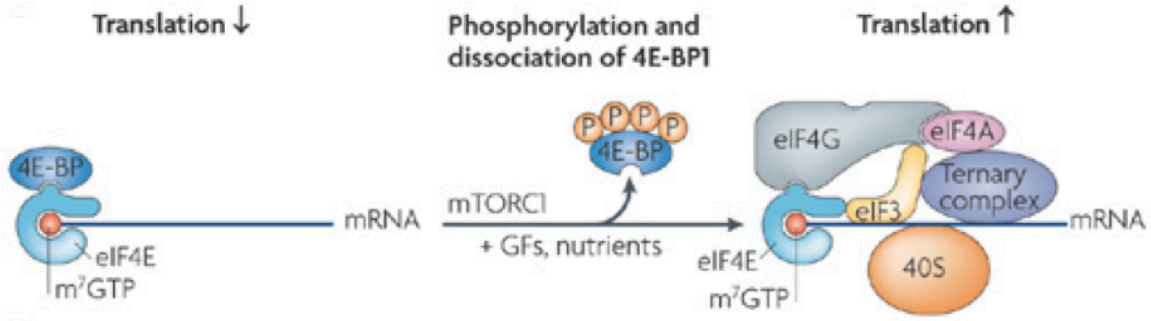


Figure 1.9: Regulation of cap-dependent translation initiation by 4E-BP: Signal transduction-mediated phosphorylation events regulate the function of eIF4E. Hypophosphorylated 4E-binding proteins (4E-BPs) bind tightly to eIF4E, thereby preventing its interaction with eIF4G and thus inhibiting translation. Mammalian target of rapamycin complex 1 (mTORC1)-mediated phosphorylation of 4E-BPs releases the 4E-BP from eIF4E, resulting in the recruitment of eIF4G to the 5' cap, and thereby allowing translation initiation to proceed (from 180)

1.7. Regulation of eIF4E/eIF4G interaction by other eIF4E binding proteins

There is a wide range of eIF4E interactive proteins other than the 4E-BPs. All of these binding partners have been shown to contain the YXXXXLϕ motif like eIF4G and the 4E-BPs. These also compete for eIF4E-1 and modulate its functions, but target specific mRNAs (20, 21, 70).

1.7.1: Maskin

The first of these to be described was an eIF4E-binding protein, maskin, in *X. laevis* which can be tethered to a specific mRNA by a 3'-UTR sequence motif and provides a mechanism for both mRNA-specific translational repression as well as cytoplasmic polyadenylation Figure 1.10 (71).

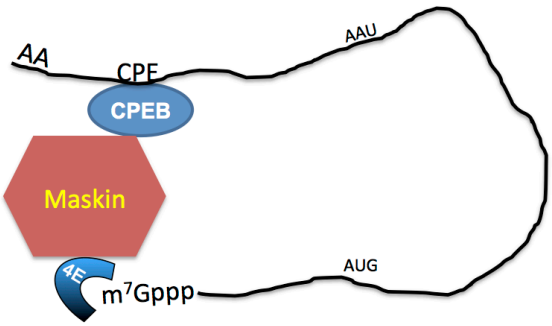


Figure 1.10: Model for translational repression by maskin

X. laevis oocytes arrested in meiotic progression contain silent maternal mRNAs with short poly(A) tails. Upon exposure to progesterone, the poly(A) tail is elongated and translation begins; a requirement for maturation of the oocytes. These "masked mRNAs" contain a cytoplasmic polyadenylation element, CPE, in their 3'-UTRs, which regulates poly(A) length by binding cytoplasmic polyadenylation binding protein, CPEB, the poly(A) polymerase, Gld2, as well as the poly(A)-specific ribonuclease, PARN. Translational suppression is due to both the short poly(A) tail and sequestration of eIF4E-1 by maskin. Progesterone initiates a signaling cascade that results in phosphorylation of CPEB, leading to dissociation of PARN, polyadenylation of mRNA by Gld2, displacement of maskin from eIF4E, and initiation of translation. Table 1.1 lists other known eIF4E interacting proteins.

Table 1.1: eIF4E-binding partners

Protein	Consequences of binding	Residues in the binding partner that interact with eIF4E
eIF4G	Recruits the eIF4A-driven unwinding machinery	KRYDREFLLGF
4E-BP1	Represses highly cap-dependent mRNA translation	IIYDRKFLMEC
p20	Represses cap-dependent translation in S. cerevisiae	IKYTIDELFQL
maskin	Represses translation of CPE-containing mRNAs	EFKLATEADFLLAA
4E-T	Transports eIF4E into the nucleus	PHRYTKEELLDIKELP
lipoxygenase 2	Competes for binding of eIF4E by eIF4G	LKKYRKEELE
vPg	Reduces eIF4E affinity for the cap and inhibits host translation	Mapped to aa 59–93 of TuMV VPg; interaction abolished by mutation of Asp-77
PGL-1	Localizes IFE-1 to P granules	
cup	Represses translation of nanos and oskar mRNAs	YTRSRLM
bicoid	Represses translation of caudal mRNA	NYNYIRPYLPNQ
BTF3	Competes for binding of eIF4(iso)4E by eIF(iso)4G	RLQSTLKRIG
brat	Represses translation of hunchback mRNA	NHL domain
gemin5	Inhibition of both cap-dependent and IRES-driven translation	LKLPFLK and YEAVELL
neuroguidin	Represses translation of CPE-containing mRNAs	YPTEKGL, YQIDKLVKT, and YVPPRLV
CYFIP1	Represses translation of mRNAs that bind FMRP	LLLDKRKRSEC
angel1	Interacts with eIF4E1 in ER and golgi	RRKYGRDFLL(Hs),KIYTRQQLL(Xe)
GIGYF2/F1	Interacts with eIF4E2 in mouse	DYRYGREEMLAL/DYRYGREEMLAL

From (20, 38, 74)

1.7.2. eIF4E transporter (4E-T) and related proteins

4E-T is a large and highly conserved protein in vertebrates (985 amino acids in humans) that harbors a canonical eIF4E-binding site at its N-terminus. It was initially characterized as a nucleocytoplasmic shuttling protein, with defined NLS (nuclear localization sequence) and NES (nuclear export sequence) that mediates eIF4E nuclear import by a piggy-back mechanism (72). 4E-T is an abundant component of the large CPEB/mRNP (mribonucleoprotein) translational repression complex in *Xenopus* oocytes, which resemble processing bodies (P-bodies), and also includes the RNA helicase Xp54/p54/DDX6, the RNA-binding proteins Pat1a and RAP55 (Lsm14) and eIF4E1B (21, 27). 4E-T proteins are notably highly expressed in oocytes and ovaries in *Xenopus*, fruit flies and nematodes, in which they repress the translation of specific mRNAs in conjunction with 3'-UTR RNA-binding proteins and are typically found in large RNP aggregates.

A range of other eIF4E binding proteins have been described that target specific mRNA translation pathways through disruption of the interaction of eIF4E/eIF4G and which involve interaction directly or indirectly with motifs in the 3’-UTR (20, 21, 38, 40, 70, 73, 74).

1.8. Why study eIF4E function in zebrafish?

1.8.1. Zebrafish as a model system

Over the past decades, zebrafish has become a preeminent vertebrate model system for clarification of the roles of specific genes, signaling pathways in development and especially the identification of new drug targets for human disease. There is a substantial historical database regarding basic developmental biology, toxicology, and gene transfer. Zebrafish can be used in forward genetic screens and reverse genetic techniques; genes can be knocked down with morpholinos (75), or knocked out with high efficiency using CRISPR/Cas technology (76, 77). CRISPR (clustered regularly interspaced short palindromic repeats) coupled with guide RNAs and the Cas9 protein can cut at any desired location in the genome. CRISPR/Cas technology has also made development of transgenic zebrafish much easier (78). The benefits of the use of zebrafish as a model system include its transparent and accessible embryos, cost-effective mutagenesis screening, and ease of maintenance and breeding, but also the availability of genomic data, fish mutant strains, and other on-line resources such as ZFIN (79). With the completed genome available, zebrafish has become a powerful model system for clarifying mechanisms in development, differentiation, toxicity, disease, and resistance to infection (79).

Transgenic zebrafish are being used to develop models of human disease (80-82). The other benefits for the use of zebrafish as a model system are the availability of genomic data, extensive resources, the ease of maintenance and breeding (http://zfin.org/). In addition, zebrafish knockouts for some of the eIF4Es

are now available, including eIF4E-1B and eIF4E-1A. There is also a knockout for 4E-BP3l (https://www.sanger.ac.uk/sanger/Zebrafish). Zebrafish females are capable of producing large quantities of high quality eggs daily, which are transparent, fast developing, and are easily manipulated (83).

1.8.2. The zebrafish genome encodes six eIF4E family members

There is now conclusive evidence that whole genome duplication occurred in ray-finned fish coincident with radiation of teleost species, followed by reciprocal gene loss (reviewed (84)). While most gene pairs formed by WGD are deleted, rapid functional divergence provides an explanation for duplicate gene retention (85). Divergence of gene function between duplicates has been reported in many studies (86, 87). Such neofunctionalization may account for the fact that the zebrafish genome has six eIF4E genes; three that express Class I eIF4Es, termed eIF4E-1A, -1B and -1C, two that express Class II eIF4E, eIF4E-2A and -2B, and one Class III, eIF4E-3. Figure 1.11 shows the multiple alignments of the zebrafish eIF4Es.

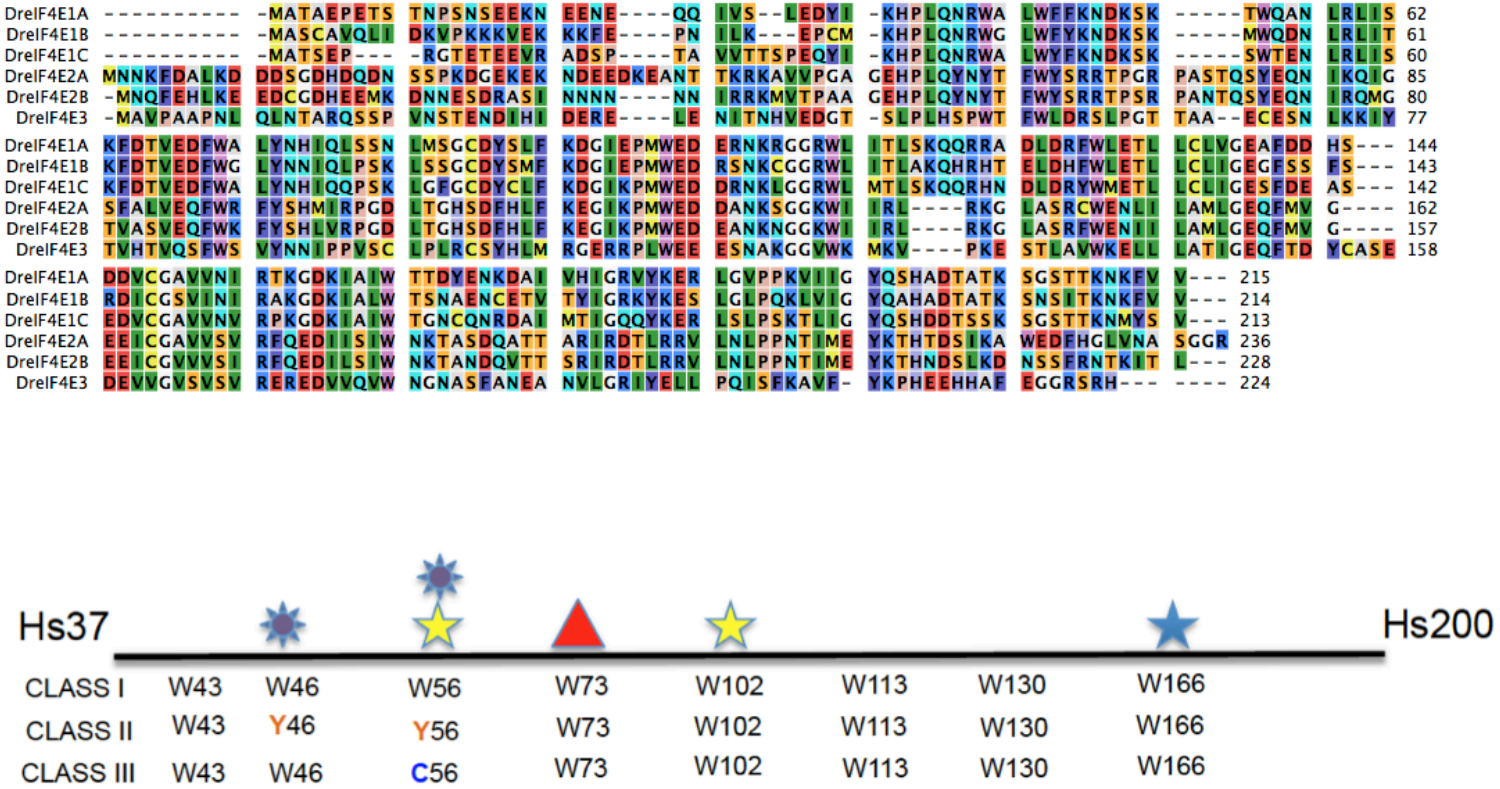


Figure 1.11: Multiple alignment of zebrafish eIF4E family members: a) Multiple alignments of the full sequences of deuterostome eIF4E family members. b) Schematic representation of core eIF4E region between H37 and H200. The conserved Trp(W) residues and indicated by location at W43,W46,W56,W73,W102,W113,W130,W166 . Yellow stars are above Trps involved in cap-binding. The blue star at W166 represents the Trp that binds the m^7-methyl moiety on the cap. Purple asterisks indicate the sites of Trp substitutions in Class II and III (Class II members have Trp→Tyr/Phe/Leu and Trp→Tyr/Phe substitutions relative to Trp-43 and Trp-56 respectively of *H. sapiens* eIF4E; Class III have Trp→Cys relative to Trp-56. The red triangle indicates the TrpW73 in the eIF4G/4E-BP binding region

Table 1.2 provides identity and similarity comparisons of zebrafish family members.

Table 1.2: Identity and Similarity comparisons of eIF4 family members in human(Hs) and zebrafish(Dr)

Similarity(%)

	DreIF4E 1A	DreIF4E 1B	DreIF4E 1C	DreIF4E 2A	DreIF4E 2B	DreIF4 E3	HseIF4E 1A	HseIF4E 1B	HseIF4 E2	HseIF4 E3
DreIF4E1 A		85.1	84.5	50.8	49.7	41.7	97.1	89.7	50.8	41.7
DreIF4E1 B	73.1		78.8	49.1	48	40	85.1	81.7	49.1	40
DreIF4E1 C	77.1	68.5		50.2	50.2	42.8	84.5	80.5	50.2	42.8
DreIF4E2 A	34.8	33.1	38.2		93.7	49.1	51.4	52	99.4	49.1
DreIF4E2 B	34.8	33.1	35.4	88.5		48.5	50.2	50.8	94.2	48.5
DreIF4E3	29.2	29.1	29.7	30.8	29.7		42.2	42.2	49.1	90.8
HseIF4EE1 A	89.7	76	78.2	35.4	34.2	31.4		87.4	51.4	42.2
HseIF4E 1B	75.4	72	69.7	35.4	34.8	29.7	74.8		52	4.2
HseIF4E 2	35.4	33.7	38.8	97.1	89.1	32	36	36		49.1
HseIF4E 3	27.2	28	28	31.4	30.2	84.5	29.7	27.4	31.4	

Identity(%)

The Jagus laboratory has confirmed the deduced sequences of the zebrafish eIF4E family members and has characterized two, eIF4E-1A- and -1B. eIF4E-1A has been concluded to represent a prototypical translation factor on the basis of its identity/similarity to human/mouse eIF4E-1, its ubiquitous expression, its ability to bind ^{7}mGTP-Sepharose, interact with eIF4G and 4E-BP, as well as complement a yeast eIF4E knock out system (26, 88). Zebrafish eIF4E-1B does not interact with eIF4G or the 4E-BPs and has been reported to be expressed only in ovary, testis and at low levels in muscle (26). Furthermore, domain swap experiments in zebrafish eIF4E-1B have shown that the inability of eIF4E-1B to support protein synthesis is a characteristic of the conserved core (26). The characteristics of the remaining four eIF4E family members have not previously

been investigated. The confirmed sequences for eIF4E-1A, and eIF4E-1B are deposited in GenBank as accession numbers numbers: NM_131733.1 (eIF4E-1A), NM_131454.1 (eIF4E-1B) (26). Other Jagus laboratory members have isolated the cDNAs and the sequences of eIF4E-1C and eIF4E-3, GenBank accession numbers, NM_001017851.2 (eIF4E-1C) and NM_001004589.1 (eIF4E-3). I have isolated the cDNA and confirmed the sequences of eIF4E-2A and eIF4E-2B, and placed in GenBank accession numbers, AGW99949.1 and AGW99950.1, respectively.

1.9. Focus and objectives

My aim in this dissertation is to determine the phylogenetic origin of the zebrafish eIF4E family members and their relationship to the eIF4Es of lower deuterostomes and the tetrapods. In addition, I aim to characterize zebrafish eIF4E-1C, eIF4E-2A, eIF4E-2B and eIF4E-3 by comparing their activities in a variety of *in vitro* assays, as well as their ability to complement an *S. cerevisiae strain* conditionally depleted of eIF4E. I will also undertake expression analysis at the transcript and protein levels.

The hypotheses to be examined are: 1) Because it has high sequence identity to the previously described eIF4E-1A, eIF4E-1C will function as a translational initiation factor; 2) Because of the evolutionary persistence of the two cognate proteins, eIF4E-2A and -2B, this suggests neofunctionalization had occurred in one of these to give an eIF4E family member with distinct characteristics; and 3) Because of apparently conflicting reports, the function and characteristics of eIF4E-3 remain unclear.

The results presented here will provide the basis for future studies that dissect the role of the zebrafish eIF4E family members in the regulation of protein expression.

Chapter 2: Molecular and phylogenetic insights of eukaryotic translational initiation factor 4E family members in teleosts

2.1. Abstract

In addition to the prototypical translation initiation factor eIF4E, eukaryotes have evolved sequence–related variants with distinct features, some of which have been shown to negatively regulate translation of particular mRNAs. I present here my perspective on the evolution of the eIF4E family in deuterostomes. Metazoan eIF4E family members have been divided into three classes, with Class I containing the canonical cap-binding protein eIF4E1. eIF4E-1 binds eIF4G to initiate translation, a process inhibited by eIF4E binding proteins such as the 4E-BPs and 4E-T that prevent the interaction between eIF4E and eIF4G by competing for the same binding site, YXXXXLΦ. All deuterostomes have at least one representative of Class I, Class II and Class III eIF4E family members. Early deuterostomes such as sea urchins, tunicates, and lancelets have only one eIF4E family member in each of the three classes; eIF4E-1, eIF4E-2 and eIF4E-3. A member of the Elasmobranchii, the elephant shark (*Callorhinchus milii*) and the Sarcopterygii, coelacanth (*Latimeria chalumnae)* have duplicated Class I eIF4Es to give eIF4E-1A, eIF4E-1B, and eIF4E-1C. Prior to the emergence of the tetrapod branch, a duplication of Class II eIF4Es occurred, becoming eIF4E-2A and -2B. eIF4E-2B was retained by amphibians (*Xenopus* spp) and teleosts, but was lost in coelacanths and amniotes. After the teleost-specific whole genome duplication event, 320-350 mya, eIF4E-1A, -1C, -2A, -2B and -3 were consistently maintained by the ray-finned fish, the salmonids, and gadiformes.

Percomorphs acquired a new cognate of Class I family member, designated eIF4E-1A-like. However, eIF4E-1B appears to disappear after the salmoniformes split. This study has shown that duplication within the different classes of eIF4E family members occurred early in vertebrate evolution with subsequent asymmetric losses in different vertebrate classes.

2.1. Introduction

2.1.2. Deuterostome phylogeny

The Cambrian explosion, also called the Cambrian radiation, was the relatively short evolutionary event, beginning around 540 mya in the Cambrian Period, during which most major animal phyla appeared as indicated by the fossil record (89). Lasting for about the next 20-25 million years, this explosion resulted in the divergence of most modern metazoan phyla (90, 91). The earliest generally accepted deuterostome fossils, those of echinoderms, appeared in the Late Atdabanian (Cambrian, 3rd Stage) (92). The deuterostome superphylum consists of three phyla: echinoderms, hemichordates and chordates. Three subphyla are recognized within the chordates themselves; the urochordates (including the ascidians and larvaceans) the cephalochordates (lancelets) and the vertebrates, including fish and tetrapods. Figure 2.1 provides an illustration of deuterostome phylogeny (93).

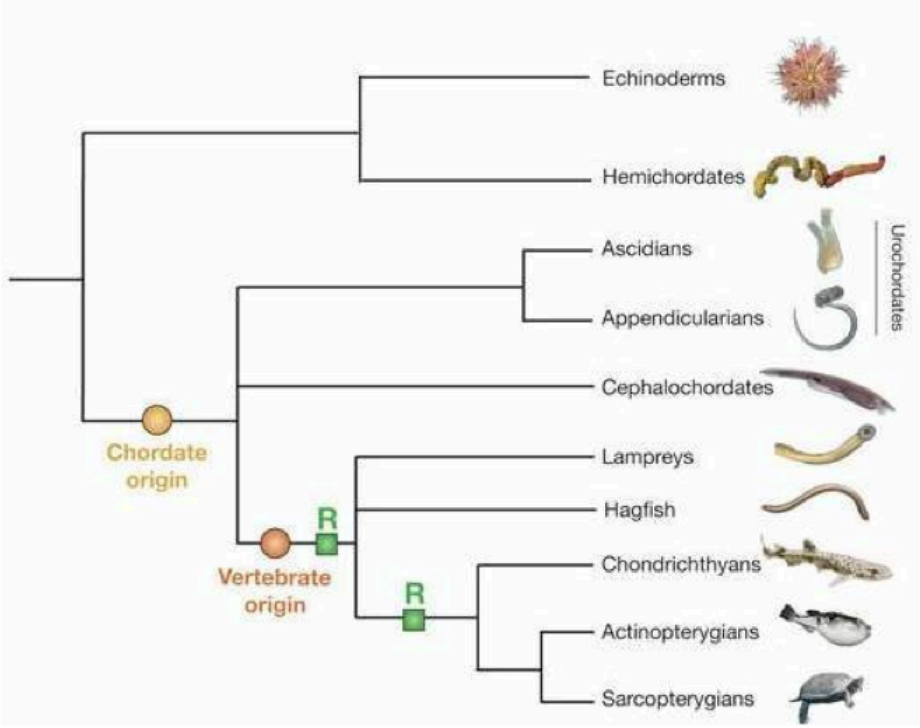


Figure 2.1: Deuterostome phylogenetic tree: Generalized phylogenetic tree of deuterostome evolutionary progression. The circles indicate where the chordate (yellow) and vertebrate (orange) origin points occurred respectively. The R box (green) is the representation of two rounds of whole genome duplication events early in vertebrate evolution; R1 at ~550 mya, R2 at 500 mya (from 93).

Phylogenetic analysis based on assembled sequences of more than 200 nuclear-encoded proteins support the pairing of echinoderms with hemichordates corroborating morphological interpretations of larval similarities between these two groups (94).

Gene duplication is considered to be a major force of evolution (95) because new copies may acquire new functions by mutation (known as neofunctionalization) (96). It is generally accepted that two rounds of whole-genome duplication occurred during the evolution of vertebrates from their deuterostome ancestors before the divergence of gnathostomes between 500 and 550 mya (97). In comparison with tetrapods, the ray-finned fishes underwent an extra round of

whole-genome duplication, which caused the teleost radiation (98). The last common ancestor of all vertebrates was the common ancestor of the Actinopterygii (ray-finned fishes) and Sarcopterygii (lobe-finned fishes and tetrapods), which probably lived during the Silurian period, approximately 420 mya (99-101). Within vertebrates, lampreys and hagfish are closely related (Cyclostomata) and are the closest relatives of jawed vertebrates, the gnathostomes. Molecular analyses suggest that the coelacanth and lungfish form a group that is the closest living relative of tetrapods and that cartilaginous fish are the most basal gnathostomes. Within the ray-finned fish, the dominant extant group is the Teleotstii comprised of over 20,000 species. The teleost lineage splits from basal ray-finned fishes and started to diverge after a whole genome duplication event that took place 320–350 mya and is referred to as the teleost specific whole genome duplication, TGD (102, 103). Molecular phylogeny suggests the initial divergence of the Teleostei from basal Actinopterygii occurred about 280 mya (101). The teleosts began a major evolutionary radiation in the Triassic, about 200 mya, and have since undergone massive diversification in morphology, physiology, and habitat. Their genomes did not remain static and they are still evolving. The evolutionary divergence and extreme diversity teleosts provide are now represented by over ten genomes that reflect all the structure–function combinations that have survived during the last 400 million years (104). The teleosts are characterized by many derived characteristics that are absent in primitive ray-fins such as gar, sturgeon and paddlefish. Teleosts are thus remote from the common actinopterygian/sarcopterygian ancestor. Within the teleosts,

the Ostariophysi (such as zebrafish) retain many primitive characteristics and occupy a relatively basal position (106). Thus the zebrafish is a rather generalized teleost and can, in most cases, be used to represent the ‘‘primitive’’ or ‘‘ancestral’’ condition in comparison with more recently evolving teleosts such as the percomorphs medaka, stickleback, tilapia and fugu (105-107). However, with an evolutionary separation of less than 150 million years, the zebrafish is still closer to the more recently evolved fish species than any mammalian model organism such as the mouse, whose common ancestor with the teleosts lived around 400 mya (107).

2.2. The radiation of the vertebrates is reflected in their eIF4E family members

All deuterostomes have at least one representative of Class I, Class II and Class III eIF4Es, eIF4E-1, eIF4E-2 and eIF4E-3. In contrast mammals have an additional eIF4E-1 cognate, eIF4E-1B, that functions to down-regulate translation of mRNAs with cytoplasmic polyadenylation elements (CPEs) in the 3'-UTR (15, 26, 27, 30). Zebrafish have two eIF4E-1 cognates, eIF4E-1B and -1C, as well as a Class II cognate, eIF4E-2B (15, 26). Figure 2.2 shows a simplified phylogeny of deuterostomes, emphasizing teleost fish adapted from (105), and indicates the occurrence of different eIF4E family members. The availability of fully sequenced genomes from many deuterostome species provides an unprecedented opportunity to systematically evaluate the origins and evolution of protein families such as the eIF4E family, shedding new light on the old question of how organismal complexity arose.

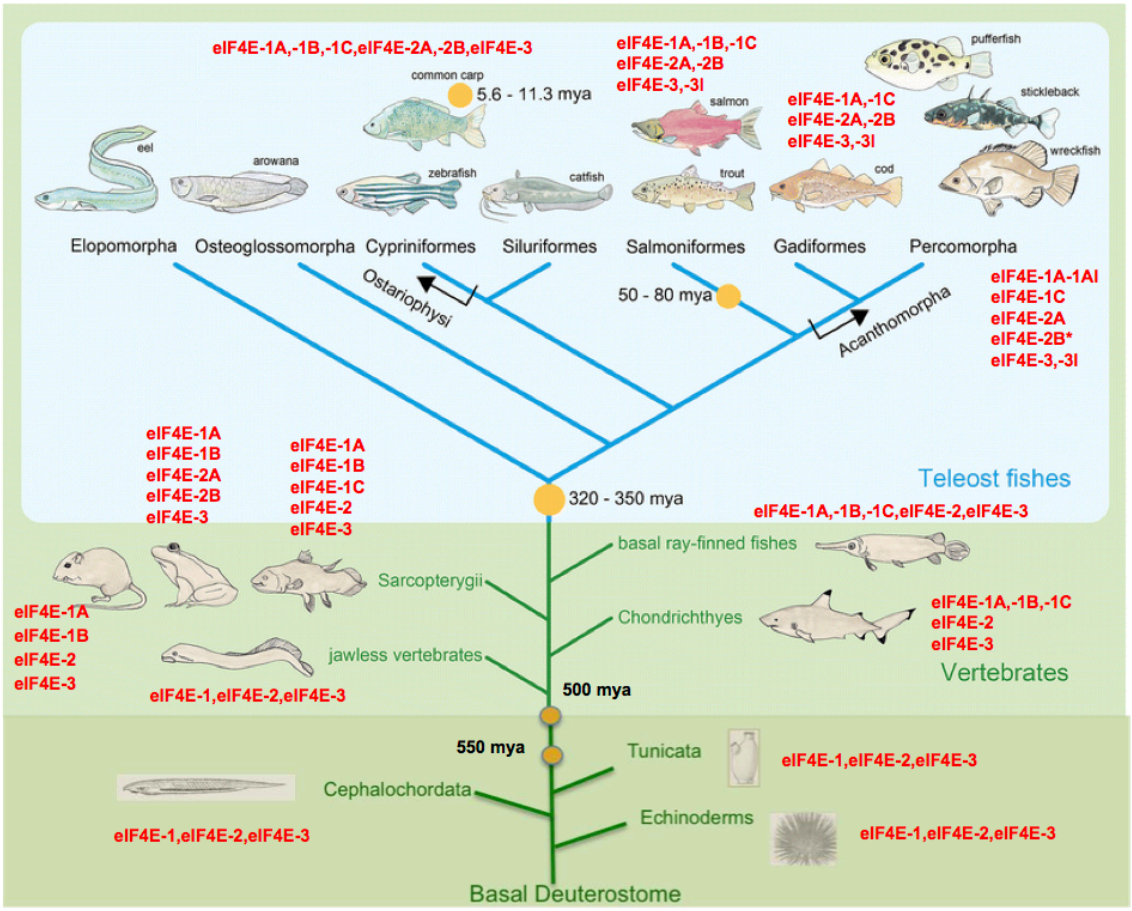


Figure 2.2: Phylogenetic tree of deuterostome: Generalized phylogenetic tree from basal deuterostomes to higher teleosts (from 105). Multiple whole genome duplication events (WGD) are indicated in yellow circles at time points of ~550, 500, 320-350, 50-80 and 5.6-11.3 (mya). The addition to figure (bottom); Basal deuterostome panel with Echinoderms, Tunicate, and Cephalochordata and time points ~550 and 500 mya. Translation initiation factor eIF4E cognate forms were overlaid onto original schematic and placed near the representative branch and/or subfamily member (highlighted in red). * Denotes teleosts family that may not have eIF4E-2B

2.3. Materials and Methods

2.3.1. Database Searches and Phylogenetic Analysis

In order to investigate eIF4E family members in deuterostomes, zebrafish eIF4E-1A (NM_131733.1), eIF4E-1B (NM_131454.1), eIF4E-1C (NM_001017851.2), eIF4E-2A (AGW99949.1), eIF4E-2B (AGW99950.1), and eIF4E-3 (NM_001004589.1) were used as templates for BlastP queries at the National Center for Biotechnology Information (NCBI) website (http://www.ncbi.nlm.nih.gov/mapview/) for sea urchin (*Strongylocentrotus purpuratus*), sea squirt (*Ciona intestinalis*), zebrafish (*Danio rerio*), and human (*Homo sapiens*) genomes. Cross references were achieved by use of alternative databases which included; Ensembl (http://www.ensembl.org/index.html) for spotted gar (*Lepisosteus oculatus)* medaka (*Oryzias latipes*), stickleback (*Gasterosteus aculeatus*), Takifugu (*Takifugu rubripes*), and Tetraodon (*Tetraodon nigroviridis*) genomes: the Institute of Molecular and Cell Biology(IMCB) elephant shark genome http://esharkgenome.imcb.a-star.edu.sg; the coelacanth genome project site (http://coelacanth.nig.ac.jp/index.php), the salmonDB genome database(http://salmondb.cmm.uchile.cl) and the Joint Genome Institute (JGI) for the *Branchiostoma floridea* genome (http://genome.jgi-psf.org); the HMMR database(http://hmmer.janelia.org). Each eIF4E sequence was verified using the Genbank BLAST tool and aligned by the MUSCLE algorithm included in the CLC workbench (CLCBio CLC Genomics Workbench 7.0.3 (http://www.clcbio.com). The phylogenetic analysis used Le Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier

(LIRMM) tool Phylogeny.fr suite of programs including Gblocks, PhyML, and TreeDyn with 100 iterations (108). The reliability of the tree was measured by bootstrap analysis. Gene loci designations and orientations were determined utilizing a combination of both the NCBI gene database and the Ensembl gene region of interest function. All accession numbers, additional database designations, and details on sequences are provided in Appendix Table A2.1.

2.4 Deuterostome Class I eIF4E family members

2.4.1. Phylogenetic analysis of Class I eIF4E family members

The evidence of gene duplication is apparent from the number of orthologues of each eIF4E class across the deuterostomes. Phylogenetic analysis divided these proteins into three clusters corresponding to Class I, Class II and Class III eIF4E family members (Figure 2.3).

Ancestral members of the vertebrates; the protochordate tunicate *Ciona intestinalis*, the cephalochordate lancelet, *Branchiostoma floridae*, and the echinoderm sea urchin *Stongylocentrotus purpuratus*, have only one Class I eIF4E cognate. These eIF4Es form a distinct clade outside of the eIF4E-1A,-1B, and -1C designations. The eIF4E of lamprey, *Petromyzon marinus*, appears to reside on a separate branch, closest to the eIF4E-1A clade. Overall in each cluster, the tetrapod and teleost eIF4Es tend to group together. Further examination shows there is a clear separation of nodes between the early teleosts such as zebrafish (Dr), carp and cavefish (Am) when compared to the more recently evolved teleosts such as cod (Gm), tilapia (On), medaka (Ol) and

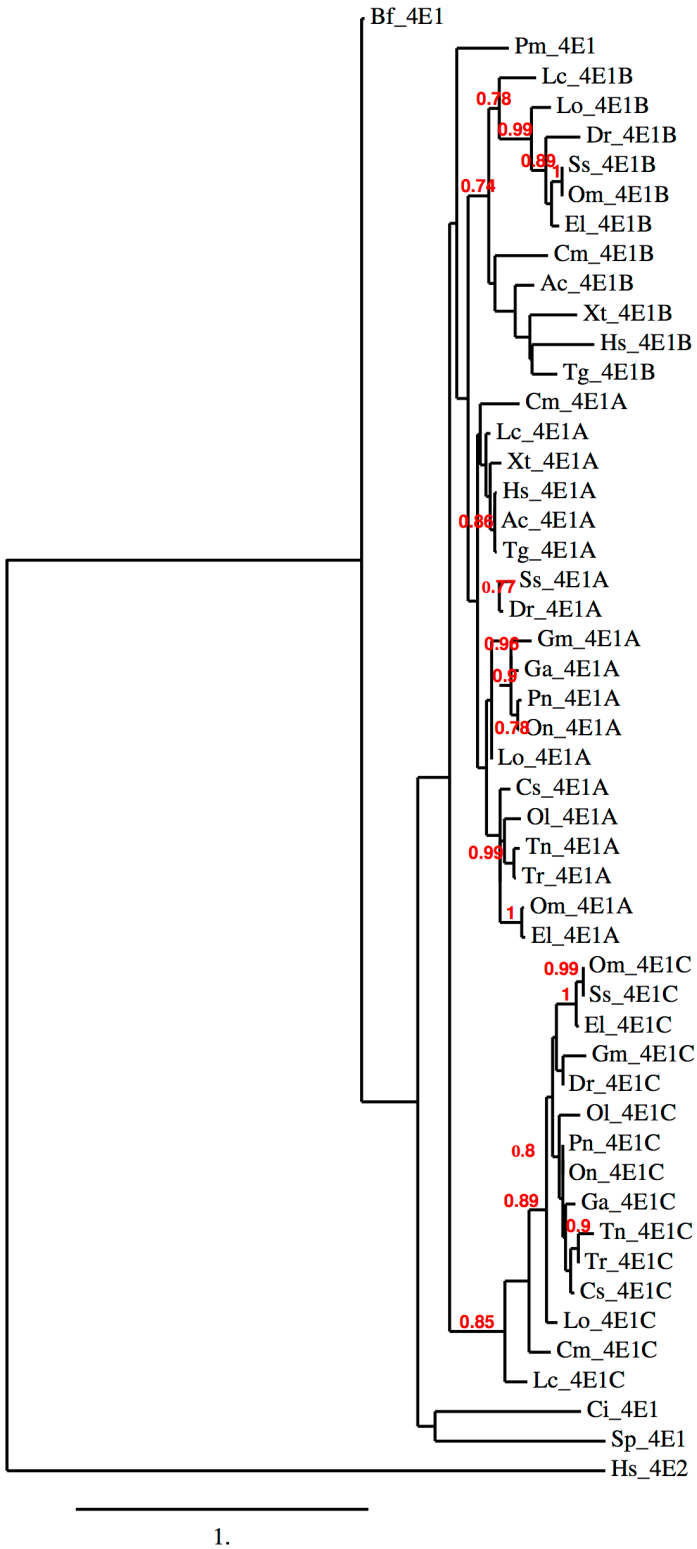


Figure 2.3: Phylogenetic analysis of deuterostome Class I eIF4Es: Phylogenetic analysis based on the conserved core and C-terminal regions from muscle alignments. The numbers on the branches are confidence limits (expressed as percentages) estimated from a bootstrap analysis with 100 replicates (above 60 % are indicated). Bar 1.0 indicates 1.0 substitutions per nucleotide position. Human eIF4E2 is used as out group.

puffer fish (Tn). Only eIF4Es from the speckled gar (Lo) and coelacanth (Lc) stand outside each cluster. Within the eIF4E-1C teleost sub-cluster, the zebrafish (Dr) is closely related to the cod (Gm). However, zebrafish eIF4E-1A and -1B align with the salmon (Ss) and trout (Om). The Class I eIF4E cognates, eIF4E-1B, and eIF4E-1C are thought to have arisen from one or more whole genome duplications. Because eIF4E-1C is found in teleosts and not in tetrapods, our original supposition was that eIF4E-1C arose as the result of the TWGD. However, both the elephant shark, *Callorhinchus milii* (Cm) and the coelacanth, *Latimeria chalumnae* (Lc) have all three sub-classes. This implies that the duplications must have occurred prior to the branching of the chondrichthyes. However, although all teleosts have retained eIF4E-1C, it has been lost in tetrapods.

The evolution of eIF4E-1B seems more complicated. The presence of eIF4E-1B in the elephant shark suggests an early origin in gnathostomes. Inspection of the teleost genomes available, has uncovered eIF4E-1B in basal ray-finned fish, such as spotted gar, *L. oculatus (Lo),* as well as in zebrafish, and rainbow trout, (Om*)*. However, it has not been found in the genomes of more recently evolved fish such as the three-spined stickleback, *G. aculeatus,* and pufferfish, *T.* nigroviridis.

2.4.2. Gene loci for the Class I eIF4Es

The gene loci for the Class I eIF4E cluster was examined and the proximal genes that overlap are highlighted (Table 2.1).

Table 2.1 Gene loci of the eIF4E-1A family members												
human	tspan5	rp11	btf313	**eIF4E1A**	tbcap3	mir3684	metap1	fam177a	ndufs5p4	abt1p1	adh5	adh4
Xenopus	adh1c	h2afz	dnajb14	mttp.1	dapp1	adh7	adh18	adh1a	adh5	metap1	**eIF4E1A**	tspan5
shark	rap1gds1	tspan5	**eIF4E1A**	metap1	UC	UC	shb	tdrd7	tmod1	tstd2	ncbp1	
coelacanth	rap1gds1	tspan5	**eIF4E1A**									
gar	**eIF4E1A**	metap1	adh5	adh8b	dnajb14	cf1	casp61l	ccdc109b	lef1			
zebrafish	metap1	adh8b	adh8a	c13h4orf32	**eIF4E1A**	lingo2	c13h0orf72					
pike	lingo3	C4orf32	adh1l	metap1	**eIF4E1A**	lingo2	ch9orf72	kiaa1109				
tongue sole	adh3cll	metap1	**eIF4E1A**	lingo3	c1h9orf72	kiaa1109						
fugu	metap1	**eIF4E1A**	lingo3	C9orf72	kiaa1109	kiaa1109l	kiaa1109l	tpol	capn1			

Table 2.1: Gene loci of the eIF4E-1A family members: Full suite of genes represented are proximal to the eIF4E within 0.1-0.4 mb on contig or chromosome. Color scheme is coordinated with identical genes. eIF4E members are in bold. UC (uncharacterized)

Additional genomes have been newly completed that provide detailed coverage of ancestral lineages (shark, coelacanth, spotted gar) as well as the more recently evolved percomorphs, such as the tongue sole (Cs). Upgrades to the annotations of existing genomes have also supplied an enhanced comprehensive picture of what genes are located near/far proximal to the eIF4E members. In the eIF4E-1A gene loci, the predominant gene appears to be metap1 (methionyl aminopeptidase**)**. My gene location analysis has also revealed that the teleosts that retain an eIF4E-1A- like family member have one universal signature gene, the Gar1 (ribonucleoprotein). It was previously reported that eIF4E-1B in zebrafish is not orthologous to the tetrapod form because the locus is not conserved (30). This analysis was done before so many genomes were available and before they were so well annotated. However, in the spotted

gar *L. oculatus*, the *Eif4e1b* locus is the same as that found in the tetrapods as is the *Eif4e1b* locus in elephant shark. Since convergent evolution seems an improbable explanation of this, it seems possible that multiple *eif4e1b-like* loci existed in the common ancestor of Actinopterygii and Sarcopterygii, with some of them asymmetrically retained in Actinopterygii such as the *eif4e1b* locus in zebrafish, while other ancestral *eif4e1b* genes gave rise to *eif4e1b* of Tetrapoda (30). In the *eif4e1b* loci, tspan17 tetraspanin) and sncb (synuclein) genes are conserved in tetrapod, basal ray fishes, and chondrichthyes. The representatives I have provided for eIF4E-1B in teleost, zebrafish and northern pike (El), have only the casr (calcium sensing receptor) gene in common (Table 2.2).

Table 2.2 Gene loci of the eIF4E-1B family members

human	rnf44	cdhr2	gprin1	mir4281	sncb	**eIF4E1B**	tspan17			
Xenopus	cdhr2	sncb	**eIF4E1B**	tspan17	unc5a	hk2	sh2d4b	znf346		
shark	faf2	cltb	cdhr2	gprin1	sncb	**eIF4E1B**	tspan17	UC	unc5a	pdlim7
coelacanth	rnf44	r	cdhr2	gprin1	scncb	**eIF4E-1B**	tspan17			
gar	anxa6	tnip1	gpx3	dctn4	synpo	tspan17	**eIF4E1B**	mchr2	ctnna1	lrrtm2
zebrafish	wasf3a	gtf3aa	mtif3	gsx1	abhd10a	tagln3a	zgc:152816	zgc:175280	**eIF4E1B**	casr
pike	matr3l	slc7a2l	**eIF4E1B**	casr	UC					

Table 2.2: Gene loci of the eIF4E-1B family members: Full suite of genes represented are proximal to the eIF4E within 0.1-0.4 mb on contig or chromosome. Color scheme is coordinated with identical genes. eIF4E members are in bold. UC(uncharacterized)

Worth noting is the obvious deviation between the tetrapod and teleost *Eif4e1b* gene loci as indicated by Evsikov (30). Additional prior analysis by Evsikov included a schematic for what they designated *Eif4e1_1*, *Eif4e1_2*, and E*if4e1_3* gene loci arrangement where *Eif4e1_3* is the Eif4e1c gene. The principal gene proximal to eIF4E-1C is the tet1 (tet methylcytosine dioxygenase 1), but it

appears that the gene slc25a16 (soluble carrier family 25 mitochondrial carrier) is also well maintained up the evolutionary tree from the chrondrichthyes (Table 2.3).

Table 2.3 Gene loci of eIF4E-1C family members										
shark	hnrnph3	pbld	dna2	**eIF4E1C**	tet1	slc25a16				
coelacanth	tspan15	rufy2	hnrnph3	pbld	dna2	slc25a16	tet1	**eIF4E1C**	bloc1s2	
gar	tspan15	rufy2	hnrnph3	pbld	dna2	stox1	ccar1	**eIF4E1C**	tet1	alox5b
stickleback	stox1	ccar1	**eIF4E1C**	tet1	fam21c	alox5a	slc25a16			
zebrafish	fam21c	alox5a	slc25a16	tet1	**eIF4E1C**	ccar1	stox1			
cavefish	stox1	ccar1	**eIF4E1C**	tet1	fam21c	alox5a	slc25a16			
pike	slc25a16	tet1	**eIF4E1C**	tmp150a	nfu1					
tongue sole	slc25a16	tet1	**eIF4E1C**	tmp150a	nfu1					

Table 2.3: Gene loci of the eIF4E-1C family members: Full suite of genes represented are proximal to the eIF4E within 0.1-0.4 mb on contig or chromosome. Color scheme is coordinated with identical genes. eIF4E members are in bold.

In spotted gar, tet1 does retain proximity to eIF4E-1C.

2.5. Deuterostome Class II eIF4E family members

2.5.1. Phylogenetic analysis of Class II eIF4E family members

eIF4E family members of Class II eIF4E family members fall within two discrete clusters within the phylogenetic tree (Figure 2.5). The eIF4E designated as eIF4E-2 (2A) comprises the majority of the class II eIF4Es within the deuterostomes. The eIF4E-2B cluster is represented primarily by the ray-finned fish, but also by the amphibian *Xenopus* spp. It should be pointed out that the elephant shark eIF4E-2 is an outlier to the teleost eIF4E-2A, and the coelacanth Class II eIF4Es tend to segregate consistent with their evolutionary relationships.eIF4E-2 falls outside of the tetrapod eIF4E-2. As with the Class I eIF4Es, the eIF4E-2A in lower teleosts such as zebrafish and cavefish is related.

In comparing the eIF4E-2B cluster, it is evident there is tight grouping between the higher and lower teleosts and *Xenopus*.

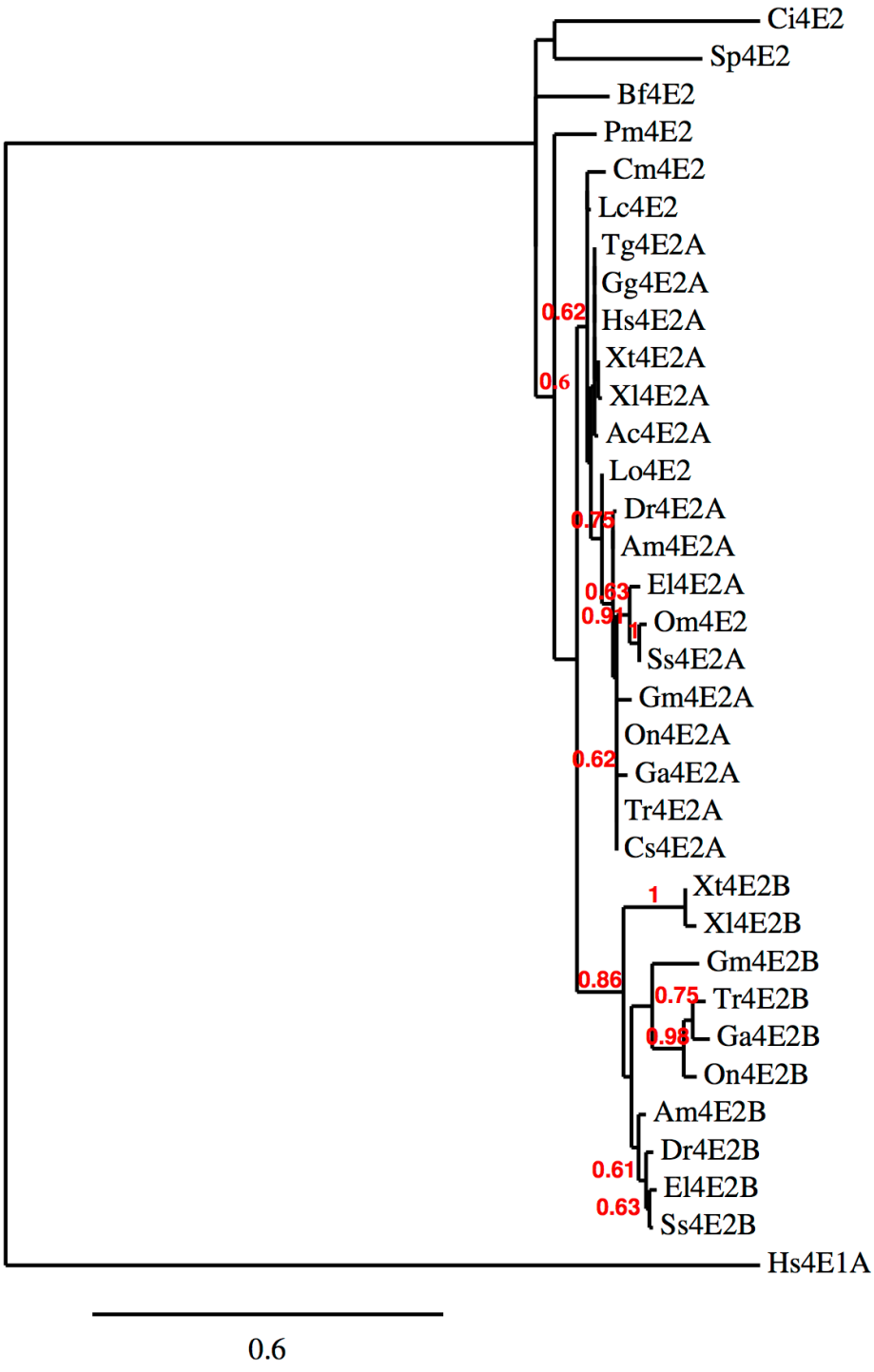


Figure 2.5: Phylogenetic analysis of deuterostome Class II eIF4Es: Phylogenetic trees based on the core and C-terminal regions from muscle alignments. The numbers on the branches are confidence limits

(expressed as percentages) estimated from a bootstrap analysis with 100 replicates (above 60 % are indicated). Bar 1.0 indicates substitutions per nucleotide position. Human eIF4E1A is used as out group.

2.5.2. Gene loci for the Class II eIF4Es

Table 2.4: Gene loci of eIF4E-2A family members											
human	chrng	prss56	chrnd	tigd1	**eIF4E2**	efhd1	gigyf2	eef1b2p7			
Xenopus	prss56	phr	chrnd	chrng	kiaa0226	**eIF4E2**	phrb				
shark	C2orf72	psmd1	htr2b	prss56	UC	chrnd	ps2l	**eIF4E2**	capn10	ecsrl	phr
coelacanth	chrnd	chrng	ps2l	**eIF4E2**	cap10l	cp450					
gar	znf862	capn10	**eIF4E2**	chrng	chrnd	vwa5b2	cops7b	prss56	alg3		
zebrafish	prss56	chrng	**eIF4E2A**	capn10	cu	mir					
stickleback	mul1a	alg3	prss56	chrng	**eIF4E2**	capn10	**gpsm2**	fndc7			
pike	cccm20	capn10	**eIF4E2**	ps2l	chrng						
tongue sole	prss56	chrng	ps2l	**eIF4E2**	capn10	ccmc26					

Table 2.4: Gene loci of the eIF4E-2A family members: Full suite of genes represented are proximal to the eIF4E within 0.1-0.4 mb on contig or chromosome. Color scheme is coordinated with identical genes. eIF4E members are in bold

When comparing the gene loci proximal genes (Table 2.4), *eif4e2a* shows a characteristic signature of the genes chrnd (cholinergic receptor nicotinic delta (muscle) and chrng (cholinergic receptor nicotinic gamma (muscle). This location is found exclusively in the tetrapod, whereas in teleosts only the chrng is found. Capn10 (calcium-activated neutral proteinase) is found in the teleost. The elephant shark and coelacanth appears to have a mixture, which may provide a strong indication of when the eIF4E2A began to diverge. The story is not as clear in the case of eIF4E-2B in teleosts, though there appears to be a distinct delineation of genes between the higher and lower teleost species (Table 2.5).

Table 2.5 Gene loci of eIF4E-2B family members							
Xenopus	mink1	wrap53	zpy1	rangrf	gp1ba	chrne	**eIF4E2B**
pike	eIF4E2B	cuorf	rgs11	tm8a	pdia2		
cavefish	eph4bl	edsp1	**eIF4E2B**	cb1l	nacht		
zebrafish	eph4b	**eIF4E2B**	act6b	cabz	spsb2	psmb6	
stickleback	ctc1	hmgb2b	atp1b2a	mogat3a	gucy2d	**eIF4E2B**	
fugu	ctc1	atp1b2a	mogat3a	**eIF4E2B**	gucy2d	sh3gl2	spag17

Table 2.5: Gene loci of the eIF4E-2B family members: Full suite of genes represented are proximal to the eIF4E within 0.1-0.4mb on contig or chromosome. Color scheme is coordinated with identical genes. eIF4E members are in bold

The *Xenopus eif4e2b* gene locus is of interest, not due to the comparison with the teleost, but because the proximal genes wrap53 and rangrf (in red)are involved in ribonucleoprotein complex formation of telomeres synthesis and protein transporter activity (http://www.genecards.org). This finding may have a relevance to our current studies on eIF4E2 (see Chapter 4).

2.6. Deuterostome Class III eIF4E family members

2.6.1. Phylogenetic analysis of Class III eIF4E family members

Class III eIF4Es are the most conserved eIF4E family members across the deuterostomes; only a single cognate is traditionally identified (Figure 2.6). Phylogenetically, all the clusters of eIF4E3 present themselves as tightly conserved units dependent on the sub-order, as was observed for the Class I. However, in the percomorph teleosts there is an additional eIF4E-3 member, designated the eIF4E-3 like, which has a gene loci organization that is distinct from eIF4E-3.

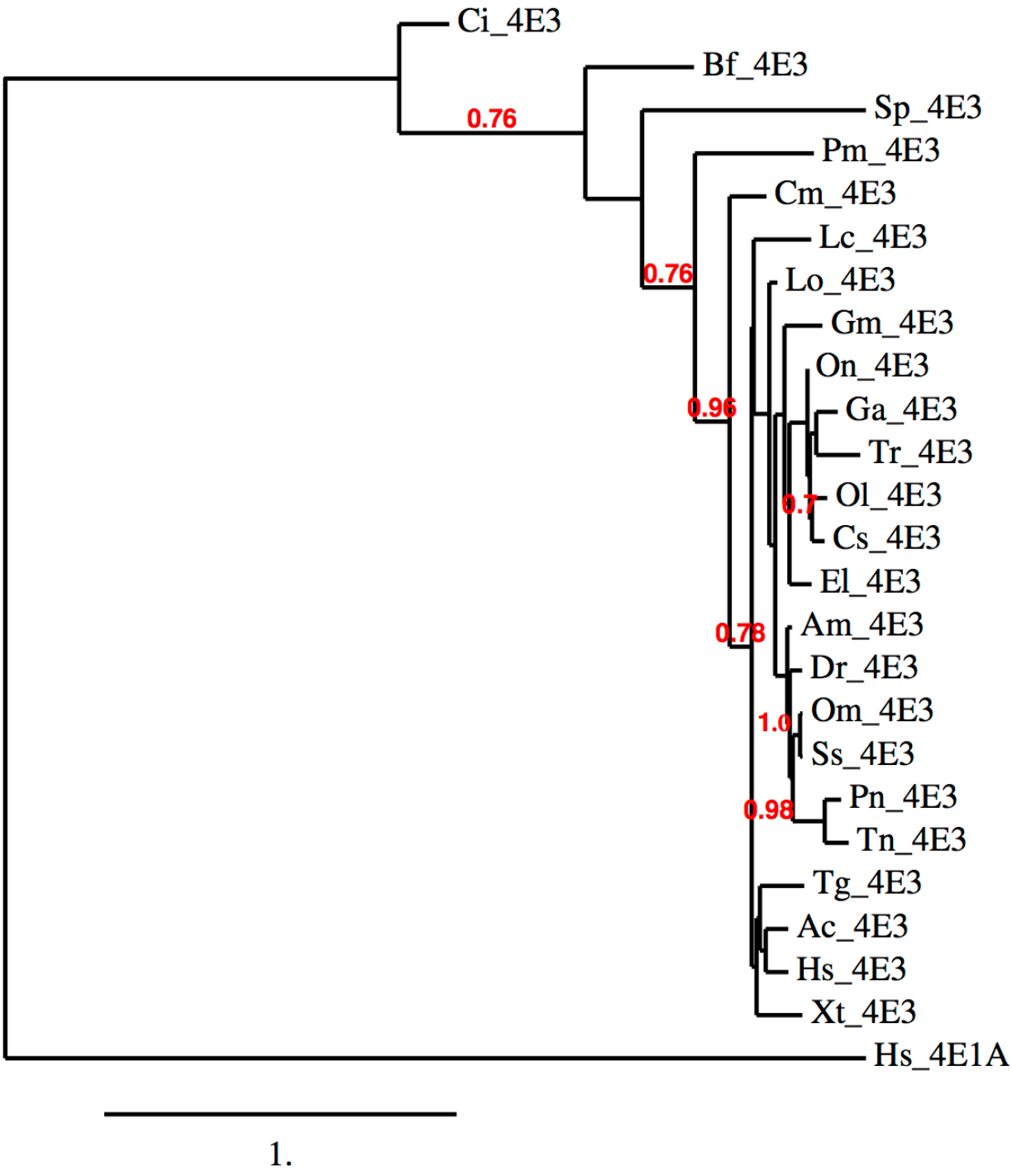


Figure 2.6: Phylogenetic analysis of deuterostome Class III eIF4Es: Phylogenetic trees based on the core and C-terminal regions from muscle alignments. The numbers on the branches are confidence limits (expressed as percentages) estimated from a bootstrap analysis with 100 replicates (above 60 % are indicated). Bar 1.0 indicates substitutions per nucleotide position. Human eIF4E1A is used as the outgroup

2.6.2. Gene loci for the Class III eIF4Es

When comparing the proximal genes from eIF4E-3 and eIF4E-3-like, it can be seen that gpr27 (G protein coupled receptor), rybp (RING1 and YY1 binding protein), and prok2 (Prokineticin) are characteristic of the *Eif4e3* loci (Table 2.6).

Table 2.6 Gene loci of eIF4E-3 family members

human	foxp1	mir1284	**eIF4E3**	gpr27	prok2				
Xenopus	**eIF4E3**	gpr27	rybp	shq1					
shark	rybp	prok1	gpr27	**eIF4E3**					
coelacanth	**eIF4E3**	gpr27	prok2						
gar	gxytl2	shq1	rybpa	prok2	gpr27	**eIF4E3**	foxp1b		
zebrafish	pdzm3a	ppp4r2a	rybp	**eIF4E3**	fox1a	tenc1a			
cavefish	krt18	eIF4ba	tenc1a	**foxp1a**	eIF4E3	rybpa	ppp4r2a	pdrn3a	
stickleback	pd2m3b	ppp4r2b	gxylt2	shq1	gpr27	**eIF4E3**	prok2	foxp1b	mitafa
tongue sole	**eIF4E3**	gpr27	prok2	foxp1	plxnal				
pike	dcrml	foxp1l	**eIF4E3**	gpr27	tMIT1l				

Table 2.6: Gene loci of the eIF4E-3 family members: Full suite of genes represented are proximal to the eIF4E within 0.1-0.4 mb on contig or chromosome. Color scheme is coordinated with identical genes. eIF4E members are in bold.

The foxp1-like (forkhead box P1) and the mitf (microphthalmia-associated transcription factor) are retained by the loci of the *Eif4e3-like* gene. However, alignment and phylogenetic analysis of these sequences does not show an obvious difference in the C-terminal regions of the eIF4E-3-like. The N terminal region does present a marked variation, but only few residues are different between eIF4E3 and eIF4E3-like in the core region or those key residues as discussed in Chapter 3.

2.7. Identification of gene loci orientation in deuterostome eIF4Es and implication for co-expression

WGD events enabled the formation of gene paralogues, which became the multitude of eIF4Es that have persisted in the evolution of the teleosts. These retained genes can provide a framework to categorize adjacent genes found in loci formation. Arrangements of gene loci are not random, and regulation of gene function may be inevitable. This co-expression of clustered genes has been documented across eukaryotic systems including human and *C. elegans* (109, 181). With the advent of accessible databases containing genomic complete annotations, it has become possible to locate a gene on a chromosome/scaffold. From this vantage point, it is possible to ascertain the orientation and distance of neighboring genes, and assign potential for expression. In mammalian systems, gene loci orientation has been examined for relative importance. It seems that when proximal genes are in the “head to head” or “HH” orientation ← →(also known as divergent transcription) gene pairs show a positive correlation for expression and genes in many such pairs share a regulatory element (110). In zebrafish, the gene orientation and co-expression has been linked to those genes pairs which display a parallel transcription, ← ← or → →. It was speculated that this occurrence is due to the genes being driven by 5’ *cis* regulatory elements or by bidirectional promoters found in zebrafish (111). Utilizing this information, the orientation of the predominant gene proximal to the eIF4Es that was discussed prior in this chapter was compared in human, shark, coelacanth, zebrafish, and northern pike. The preliminary results indicate possible co-expression of these eIF4E proximal genes in zebrafish across

classes of eIF4E, due to the parallel transcription orientation observed (highlighted in green). Conversely, there are only two potential HH orientation proximal genes in human, those from eIF4E-1A and eIF4E-3 (highlighted in orange). As in tetrapods, coelacanth may have co-expression of the *tet1* gene, proximal to the *Eif4e1c* gene of higher teleosts. Likewise, elephant shark may have co-expression of the *capn10* gene, proximal to the *Eif4e2* gene. The *Eif4E3 gene of* northern pike, elephant shark, and coelacanth all show HH orientation

Table 2.7 Comparison of gene orientation in eIF4E family members

	human			coelacanth			elephant shark			zebrafish			northern pike		
	eIF4E	gene		eIF4E	gene		eIF4E	gene		eIF4E	gene		eIF4E	gene	
eIF4E-1A	<	metap1	>				<	metap1	>	>	metap1	>	<	metap1	>
eIF4E-1B	>	sncb	<	>	sncb	<	>	sncb	<						
eIF4E-1B										>	casr	>	>	casr	>
eIF4E-1C				>	tet1	>	<	tet1	<	>	tet1	>	>	tet1	>
eIF4E-2A	>	chrnd	>	<	chrnd	>	<	chrnd	>	<	capn10	<	>	capn10	>
eIF4E-2A	>	chrng	>	<	chrng	>	<	capn10	<	<	chrng	>	>	chrng	<
eIF4E-2B										>	ehp4b	>	<	rgs11	>
eIF4E-3	<	gpr27	>	<	gpr27	>	<	gpr27	>	<	rybp	<	<	gpr27	>

<> divergent transcription (head to head) >> << parallel transcription

tetrapod | low Teleost | high teleost | orientation like tetrapod | orientation like teleost | Not found

with the *gpr27* gene that is specific to tetrapods (Table 2.7).

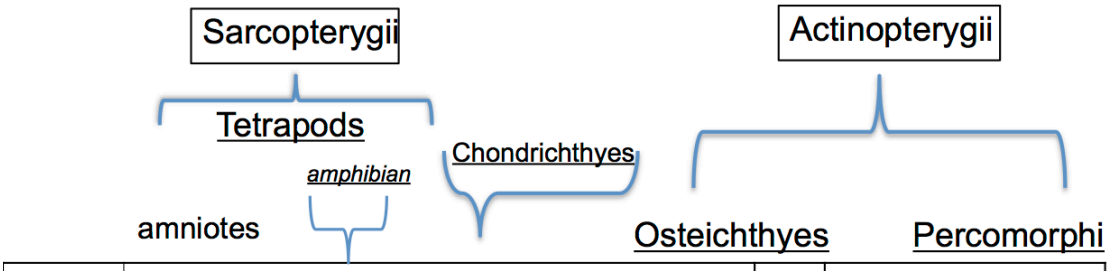
It may be of interest to determine if this gene has some evolutionary significance that caused it to be maintained in a wide variation of lineages. Though not definitive about unknown functions, analysis of possible correlation of related function in paired co-expression may provide key insights. My analysis is preliminary in scope, but it may be probable for an extensive gene survey of the

arrangement and frequency of the gene far/near proximal pairs found in relation to all *Eif4e* genes. This could be verified by analyzing expression data.

2.8. Discussion

Table 2.8 summarizes the distribution of eIF4E family members in deuterostomes.

Table 2.8 Phylogenic distribution of eIF4E family members in deuterostomes



	Hs	Tg	Ac	Xt	Lc	Cm	Lo	Dr	Am	**El**	Ss	Gm	Tr	Cs
eIF4E1A	1A	1A	1A	1A	1A	1A	1A	1A	**	**1A**	1A	1A	1A	1A
eIF4E1Al								**1Al**	**1Al**	**1Al**			**1Al**	**1Al**
eIF4E1B	1B	1B	1B	1B	1B	1B	1B	1B	1B*	**1B**	1B			
eIF4E1C					1C	1C	1C	1C	1C	**1C**	1C	1C	1C	1C
eIF4E2A	2A	2A	2A	2A	2A	2A	2A	2A	2A	**2A**	2A	2A	2A	2A
eIF4E2B				2B				2B	2B	**2B**	2B	2B	2B	2B
eIF4E3	3	3	3	3	3	3	3	3	3	**3**	3	3	3	3
eIF4E3l									**3l**	**3l**	**3l**	**3l**	**3l**	**3l**

Table 2.8: Phylogenic distribution of eIf4E family members in deuterostomes: Hs,*Homo sapiens;* Tg, *Taeniopygia guttata;* Ac, *Anolis carolinensis;* Xt , *Xenopus tropicalis;* Lc, *Latimeria chalumnae;* Cm, *Callorhinchus milii;* Lo, *Lepisosteus oculatus;* Dr, *Danio rerio;* Am, *Astyanax mexicanus;* El, *Esox lucius*; Ss *Salmo Salar;* Gm, *Gadus morhua*; Tr, *Takifugu rubripes*; Cs, *Cynoglossus semilaevis.*(**)eIF4E-1A sequence was not located in the genomic databases. (*) A partial eIF4E-1B sequence identified.

The distribution of the subclasses of eIF4E1 and eIF4E2 is consistent with the duplication of Class I and II prior to the teleost specific whole genome duplication. eIF4E-1A is prevalent across deuterostomes from echinoderms to mammals, but eIF4E-1C is lost in tetrapods. eIF4E1B has apparently been lost in the percomorph teleosts, but retained in sharks, basal ray-finned fish, lower teleosts and tetrapods. eIF4E-2B has been lost in the amniotes but retained in basal ray-finned fish, teleosts and *Xenopus*. The *Eif4e genes* of teleosts and

tetrapods display marked differences in their proximal genes. Interestingly, it is the *Eif4e* genes in elephant shark and coelacanth that appear to preserve nearly identical proximal genes to each other. The representative of basal ray finned fish, spotted gar, has proximal gene patterns similar to lower and upper teleosts, sarcopterygii and chondrichithyes across the eIF4E cognates. Northern pike (EI) are genetic wild cards of sorts, in that they have all eight known deuterostome eIF4Es. Conclusions could be drawn that the reduction of eIF4E family members accompanied the evolution of the amniotes. In the expansion of the deuterostome suite, eIF4Es preceded the diversification of the teleosts.

Chapter 3: Comparison of Class I and III eIF4E Family Members in Zebrafish (*Danio rerio*)

3.1. Abstract

Six members of the eukaryotic translational initiation factor 4 (eIF4E) family of proteins have been identified in zebrafish. Functional characteristics of zebrafish Class I eIF4Es, eIF4E-1A and eIF4E-1B *in vitro* have been studied previously, identifying eIF4E-1A as a prototypical initiation factor and eIF4E-1B as a tissue specialized translational regulation factor. Hitherto nothing has been reported on the function of zebrafish Class I eIF4E-1C or the Class III eIF4E-3. Here we describe the characterization of zebrafish eIF4E-1C and eIF4E-3. eIF4E-1C can be recognized first in jawed vertebrates and persists in teleosts. Although eIF4E-1C is present in coelacanth, a basal sarcopterygian, it has been lost in tetrapods. eIF4E-3 is found in all deuterostomes. eIF4E-1C is ubiquitously expressed like eIF4E-1A, but has higher protein expression levels than eIF4E-1A across adult tissues, during early embryogenesis, and in the zebrafish liver cell line (ZFL). We show that, like eIF4E-1A, eIF4E-1C is confirmed to function as a translational initiation factor by its ability to bind to cap analogue, interact with the scaffold protein (eIF4G), and complement a *S. cerevisiae* strain conditionally deficient in functional eIF4E. Like eIF4E-1A, eIF4E-1C also interacts with the eIF4E-binding proteins (4E-BPs). Although zebrafish eIF4E-3 binds to cap and eIF4G, it does not complement in eIF4E conditionally deficient yeast strain. Here we provide the first assessment of protein expression of eIF4E-3 in tissues and non-transformed cells that indicates that its levels are lower than that of eIF4E-1A and -1C. eIF4E-

3 is found above threshold levels only in specialized tissues such as muscle and brain.

3.2. Introduction

The importance of eukaryotic translational initiation factor eIF4E lies in its ability to recruit mRNA to the ribosome through specific and high affinity binding to eIF4G (reviewed (2, 20, 73, 112, 113)). This is accomplished by the binding of eIF4E to the 7-methylguanosine cap structure at the 5'-end of mRNA, allowing for interaction with eIF4G, eIF4A, and eIF3 (reviewed (3, 4, 69, 73, 114, 115)). This assemblage places the 40S ribosomal subunit in contact with the 5'-end of mRNA, so translation can commence. eIF4E structure and activity is highly conserved across eukaryotic lineages with the core structure representing a novel fold (2, 13). eIF4E is part of an extended gene family found exclusively in eukaryotes (15, 21, 22, 116, 117). Although the family is named for the translation initiation factor, not all members of the gene family function as such. There is an accumulation of evidence showing functional specialization of eIF4E cognate proteins, each having a particular role in the regulation of gene expression, some involved in translational initiation but others having alternate functions, including modulation or suppression of translation of particular mRNA species (15, 20, 21, 70, 116, 117).

Phylogenetic analysis has grouped eIF4Es from multicellular eukaryotes into three classes, Classes I-III, with mammals expressing two Class I eIF4Es, eIF4E-1A and -1B, one Class II, eIF4E-2 (4EHP) and one Class III, eIF4E-3 (15, 39). The eIF4E fold is characterized by an eight β-sheets that form the cap cavity,

backed by three long α-helices with a binding site for eIF4E protein partners, including eIF4G and a variety of regulatory proteins such as the 4E-BPs (16-18). The consensus sequence of the conserved core of eIF4E shows a distinctive pattern of aromatic residues Trp, Phe, and His across from N- to C-terminus (15). The contacts between the translation factor eIF4E and cap analogues involve sandwiching of the aromatic guanine residue of the cap-structure between two tryptophans in (in metazoan Class I eIF4Es), or a tryptophan and a tyrosine (in metazoan Class II eIF4Es). Additional contacts include hydrogen bonds with the N(^{7})-methylguanosine and the second nucleoside, as well as direct and water-mediated contacts with the phosphate chain (16-19). The structures of mammalian eIF4E-1A, eIF4E-2 and eIF4E-3 resolved in NMR or crystallographic studies all show that the characteristic α+β domain is representative of all three metazoan classes of eIF4E (16, 18, 42, 118).

In ray-finned fish (Actinopterygii), such whole genome duplication occurred coincident with the radiation of teleost species. This has been termed the teleost specific whole genome duplication (TGD) (102, 103). There is additional evidence that two earlier rounds of large-scale gene duplication occurred early in vertebrate evolution (86, 119). In general, while most gene pairs formed by WGD are subsequently deleted, rapid functional divergence is known to allow duplicate gene retention (84, 120-123). Such neofunctionalization may account for the fact that the zebrafish genome has additional Class I and Class II *Eif4e* genes compared to tetrapods; three that express Class I eIF4Es, termed eIF4E-1A, -1B and -1C, and two that express Class II eIF4Es, termed eIF4E-2A and -2B.

Zebrafish eIF4E-1A has been described previously as a prototypical translation initiation factor, expressed ubiquitously, able to bind m^7GTP, interact with eIF4G and the vertebrate 4E-BPs (26). eIF4E-1A also complements a *S. cerevisiae* strain containing a glucose repressible *eif4e* gene (26, 88). In contrast, zebrafish eIF4E-1B is a tissue specific translational regulation factor expressed primarily in ovary and testis (26) that is also seen in tetrapods (27, 30, 32). Although all the residues critical for 5'-cap mRNA binding and interactions with eIF4Gs or eIF4E-BPs are absolutely conserved among eIF4E-1Bs, eIF4E-1B shows only weak interactions with m^7GTP-Sepharose, eIF4G and 4E-BPs (26, 31). Conversely, eIF4E-1B is distinguishable from eIF4E1A by a set of conserved amino acid substitutions several of which are located near to cap-binding residues (31). Instead of eIF4G, eIF4E-1B interacts with the purported eIF4E transporter protein, 4E-T (27, 28, 33). Unlike eIF4E-1A, *D. rerio* eIF4E-1B cannot be exchanged for mammalian eIF4E in complementation assays using an *S.cerevisiae* strain conditionally deficient in eIF4E (26, 88). *Xenopus* eIF4E-1B is found in a complex with 4ET, CPEB and mRNAs containing 3'-UTR recognized by CPEB precluding productive binding of eIF4E-1A to eIF4G (27, 28). Interestingly, although tetrapod eIF4E-1Bs have a high identity (72.4 %) and similarity (82.8 %) index when compared to zebrafish eIF4E-1B in the conserved core region, and have a similarly restricted pattern of expression, the zebrafish *Eif4e1b* gene is not orthologous to the *Eif4e1b* locus of tetrapods (30). Since convergent evolution seems an improbable explanation of this, it is possible that multiple *Eif4e1b*-like loci existed in the common ancestor of Actinopterygii and

Tetrapoda. It may have occurred that with some of them were asymmetrically retained in Actinopterygii ,such as the *Eif4e1b* locus in zebrafish, while other ancestral *Eif4e1b* genes gave rise to *Eif4e1b* of Tetrapoda (30).

The two Class II family members of zebrafish, eIF4E-2A and -2B, are anticipated to have a similar regulatory role to Class II eIF4Es from *Drosophila*, *C. elegans* and mouse (33, 34, 38, 124-126), although some neofunctionalization should be anticipated from their evolutionary persistence. An investigation of these orthologues is the subject of a separate study (Chapter 4 & Gillespie *et al*, ms in preparation). The role of eIF4E-3, found primarily in chordates, is the least understood of the chordate eIF4E family members. Mammalian eIF4E-3 binds the m^7G cap in the absence of an aromatic sandwich, using instead a cluster of hydrophobic and charged residues in the C-terminus to make extensive contact with the cap to increase affinity (42). Only one variant of eIF4E-3 has been found in most chordates. However, in Percomorpha, the most recently evolved teleosts, such as tongue sole, *Cynoglossus semilaevis,* and pufferfish, *Tetroadon nigroviridis*, there is an eIF4E-3 cognate protein termed eIF4E3-like (Chapter 2 & Gillespie, Bachvaroff & Jagus, m/s in progress). eIF4E-3 appears to have a limited tissue distribution. Its role in the regulation of gene expression is not well established. In mammals, eIF4E-3 functions as a tumor suppressor suggesting a role in repression of mRNA utilization (42), although this role seems at odds with its ability to prevent muscle atrophy (43).

In the present study, we describe the expression and functional characteristics of eIF4E-1C and eIF4E-3 and compare them with eIF4E-1A, and eIF4E-1B. It

appears that eIF4E-1C, previously recognized but uncharacterized in the Jagus laboratory, functions as a prototypical translational initiation factor. Furthermore, eIF4E-1C appears to be the prevalent form of translation initiator expressed throughout adult zebrafish tissue, in early embryogenesis and in cultured ZFL cells. eIF4E-3 does not appear to function as a translational initiation factor; it does not rescue initiation of translation in eIF4E-deficient yeast cells, and is evident only in muscle, heart, and brain. The results presented here on eIF4E-3 are consistent with the emerging picture of eIF4E3 from other systems as having a regulatory role in mRNA recruitment in select tissues.

3.3. Materials and Methods

3.3.1. Rearing and spawning zebrafish

Adult fish were maintained at 28.5 °C in a constant flow-through system. Embryos were obtained by spontaneous spawning, maintained at 28.5 °C, and staged as described (127). Staged embryos were either immediately processed or snap-frozen and stored at -80 °C for future use.

3.3.2. Culture of ZFL cells

Cells were grown at 28 °C in L-15 medium supplemented with 10 % fetal calf serum but without sodium bicarbonate.

3.3.3. Identification of zebrafish eIF4E family members

eIF4E-1A (Genbank mRNA AF176317.1, (cds): AAG09794.1) and eIF4E-1B (UniProtKB/Swiss-Prot: Q9PW28.1) were previously described (26). The sequence for eIF4E-1C was deposited into Genbank as NP_001017851.2, and

eIF4E-3 as NP_001004589.1. The Jagus laboratory deposited the indicated *Danio rerio* eIF4Es sequences (mRNA and/or coding regions) in Genbank from 2000-2013. Other sources have provided additional sequences (BC081620.1, NP_571529.1, AAH55649.1, AAD50526.1) that are identical to the deposited Jagus laboratory sequences.

3.3.4. Identification of eIF4E family members from other deuterostomes:

The peptide sequences of eIF4E family members of the deuterostomes included in this study have been collected from on-line genomic resources including; the National Center for Biotechnology Information (NCBI) at http://www.ncbi.nlm.nih.gov, the Ensembl project (156), the HMMR database (http://hmmer.janelia.org), the Institute of Molecular and cell Biology (IMCB) elephant shark genome http://esharkgenome.imcb.a-star.edu.sg, the coelacanth genome project site (http://coelacanth.nig.ac.jp/index.php) and the Joint Genome Institute (JGI) for the *Branchiostoma floridea* genome (http://genome.jgi-psf.org). Each eIF4E sequence was verified using the Genbank BLAST tool and aligned by the MUSCLE algorithm applying the suite of software provided by CLC workbench (CLCBio CLC Genomics Workbench 7.0.3 (http://www.clcbio.com). Accession numbers and details on sequences are provided in the Appendix (Appendix Table A2.1).

3.3.5. Generation of cDNAs encoding zebrafish eIF4E family members

The generation of zebrafish eIF4E-1A and eIF4E-1B constructs have been described previously (26). cDNAs encoding zebrafish eIF4E-1C and eIF4E-3 were cloned into the *in vitro* transcription/ translation plasmid vector pCITE-4a(+)

(Novagen, EMD-Millipore, Billerica, MA, USA), using engineered NcoI and BamHI sites (primers listed in Table 3.1A)

3.3.6. Generation of constructs encoding zebrafish 4E-BPs and fragment of zebrafish eIF4G1

The generation of a zebrafish 4E-BP3-like construct has been described previously (26). Nucleotide sequences for zebrafish 4E-BPs 4E-BP1 (NP_955939.1), 4E-BP2 (NP_997968.1), 4E-BP3 (NP_001007355.1) were codon optimized for rabbit, *Oryctolagus cuniculus*, using Advanced OptimumGene™ (Genscript, Piscataway, NJ, USA). The nucleotide sequence was synthesized by Genscript, augmented with additional methionine residues and cloned into the *in vitro* transcription/translation plasmid vector pCITE-4a (+) (Novagen, EMD-Millipore, Billerica, MA, USA), using engineered Nde1 and BamHI sites. The cloning strategy adds an S-tag to the amino-terminus and uses the stop codon from the coding sequences. The nucleotide sequences for the zebrafish eIF4G-1 fragment 262-681 were codon optimized for rabbit, synthesized and cloned into pCITE4a (+) as for the 4E-BPs (Genscript, Piscataway, NJ, USA) using the same cloning strategy.

3.3.7. RNA purification, cDNA synthesis, RT-PCR and RT-qPCR

Fresh tissues, embryos, or harvested cells were homogenized by bead beating and extracted using a Purelink RNA minikit: (Ambion™ Grand Island, NY, USA). RNA was quantified on a Nanodrop 1000 (Thermo Fisher by Life Technologies Waltham, MA). Values of >2 for 260/280 and 260/230 ratios were considered to be of sufficient purity. RNA was reverse-transcribed using Superscript II reverse

transcriptase (Invitrogen, by Life Technologies, Carlsbad, CA, USA) with random hexamers in accordance with the manufacturer's instructions. The generated cDNA was used as template for RT-PCR and RT-qPCR. The quality of all cDNA preparations was assessed by end point PCR amplification.

Table 3.1A: zebrafish eIF4E cloning primers

eIF4E	Forward/Reverse	Sequence	T_m	bp
1A	F	CGAGCCATGGCGACTGCTGAACCGGAAAC	67.7	937
	R	GAAGGATCCGCACTCCCCCAATCCCCACTA	67.3	
1B	F	GCAGCCATGGCGTCGTGTGCTGTACAACTGATTGATAAAGTACCGAAG	68.3	667
	R	CCAGGATCCGCCCACTTTTAAACAACAAACT	62.4	
1C	F	ATATATCCATGGCGACTTCGGAGCCG	62	662
	R	TACAACAAAGAATATGTACTCTGTTTGAGGATCCAAGAAG	60.5	
2A	F	GGCAAACCACCATGGACAACAAATTTGAC	64.0	704
	R	GGCGGATCCCTATACGAAATCCTCCCAAGC	64.2	
2B	F	GGCAAACCACCATGGATCAGTTTGAAC	60.5	735
	R	GGCAAATTCGGATCCTCACAAAGTGATC	59.8	
3	F	ATATATCCATGGCGGTTCCTGCAGCCC	58.9	692
	R	ATATGGATCCCTAATGTCTTGAGCGA	58.2	

were performed under standard conditions using Taq DNA polymerase (Denville Scientific Inc, South Plainfield, NJ, USA). Whenever purified PCR products were transferred into plasmids, insertions were sequenced in both orientations to ensure that no errors had been introduced due to amplification. The products were resolved by TAE-agarose electrophoresis and imaged in the Typhoon 9410 Variable Mode Imager (GE Healthcare Life Sciences, Pittsburgh, PA, USA).

Primers for qPCR were designed by PearlPrimer and Primer 3 software to span exon-exon junctions (Table 3.1B). For RT-qPCR using an Applied Biosystems Fast 7500 thermal cycler (Life Technologies, Grand Island, NY, USA), cDNA from 20 ng RNA was amplified using Taqman Fast Universal PCR Mastermix (no AmpErase UNG) (Applied Biosystems, Foster, CA, USA). Thermal cycling

conditions consisted of an initial denaturation at 95 °C for 2 min followed by 40 cycles of denaturation at 95 °C for 15 sec, annealing at 60 °C for 15 sec, and extension at 72 °C for 30 sec. The reaction was completed with a melt curve to detect any spurious PCR products. Each eIF4E was encoded into cDNA plasmids to generate transcripts used to determine the absolute copy number. A standard curve was subsequently constructed from the cDNA from 25 ng of *in-vitro* transcribed RNA and utilized for extrapolation of mRNA targets of unknown concentration. Elongation factor 1A (EF1A) was used as control.

Table 3.1B: zebrafish eIF4E qPCR primers

eIF4E	Primer(5' to 3')	F/R	Position	length	T_m	Size(bp)
1A	ACTGAATGTGATTGTATAACGCCC	F	234	24	61.75	170
1A	ATGAGCAACAGATCGTGAGTC	R	64	21	60.62	
1B	CTAAGGCTCATCACCAAATTCGA	F	228	23	61.2	119
1B	CTCTATGCCATCCTTGAACATGG	R	347	23	61.58	
1C	TGAACAGTACATCAAACACCCT	F	130	22	60.07	137
1C	TTGTATAATGCCCAGAAATCTTCC	R	267	24	59.39	
2A	ACGCCCTGAAAGATGATGAC	F	16	20	60.59	124
2A	GACCACTGCCTTTCTCTTTG	R	140	20	59.23	
2B	ACAGCCAATGATCAGGTGAC	F	537	20	60.52	125
2B	GAAGCTGGAGTTATCCTTCAGAC	R	662	23	60.95	
3	TGCATCAGAGGATGAAGTGGT	F	593	21	61.69	212
3	TGCTAATGTCTTGAGCGACC	R	805	20	60.59	
EF1A	CTTCAACGCTCAGGTCATCAT	F	1091	21	52.59	261
EF1A	ACAGCAAAGCGACCAAGAGGA	R	1351	21	56.35	

3.3.8. Preparation of protein extracts from cultured cells and tissues

Cultured cells or tissues (except for ovary) were homogenized in up to 10 volumes of ice-cold buffer containing 25 mM Tris -HCl, pH 7.4, 1 mM EDTA, 1 mM EGTA, 1 mM dithiothreitol, 100 mM KCl, 0.5 % Elugent, and Complete^{TM} Protease Inhibitors (Roche Applied Science, Madison WI, USA). Homogenates were clarified by centrifugation (15 k x g, 4 °C, for 15 min). Supernatants were snap-frozen and stored in liquid N_2.

3.3.9. *In vitro* transcription and translation

^{35}S-radiolabeled proteins were translated *in vitro*, using pCITE4a constructs as templates in the rabbit reticulocyte TnT (Promega, Madison, WI, USA) coupled transcription-translation system, containing [^{35}S]-methionine as per the manufacturer's directions. 2 μl of the *in vitro* translation reaction (IVT) reaction was taken for analysis of ^{35}S-methionine incorporation by mixing to a final concentration of 5 % TCA, boiling and capturing on GF/C filter paper(EMD-Millipore, Billerica, MA, USA).

3.3.10. m^{7}GTP-Sepharose binding assay

Sepharose beads bound to 7-methyl-guanosine-triphosphate (Jena Bioscience GmbH, Jena, Germany) were blocked using 1 mg/ml soybean trypsin inhibitor (Sigma, St. Louis, MO, USA) in binding buffer (25 mM HEPES/KOH pH 7.2, 10 % glycerol, 150 mM KCl, 1 mM dithiothreitol, 1 mM D-L methionine) for 1 h at 4 °C shaking at 1400 rpm in a benchtop thermomixer 22331 (Eppendorf, Hamburg, Germany). The beads were washed twice with binding buffer and suspended in

50 % v/v binding buffer. 20 μl of each *in vitro* translation product (IVT) was diluted 10-fold with binding buffer containing 200 μM GTP and 200 μM $MgCl_2$, mixed with the bead suspension and incubated at 4 °C for 1 h with shaking at 1400 rpm. The supernatant containing the unbound fraction was recovered by centrifugation at 500 x g at 4 °C. An equivalent of 2 μl of the original IVT was used for TCA precipitation and filtered onto a GF/C membrane (Millipore, Billerica, MA, USA,). These were washed 5 times with binding buffer and the final bead-bound fraction was suspended in SDS-PAGE sample buffer. The bead suspensions were heated to 90 °C and a fraction equivalent to 2 μl of the original IVT reaction applied to GF/C filter paper. Fractions were counted in Ecoscint Original scintillation cocktail (National Diagnostics, Georgia, USA) and cpm was determined using a LS6500 Multipurpose Scintillation Counter (Beckman Coulter). IVT, unbound, and bead bound fractions were diluted in SDS-PAGE sample buffer and heated to 90 °C for 3 min. The samples were separated by 17.5 % high-Tris SDS-PAGE, transferred to PVDF membranes and visualized using a Storage Phosphor screen (Molecular Dynamics, GE Healthcare Life Sciences, Pittsburgh, PA, USA) and imaged with a Typhoon 9410 Variable Mode Imager (GE Healthcare, Healthcare Life Sciences, Pittsburgh, PA, USA).

3.3.11. Protein-protein interaction assays

For protein interaction assays, a fragment of zebrafish eIF4GI from amino acid 262-681, containing the eIF4E-binding domain was cloned into pCITE4a. The zebrafish 4E-BP was co-translated with either S-tagged eIF4E-1A,-1B,-1C or -3 in 35 μl reactions for 60 min at 30 °C. Reactions were diluted with 10 volumes of

S-binding/washing buffer and incubated with 50 μl of S-protein agarose (Novagen, Madison, WI, USA) for 60 min at 10 °C. S-protein-agarose beads were recovered by centrifugation and washed 5 times with buffer (1 ml each), prior to elution with SDS-PAGE sample buffer. Samples of fractions, equivalent to 2 μl of the initial translation reactions, were analyzed by high-Tris SDS-PAGE, transferred to PVDF membrane and labeled proteins visualized using a Storage Phosphor screen (Molecular Dynamics, GE Healthcare Life Sciences, Pittsburgh, PA, USA) and imaged with a Typhoon 9410 Variable Mode Imager (GE Healthcare, Healthcare Life Sciences, Pittsburgh, PA, USA).

3.3.12. Production of recombinant eIF4Es from *E. coli*

Zebrafish eIF4E family members were sub-cloned into pET11d (Novagen, EMD Millipore, Billerica, MA, USA) to give untagged proteins. Constructs were transfected into Rosetta™ (DE3)-pLysS competent cells (EMD Miliipore, Billerica, MA, USA) and expressed essentially as described (26). 10-ml cultures were grown in LB, 100 μg/ml carbenicillin, 34 μg/ml chloramphenicol, overnight at 37 °C with shaking (220 rpm). Cells were harvested, resuspended in fresh medium, diluted to an optical density (OD) of 0.1 and grown to an OD of 0.5. Expression was induced with isopropyl β-D-1-thiogalactopyranoside (IPTG) at a concentration of 1 mM for 2.5 h. Cells were harvested by centrifugation at 10,000 x g for 5 min and lysed in 10 μg/μl lysozyme, 25 mM HEPES-KOH, pH 7.2, 100 mM KCl, 10 % glycerol, 1 mM EDTA, 1 mM EGTA, 0.5 % Elugent (Calbiochem La Jolla, CA, USA). The supernatant and/or protein pellet were isolated after

DNAse treatment. Expression was assessed after SDS-PAGE fractionation followed by staining or immunoblotting.

3.3.13. Development and validation of affinity-purified antibodies

The best antigenic regions to use for immunization were determined using the Genscript OptimumAntigen™ Design Tool. The Genscript™ company synthesized the suite of eIF4E antigenic peptides (Table 3.2) .

Table 3.2 : eIF4E Genscript antibodies		
eIF4E	**peptide sequence**	**region**
eIF4E-1A	HADTATKSGSTTKNKFVVC*	C terminus
eIF4E-1A	AEPETSTNPSNSEEC*	N terminus
eIF4E-1B	VPKKKVEKKKFEPNC*	N terminus
eIF4E-1C	TSEPRGTRTEEVRAC*	N terminus
eIF4E-2A	QDNSSPKDGEKEKNC*	N terminus
eIF4E-2B	EMKDNNESDRASINC*	N terminus
eIF4E-3	PHEEHHAFEGGRSRHC*	C terminus

Table 3.2: Peptides for antibody development zebrafish eIF4Es
* indicated cysteine addition for antigenic processing

There was an additional cysteine residue added at the C-terminus to allow for conjugation to the KLH adjuvant. Antibodies were raised in New Zealand white rabbits. Specific antibodies were isolated by affinity purification using the synthesized peptide. Antibodies were tested for specificity and cross-reactivity by an ELISA assay and western blot analysis using the peptide used to generate the antibody and the recombinant protein of each eIF4E, respectively. The specificity of each antibody was validated using recombinant eIF4Es and tested for cross

reactivity against all recombinant zebrafish eIF4Es (Figure 3.1).

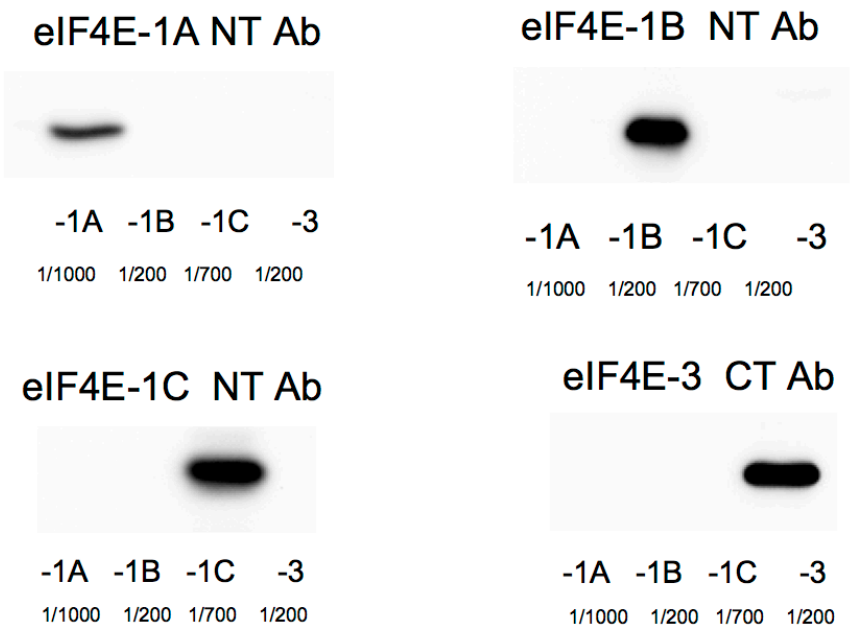


Figure 3.1: Specificity of antibodies for eIF4E-1A, -1B, -1C, eIF4E-3: dilutions indicated in x/y (x= µl of protein, y= µl volume SDS page sample buffer)

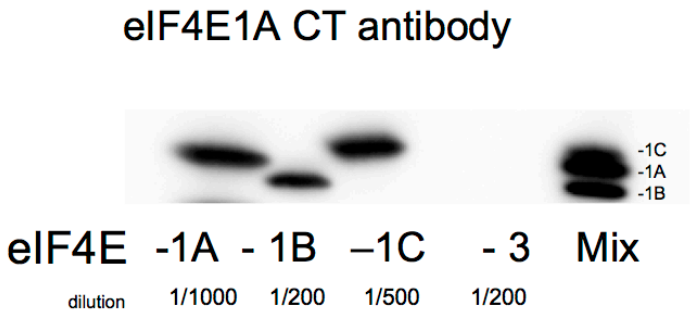
Antibody dilutions used for immunoblotting were adjusted to reflect the avidity and titer.

3.3.14. SDS-PAGE and immunoblotting

Proteins were fractionated by 17.5 % high-Tris SDS-PAGE as described (26, 128), and were electro-transferred to PVDF membrane and subjected to blot analysis using the custom polyclonal antibodies followed by goat anti-rabbit secondary antibody coupled to HRP for an chemiluminescence reaction.

Chemiluminescence was detected using the ProteinSimple Fluorochem E with quantification using AlphaImager software. When used with full size gels (16 x 18

cm), the SDS-PAGE conditions allowed resolution of all three Classes I eIF4Es (Figure 3.2)



.

Figure 3.2: Separation of Class 1 eIF4Es by 17.5 % high-Tris SDS-PAGE

3.3.15. Quantification of eIF4E levels

Expression of each eIF4E was determined by immunoblotting using standard procedures. Comparison of signal from equal loading of each recombinant protein allowed avidity of each antibody to be established. The ECL signal was normalized by avidity and the relative levels of each eIF4E determined. The relative levels of eIF4E-1A, -1B, -1C and eIF4E-3 were determined from tissue samples of muscle, brain, and ovary. Quantification by saturated pixel (SD) intensity was measured by Alphaimager™ software. A boxplot was generated from multiple samples, in which the line within the blot represents the median; the box length corresponds to the interquartile range, with bars bracketing the smallest and largest observed protein levels.

3.3.16. Complementation assays in *S. cerevisiae*

Each of the zebrafish eIF4Es were sub-cloned into the URA-selectable yeast expression vector pRS416GPD at BamH1 and Xbal sites (129) and transformed

into the *S. cerevisiae* strain JOS003 using a modified lithium acetate/salmon sperm carrier DNA/PEG method (130). JS003 is a LEU-selectable strain from which the endogenous *EIF4E* gene has been replaced by homologous recombination with a KanMX4 cassette making it resistant to G418 (88). JOS003 cells lack an endogenous yeast *eif4e* gene and express human eIF4E-1 under the control of the galactose-dependent and glucose-repressible GAL1 promoter. As a consequence, JOS003 cells are able to survive in medium containing galactose as carbon source but are not viable in medium containing glucose due to depletion of human eIF4E-1. Growth of JOS003 in glucose can be mediated by ectopic expression of a functional eIF4E in pRS416GPD at BamH1 and Xbal sites, the regulation of which is under the control of a glyceraldehyde-3-phosphate (GPD) promoter active in the presence of glucose. This system has been used previously to investigate the ability of heterologous eIF4Es to function in translation by rescuing growth in the presence of glucose (88). The transformed yeast were spot plated on synthetic deficient (SD) media lacking uracil and leucine and containing 200 μg/ml G418, with either galactose or glucose. Plates were incubated at 30 °C for 3-4 days, and growth was assessed visually by colony formation. Growth on plates containing glucose indicates the ability of an ectopic *eif4e* gene to complement eIF4E deficiency. To verify that the zebrafish eIF4Es were expressed as protein in yeast, protein extracts were prepared using the TCA extraction/bead homogenization method, essentially as described by the Keogh laboratory (131). Zebrafish eIF4Es were visualized by SDS-PAGE electrophoresis and immunoblotting by standard procedures.

3.3.17. Recovery of eIF4Es from zebrafish ovary extracts by methyl-^{7}GTP-Sepharose binding

Multiple zebrafish ovaries were combined, and subjected to 10 volumes of mild disruption buffer, 0.35 M sucrose, 25 mM HEPES-KCl, pH 7.2, 1.5 mM $MgCl_2$, 250 μg/ml lysolecithin, 1 mM spermidine, 1 mM DTT, protease inhibitor pill and homogenized briefly using the Kinematica Brinkmann Polytron PT 3000 (Brohemia NY, USA). Vitellogenin was released by this centrifugation at 1000 rpm (228 x g) for 10 min at 4 ^{o}C. The resultant pellet was washed in 10 volumes pellet rinse buffer, 0.35 M sucrose, 25 mM HEPES-KCl, pH 7.2, 1.5 mM $MgCl_2$, 1 mM spermidine, and recovered by centrifugation at 1000 rpm (228 x g) for 10 min at 4 ^{o}C. The pellet was resuspended in 10 vol pellet solubilization buffer, 140 mM KCl, 50 mM HEPES-KCl, pH 7.2, 5 mM EGTA, 1 mM spermidine, 0.1 % Elugent, 0.5 % Na deoxycholate, 10 % glycerol, protease inhibitor pill, vortexed and left on ice for 5 min. The supernatant was clarified by centrifugation at 10,000 x g x 5 min at 4 ^{o}C and stored in liquid N_2. 200 μl of this extract was bound to 25 μl of m^{7}GTP-Sepharose beads. Non-specific binding sites were blocked by washing with binding buffer, (25 mM HEPES/KOH pH 7.2, 10 % glycerol, 150 mM KCl, 1 mM dithiothreitol) ,that contained 1 mg/ml SBTI (soybean trypsin inhibitor). Extracts were incubated at 4 ^{o}C with agitation (1400 rpm) for 1 h. The supernatant containing the unbound fraction was recovered by centrifugation at 500 x g at 4 °C. The cap-analogue beads were washed 5 times with binding buffer and the final bead-bound fraction was suspended in SDS-PAGE sample buffer. Protein precipitated with 2 volumes acetone from the combined washes

overnight at -20 ^{o}C. Equivalent volumes of fractions representing 20 μl of ovary extract were used for SDS-PAGE and immunoblotting analyses.

3.4. Results

3.4.1. Sequence comparisons of eIF4E-1A, -1B, -1C and eIF4E-3

The multiple alignments of the zebrafish eIF4E family members can be seen in Chapter 1, Figure 1.11. Table 3.3 illustrates the predicted physical characteristics of the zebrafish Class I and Class III eIF4E family members.

Table 3.3: Characteristics of zebrafish eIF4E family members

eIF4Es	Gene ID	cds(bp)	#aas	pI	MW(kDa)	Chrom	Location	#exons	#Met
eIF4E-1A	79380	648	216	5.6	24.7	14	NC_007125.6	8	3
eIF4E-1B	30738	644	215	9.1	24.6	5	NC_007116.6	7	5
eIF4E-1C	550549	641	214	6.1	24.4	13	NC_007124.6	7	6
eIF4E-2A	541523	711	237	6	27	2	NC_007113.6	6	6
eIF4E-2B	393732	687	229	7	26.7	23	NC_007134.6	6	8
eIF4E-3	447850	674	225	5.4	25.3	23	NC_007134.6	7	3

Cds(bp): coding sequence base pair, #aas: number of amino acids,pI: isoelectric point MW: molecular weight, Chrom: chromosome,Met: methionine

eIF4E-3 is slightly larger than the Class I eIF4Es. All except eIF4E-1B have an acidic isoelectric point, and each resides on a different chromosome. To facilitate comparison between the zebrafish Class I eIF4Es with themselves and with human eIF4Es, the numbering of amino acids discussed in the text is as per the equivalent amino acid position in human eIF4E-1. The N-termini of eIF4E family members show the greatest variability with only 5-15 % identity between each. There are significant differences in the N-terminal domains of eIF4E-1A and -1C; a shorter N-terminal domain in eIF4E-1C without the multiple glutamic acid and glutamine residues found in eIF4E-1A. These differences suggest that perhaps eIF4E-1C plays a subfunctional role, providing translational initiation under

specific conditions. The basic isoelectric point of eIF4E-1B reflects a lysine-rich region in the N-terminus (26). Table 1.2 shows the identities and similarities (based on PAM 250 matrix) between the amino acid sequences representing the core regions of the zebrafish eIF4E family members. Comparisons of the amino acid sequences representing the core regions of zebrafish eIF4E family members reveal that they share ~35–40 % identity and ~60–65% similarity with one another.

Figure 3.3 represents sequence logos that were created from the alignments of the core sequences of eIF4E-1A, -1B, -1C and eIF4E-3 from a range of gnathostome and tetrapod species. A sequence logo is a graphical technique for displaying a summary of a set of aligned sequences (132, 133). Logos compare an overlay of multiple sequences based on the frequency of amino acid residues (height) and the charge to highlight similarities and differences between sequences. For the eIF4E Class I suite, a total of nine sequences of each cognate protein each were aligned. A list of the tetrapods and teleosts are given in the Appendix (Table A3.1). Echinoderm, chordate and agnathan eIF4Es were not included in this analysis because they encode only one cognate of each eIF4E from each class. The multiple alignments of the full sequences of the sequences analyzed in the logos are shown in Appendix (Figure A2.1).

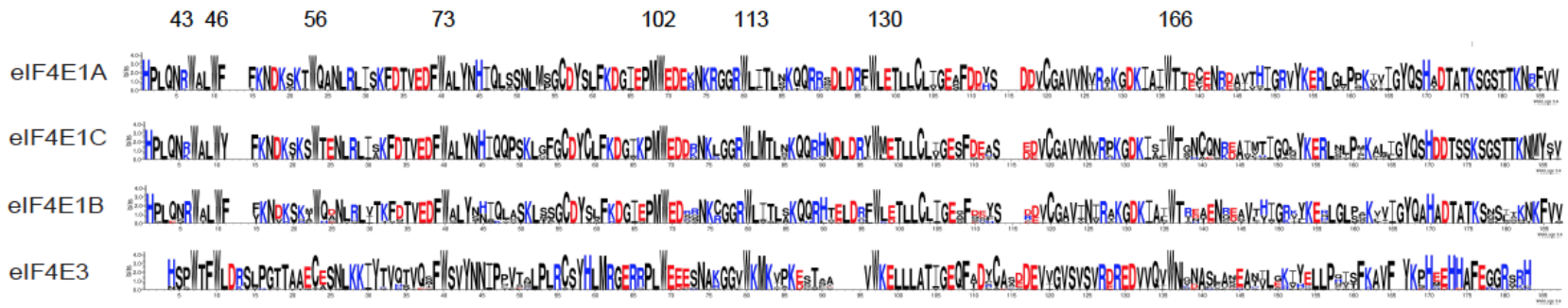


Figure 3.3: Logo of eIF4E-1A, -1C, -1B and eIF4E-3 alignments: The core region and C-terminal regions of eIF4Es from 11 species of teleosts and tetrapods are represented as logos. The charge is indicated as positive (blue) negative (red) or uncharged (black).

Reflecting their classification into Class I, eIF4E-1A, -1B, -1C, have tryptoph
(W) at positions equivalent to W43, W46, W56, W73, W102, W113, W130 a
W166 in human eIF4E-1A. Similarly, all three Class I eIF4Es have the positi
charged residues equivalent to R112, R157 and K162 that form salt bridges
the triphosphate of the cap. The logos highlight a signature residue pattern
surrounding His-170 that distinguishes the eIF4E Class I subtypes from eac
other and from eIF4E-3. eIF4E-1A has the motif “SHAD”, eIF4E-1B has “AH
eIF4E-1C has “SHDD” and eIF4E-3 has PHEEHH”. Using this distinction, it
possible to screen for the presence of a particular Class or subclass eIF4E
quickly across genomic databases. There are only a few differences in the
sequence of eIF4E-1C compared to eIF4E-1A; these include the substitutio
F47Y, T55S, Q57T, A58E, L81Q, S82P, S87F, S92C, E99K, R109L, A201D
A204S, and T205S.

3.4.1b. Zebrafish eIF4E-1B

Zebrafish eIF4E-1B has all the substitutions reported for *Xenopus* eIF4E-1B
have been shown to reduce binding to cap analogue (Figure 3.4).

Figure 3.4: Alignment of human eIF4E-1A, eIF4E-1B, and zebrafish eIF4E-1B

These include a Met to Ser/Thr substitution at the position equivalent to M8
human eIF4E-1A, negatively charged amino acid substitutions for the acidic
residues just C-terminal to W102, the Ser to Ala substitution in the eIF4E-1E
distinguishing motif “AHAD”, and the Leu to Thr substitution at the position
equivalent to T211 in human eIF4E-1A (31). The substitution of serine and
arginine in eIF4E1B at positions corresponding to glutamine and lysine in hu
eIF4E-1A may directly influence the position of Trp102 (involved in cap-bind
modifying the stacking interaction with the cap. Similarly, the substitution of
Ala199 for Ser may induce changes in the orientation of the indole ring of Tr
by influencing the position of His200 located close to Trp102 in the 3-D struc
Replacement of Thr in position 210 and 211 by Leu and Ser in zebrafish eIF
1B is also likely to be important, because they are located in the C-terminal
responsible for binding the phosphate chain and second cap nucleoside.

3.4.1c. Zebrafish eIF4E-3

eIF4E-3 deviates from the Class I translational initiation factors by the
substitution of cysteine at the position equivalent to W56 (Figure 3.5).

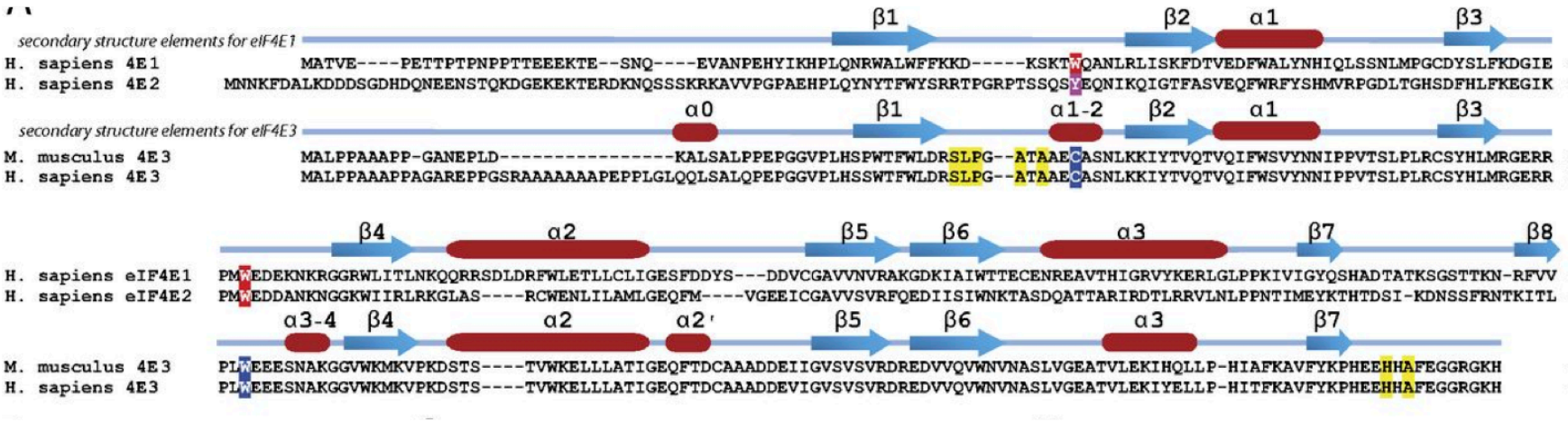


Figure 3.5: Amino acid sequence and secondary structure of eIF4E-3: showing important residues for cap-binding highlighted in yellow (from 42).

eIF4E-3 is capable of binding to cap, though its affinity is lower than that of Class I eIF4Es (42). Although mouse eIF4E-1A is dependent on the tryptophans at W56 and W102 for optimal π-stacking, and W166 to recognize the methyl^{7}GTP moiety of the cap, structural analysis of human eIF4E-3 implicates the residues C52 and W98 (equivalent to W56 and W102 of human eIF4E-1A) as playing important roles in eIF4E-3 cap-binding. In human eIF4E-3 the amino acid C52 is the residue equivalent to W56 in human eIF4E-1A. C52 forms part of a helix in the S1–S2 loop (designated α1–2) in both the apo (unbound) and m^{7}GDP forms of human eIF4E-3. This pre-formed helix is thought to play a key role in cap recognition since mutation of the S43, A47, A49, H194, and H197 of mouse eIF4E-3 reduces cap-binding. eIF4E-3 seems to recruit these additional contacts in order to offset the decline in binding energies due to the deficiency of the second aromatic residue, the Trp to Cys substitution and associated π-packing (42). Zebrafish eIF4E-3 has been verified to have all the signature residues described for the binding of human eIF4E-3 to cap analogue terminus (Figure 3.5) (42).

3.4.2. Phylogenetic analysis of deuterostome Class I eIF4Es

To investigate the origin of eIF4E-1B and eIF4E-1C, a phylogenetic analysis was undertaken of deuterostome Class I eIF4E family members. The tree subsequently constructed (see Chapter 2, Figure 2.3) indicates that each Class I sub-type of eIF4E comprises a unique clade. Ancestral members of the vertebrates; the protochordate tunicate *Ciona intestinalis*, the cephalochordate lancelet, *Branchiostoma floridae*, and the echinoderm sea urchin

Stongylocentrotus purpuratus, have only one Class I eIF4E cognate. These eIF4Es form a distinct clade outside of the eIF4E-1A,-1B, and -1C designations. The eIF4E of lamprey, *Petromyzon marinus*, appears to reside on a separate branch, closest to the eIF4E-1A clade. The Class I eIF4E cognates, eIF4E-1B, and eIF4E-1C are thought to have arisen from one or more whole genome duplications. Because eIF4E-1C is found in teleosts and not in tetrapods, our original supposition was that eIF4E-1C arose as the result of the TGD. However, with the recent availability of the genomes of many fish, it has become clear that the elephant shark, *Callorhinchus milii*, and the coelacanth, *Latimeria chalumnae* , have all three Class I eIF4Es. This implies that the duplications must have occurred prior to the branching of the chondrichthyes. However, while all teleosts have retained eIF4E-1C, it has been lost in tetrapods.

The phylogenetic analysis of vertebrate eIF4E-1Bs can be seen in Chapter 2, Figure 2.3. The presence of eIF4E-1B in the elephant shark suggests an early origin in gnathostomes. Inspection of the teleost genomes available have uncovered eIF4E-1B in basal ray-finned fish, such as speckled gar, *L. oculatus,* as well as in zebrafish, and rainbow trout, *O. mykiss*. eIF4E-1B has not been found in the genomes of more recently evolved fish such as the three-spined stickleback, *G. aculeatus,* and pufferfish, *T. nigroviridis*. Furthermore, it was previously reported that eIF4E-1B in zebrafish is not orthologous to the tetrapod form because the locus is not conserved (30). Interestingly, in *L. oculatus*, the *e-Eif4e1b* locus is the same as that found in the tetrapods. Since convergent evolution seems an improbable explanation of this, it is possible that multiple

Eif4e1b-like loci existed in the common ancestor of Actinopterygii and Tetrapoda. *Eif4e1b* genes may have been asymmetrically retained in Actinopterygii, such as the *Eif4e1b* locus in zebrafish, while other ancestral *Eif4e1b* genes gave rise to *Eif4e1b* of Tetrapoda (30).

3.4.3. Expression and quantitation of eIF4E-1A, -1B, -1C and eIF4E-3 in zebrafish tissues, ZFL cells and early embryos

It was anticipated that analysis of the levels and distribution of eIF4E family members in cultured cells, different tissues, and developmental stages would be indicative of the relative importance of each form and could assist in directing the functional analyses of each. In particular, the spatio-temporal patterns of expression could indicate whether increased/decreased expression of one form of eIF4E is linked to a particular differentiated state or developmental event. Analysis of the expression patterns of zebrafish eIF4E-1A and eIF4E-1B by end-point RT-PCR had previously shown that eIF4E-1A transcript is expressed ubiquitously, but eIF4E-1B is expressed only in muscle, ovary, and testis and in embryos up to the 21-somite stage of development (26). Coupled with the failure of eIF4E-1B to function in several eIF4E-1-specific assay systems, this pointed to a tissue/developmental stage-specific regulatory role. This was later confirmed by the findings of the Standart lab (27, 28, 32).

3.4.3a) Transcript levels

To assess where the eIF4E family members are expressed during early development and in different tissues, we looked at transcript levels of each in a variety of tissues using RT-qPCR (Figures 3.6A).

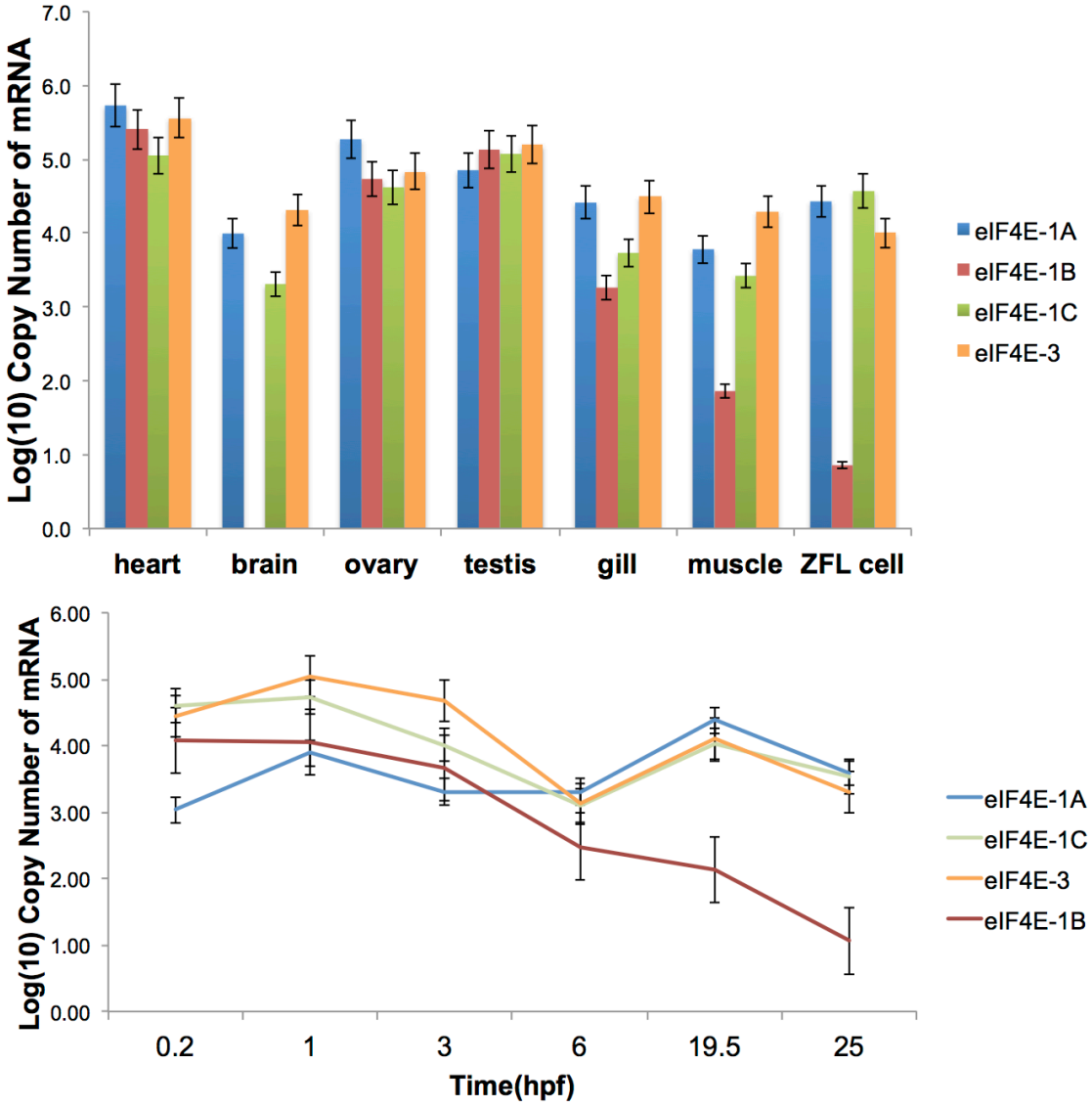


Figure 3.6: Transcript levels of eIF4E family members in adult tissues and early developmental stages: Top panel (A)Transcript levels from adult zebrafish tissues (top) and in embryos at various times post-fertilization (bottom panel)(B) were determined by RT-qPCR using cDNA generated from 25 ng RNA.

eIF4E-1A, eIF4E-1C, and eIF4E-3 transcripts were detected in all tissues examined. Except for eIF4E-1B, transcript levels for all eIF4Es ranged from 10^3 to 10^6 copies per 25 ng RNA, with the highest transcript levels of all six eIF4Es in heart, ovary, and testis (Table 3.4).

Table 3.4: Transcript levels of zebrafish eIF4E family members (copy number of mRNA per 25 ng RNA)

Tissue	**eIF4E-1A**	**eIF4E-1B**	**eIF4E-1C**	**eIF4E-3**
heart	5×10^5	2.5×10^5	1.25×10^5	3.98×10^5
brain	1×10^4	0.0	1.9×10^3	1.99×10^4
ovary	1.99×10^5	5×10^4	3.9×10^4	6.3×10^4
testis	7.9×10^4	1.25×10^5	1.25×10^5	1.58×10^5
gill	2.5×10^4	1.99×10^3	5×10^3	3.16×10^4
muscle	6.3×10^3	79	2.5×10^3	1.99×10^4
ZFL cell	1.1×10^3	0.81	3.9×10^4	1×10^4
Embryo hpf	**eIF4E-1A**	**eIF4E-1B**	**eIF4E-1C**	**eIF4E-3**
0.2	1.3×10^3	1.25×10^4	3.9×10^4	2.75×10^4
1.0	7.76×10^3	1.12×10^5	5.37×10^4	1.12×10^5
3.0	1.99×10^3	4.57×10^3	1.02×10^4	4.67×10^4
6.0	1.99×10^3	3×10^2	1.3×10^3	1.3×10^3
19.5	2.4×10^4	134	1×10^4	1.25×10^4
25	3.9×10^4	91	3.4×10^3	1.95×10^3

In the adult liver cell line, ZFL, trancript levels for all eIF4Es, except for eIF4E-1B, ranged from 1×10^4 to 3×10^4 copies/25 ng RNA. This suggests there is no real tissue specific expression, except for eIF4E-1B. eIF4E-1B transcript levels vary dramatically between tissue types. By endpoint PCR, eIF4E-1B is only seen in ovary, testis, muscle, and heart. In RT-qPCR, it is seen in most tissues except for brain. Transcript levels for eIF4E-1B were highest in ovary, testis, and heart tissue (2.5×10^5 copies/25ng RNA), though lowest in ZFL cells (7.2×10^0 copies/25ng RNA), and below detection limits in brain. Transcript levels of the zebrafish eIF4Es were also determined for embryos at different developmental

stages (Figure 3.6B). eIF4E-1C and eIF4E-3 transcripts were detected at the highest levels shortly after fertilization after which they decline through the maternal-zygotic transition(MZT). eIF4E-1A mRNA transcripts were detected at the lowest level of expression from the zygote (0.2 hpf) to the gastrula (6 hpf). Transcript levels for eIF4E-1A, -1C, and eIF4E-3 were lowest shortly after the MZT, and then began to increase. eIF4E-1B transcripts levels steadily declined post fertilization. The overall conclusions of the expression patterns of eIF4E family members suggest that eIF4E-1A and eIF4E-1C are ubiquitously expressed, as expected for an essential translation factor with some cell type-specific modulation of expression of eIF4E-1A, -1B, -1C, and -3.

3.4.3b) Protein levels: eIF4E-1A and -1C are ubiquitously expressed

In order to assess the extent of eIF4E protein expression in tissues, antibodies were custom developed. Signature peptide sequences can be identified in the N-terminus. eIF4E-1A,-1B,- and -1C antibodies derived from this region and confirmed for antigenic specificity by using recombinant proteins prior to tissue

analysis (Figure 3.7).

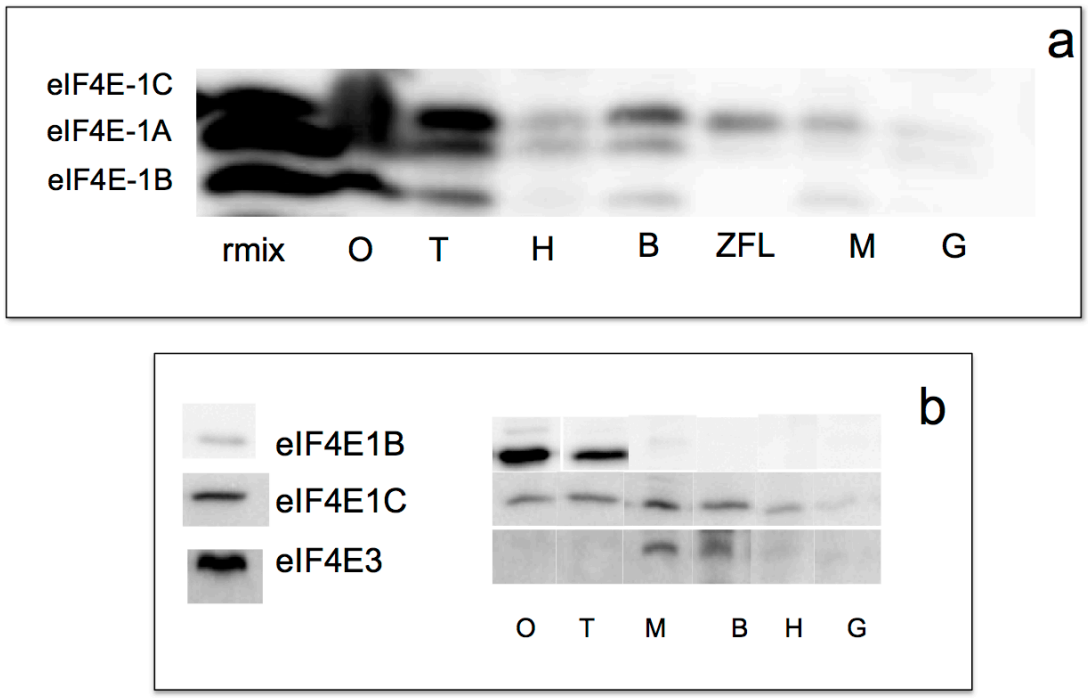


Figure 3.7: Expression of eIF4E class I and III in zebrafish tissues: (A) Extracts from adult zebrafish tissues and ZFL cells, were subjected to high-Tris SDS-PAGE at 50 V for 17 h prior to being transferred to PVDF membranes and probed with antibody to eIF4E-1A C-terminus. (B) Expression observed using the specific eIF4E antibodies developed to eIF4E-1C, eIF4E-1B, and eIF4E-3 in zebrafish tissues. The tissues represented by letter are testis (T), ovary (O), heart (H), muscle (M), gill (G), brain (B) and ZFL cells (Z).

The antibody to eIF4E-1A has significantly lower in avidity. Fortuitously, the C-terminus derived antibody for eIF4E-1A also recognized eIF4E-1A and eIF4E-1B, because of that high sequence identity, and could be used to assess the levels of all three Class I eIF4E simultaneously. When this C-terminal antibody was used in parallel studies with the N-terminal antibody, the results indicated that the eIF4E-1A and eIF4E-1C were prevalent in most tissues types (Figure 3.7a and 3.7b). eIF4E-1A and -1C display variable expression across different tissues, but the eIF4E-1C protein was expressed across all the tissues (Figure 3.7b). In ZFL cells, only the expression of eIF4E-1C could be detected consistently at higher

levels (data not shown). eIF4E-1A and -1B previously were shown to be expressed in three tissues; ovary, testis ,and muscle (26). Clear expression of eIF4E-3 was observed in brain and muscle. On the basis of this, ovary, muscle and brain were selected for quantification purposes. Determination of relative levels of expression of each eIF4E across the selected tissues was achieved by immunoblot analysis using the Alphaimager™ program. The resultant saturated pixel density values were normalized against an eIF4E standard (Figure 3.8).

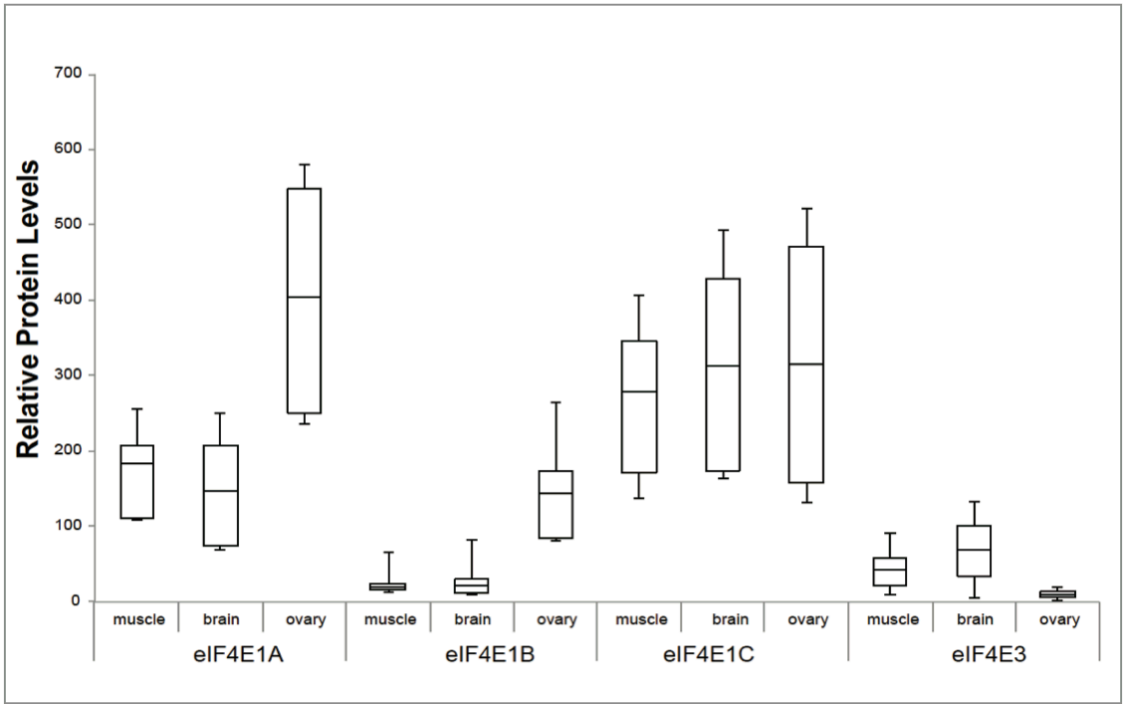


Figure 3.8: Quantification of zebrafish eIF4E expression in muscle, brain, and ovary tissue: Avidity of antibody was assessed using dilutions of eIF4E recombinant protein dilutions on gel and then for relative levels by use of pixel saturation intensity from Alpha imager. Samples of 2 to 6 representative blots were analyzed for creation of a boxplot. The minima/maximal values are bracketed

By comparing a selection of samples, the boxplot median values indicate that levels of eIF4E-1A and eIF4E-1C are considerably higher than other eIF4E family members. eIF4E-1A displayed the highest expression level in ovary. eIF4E-1C levels were higher than the eIF4E-1A in muscle and brain. In ovary, eIF4E-1B

levels were approximately a third the level of eIF4E-1A and approximately 50 % of the level of eIF4E-1C. Low levels of eIF4E-1B could be seen in skeletal muscle and brain.

3.4.4. eIF4E3 is expressed in muscle and brain of adult zebrafish

In mouse, eIF4E3 transcripts have been reported in skeletal muscle, lung, and heart using a Northern blot analysis (39). In this current study, western analysis confirmed that eIF4E-3 is observed at the level of protein in skeletal muscle and heart tissue. The highest levels of eIF4E-3 are seen in brain, although at only 20 % of the level of eIF4E-1C and at approximately 30 % the level of eIF4E-1A. A recent proteomic analysis of zebrafish has supported this result by confirming that eIF4E-3 is present in brain tissue, although other tissues tested showed negligible expression levels for eIF4E3 (134).

3.4.5. eIF4E shows increased expression across zebrafish embryonic development

After assessment of the eIF4E transcript levels of eIF4E in embryos, westerns blots were performed to analyze Class I and III eIF4E family members during early development. A stepwise methodology was employed to remove the chorion and de-yolk the samples (135). In particular, the removal of the vitellogenin fraction was critical, since its presence obscures the eIF4Es due to

the similarity in molecular weight. Figure 3.9

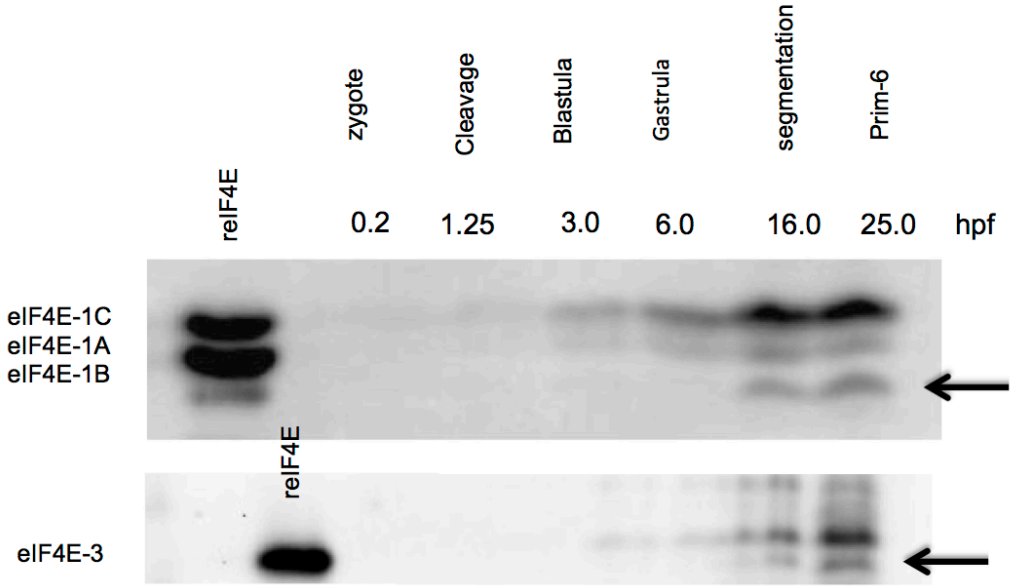


Figure 3.9: Expression of eIF4E in zebrafish embryonic development: Embryos were collected at the indicated hour post fertilization (hpf); assessed microscopically, flash frozen, and protein extracts prepared as described. Proteins were fractionated by high-Tris SDS-PAGE for 2 h at 200 V. After transfer to PVDF membrane, proteins were probed using indicated specific antibody to eIF4E-1A C terminus and the eIF4E-3 C-terminal respectively.

shows eIF4E family member expression from 0.2 (zygote) to 25 hpf (prim6) comparing equivalent numbers of embryos. Unfortunately, protein recovery was poor particularly at early time points and protein loaded increased from 0.2 -16 hpf. eIF4E-1A,-1C,-1B, and -3 can be seen at 3-6 hpf (blastula). The levels of eIF4E-1A and eIF4E-1C are expressed proportionally from 3-25 hpf, although levels of eIF4E-1C are consistently higher. Although eIF4E-1B transcript levels fall continuously during early development, with lowest levels at 25 hpf, eIF4E-1B protein levels begin to increase between 16-25 hpf, coincident with somitogenesis. eIF4E-1B transcript and protein levels are not coordinately regulated, suggesting regulated mRNA recruitment or protein turnover, or both.

3.4.6. Zebrafish eIF4E-1C, but not eIF4E-3, is functionally equivalent to human eIF4E-1

Although there is considerable sequence divergence between human eIF4E-1 and *S. cerevisiae* eIF4E (31 % identity), the mammalian factor can sustain growth of yeast deficient in eIF4E. The previously developed yeast strain, JOS003 (88), was used to compare the functionality of eIF4E-1C and eIF4E-3 with eIF4E-1A and eIF4E-1B. The JOS003 strain lacks the endogenous yeast *eIF4E* gene and expresses human eIF4E-1 inserted in the pRS415 leu(-) vector under the control of the galactose-dependent and glucose-repressible GAL1 promoter. As a consequence, strain JOS003 is able to survive in medium containing galactose as carbon source but is not viable in medium containing glucose due to depletion of the human eIF4E-1. Growth of JOS003 in glucose can be mediated by ectopic expression of a functional eIF4E, the regulation of which is under the control of a promoter in the pRS416 ura(-) vector, which is active in the presence of glucose. The cDNAs encoding the zebrafish eIF4E cognates were cloned into pRS416, allowing expression from the constitutively active glyceraldehyde-3-phosphate dehydrogenase (GPD) promoter. Following transfection and selection on media lacking uracil, the yeast cells containing control vector, or vectors for the expression of eIF4E-1A, -1B, -1C or eIF4E-3, were streaked on selective plates; Synthetic medium (SC) –Ura, -Leu containing

either galactose or glucose as carbon source (Figure 3.10).

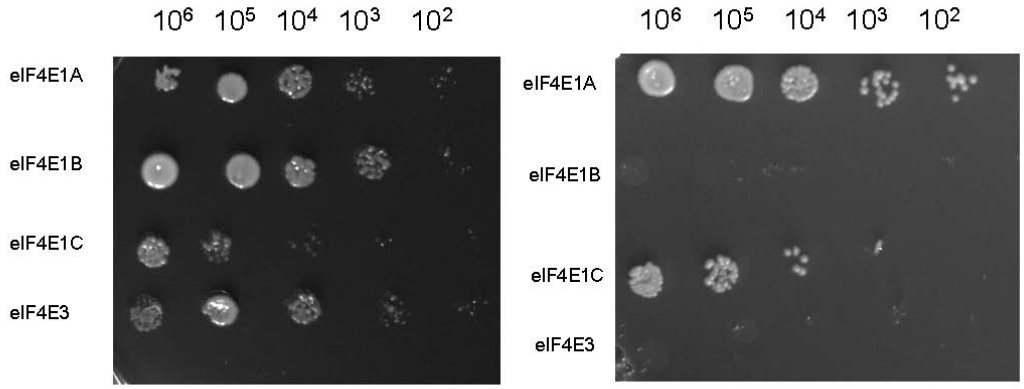


Figure 3.10: Ability of zebrafish eIF4Es to rescue the growth of *S. cerevisiae*, JOS003: *The S. cerevisiae* strain, JOS003, (88) was transformed with the Ura-selectable vector, pRS416GPD, containing cDNAs encoding one of the following products: eIF4E-1A, eIF4E-1B, eIF4E-1C and eIF4E-3, as indicated. Following selection on SC medium with galactose lacking uracil and leucine, yeast from the resulting single colonies were diluted 10^{-1} to 10^{-7} fold and transferred onto YP-agar media containing G418 and either glucose (*left*) or galactose (*right*). Growth was assessed after 48 h.

As previously reported, eIF4E-1A is capable of complementation, while eIF4E-1B is not (26). It is evident that eIF4E-1C, but not eIF4E-3 is able to rescue the JOS003 strain under conditions in which human eIF4E-1 is depleted. Expression of each eIF4E was verified by immunoblot analysis using antibodies specific to each eIF4E (results not shown). These results demonstrate that zebrafish eIF4E-1C is functionally equivalent to a tetrapod prototypical Class I eIF4Es.

3.4.7. eIF4E-1A and eIFE-1C and eIF4E-3 bind to m^{7}GTP cap analogue

Recombinant eIF4Es proteins were synthesized via production of ^{35}S-radiolabeled proteins translated *in vitro* and the relevant pCITE4a constructs were used as templates in the rabbit reticulocyte-coupled transcription-translation system, containing [^{35}S]Met , essentially as described previously (26). The resultant pools were mixed with m^{7}GTP-Sepharose bead slurry, and the total, unbound and bound fractions were analyzed by SDS-PAGE and immunoblotting

(Figure 3.11).

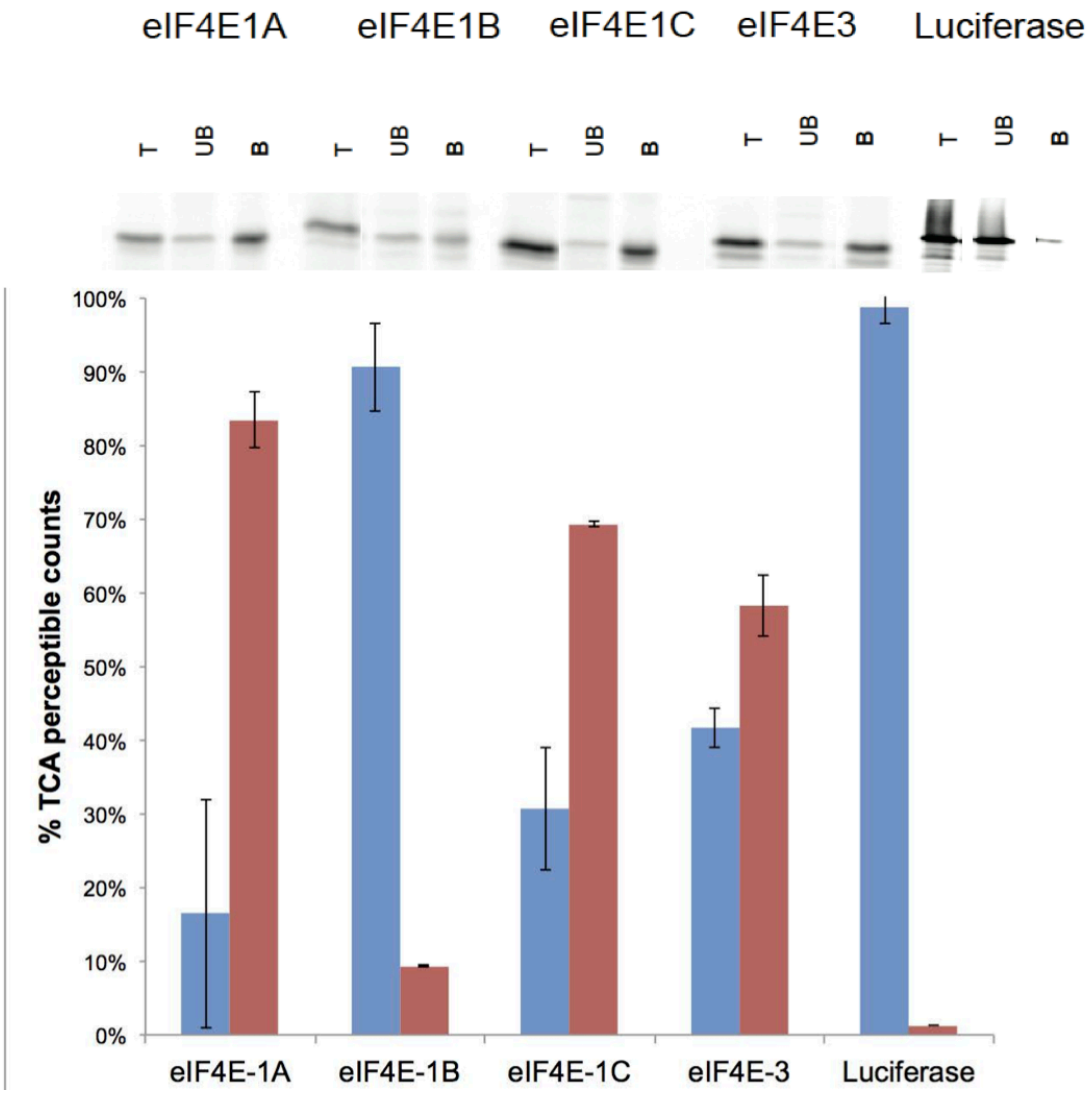


Figure 3.11: m^{7}GTP binding activity of zebrafish Class I and III eIF4Es: eIF4Es were translated in the reticulocyte cell-free translation system, in the presence of [^{35}S]Met. The proteins were bound to m^{7}GTP-Sepharose beads and analyzed by high-Tris SDS-PAGE and autoradiography. Each sample is equivalent to an equal volume of the translation reaction. The total (T), unbound (U), and bead bound fractions (B) are labeled as indicated. Trichloracetic acid precipitated samples of each fraction were analyzed by scintillation cpm counts and represented as bound versus unbound expressed as a percentage of the incorporated cpm total. The bound fraction (Red) and unbound(Blue).The binding of ^{35}S Met luciferase was included as a negative control. The proteins were also analyzed by high-Tris SDS-PAGE.

For the investigation of native eIF4E, ovary extract was bound to m^7GTP-Sepharose, followed by extensive washing and elution with excess m^7GTP. Eluted proteins were resolved by high-Tris SDS-PAGE prior to immunoblotting and visualization of eIF4E-1A, -1B, and -1C. The ovary extract contains eIF4E-1A, eIF4E-1B and eIF4E-1C, but not eIF4E-3. eIF4E-1A and eIF4E-1C, but not eIF4E-1B, bound to the m^7GTP-matrix and was specifically eluted with m^7GTP (Figure 3.12).

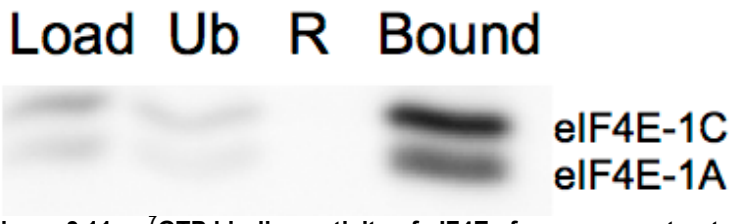


Figure 3.11: m^7GTP binding activity of eIF4Es from ovary extract

A previous report from the Jagus laboratory claimed that only eIF4E-1A from ovary bound to m^7GTP-Sepharose (26). However, at that time, eIF4E-1C had not been identified. The antibody used cross-reacts with eIF4E-1A and -1C. Only one protein was observed, but the gel electrophoresis conditions used would not have separated eIF4E-1A and -1C.

3.4.8. eIF4E-1A, -1C and eIF4E-3 interact with zebrafish eIF4GI *in vitro*

The platform protein, eIF4G, binds to eIF4E and will compete with the 4E binding proteins (4E-BPs) for a common binding site -YXXXXLϕ) located within the conserved core region of the eIF4Es (19, 65, 136-139). A polypeptide corresponding to residues 262-681 of zebrafish eIF4GI (molecular mass ~45 kDa), which brackets the eIF4E-1 interaction domain, was co-translated with S-

tagged variants of eIF4Es in a reticulocyte cell free translation system in the presence of ^{35}S Met (Figure 3.13).

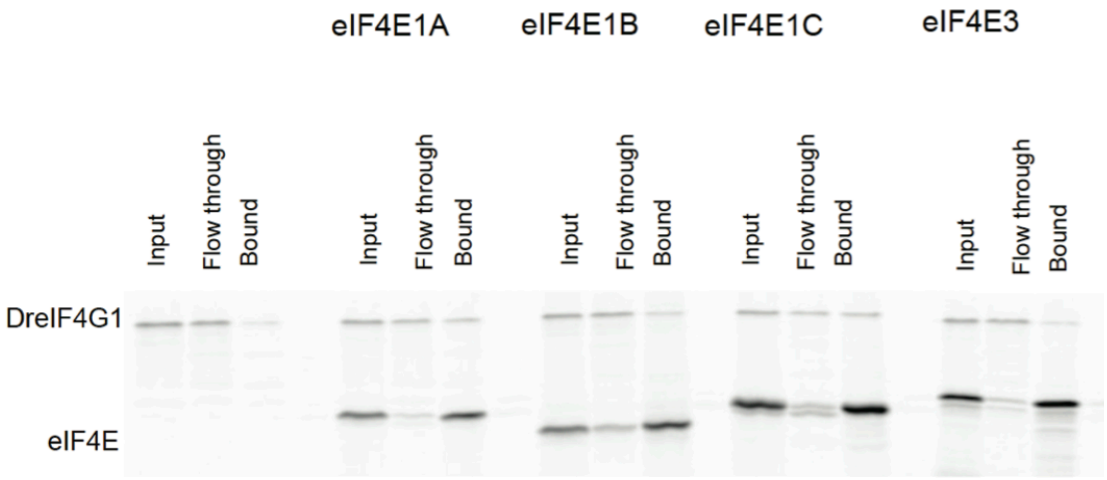


Figure 3.13: Interaction of zebrafish eIF4G with Class I and III eIF4Es: The Mwt(kDa) of eIF4E is ~24-27 and eIF4G is ~51 kDa*(eIF4G migrates as ~ 100 kDa under the SDS page conditions (39)).

Reaction mixes were incubated with S-protein-agarose. Following extensive washing, all proteins, which bound to the matrix, were eluted with SDS-PAGE sample buffer. Fractions were resolved by high-Tris SDS-PAGE and analyzed by Typhoon Storm imaging. Whereas zebrafish the eIF4GI fragment co-purified with eIF4E-1A and eIF4E-1C, eIF4E-1B failed to interact with the same polypeptide. eIF4E-3 is bound to eIF4G, but more weakly than eIF4E-1A and -1C. These data, coupled with the yeast complementation data, confirms the findings of the original study that zebrafish eIF4E-1A is able to interact with human eIF4GI *in vitro* and with yeast eIF4G *in vivo*. eIF4E-1C was also observed to bind zebrafish eIF4G1 supporting its role as a translation initiation factor. In contrast, eIF4E-1B has a low affinity for both human and zebrafish eIF4GI and thus is unlikely to function as an efficient *in vivo* competitor of eIF4E-1A or eIF4E-1C.

3.4.9. eIF4E-1A and eIF4E-1C interact with the zebrafish 4E-BPs

The binding partners of eIF4E are designated 4E binding proteins, 4E-BPs, and act to regulate translation through phosphorylation and the mTOR pathways (21, 65, 136, 137, 140-143). The 4E-BPs bind to eIF4E through common motifs and it is anticipated that zebrafish eIF4E-1A and eIF4E-1C, but not eIF4E-1B or eIF4E-3, would be targeted by the 4E-BP repressors. However, there remained the possibility that a homologue of eIF4E that is deficient in both cap-binding activity and eIF4G interaction could potentially bind to 4E-BPs and work as a translational de-repressor. There are four variants of the zebrafish 4E binding proteins, which are designated as 1,2,3 and 3-like (26). In GenBank, the designation of 4E-BP1, -2 and -3 is a consistent nomenclature across the mammalian systems, but only the Actinopterygii appears to have an additional 4E-BP3-like type. To assess the functionality of the cloned zebrafish 4E-BPs, *in vitro* interaction assays with S-tagged variants of eIF4E-1A, -1B, -1C, eIF4E-3 was performed using the same bead binding strategy as for the eIF4E/eIF4G interaction assay previously described in Section 3.4.8. After synthesis, reactions were incubated with S-protein-agarose. Following extensive washing, proteins bound to the matrix were eluted with SDS-PAGE sample buffer (Figure 3.13).

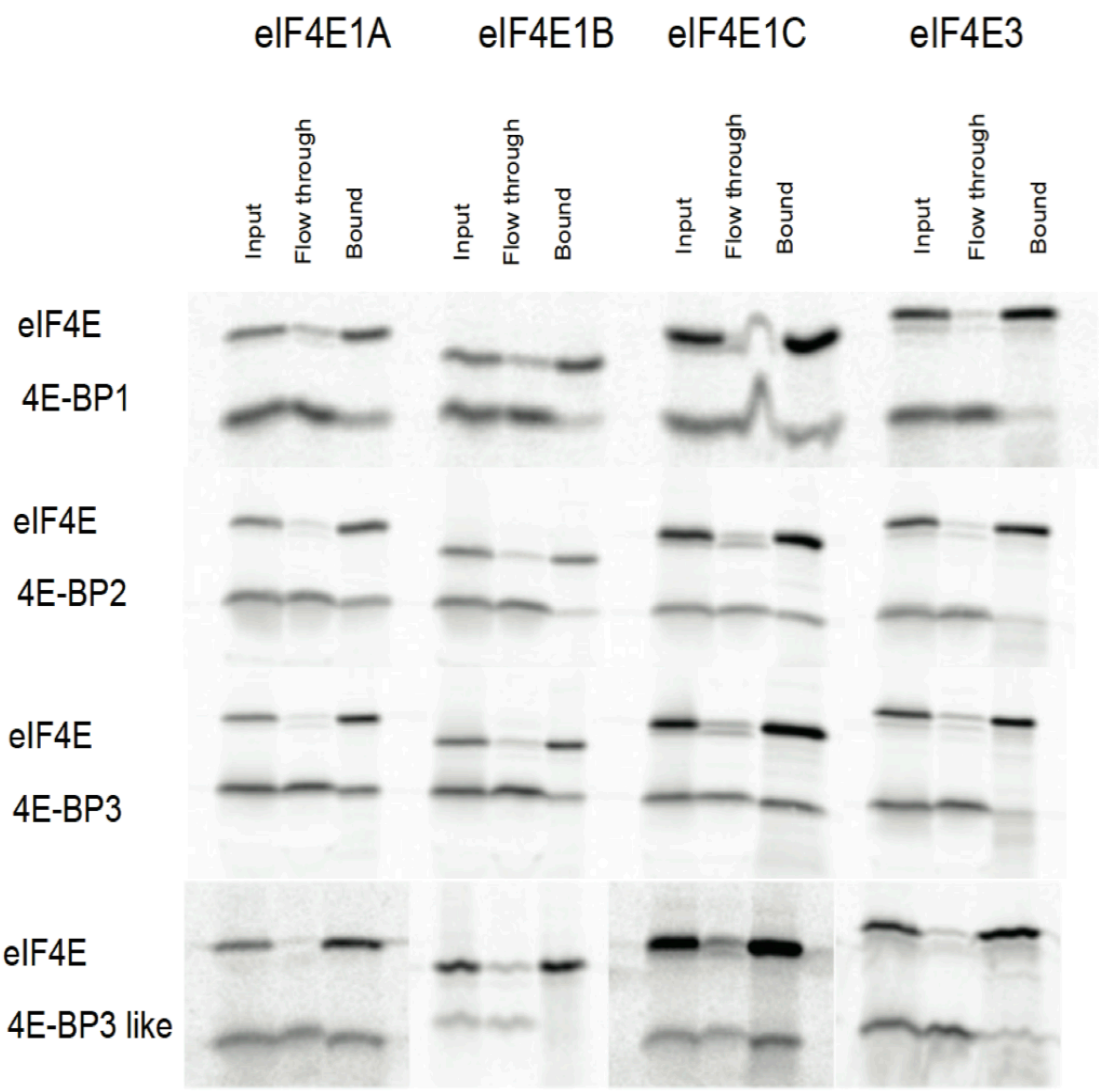


Figure 3.14: Interaction of zebrafish 4E-BPs with zebrafish Class I and III eIF4Es: The Mwt (kDa) of the 4E-BPs is ~15 and eIF4Es are~24-27.

The data showed that zebrafish 4E-BP was enriched in the fraction of bound proteins in the presence of eIF4E-1A and eIF4E-1C, but not eIF4E-1B or eIF4E-3. This evidence supports the conservation of the 4E-BP mediated translational repression pathway in zebrafish. However, neither the eIF4E-1B nor the eIF4E3 binds to any zebrafish 4E-BP variant with an affinity that would be consistent with a role as a de-repressor of 4E-BP-mediated inhibition of translation.

3.5. Discussion

This present study has focused on the unique ray finned species, *Danio rerio,* from the teleost superfamily Ostariophysi, whose eIF4E members include three Class I eIF4Es (eIF4E-1A, -1B, -1C), two Class II eIF4Es (eIF4E-2A, eIF4E-2B) and a single Class 3 eIF4E (eIF4E-3). There are currently twenty fish genomes available at NCBI, http://www.ncbi.nlm.nih.gov/genome/annotation_euk/all/)and many more in the pipeline for annotation. Analysis of eIF4E sequences from these twenty species, as well as the echinoderm, tunicate and cephalocordate sequences has allowed a glimpse of origins and evolution of the eIF4E family. In particular, access to protein sequences generated by genomic annotation from genomes of the coelacanth, elephant shark, lamprey and basal ray-finned fish, has provided a means to speculate on when the duplications occurred. The distribution of the subclasses of eIF4E1 is consistent with the duplication of Class I prior to the teleost specific whole genome duplication, so probably one of the whole genome duplications thought to have occurred at ~500 (2R) mya and 550 (1R) mya. Although there is some uncertainty on whether these duplications occurred before or after the separation of agnathans and gnathostomes, Kuraku and colleagues have suggested that the data favor the scenario whereby both the 1R and 2R WGD events occurred prior to the lamprey-gnathostome split, based on analysis of selected families of gene duplicates, (144, 145). This scenario would predict that lamprey should also have eIF4E-1A, -1B, and -1C. However, it seems that lamprey have thrown out more and different duplications than the gnathostomes (97, 144-146).

The diversification of the Class I eIF4Es are interesting because the product of one gene, eIF4E-1B, has neofunctionalized to become a tissue specific regulator of mRNA recruitment. The other, eIF4E-1C, appears to have retained function as a prototypical initiation factor. In view of the fact that eIF4E-1A and -1C have been conserved for 500-550 mya, it would seem likely that some subfunctionalization has occurred but was not apparent in the studies here. Only further work with zebrafish themselves is likely to shed light on this. The question that arises is whether both are essential or whether either one alone can support normal growth and development in zebrafish. eIF4E-1A is prevalent across deuterostomes from echinoderms to mammals. eIF4E-1C is first seen shark and retained in basal ray-finned fish, teleosts and coelacanth. However, eIF4E-1C is lost in tetrapods. eIF4E-1B is a chordate specific eIF4E, although eIF4E family members with convergent characteristics have been found in Drosophila. eIF4E-1B is also first seen in shark and is retained in basal ray-finned fish, lower teleosts, and tetrapods, but has apparently been lost in higher spiny ray fish known as the percomorph teleosts. It will be of interest to determine how the recruitment of CPE-containing mRNAs is regulated during meiosis in these fish. eIF4E-3 is the most conserved of the eIF4E classes. Only one form of eIF4E-3 had been discussed in the literature from primarily tetrapod research. My current analysis of gene loci has revealed that a cognate protein that is referred to as eIF4E-3-like appears in the percomorph teleosts. The function of eIF4E-3 is still uncertain. In mice, it has been shown that the microRNAs, miRNA-206 and miRNA-21 are sufficient and required for muscle wasting during catabolic

conditions (43). *In silico* and *in vivo* approaches have identified transcription factor YY1 and the translational initiator factor eIF4E3 as downstream targets of these miRNAs. This suggests that eIF4E-3 is involved in muscle protein synthesis. Conversely, eIF4E-3 has been reported to suppress translation of a subgroup of mRNAs associated with oncogenesis including VEGF, c-myc and cyclin D1 in mouse NIH3T3 cells (42, 147). This implies that eIF4E3 is not involved in forming active translation complexes but rather forms inactive complexes sequestering the mRNA away from the active translation machinery. The knockout of this gene in zebrafish should allow for the study of eIF4E-3 function in muscle development and growth, as well as its role in mRNA recruitment.

This is the first description of the functional characteristics and expression of zebrafish eIF4E-1C and eIF4E-3 ,and will provide the basis for ongoing studies of their roles in the translational regulation of gene expression in the zebrafish.

Chapter 4: Class II eIF4E Family Members in Zebrafish (*Danio rerio*): Neofunctionalization of eIF4E-2B

4.1. Abstract

The translation initiation factor, eIF4E, is an essential component of the eukaryotic translation machinery that binds to the 5'-cap of mRNAs and promotes recruitment to the small ribosomal subunit. Prototypical eIF4E falls into Class I of the metazoan eIF4E family. In contrast, Class II eIF4E family members have been found to down-regulate the translation of specific mRNAs by tethering the 5' and 3' ends and preventing the interaction of the translation factor eIF4E and eIF4G with the 5'-cap structure. The zebrafish, *Danio rerio*, has two Class II eIF4Es, designated eIF4E-2A and eIF4E-2B. eIF4E-2A is found across vertebrates, but the cognate protein, eIF4E-2B, is only seen in basal ray-finned fish, teleosts and the amphibian genus *Xenopus.* The genes located in close proximity to the eIF4E-2A locus appear to be conserved across teleosts and tetrapods, but the eIF4E-2B genetic loci are more variable. This suggests that eIF4E-2A is the ancestral form, whereas the eIF4E-2B cognate may have resulted from a genomic duplication event. The retention of these two cognates suggests that neofunctionalization may have occurred. Here we compare the characteristics of zebrafish eIF4E-2A and -2B. Zebrafish eIF4E-2A and -2B both bind to cap analogue, are unable to interact with zebrafish eIF4G,and bind poorly to the 4E-BPs. Zebrafish eIF4E-2B and -2A can be distinguished from eIF4E-1A by its ability to bind trimethyl GTP (TMG) and to complement a *S. cerevisiae* strain conditionally deficient in functional eIF4E.

4.2. Introduction

Prototypical eIF4E is important for its essential role in recruitment of mRNA to the small ribosomal subunit through a complex involving the poly(A) binding protein (PABP), eIF4G, eIF4A and eIF3 (reviewed, 1-5). Prototypical eIF4E begins the recruitment process by binding to the 5'-m^7Gppp cap of mRNA. The 'closed-loop' model of translation initiation hypothesizes that interactions of the cap-binding eukaryotic initiation factor eIF4E, eIF4G and PABP hold the 5' and 3' ends of mRNA in close proximity and promote recruitment of the small ribosomal subunit to the mRNA 5' end (4, 11). The anchoring of eIF4E and eIF4G to the 3'-poly(A) tail ensures that they will remain tethered to the mRNA and increase the efficiency of subsequent rounds of initiation.

Phylogenetic analysis of the translation initiation factor eIF4E is part of a family of proteins (15, 21, 22, 116, 117). Most eIF4E family members do not function as translational initiation factors, but as regulators of mRNA recruitment (15, 70). With the exception of eIF4Es from protists, all eIF4Es can be grouped into one of three classes, Class I, Class II, Class III (15). The structures of mammalian Class I (eIF4E-1A), Class II (eIF4E-2) and Class III (eIF4E-3) all show the characteristic α+β domain as resolved in NMR or crystallographic studies (16, 18, 42, 118). Class I members from Viridiplantae, Metazoa, and Fungi carry Trp residues equivalent to W43, W46, W56, W73, W102, W113, W130, and W166 of *H. sapiens* eIF4E-1**.** eIF4E sandwiches the m^7G cap via tryptophan residues, W56 and W102, and binds the consensus YXXXXL**Φ** sequence in eIF4G (in which **Φ** is hydrophobic and X is any amino acid) on its convex side (16, 18).

Prototypical eIF4Es bind eIF4G through the consensus motif S/TVE/DE/DFW in which the Trp is W73 in mouse eIF4E-1A. Other eIF4E family members show functional specialization and operate as regulators of initiation (reviewed (15, 20, 21, 70, 116, 117)).

Class II eIF4E family members, the eIF4E-2s (also called 4EHP (34)) have been shown to regulate specific mRNA recruitment in *Drosophila* (35), *C. elegans* (36) and mouse (37, 38). Class II members possess W→Y/F/L and W→Y/F substitutions relative to W43 and W56 of *H. sapiens* eIF4E, respectively (15). There is no eIF4E-2 interaction with eIF4G ,and binding to 4E-BPs is relatively weak (39, 118, 148). Mouse eIF4E-2 has a 30-fold lower affinity for the cap analogue, m^7GTP (118, 149). This means that eIF4E-2 alone, will not compete with eIF4E1 for mRNA effectively, but may do so with a partner protein. The lower affinity of mouse eIF4E-2 for m^7GTP is largely due to an extension of the loop, which creates the ligand binding site, and thus negatively affects formation of the three stacked aromatic rings, Trp124/m^7G/Tyr78. In addition, mouse eIF4E-2 has different arrangements of basic amino acids interacting with the phosphate chain of the cap (118, 149). The *Drosophila* homologue, d4EHP (eIF4E-8) binds Bicoid, an RNA-binding protein that recognizes a 3' UTR element in *caudal* mRNA to specifically repress its translation (35, 41). Similarly, in mouse, eIF4E-2 (4EHP) binds cytoplasmic Prep1 inhibiting *Hoxb4* translation (125). Recently, Morita *et al*. showed that mouse eIF4E-2, forms a translational repressor complex with Grb10-interacting GYF protein 2 (GIGYF2) and zinc finger protein 598 (38). eIF4E-2 is essential for mammalian development;

eIF4E2$^{-/-}$ mice are not viable, with the embryos dying perinatally (38). Mouse eIF4E-2 also interacts with an eIF4E-binding protein, eIF4E transporter protein (4E-T), which has been shown to inhibit cap-dependent translation (72). 4E-T is a component of processing bodies (P-bodies) and a nucleocytoplasmic protein that transports eIF4E into nuclei (150-152). P-bodies are distinct cytoplasmic foci containing mRNA, microRNAs, mRNA decay enzymes, and RNA-binding proteins/translational repressors but not ribosomes, and are understood to participate in mRNA decay and in reversible translational repression including that mediated by microRNAs (153-155).

My study focuses on a comparison of the functional characteristics and expression of zebrafish eIF4E-2A and -2B. Consistent with the retention of both cognate forms for over ~500 mya, it appears that eIF4E-2B has undergone neofunctionalization. Zebrafish eIF4E-2B can be distinguished from eIF4E-2A by its ability to bind with greater affinity to trimethyl GTP (TMG), and its ability to complement a *S. cerevisiae* strain conditionally deficient in functional eIF4E.

4.3. Materials and Methods

4.3.1. Identification of zebrafish eIF4E family members

The eIF4E-1A (GenBank mRNA AF176317.1, (cds): AAG09794.1) and eIF4E-1B (UniProtKB/Swiss-Prot: Q9PW28.1) were previously described (26). Sequence for eIF4E-2A and eIF4E-2B were deposited into GenBank as AGW99949.1 andAGW99950.1, respectively. The indicated *Danio rerio* eIF4E sequences (cDNA and/or coding region cDNA) are currently residing in GenBank and were deposited from 2000-2013 by the Jagus laboratory.

4.3.2. Identification of eIF4E family members from other deuterostomes

The amino acid sequences of eIF4E family members of the deuterostomes included in this study have been collected from on-line genomic resources including; the National Center for Biotechnology Information (NCBI) at http://www.ncbi.nlm.nih.gov, the Ensembl project (156), the HMMR database (http://hmmer.janelia.org), the Institute of Molecular and Cell Biology (IMCB) elephant shark genome (http://esharkgenome.imcb.a-star.edu.sg), the coelacanth genome project site (http://coelacanth.nig.ac.jp/index.php) and the Joint Genome Institute (JGI) for the *Branchiostoma floridea* genome (http://genome.jgi-psf.org). Each eIF4E sequence was verified using the Genbank BLAST tool and aligned by the MUSCLE algorithm applying the suite of software provided by CLC workbench (CLCBio CLC Genomics Workbench 7.0.3 (http://www.clcbio.com). Accession numbers and details on sequences are provided in the Appendix (Appendix Table A2.1).

4.3.3. Generation of cDNAs encoding zebrafish eIF4E family members

cDNAs encoding zebrafish eIF4E-1C and eIF4E-3 were amplified by RT-PCR from RNA from ZFL cells and cloned into the *in vitro* transcription/translation plasmid vector pCITE-4a(+) (Novagen, EMD-Millipore, Billerica, MA, USA), using engineered NcoI and BamHI sites (primers listed in Chapter 3,Table 3.1). From here they were transferred to other vectors such as pET11d and the yeast pRS416GPD. Because eIF4E-2B gave such an unexpected result in complementing a yeast strain conditionally deficient in eIF4E, the eIF4E-2B

cDNA was codon optimized for yeast, synthesized and cloned into pRS416GPD by GenScript (Piscataway, NJ, USA).

4.3.4. Generation of constructs encoding zebrafish 4E-BPs and fragment of zebrafish eIF4GI

The constructs for zebrafish 4E-BP3-like were described previously (26). Nucleotide sequences for zebrafish 4E-BPs 4E-BP1 (NP_955939.1) 4E-BP2 (NP_997968.1) 4E-BP3 (NP_001007355.1) were codon optimized for rabbit, *Oryctolagus cuniculus*, using Advanced OptimumGene™ (Genscript, Piscataway, NJ, USA). The nucleotide sequence was synthesized by Genscript, augmented with additional methionine residues and cloned into the *in vitro* transcription/translation plasmid vector pCITE-4a(+) (Novagen(EMD-Millipore) Billerica, MA, USA), using the NcoI and BamHI sites. This cloning strategy adds an S-tag to the amino-terminus and includes a stop codon at the carboxy terminus. The nucleotide sequences for the zebrafish eIF4G-1 fragment aa 262-681 were codon optimized for rabbit, generated and cloned into pCITE4a(+).

4.3.5. RNA purification, cDNA synthesis, RT-PCR and RT-qPCR

Fresh tissue, embryos, or harvested cells was homogenized by bead beating and extracted via kit Purelink RNA minikit: (Ambion™ Grand Island, NY, USA). RNA was quantified on a Nanodrop 1000 (Thermo Fisher by Life Technologies Waltham, MA). Values of >2 for 260/280 and 260/230 ratios were considered to be of sufficient purity. RNA was reverse-transcribed using Superscript II reverse transcriptase (Invitrogen, Life Technologies, Carlsbad, CA, USA) with random hexamers in accordance with the manufacturer's instructions. Generated cDNA

was used as template for RT-PCR and RT-qPCR. The quality of all cDNA preparations was assessed by end point PCR amplification. Amplification reactions (primers listed in Chapter 3,Table 3.1A) were performed under standard conditions using Taq (Denville Scientific Inc, South Plainfield, NJ, USA) DNA polymerases. Whenever purified PCR products were transferred into plasmids, insertions were sequenced in both orientations to ensure that no errors had been introduced due to amplification. The products were resolved by TAE-agarose electrophoresis and recorded in a fluorimager (Amersham Biosciences, Pittsburgh, PA, USA).

Primers for qPCR were designed by PearlPrimer and Primer 3 software to span exon-exon junctions (listed in Chapter 3,Table 3.1B). For RT-qPCR using an Applied Biosystems (Life Technologies) Fast 7500 thermal cycler, cDNA from 20 ng RNA was amplified using Taqman Fast Universal PCR Mastermix (no AmpErase UNG) (Applied Biosystems Foster, CA, USA). Thermal cycling conditions consisted of an initial denaturation at 95 °C for 2 min followed by 40 cycles of denaturation at 95 °C for 15 sec, annealing and fluorescent data collection at 60 °C for 15 sec, and extension at 72 °C for 30 sec. The reaction was completed with a melt curve to determine the presence of spurious PCR products. Cycle thresholds and baselines were determined manually and quantities were normalized by absolute quantification using linearized plasmid DNA.

4.3.6. Preparation of protein extracts from cultured cells and tissues

The cultured cells or tissues were homogenized in up to 10 volumes of ice-cold buffer containing 25 mM Tris -HCl, pH 7.4, 1 mM EDTA, 1 mM EGTA, 1 mM dithiothreitol, 100 mM KCl, 0.5 % Elugent, and Complete^{TM} Protease Inhibitors (Roche Applied Science, Madison WI, USA). Homogenates were clarified by centrifugation (15 k x g, 4 °C, for 15 min). Supernatants were frozen and stored in liquid N_2.

4.3.7. *In vitro* transcription and translation

^{35}S-radiolabeled proteins were translated *in vitro*, using pCITE4a constructs as templates in the rabbit reticulocyte TnT (Promega, Madison, WI, USA) coupled transcription-translation system, containing ^{35}S Met as described as recommended by the manufacturer.

4.3.8. Protein binding assays

For protein interaction assays, the fragment of zebrafish eIF4GI containing the eIF4E-binding domain was cloned into pCITE4aDr4GI4EBD. The zebrafish 4E-BP was co-translated with either S-tagged eIF4E-1A,-1B,-1c or -3 in 35 µl reactions for 60 min at 30 °C. Reactions were diluted with 10 volumes of S-binding/washing buffer and incubated with 50 µl of S-protein agarose (Novagen, Madison, WI, USA) for 60 min at 10 °C. S-protein-agarose beads were recovered by centrifugation and washed 5 times with buffer (1 ml each), prior to elution with SDS-PAGE sample buffer. Samples of fractions, equivalent to 2 µl of the initial

translation reactions, were analyzed by SDS-PAGE and scanned for radioactivity using a Typhoon Storm (26).

4.3.9. Production of recombinant eIF4Es and 4E-BPs from *E. coli*

pET11deIF4E constructs were transfected into Rosetta™(DE3)-pLysS competent cells (EMD Miliipore, Billerica, MA, USA) and expressed as described (26). 10-ml cultures, grown in LB, 100 µg/ml carbenicillin, 34 µg/ml chloramphenicol, overnight at 37 °C. This culture was diluted to an optical density (OD) of 0.1 in LB/carbenicillin/chloramphenicol and grown to an OD of 0.5. Expression was induced with isopropyl β-D-1-thiogalactopyranoside (IPTG) at a concentration of 1 mM. This culture was shaken at 37 °C at 220 rpm for 2.5 h, and harvested by centrifugation at 10,000 x g for 5 min. Cells were lysed (10 µg/µl lysozyme, 25 mM HEPES-KOH pH 7.2, 100 mM KCl, 10 % glycerol, 1mM EDTA, 1mM EGTA, 0.5% Elugent (Calbiochem La Jolla, CA, USA) and the supernatant and/or protein pellet isolated after DNAse treatment. Expression was assessed by SDS-PAGE fractionation followed by staining or immunoblotting.

4.3.10. Development and validation of affinity purified antibodies

The amino acid sequences of each eIF4E zebrafish were submitted to the Genscript OptimumAntigen™ Design Tool to determine the best antigenic regions to use for immunization. Genscript synthesized each antigenic peptide (See Table 3.2) and added an additional cysteine residue to allow for conjugation to the KLH adjuvant. These were used for immunization of New Zealand white rabbits. Specific antibodies were isolated from the resulting serum by affinity purification using the synthesized peptide as bait. Antibodies were tested for

specificity and cross-reactivity by an ELISA assay and western blot analysis using the peptide used to generate the antibody and the recombinant protein of each eIF4E, respectively. The specificity of each antibody was validated using recombinant eIF4Es and tested for cross reactivity against all recombinant zebrafish eIF4Es. Antibody dilutions used for immunoblotting were adjusted to reflect the avidity and titer.

4.3.11. SDS-PAGE and immunoblotting

Proteins fractionated by 17.5 % high-Tris SDS-PAGE as described (26) were electro-transferred to PVDF membranes and subjected to immunoblot analysis using our custom polyclonal antibodies followed by goat anti-rabbit secondary antibody coupled to HRP and coupled with chemiluminescence. Chemiluminescence was detected using the ProteinSimple Fluorochem E with quantification using AlphaImager software.

4.3.11. Quantification of eIF4E levels

Unfortunately, our antibody for zebrafish eIF4E-2A gave a very poor signal. I was able to look at eIF4E-2B levels only and the combined levels of eIF4E-2A and -2B. Expression of each eIF4E was determined by immunoblotting using standard procedures. Comparison of signal from equal loading of each recombinant protein allowed avidity of each antibody to be established. The ECL signal was normalized by avidity and the relative levels of each eIF4E determined. The relative levels of eIF4E-2B and eIF4E-1A, -1B, -1C and eIF4E-3 assessed from tissue samples of muscle, brain, and ovary. Quantification by saturated pixel (SD) intensity was measured by Alphaimager™ software.

4.3.12. Complementation assays in *S. cerevisiae*

The yeast expression vector, pRS416GPD, separately containing each of the zebrafish *Eif4es* was transformed into *S. cerevisiae* strain JOS003 (88) using modified LiAc/SS carrier DNA/PEG Method (130, 157). JOS003 is a strain from which the endogenous *EIF4E* gene has been replaced by homologous recombination with a KanMX4 cassette. This makes the strain resistant to G418. It also expresses the human *EIF4E-1* gene behind a glucose-sensitive promoter on a plasmid conferring the ability to grow on uracil-deficient media. This system has been used previously to investigate the ability of heterologous eIF4Es to function in translation by rescuing growth in the presence of glucose (88). The transformed yeast were spot plated on synthetic deficient (SD) media lacking uracil and leucine and containing 200 µg/ml G418, with either galactose or glucose. Plates were incubated at 30 °C for 3-4 days, and growth was assessed visually by colony formation. Growth on plates containing dextrose indicates the ability of an ectopic *EIF4E* gene to complement eIF4E deficiency.

4.3.13. m^7GTP-Sepharose and TMG-Sepharose binding assay

Sepharose beads bound to 7-methyl-guanosine-triphosphate (Jena Bioscience GmbH, Jena, Germany) were blocked using 1 mg/ml soybean trypsin inhibitor (Sigma, T9128) in binding buffer (25 mM HEPES/KOH pH 7.2, 10 % glycerol, 150 mM KCl, 1 mM dithiothreitol, 1 mM D-L methionine) for 1 h at 4 °C shaking at 1400 rpm in a benchtop thermomixer 22331(Eppendorf). The beads were washed twice with binding buffer without soybean trypsin inhibitor and suspended in 50 % v/v binding buffer. 20 µl of each In Vitro Translation (IVT)

product was diluted 10-fold with binding buffer containing 200 μM GTP and 200 μM $MgCl_2$ and mixed with the bead suspension. Binding was allowed to occur at 4 °C for 1 h shaking at 1400 rpm. The supernatant containing the unbound fraction was recovered by centrifugation at 500 x g at 4 °C. An equivalent of 1 μl of the original IVT was used for TCA precipitation and filtered onto a GF/C membrane (Millipore). The cap-analogue beads were washed 5 times with binding buffer and the final bead fraction was suspended in SDS-PAGE sample buffer. The bead suspensions were heated to 90 °C and a fraction equivalent to 1 μl of the original IVT reaction applied to GF/C filter paper. Fractions were counted in Ecoscint Original scintillation cocktail (National Diagnostics, Georgia, USA) and CPM determined using a LS6500 Multipurpose Scintillation Counter (Beckman Coulter). IVT, unbound, and bead bound fractions were diluted in SDS-PAGE sample buffer and heated to 90 °C for 3 min. The samples were separated by 17.5 % high-Tris SDS-PAGE. Bead binding assays were also conducted with TMG-agarose, a gift from Dr. Ed Darzynkiewicz, University of Warsaw, Poland.

4.4. Results

4.4.1. The zebrafish genome encodes two Class II eIF4E family members: eIF4E-2A and eIF4E-2B

eIF4E Class II members are distinguished from Class I eIF4Es primarily by two Trp to Tyr residue substitutions in the conserved core (at residues equivalent to W46, W56 of mouse eIF4E-1). In zebrafish there are two eIF4E-2 cognates, designated as eIF4E-2A and eIF4E-2B. These are 237 and 229 amino acids in

length, respectively (see Chapter 3,Table 3.3). eIF4E-2A and -2B are 88.5 %.identical. Both have 34.8 % identity to zebrafish eIF4E-1A and -1C, and 38.2 % / 35.4 % identity to zebrafish eIF4E-1A and -1C, respectively (see Chapter 1, Table 1.2).

Although the Class II eIF4Es have similar core peptides sequences, the N- and C- termini differ significantly (Figure 4.1 and Appendix Figure A2.2). eIF4E-2B displays an unusual string of asparagine (N) residues, whereas the eIF4E-2A has more negative residues (aspartic acid (D) and glutamic acid (E)) in the N-terminal region. Appendix Figure A2.2 shows a multiple alignment that most of the residues involved in binding the m^7GTP in eIF4E-1 are conserved across species in eIF4E-2. Tyr77, Trp123 and Glu124 residues in zebrafish eIF4E-2A and Tyr71, Trp118, and Glu119 in zebrafish eIF4E-2B correspond to Trp56, Trp 123 and Glu124 in mouse eIF4E-1, the residues that interact with the guanine moiety of the m^7GTP. These amino acids, at equivalent positions, are invariant in all deuterostome eIF4E-2s. Similarly, Lys133, Arg173 in zebrafish eIF4E-2A and Lys128, Arg168 in zebrafish eIF4E-2B are equivalent to Arg112, Arg157 in mouse eIF4E-1, which form salt bridges with the triphosphate of the cap. These amino acids, at equivalent positions, are invariant in all deuterostome eIF4E-2s.

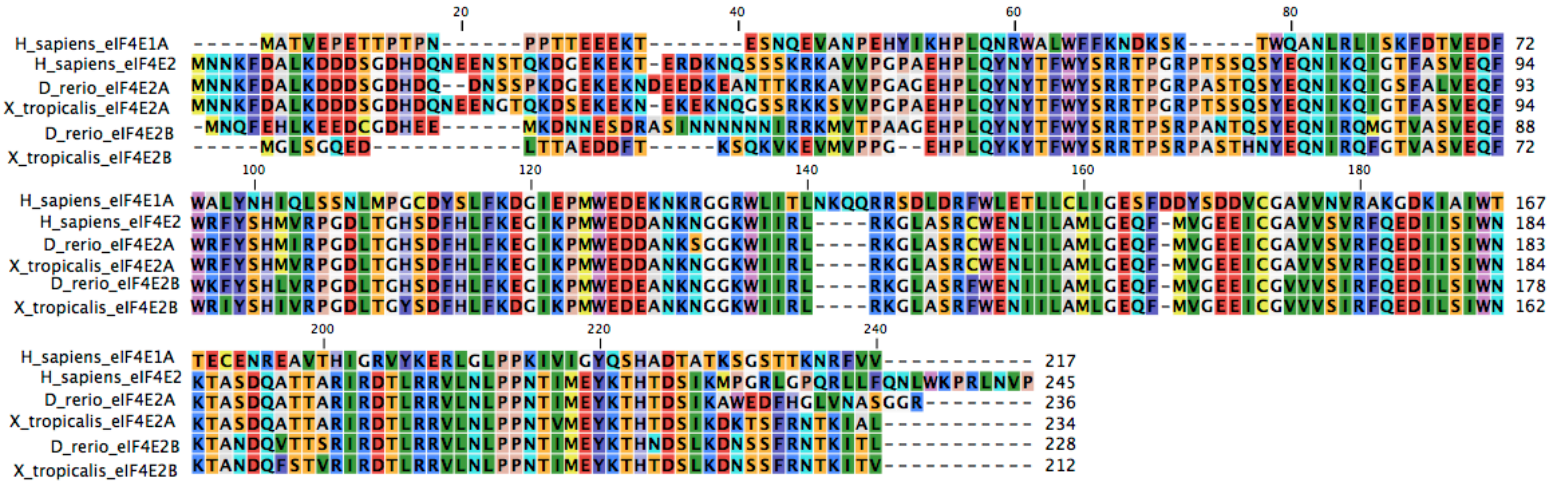


Figure 4.1: multiple alignment of HseIF4E-1A with Hs eIF4E-2, Dr eIF4E-2A, -2B, Xl eIF4E-2A, -2B

At positions equivalent to the negatively charged Lys162 in mouse eIF4E-1, eIF2B-2A and -2B have Ile at positions amino acid residue positions 178 and 173, respectively. This serves to reduce the salt bridges with the triphosphat the cap. These amino acids, at equivalent positions, are invariant in all deuterostome eIF4E-2s. The loop-β1β2, which forms one wall of the m^7GTP binding site is five residues longer in deuterostome Class II eIIF4Es and con an additional short α-helix enclosing more of the ligand binding site (118). Th flexible loop on which the cap-binding Tyr residue resides is also of variable length amongst the Class II eIF4Es from different species. All these characteristics appear to reduce the affinity of the Class II eIF4Es to the cap structure (118). There are many amino acids in the core region that differ between zebrafish eIF4E-2A and -2B. However, none of them occurs at the identified as important for affinity to cap described above. There are only thr positions in the core structure at which eIF4E-2A, but not eIF4E-2B, differs f eIF4E-1A. The first of these is Met101 in zebrafish eIF4E-2A and Leu101 in eIF4E-2B, equivalent to Ile79 in human eIF4E-1A and Ile in zebrafish eIF4E- The second is Cys148 in eIF4E-2A and Phe141 in eIF4E-2B, equivalent to Phe129 in human eIF4E-1A and Phe in zebrafish eIF4E-1A. The third is Ala in zebrafish eIF4E-2A and Ser188 in zebrafish eIF4E-2B, equivalent to Thr1 human eIF4E-1A and Ser in zebrafish eIF4E-1A. The question then arises, c such minimal differences in the core sequences be sufficient for neofunctionalization?

The phylogenetic analysis of deuterostome Class II eIF4Es can be seen in Chapter 2, Figure 2.4. eIF4E family members of Class II fall within two discrete clusters within the phylogenetic tree. The eIF4E designated as eIF4E-2 (2A) comprises the majority of the class II eIF4Es within the deuterostomes. The eIF4E-2B cluster is represented primarily by the ray-finned fish, but also by the amphibian *Xenopus* spp. As with the Class I eIF4Es, the Class II eIF4Es tend to segregate consistent with their evolutionary relationships. In comparing the eIF4E-2B cluster, it is evident there is tight grouping between the higher and lower teleosts and *Xenopus*. Only one variant of Class II was initially present in the early deuterostomes; the cognate protein eIF4E-2B first appears in the basal ray-finned fish, has been lost in coelacanth and amniotes but retained in amphibians. All teleosts so far examined have eIF4E-2A and most have eIF4E-2B (see also Chapter 2, Figure 2.2).

4.4.2. eIF4E-2A and eIF4E-2B transcript levels To assess where the eIF4E-2 family members are expressed in zebrafish embryonic series and tissues, I looked for the transcripts of each in a variety of tissues using RT-qPCR (Figures 4.2A and 4.2B).

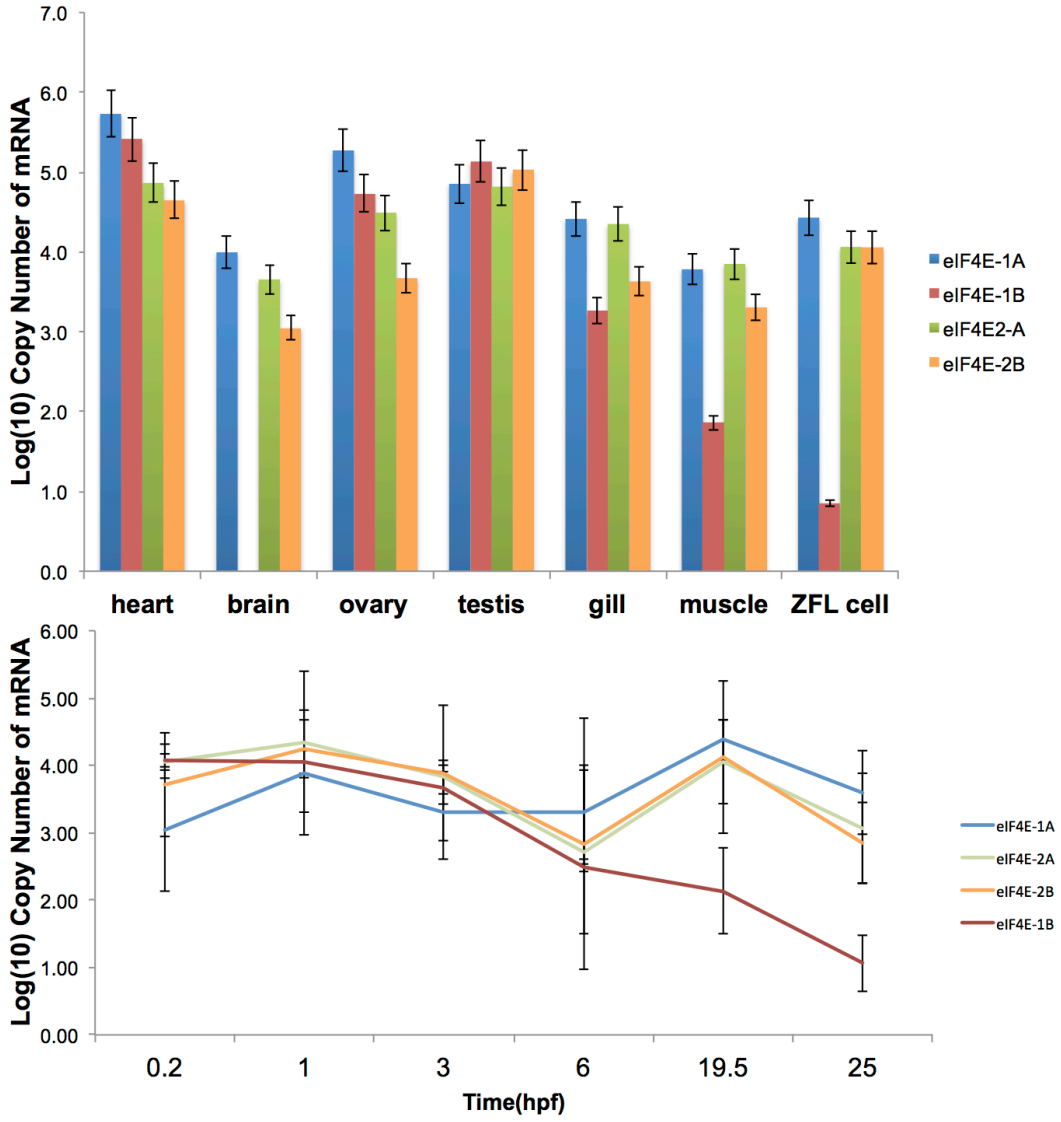


Figure 4.2: Transcript levels of Class II eIF4Es in adult tissues and early developmental stages: Top panel (A)Transcript levels from adult zebrafish tissues (top) and in embryos at various times post-fertilization (bottom) (B)were determined by RT-qPCR using cDNA generated from 25 ng RNA

eIF4E-2A and eIF4E-2B transcripts were detected in all tissues examined, with the highest transcript levels of eIF4E-2A transcripts in heart, ovary, and testis and the highest levels of eIF4E-2B transcripts in heart and testis (Figure 4.2A, Table 4.1). In most tissues, and in ZFL cells, transcript levels for the eIF4E-2A and -2B appeared to be about the same. Copy numbers of eIF4E-2A transcripts vary from 4.5 x 10^3 (brain) to 70 x 10^3 (heart), based on 25 ng RNA samples.

Transcript levels for eIF4E-2B ranged from 2 x 10^3 (muscle) to 100 x 10^3 (testis). These numbers compare to 6 x 10^3 (muscle) to 537 x 10^3 (ovary) for eIF4E1A. Transcript levels were also determined for embryos at different developmental stages (Figure 4.2B, Table 4.1).

Table 4.1:Transcript levels of zebrafish eIF4E family members (copy number of mRNA per 25 ng RNA)

Tissue	**eIF4E-1A**	**eIF4E-1C**	**eIF4E-2A**	**eIF4E-2B**
heart	$5.37x10^5$	$1.13x10^5$	$7.31x10^4$	$4.51x10^4$
brain	$9.89x10^3$	$2.05x10^3$	$4.52x10^3$	$1.12x10^3$
ovary	$1.87x10^5$	$4.18x10^4$	$3.10x10^4$	$4.71x10^3$
testis	$7.13x10^4$	$1.19x10^5$	$6.58x10^4$	$1.06x10^5$
gill	$2.59x10^4$	$5.40x10^3$	$2.24x10^4$	$4.32x10^3$
muscle	$6.05x10^3$	$2.64x10^3$	$7.07x10^3$	$2.03x10^3$
ZFL cell	$2.69x10^4$	$3.71x10^4$	$1.16x10^4$	$1.14x10^4$
Embryo hpf	**eIF4E-1A**	**eIF4E-1C**	**eIF4E-2A**	**eIF4E-2B**
0.2	$1.09x10^3$	$4.09x10^3$	$1.15x10^4$	$5.14x10^3$
1.0	$7.76x10^3$	$5.37x10^4$	$2.21x10^4$	$1.78x10^4$
3.0	$2.02x10^3$	$1.03x10^4$	$6.75x10^3$	$7.65x10^3$
6.0	$2.02x10^3$	$1.26x10^3$	$5.17x10^2$	$6.78x10^2$
19.5	$2.40x10^4$	$1.04x10^4$	$1.12x10^4$	$1.32x10^4$
25	$3.97x10^3$	$3.36x10^3$	$1.17x10^3$	$7.11x10^2$

Both eIF4E-2A and eIF4E-2B transcripts were detected at the highest level shortly after fertilization. Transcript levels for eIF4E-2A and eIF4E-2B were lowest shortly after the maternal to zygotic transition (MZT), and then began to increase up to 16 hpf. The overall conclusions of the expression patterns of eIF4E family members suggest that eIF4E-2A and eIF4E-2B are ubiquitously expressed and follow the same pattern as the transcripts for genes over this period.

4.4.3. eIF4E-2A and -2B protein levels

Unfortunately, the antibody developed for zebrafish eIF4E-2A gave a very poor signal. A consultant at Genscript informed me that developing an additional antibody to the eIF4E-2A in the C terminus could incur further poor performance due to problems with glycosylation at that site. Conversely, the antibody for eIF4E-2B has a comparable avidity to that for eIF4E-1C and eIF4E-3 (data not shown). eIF4E-2B expression can be seen slightly in brain, and definitively in gill (Figure 4.3).

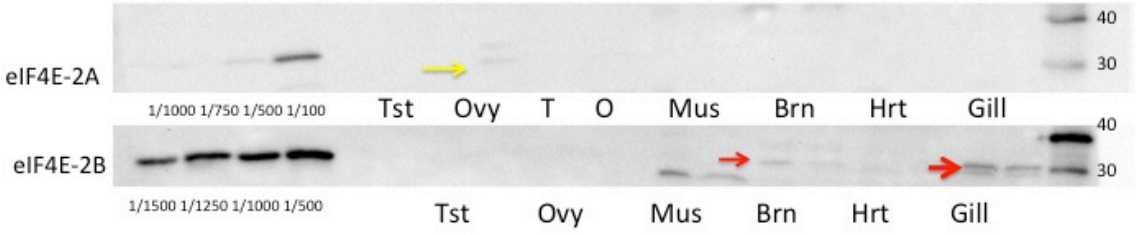


Figure 4.3: eIF4E-2A and -2B protein levels: Ovy, O: (ovary); Tst, T: (testis); Mus: (muscle); Brn: (brain); Hrt: (heart); Gill. Samples in duplicate lanes, or individual when noted.

eIF4E-2B levels in brain may be very low, but it exceeds eIF4E-1C in expression levels in gill. eIF4E-2B levels are undetectable in testis and ovary and barely detectable in heart. Unlike at the transcript level, eIF4E-2B was undetectable in embryos at any stage. This obvious disconnect between transcript and protein level of the Class II eIF4Es suggests that their level is regulated by low translational efficiency or high protein turnover. Both eIF4E-2A and -2B proteins have putative SUMO sites.

4.4.4. Zebrafish eIF4E-2B will rescue eIF4E function in a yeast strain conditionally deficient in eIF4E-1

The JOS003 strain lacks the endogenous yeast *eIF4E* gene and expresses human eIF4E-1 inserted in the pRS415 leu (-) vector under the control of the galactose-dependent and glucose-repressible GAL1 promoter (88). As a consequence, strain JOS003 is able to survive in medium containing galactose as carbon source but is not viable in medium containing glucose due to depletion of the human eIF4E-1. Growth of JOS003 in glucose can be mediated by ectopic expression of a functional eIF4E, the regulation of which is under the control of a promoter in the pRS416 ura (-) vector, which is active in the presence of glucose (88). The cDNAs encoding the zebrafish eIF4E-2 cognates were cloned into pRS416, allowing expression from the constitutively active glyceraldehyde-3-phosphate dehydrogenase (GPD) promoter. Following transfection and selection on media lacking uracil, the yeast cells containing control vector, or vectors for the expression of eIF4E-1A, -1B, -2A or eIF4E-2B, were streaked on selective plates; Synthetic media (SC) –Ura, -Leu containing either galactose or glucose as carbon source (Figure 4.4). As previously reported, eIF4E-1A is capable of complementation, while eIF4E-1B is not (26). Though unexpected, it is evident that eIF4E-2B is able to rescue the JOS003 strain under conditions in which human eIF4E-1 is depleted. Expression of each the eIF4E-2A and eIF4E-2B was verified by western blot analysis using antibodies specific to each eIF4E (data not shown) and each sequence identity was confirmed by PCR analysis and DNA sequencing (data not shown).

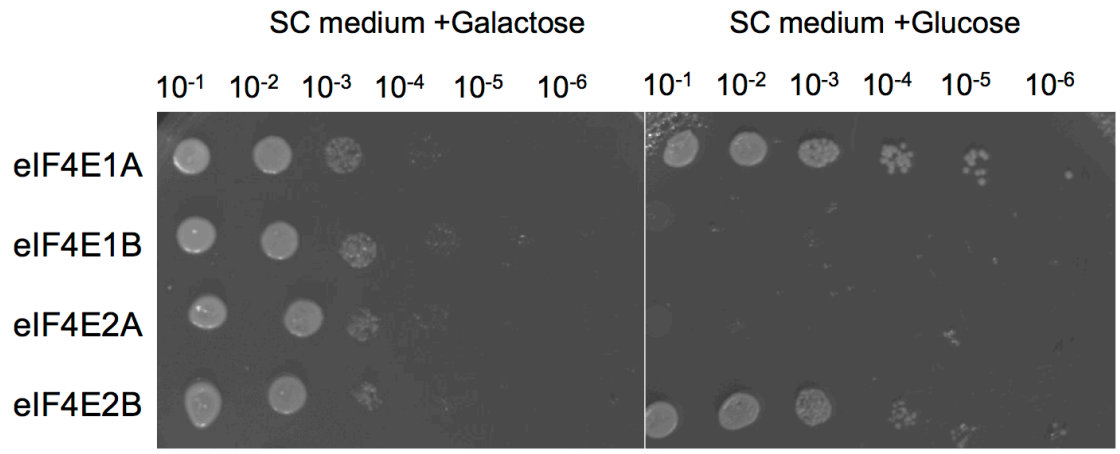


Figure 4.4: Ability of zebrafish Class II eIF4Es to rescue the growth of *S. cerevisiae*, JOS003: *The S. cerevisiae* strain, JOS003,(88) was transformed with the Ura-selectable vector, pRS416GPD, containing cDNAs encoding one of the following products: eIF4E-1A, eIF4E-1B, eIF4E-2A and eIF4E-2B, as indicated. Following selection on SC medium with galactose lacking uracil and leucine, yeast from the resulting single colonies were diluted 10^{-1} to 10^{-7} fold and transferred onto YP-agar media containing G418 and either glucose (*left*) or galactose (*right*). Growth was assessed after 48 h.

These results show that zebrafish eIF4E-2B is capable of supporting protein synthesis in yeast like prototypical eIF4E-1. Because eIF4E-2B gave such an unexpected result, the pRS416GPD/eIF4E-2B cDNA was re-made using synthetic cDNA, codon optimized for yeast, and cloned into pRS416GPD by Genscript. This construct also complemented the yeast strain grown in the presence of glucose.

4.4.5. eIF4E-2A and eIF4E-2B bind to m^7GTP cap analogue

Mammalian eIF4E-2 was previously examined for binding affinity to the m^7GTP cap analogue in bead binding assays, which showed that it was able to bind *in vitro* (26). This study was replicated using recombinant zebrafish eIF4E-2A and -2B synthesized *in vitro* using pCITE4a constructs in the rabbit reticulocyte coupled transcription-translation system, TnT, containing ^{35}methionine as

described previously (26). Zebrafish eIF4E-1A and -1B were used as positive and negative controls respectively. Luciferase was also used as a negative control. The TnT incubations were diluted and mixed with a m^7GTP-Sepharose bead slurry, and the total, unbound, and bound fractions were analyzed by SDS-PAGE, immunoblotting, and scanned for radioactivity using a Typhoon Storm. Eluted proteins were resolved by high-Tris SDS-PAGE prior to immunoblotting and visualization of the eIF4E-2A and eIF4E2B proteins. The results obtained indicate that both eIF4E-2A and eIF4E-2B will bind to the m^7GTP beads (Figure 4.5).

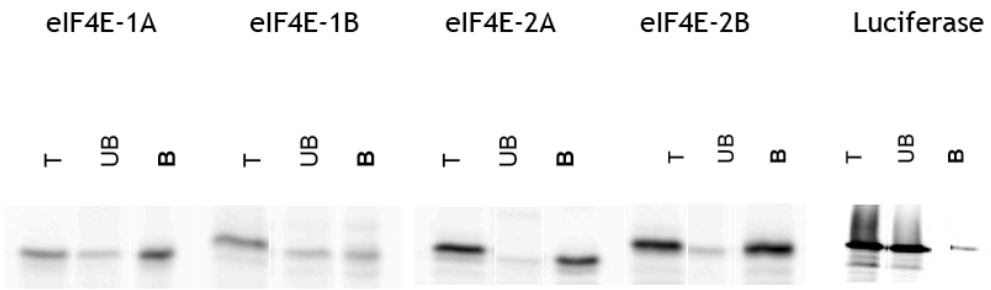


Figure 4.5: m^7GTP binding activity of zebrafish Class II eIF4Es

4.4.6. eIF4E-2B binds to $m^{7,7,7}$GTP cap analogue (TMG)

In the initial paper describing the human form of eIF4E2 (4EHP), Rom discussed the possibility that eIF4E-2 may be involved in recognition of the hypermethylated 2,2,7-trimethyguanosine (TMG) of small nuclear RNAs for import into or export from the nucleus (148). TMG cap structures are characteristic of small nuclear and nucleolar RNAs that program pre-mRNA splicing (U1, U2, U4, and U5 snRNAs) and pre-rRNA processing (U3 and U8 snRNAs) (158). These small nuclear RNAs (snRNAs) are known to exit the nucleus in human cells and their

export is enhanced by their cap structures (159). Because 4E transporter (4E-T) binds to eIF4E-2 and eIF4E-2 shuttles through nuclei (33), zebrafish eIF4E-2A and -2B were assayed for binding to TMG-agarose. The binding assay was conducted using TMG-agarose beads kindly provided by Dr. Ed Darzynkiewicz. Comparing the TCA-precipitable protein in the total, unbound, and bound fractions then assessed binding to these specialized beads. *C. elegans* IFE-1, a Class I eIF4E that is known to bind to the TMG caps of *C. elegans* mRNAs, was included to provide a positive control for TMG binding. The results indicate that eIF4E-2B binds TMG at a comparable level to *C. elegans* IFE-1s (Figure 4.6). These results could reflect potential for neofunctionalization of eIF4E-2B. However, I will need to properly confirm these results by utilizing a more sensitive assay for affinity, such as SPR, to quantify the interaction of eIF4E-2B and TMG.

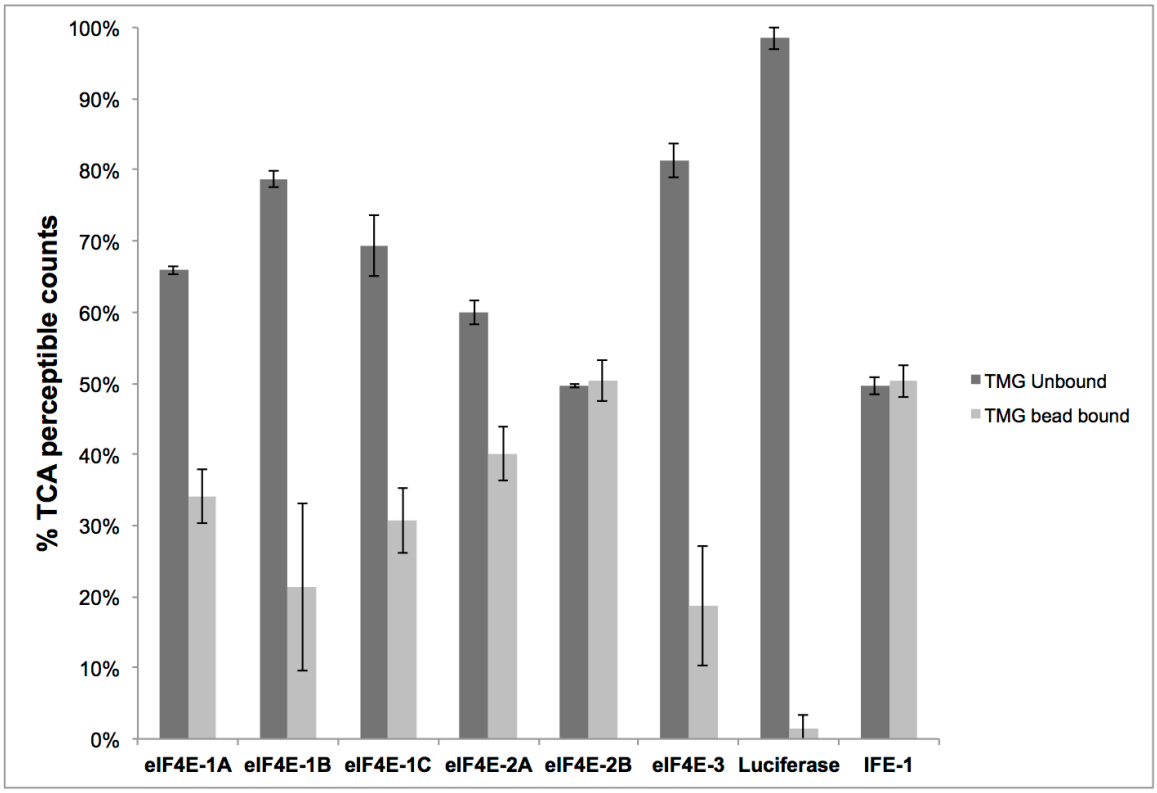


Figure 4.6: TMG binding activity of zebrafish Class II eIF4Es

4.4.7. Neither eIF4E-2A nor eIF4E-2B interact with zebrafish eIF4GI *in vitro*

The nucleotide sequences for the zebrafish eIF4G1 corresponding to amino acids 262-681 were codon optimized for rabbit, synthesized and cloned into pCITE4a (+) (Genscript). This region brackets the eIF4E interaction domain. The eIF4E fragment was co-translated with S-tagged variants of eIF4E-2A and -2B in a reticulocyte cell free translation system in the presence of ^{35}S methionine (Figure 4.7). After synthesis, reactions were incubated with S-protein-agarose. Following extensive washing, proteins that bound to the matrix, were eluted with SDS-PAGE sample buffer. Zebrafish eIF4E-1A was used as a positive control; eIF4E-1B was used as a negative control.

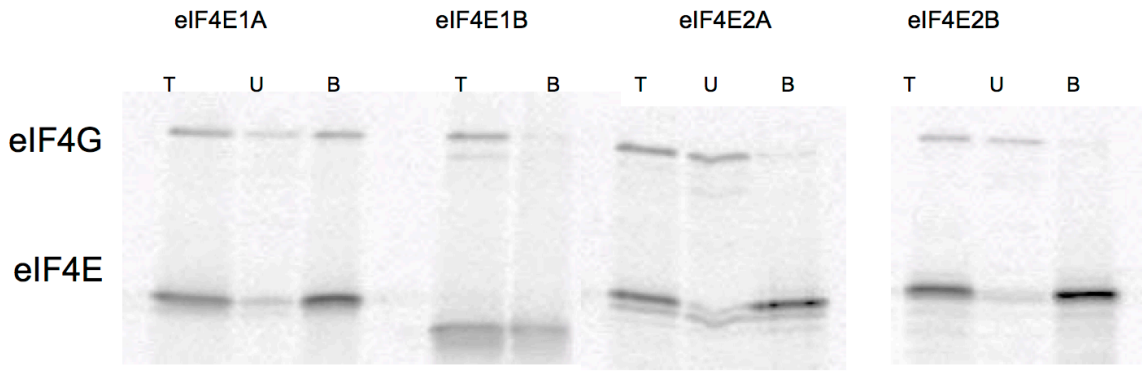


Figure 4.7: Interaction of zebrafish eIF4G with Class I and II eIF4Es: The Mwt(kDa) of eIF4E is ~24-27 and eIF4G is ~51 kDa*(eIF4G migrates as ~ 100 kDa under the SDS page conditions, 39).

Proteins were fractionated by high-Tris SDS-PAGE and an image generated using the Typhoon Storm. Whereas the zebrafish eIF4GI fragment co-purified with eIF4E-1A, neither eIF4E-2A nor eIF4E2B bound to eIF4G. As expected, zebrafish eIF4E-1B also failed to interact with eIF4G. The inability of zebrafish eIF4E-2B to bind to zebrafish eIF4G seems to be at odds with its ability to complement the yeast strain conditionally deficient in eIF4E. This will require further study.

4.4.8. eIF4E-2A and eIF4E-2B interact poorly with the 4E-BPs

The binding partners of vertebrate eIF4E the 4E binding proteins (4E-BPs) act to regulate translation through phosphorylation and the mTOR pathways (reviewed (160-162)). Most vertebrates have three 4E-BPs, 4E-BP1, 4E-BP2 and 4E-BP3. The Actinopterygii have an additional 4E-BP3-like (26) giving four variants of the zebrafish 4E binding proteins, which are designated as 4E-BP1, 2, 3 and 3-like (3L). The 4E-BPs bind to eIF4E through the consensus motif, YXXXXLϕ, also the binding site for eIF4G. Previous studies have shown that 4E-BPs bind poorly to

the Class II eIF4Es (39). In deuterostome Class II eIF4Es the eIF4G/4E-BP binding motif, TVEDFW of the Class I eIF4Es, has been changed to SVEQFW (Appendix Figure A2.2). The change from the negatively charged aspartate to polar, uncharged glutamine likely to affect the conformation of this motif substantially and likely accounts for the inability of eIF4G to bind to the Class II eIF4Es. To assess the functionality of the eIF4E-2A and -2B to bind to the 4E-BPs, in *vitro* interaction assays with S-tagged variants of ^{35}S-labeled eIF4E-2A, -2B was performed in the same method as the eIF4E/eIF4G interaction assay described in Section 4.4.7. Zebrafish eIF4E-1A and -1B were used as positive and negative controls. After synthesis, reactions were incubated with S-protein-agarose. Following extensive washing, proteins bound to the matrix were eluted with SDS-PAGE sample buffer (Figure 4.8). The data showed that unlike eIF4E-1A, zebrafish eIF4E-2A and -2B bound the 4E-BPs poorly. The exception to this was that both eIF4E-2A and -2B bound to the teleost specific 4E-BP3l. However, no differentiation was seen between the activity of eIF4E-2A and -2B.

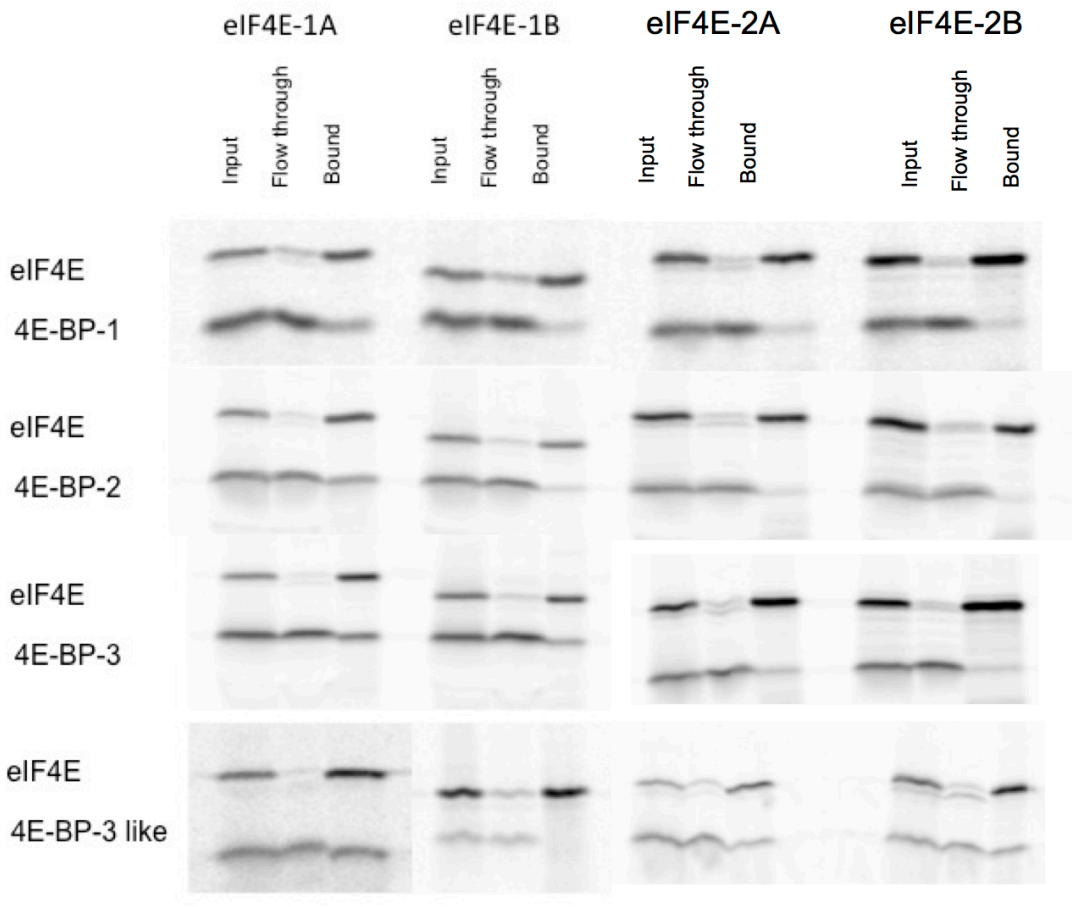


Figure 4.8: Interaction of zebrafish 4E-BPs with zebrafish Class I and II eIF4Es: The Mwt (kDa) of the 4E-BPs is ~15 and eIF4Es are~24-27.

5. Discussion

This represents the first study to investigate the origin of the duplication of the deuterostome Class I eIF4Es, as well as the neofunctionalization of the vertebrate eIF4E-2Bs. The ability of zebrafish eIF4E-2B to complement the yeast strain conditionally deficient in eIF4E was most surprising since both eIF4E-2A and -2B have the amino acid substitutions thought to account for its more than 100-fold lower affinity for the m^7GTP cap (118). For complementation to occur, eIF4E-2B must support protein synthesis and the sequence of eIF4E-2B does

not seem compatible with this. This result was of concern, so I had an alternate yeast construct prepared by Genscript using a chemically synthesized cDNA. And again, the results were the same. In fact, there are only three positions in the core structure at which eIF4E-2A, but not eIF4E-2B, differs from eIF4E-1A. From their location, it seems unlikely this would make a difference in their ability to support protein synthesis. The two proteins do differ markedly in the N-terminal and C-terminal regions. In looking at the multiple alignment of the deuterostome Class II eIF4Es in the Appendix Figure A2.2, it can be seen that the eIF4E-2Bs have consistently shorter N-terminal regions that include Q-rich regions. Q- or Q/N-rich regions have been reported to induce aggregation-prone proteins and have been suggested to play a role in the accumulation of proteins in P-bodies (163). However, although this is a significant difference between eIF4E-2A and -2B, it is unlikely to account for its ability to promote protein synthesis. Similarly, there are differences in C-terminal motifs; the eIF4E-2As of teleosts have a motif, KAWEDFH, whereas the eIF4E-2Bs have KDNSSF. The motif in eIF4E-2A is shared with the single eIF4E-2 in lamprey. The motif in eIF4E-2B is shared with the single eIF4E-2s in the cephalochordates (Amphioxus), the tunicate, *C. intestinalis*, elephant shark and the coelacanth.

The ability of eIF4E-2B to bind TMG may be significant under certain biological conditions such as embryogenesis, where the TMG cap is a signal for snRNP re-import into the nucleus for formation of immature snRNPs into Cajal bodies (CB) (164) Since eIF4E-2B is predicted to be of nuclear localization via PSORTII

algorithms (data not shown), its role in the SMN complex would still need to be experimentally verified for biological significance in zebrafish.

In looking toward future research, I plan to substitute the three discussed amino acid differences in eIF4E-2A, with those in eIF4E-2B to determine if this will change its ability to bind. In addition, I plan to switch the N- and C-terminal domains of eIF4E-2A and -2B and look at the effects on the ability of the proteins to complement the eIF4E-deficient yeast strain. In this way, I hope to be able to establish alternate functions.

Chapter Five: Summary and Future Directions

5.1. Summary/Conclusions

In this chapter, I reaffirm the research objectives and review the contributions and significance of my work in meeting my objectives of defining the phylogenetic origins of the teleost eIF4E family members, demonstrating that eIF4E-1C functions as a prototypical initiation factor, and showing that eIF4E-3 is playing a regulatory role in the regulation of gene expression in muscle and brain. I also discuss the questions arising from my work to date and suggest approaches that can be used to answer these questions.

The zebrafish, *Danio rerio*, has proved to be a useful model system for studying the evolution and differing roles of eIF4E family members in deuterostomes. The recent expansion of accessible completed deuterostome genomes in online databases has enabled the investigation of deuterostome eIF4Es from echinoderms through mammals. In particular, access to protein sequences generated from the genomes of sea urchin, tunicates, lancelets *(Amphioxus*), lamprey, elephant shark, coelacanth, and several teleosts, along with many tetrapod species has provided an illustration of the duplication of eIF4Es through multiple whole genome duplications, neo-functionalization and asymmetric deletion of eIF4Es among the different vertebrate classes. eIF4E-1B represents a duplication of an ancestral deuterostome eIF4E-1 that is expressed primarily in ovary and testis and has acquired a regulatory function. It can first be recognized in the elephant shark suggesting that the duplication must have occurred prior to the branching of the chondrichthyes and persists throughout the tetrapod lineage.

eIF4E-1B is not seen in the higher spiny ray fish of the Actinopterygii, such as Tetradonts (*Takifugu rubripes*) and Percoids (*Oreochromis niloticus*), although such species have the same needs for the regulation of mRNA recruitment in circumstances such as completion of meiosis that eIF4E-1B is known to accomplish (27, 30, 31).

Unlike the distinct neofunctionalization seen in eIF4E-1B, duplication of eIF4E-1 to give eIF4E-1A and -1C has seemingly provided two forms that function as translation initiation factors. The duplication must have occurred prior to the branching of the chondrichthyes because the elephant shark, *Callorhinchus milii*, has both eIF4E-1A and -1C. Both eIF4E-1A and -1C are conserved across the Actinopterygii. The lobe-finned fish represented by the coelacanth, *Latimeria chalumnae* , is basal to the tetrapods and also has both eIF4E-1A and -1C. However, eIF4E-1C has been lost in tetrapods. Since both eIF4E-1A and -1C function as the translation initiation factor, it is currently unclear what selective advantage can be attributed to the retention of both forms in gnathostomes.

This present study utilizes the ray-finned species, *Danio rerio,* from the superfamily Ostariophysi, whose eIF4E members include each of the three classes of eIF4E found in metazoans; three Class I (eIF4E-1A, -1B, -1C), two Class II (eIF4E-2A, eIF4E-2B) and a single Class 3 (eIF4E-3). One of the advantages of working with the zebrafish model system is the online availability of a completed genome database for comparison of protein sequences and genetic loci. When comparing this data with other teleost fish models, it provides a reliable framework upon which to construct hypotheses concerning whole

genome duplication events and neofunctionalization. My purpose was to examine the origin of the eIF4E family members, to characterize their function in *in vitro* and complementation assays, and compare their expression in tissues and developmental stages. Overall, the results support the Jagus laboratory's previous recognition of eIF4E-1A as a translation initiation factor, and eIF4E-1B as a regulatory factor in specialized tissues. My contribution is the first account of the functional characterization and expression of eIF4E-1C and eIF4E-3.

With development of cross reactive and specific antibodies to zebrafish eIF4Es, it became possible to compare expression across different tissues and developmental stages at the level of protein. In zebrafish, both eIF4E-1A and-1C are ubiquitously expressed, although eIF4E-1C is the predominant form. The ability to monitor expression of the zebrafish eIF4E family members at both the protein and transcript levels demonstrated that, except for eIF4E-1B and eIF4E-3, transcript and protein levels vary coordinately. Though relatively consistent transcript levels of the eIF4E Class II proteins, eIF4E-2A and eIF4E-2B were observed across all tissues, protein expression was only significant in the gill and brain tissues, and not until 25 hpf in embryos. At the level of protein, eIF4E-3 is only detected in muscle and brain. This investigation is the first to detect eIF4E-3 at the protein level in normal tissues.

The duplication of eIF4E-2 seen in zebrafish seems to be a teleost-specific attribute, with the exception of *Xenopus* species. I have been able to provide evidence of neofunctionalization of eIF4E-2B. The surprising finding is that the teleost specific form, eIF4E-2B, is able to complement the growth of an eIF4E-

deficient yeast strain. This was unexpected, since it is not able to bind to eIF4G in *in vitro* assays (Chapter 4). It could be speculated that this variant of eIF4E-2 will have the ability to interact with the yeast eIF3 and thereby facilitating the binding to the yeast eIF4G and subsequently initiation**.** eIF4E-2 is capable of having enhanced cap-binding ability when it is ISGylated with ISG15 through interaction with a binding partner (63), therefore it is possible that a protein partner may be operating in conjunction to allow the eIF4E-2B to function as a translation initiation factor. One difference between eIF4E-2A and eIF4E-2B is that eIF4E-2B is predicted to have a nuclear location (83 %) and eIF4E-2A is predicted to be primarily cytoplasmic. However, this would not account for its ability to function as a translation initiation factor.

5.2. Future Directions

The research I have summarized has raised a number of questions that future research needs to address.

5.2.1. What are the affinities of zebrafish eIF4E family members for cap analogues?

Traditionally, the measurement of the relative binding affinity of the cap structure to eIF4Es has been implemented using an *in vitro* binding assay to cap analogues on beads. This method only provides resolution of the strong binders and cannot be used quantitatively. Another methodology that provides both association and dissociation constants using surface plasmon resonance (SPR) will be employed using customized chip formats of the CM5 with the GST conjugated eIF4Es and the Biacore™ T200 instrumentation. The advantage of

this technology is in the presentation of an immobilized functional eIF4E on a re-usable chip to a selection of cap analogues and observing the binding affinity in real time to give dissociation constants. In particular, this usage will enable a definitive comparison of the zebrafish eIF4E-2 members when challenged with alternative forms of the cap structure and may confirm TMG binding affinity.

5.2.2. What is the effect of substituting eIF4E-2B-specific amino acids in eIF4E-2A?

The amino acids thought to give eIF4E-2B its ability to complement the yeast strain conditionally deficient in eIF4E will be substituted in eIF4E-2A, both separately and together. The effect on activity will be determined by the ability of the eIF4E-2A variants to support growth in the yeast strain conditionally deficient in eIF4E. In addition, since the N- and C-terminals of eIF4E-2A and -2B are very different, the effect of domain swapping will also be investigated.

5.2.3. What are the binding partners for zebrafish eIF4Es?

There are numerous examples of binding partner proteins to eIF4E family members, in addition to the 4E-BPs, that have been identified for eukaryotic organisms including human, *Xenopus*, mouse, and *Drosophila*. To that end, a reporter system (pGEX-GST fusion) will be commercially created and tailored for each zebrafish eIF4E. Cell lysates from selected tissues and embryo series will be run through eIF4E-GST affinity columns to purify eIF4E interacting proteins. The eluted fraction of captured proteins will then be analyzed by LC/Mass spectrophotometry for determination of amino acid sequence. This data would then be compiled for identification by means of online databases of protein motif

and family classification. The identity of these proteins will provide the means to compare these proteins to existing eIF4E binding partners such as 4E-T. When optimized, the Biacore™ system will be an additional method to pinpoint and collect potential binding partners from cell lysates with greater ease.

5.2.4. Where do eIF4E family members localize in cells, tissues and during development?

Though expression of eIF4E family members has been shown to occur in various tissues and embryonic stages, the exact localization of expression needs to be determined. This could be accomplished by means of fluorescence *in situ* hybridization (FISH), utilizing all the available antibodies of eIF4Es as probes. The expectation is that the Class eIF4E-1A; -1C and I eIF4Es will be ubiquitously expressed throughout tissue and embryonic stages due to the expression studies results. Differences in nuclear versus cytoplasmic localization may be resolved since eIF4E-1 and eIF4E-2 have been shown to play a role in shuttling mRNAs from the nucleus to the cytoplasm (33, 165) and eIF4E-1C is predicted to be primarily nuclear. My results also predict that eIF4E-1B, the eIF4E Class II eIF4Es and eIF4E-3 should be observed predominantly in select tissue such as muscle, brain, gill and different time points of embryo development. The results of these studies will be used to inform knockdown or knockout experiments.

5.2.5. How will knock-down and/or knockout of eIF4E family members affect phenotypes of zebrafish?

The simple questions I ask with knockdowns or knockouts include: 1) are both eIF4E-1A and eIF4E-1C essential? 2) does deletion of eIF4E-1B affect formation

of germ cells or fertility of adults? 3) what phenotypes will a knockout of the Class II eIF4Es have? and 4) is eIF4E-3 essential for muscle development? The effects of gene knock-downs and knock-outs of eIF4E family members will be determined with an emphasis on eIF4E-1B, eIF4E-2A, -2B and eIF4E-3. For "knocking down" the expression of a gene of interest in zebrafish, the preferred method has become morpholinos, whereas the newest choice in zebrafish for deleting a gene is the prokaryotic derived Clustered Regularly Interspaced Short Palindromic repeats (CRISPR)/cas9 system. Morpholinos will be implemented to observe the effect of eIF4E Class II and III knock downs during development. In order to deal with possible "off target" effects, either a control non-target morpholino will be co-injected with the eIF4E target, or a "rescue" mRNA experiment with co-injection of a mRNA of the intended gene to restore to wildtype will be subsequently conducted (166). However, the CRISPR/cas9 system may provide a more efficient means in later stages of embryonic development to complement the use of morpholinos. Two of the newest modifications of the CRISPR system have been developed. One involves dual use of RNA guided endonuclease (RGENs) Cas9 derived from Streptococcus (167), and the other is vector based CRISPR methodology targeting a specific tissue type and inducing a fluorescence phenotype in the F1 generation (168).

5.2.6. Do the class II eIF4E-2A and -2B N terminal regions affect binding to potential partners?

There are two distinct variants of Class II eIF4Es, designated eIF4E-2A and eIF4E-2B, which display unique N and C terminus regions that may play a role in snRNP binding. To assess if the terminal sequences play a role in cap and/or 4E

binding partner interaction, a mutagenesis study will be implemented. The alteration of the N terminus by switching the asparagine rich (N) eIF4E-2B to the eIF4E-2A, -1A, and -3 will determine if there is a functional aspect to this region that enhances cap-binding or interaction with protein partners.

5.2.7. Does the zebrafish eIF4E-2B react with components of the RNA granules?

eIF4E-2 (also termed 4E-HP) has been observed to interact with the 4E-Transporter protein(4E-T), P body aggregations, nuclear bodies (NB), and spliceosomal snRNP assembly in Cajal bodies containing Coilin which are all components of the RNA granule complex (28, 150, 164, 169-174). Trimethylguanosine (TMG)-capped snRNAs are present in the zebrafish nucleus as part of this assembly and enable nuclear transport (172). In Chapter 3, zebrafish eIF4E-2B was shown to bind with TMG-bound beads, similar in ability to the positive control *C. elegans* IFE-1. If the proposed Biacore studies confirm that eIF4E-2B interacts with TMG, then this study will be conducted to determine if an additional component is involved. To confirm whether zebrafish eIF4E-2B is capable of interacting with TMG and 4E-T specifically, antibodies to TMG and 4E-T could be utilized in immunoprecipitation studies. To prove that Cajal bodies and snRNAs are present in early embryonic stages, immunofluorescence studies from 0-5 hpf embryos that are enriched in Cajal bodies and snRNA will be performed (172). The zebrafish tissues of gill have been observed to express eIF4E-2B proteins, so would be a good source material to utilize in far western blotting. The eIF4E pGEX-GST tagged system would serve as the probe, after

mixing the tissue lysate then blotting with the eIF4E-2A and -2B antibodies (eIF4E-1B is used as a negative control).

5.2.8. Do zebrafish eIF4E-2A and -2B undergo ISGylation or sumoylation?

Under stress conditions, modifications involving eIF4E proteins may occur by interaction with Ubiquitin (Ub). As it pertains to eIF4E, the ISG15 appears to play a role in potential interaction with human eIF4E class II (4EHP) and may enhance cap-binding (63). Sumoylation is involved in various cellular processes, such as nuclear-cytoplasmic transport, transcriptional regulation, apoptosis, protein stability, response to stress, and progression through the cell cycle (175). SUMO-1 has been shown to modify eIF4E by conjugation and promotes eIF4F complex formation (50). Prior studies have observed that the presence of key lysine residues in an C terminal LRLRGG motif indicates when a protein is capable of being ISGylated (56) . The motif ψ-K-X-E functions as the same for sumoylation (176). Using predictive algorithims from a suite of online resources indicates that zebrafish eIF4E-2A and -2B both contain the necessary motifs for sumoylation (GP-SUMO). eIF4E-2B has a higher potential for ubiquitin interaction (UbPred), and possess a nuclear export signal peptide (NES). ISGylated mammalian eIF4E-2 (4EHP) will bind cap with enhanced affinity when observing the m^7GTP pull down results and modified/non-modified ISG15 4EHP fusion protein experiments (63). Subsequent experimentation with zebrafish eIF4E-2A and -2B would include cell lysate pull downs with the pGEX-GST fusion system and immunoblotting analysis with anti-ISG15, anti-SUMO-1 and anti-SUMO 2-3.

Overall statement: My work has revealed that the duplications of eIF4E family members arose early in vertebrate evolution, with either the 1R or 2R whole genome duplications. The evolutionary picture seems to be that the eIF4E family expanded to reach its zenith in northern pike (eight eIF4E family members) and then contracted in amniotes and to some extent in the percomorph teleosts. eIF4E-1C can be found in basal actinopterygians, teleosts lamprey, but has been lost in tetrapods. Tetrapods and lower teleosts have retained eIF4E-1B, but it has been lost in the percomorph teleosts. Subfunctionalization of eIF4E-1A and -1C is suspected although not apparent from my investigations.

Overall, I can say the following about the function of each zebrafish eIF4E family member. eIF4E-1A and eIF4E-1C definitively serve as initiation factors, as demonstrated by the ability of both to rescue growth of a yeast strain conditionally deficient in eIF4E. This is reflected in their ubiquitous expression. The differences in expression levels of each from tissue to tissue are suggestive of subfunctionalization. eIF4E-1B does not function as a translation initiation factor, but plays a regulatory role in the recruitment of CPE-containing mRNAs in specialized tissues such as ovary and testis (27). However, it is not clear how the more recently evolved teleosts such as pufferfish, stickleback, tilapia, regulate this function. Since eIF4E-2A and eIF4E-3 are not able to rescue growth of a yeast strain conditionally deficient in eIF4E, they are likely to play regulatory roles, by mechanisms still to be elucidated. eIF4E-3 expression is only detectable in muscle and brain. The ability of microRNA-206 and microRNA-21 to promote muscle atrophy, as well as target eIF4E-3, suggests that eIF4E-3 plays a role in

muscle development and maintenance (43). The ability of eIF4E-2B to rescue growth of a yeast strain conditionally deficient in eIF4E and to bind TMG are both clear indications of neofunctionalization, although its role has not been defined. It is also not clear how eIF4E-2B is able to support growth in yeast.

Complete understanding of the alternative functions of the zebrafish eIF4E family members in the regulation of gene expression will provide useful insights into the understanding of cell proliferation, cellular stressors, and the regulated utilization of mRNAs during development.

Bibliography

1. Raught B, and Gingras AC (1999) eIF4E activity is regulated at multiple levels. *Int J Biochem Cell Biol* 31:43-57.

2. von der Haar T, Gross JD, Wagner G, and McCarthy JE (2004) The mRNA cap-binding protein eIF4E in post-transcriptional gene expression. *Nat Struct Mol Biol* 11:503-11.

3. Sonenberg N, and Hinnebusch AG (2009) Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell* 136:731-45.

4. Jackson RJ, Hellen CU, and Pestova TV (2010) The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol* 11:113-27.

5. Topisirovic I, Svitkin YV, Sonenberg N, and Shatkin AJ (2011) Cap and cap-binding proteins in the control of gene expression. *Wiley Interdiscip Rev RNA* 2:277-98.

6. Prévôt D, Darlix JL, and Ohlmann T (2003) Conducting the initiation of protein synthesis: the role of eIF4G. *Biol Cell* 95:141-56.

7. LeFebvre AK, Korneeva NL, Trutschl M, Cvek U, Duzan RD, Bradley CA, Hershey JW, and Rhoads RE (2006) Translation initiation factor eIF4G-1 binds to eIF3 through the eIF3e subunit. *J Biol Chem* 281:22917-32.

8. Yanagiya A, Svitkin YV, Shibata S, Mikami S, Imataka H, and Sonenberg N (2009) Requirement of RNA binding of mammalian eukaryotic translation initiation factor 4GI (eIF4GI) for efficient interaction of eIF4E with the mRNA cap. *Mol Cell Biol* 29:1661-9.

9. Imataka H, Gradi A, and Sonenberg N (1998) A newly identified N-terminal amino acid sequence of human eIF4G binds poly(A)-binding protein and functions in poly(A)-dependent translation. *EMBO J* 17:7480-9.

10. Sachs AB (2000) in *Translational Control of Gene Expression* (Cold Spring Harbor, NY), p 451-480.

11. Kahvejian A, Roy G, and Sonenberg N (2001) The mRNA closed-loop model: the function of PABP and PABP-interacting proteins in mRNA translation . *Cold Spring Harbor Symp Quant Biol* 66:293-300.

12. López-Lastra M, Rivas A, and Barría MI (2005) Protein synthesis in eukaryotes: the growing biological relevance of cap-independent translation initiation. *Biol Res* 38:121-46.

13. von der Haar T, Hughes JM, Manjarul Karim M, Ptushkina M, and McCarthy JE (2002) Translation initiation and surface plasmon resonance: new technology applied to old questions. *Biochem Soc Trans* 30:155-62.

14. Jones GD, Williams EP, Place AR, Jagus R, and Bachvaroff TR (2015) The alveolate translation initiation factor 4E family reveals a custom toolkit for translational control in core dinoflagellates. *BMC Evol Biol* 15:14.

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

16. Marcotrigiano J, Gingras AC, Sonenberg N, and Burley SK (1997) Cocrystal structure of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP. *Cell* 89:951-61.

17. Matsuo H, Li H, McGuire AM, Fletcher CM, Gingras AC, Sonenberg N, and Wagner G (1997) Structure of translation factor eIF4E bound to m7GDP and interaction with 4E-binding protein. *Nat Struct Biol* 4:717-24.

18. Tomoo K, Shen X, Okabe K, Nozoe Y, Fukuhara S, Morino S, Sasaki M, Taniguchi T, Miyagawa H, Kitamura K, Miura K, and Ishida T (2003) Structural features of human initiation factor 4E, studied by X-ray crystal analyses and molecular dynamics simulations. *J Mol Biol* 328:365-83.

19. Niedzwiecka A, Marcotrigiano J, Stepinski J, Jankowska-Anyszka M, Wyslouch-Cieszynska A, Dadlez M, Gingras AC, Mak P, Darzynkiewicz E, Sonenberg N, Burley SK, and Stolarski R (2002) Biophysical studies of eIF4E cap-binding protein: recognition of mRNA 5' cap structure and synthetic fragments of eIF4G and 4E-BP1 proteins. *J Mol Biol* 319:615-35.

20. Rhoads RE (2009) eIF4E: new family members, new binding partners, new roles. *J Biol Chem* 284:16711-5.

21. Kamenska A, Simpson C, and Standart N (2014) eIF4E-binding proteins: new factors, new locations, new roles. *Biochem Soc Trans* 42:1238-45.

22. Jagus R, Bachvaroff TR, Joshi B, and Place AR (2012) Diversity of Eukaryotic Translational Initiation Factor eIF4E in Protists. *Comp Funct Genomics* 2012:134839.

23. Ptushkina M, von der Haar T, Karim MM, Hughes JM, and McCarthy JE (1999) Repressor binding to a dorsal regulatory site traps human eIF4E in a high cap-affinity state. *EMBO J* 18:4068-75.

24. Jankowska-Anyszka M, Lamphear BJ, Aamodt EJ, Harrington T, Darzynkiewicz E, Stolarski R, and 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* 273:10538-42.

25. Keiper BD, Lamphear BJ, Deshpande AM, Jankowska-Anyszka M, Aamodt EJ, Blumenthal T, and Rhoads RE (2000) Functional characterization of five eIF4E isoforms in *Caenorhabditis elegans. J Biol Chem* 275:10590-6.

26. Robalino J, Joshi B, Fahrenkrug SC, and Jagus R (2004) Two zebrafish eIF4E family members are differentially expressed and functionally divergent. *J Biol Chem* 279:10532-41.

27. Minshall N, Reiter MH, Weil D, and Standart N (2007) CPEB interacts with an ovary-specific eIF4E and 4E-T in early *Xenopus* oocytes. *J Biol Chem* 282:37389-401.

28. Standart N, and Minshall N (2008) Translational control in early development: CPEB, P-bodies and germinal granules. *Biochem Soc Trans* 36:671-6.

29. Hernández G, Han H, Gandin V, Fabian L, Ferreira T, Zuberek J, Sonenberg N, Brill JA, and Lasko P (2012) Eukaryotic initiation factor 4E-3 is essential for meiotic chromosome segregation, cytokinesis and male fertility in *Drosophila*. *Development* 139:3211-20.

30. Evsikov AV, and Marín de Evsikova C (2009) Evolutionary origin and phylogenetic analysis of the novel oocyte-specific eukaryotic translation initiation factor 4E in Tetrapoda. *Dev Genes Evol* 219:111-8.

31. Kubacka D, Miguel RN, Minshall N, Darzynkiewicz E, Standart N, and Zuberek J (2015) Distinct Features of Cap-binding by eIF4E1b Proteins. *J Mol Biol* 427:387-405.

32. Minshall N, Allison R, Marnef A, Wilczynska A, and Standart N (2010) Translational control assessed using the tethered function assay in *Xenopus* oocytes. *Methods* 51:165-9.

33. 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.

34. Rom E, Kim HC, Gingras A-, Marcotrigiano J, Favre D, Olsen H, Burley SK, and Sonenberg N (1998) Cloning and Characterization of 4EHP, a Novel Mammalian eIF4E-related Cap-binding Protein. *Journal of Biological Chemistry* 273:13104-13109.

35. Cho PF, Poulin F, Cho-Park YA, Cho-Park IB, Chicoine JD, 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* 121:411-23.

36. Dinkova TD, Keiper BD, Korneeva NL, Aamodt EJ, and Rhoads RE (2005) Translation of a small subset of *Caenorhabditis elegans* mRNAs is dependent on a specific eukaryotic translation initiation factor 4E isoform. *Mol Cell Biol* 25:100-13.

37. Villaescusa JC, Verrotti AC, Ferretti E, Farookhi R, and Blasi F (2004) Expression of Hox cofactor genes during mouse ovarian follicular development and oocyte maturation. *Gene* 330:1-7.

38. Morita M, Ler LW, Fabian MR, Siddiqui N, Mullin M, Henderson VC, Alain T, Fonseca BD, Karashchuk G, Bennett CF, Kabuta T, Higashi S, Larsson O, Topisirovic I, Smith RJ, Gingras AC, and Sonenberg N (2012) A novel 4EHP-GIGYF2 translational repressor complex is essential for mammalian development. *Mol Cell Biol* 32:3585-93.

39. Joshi B, Cameron A, and Jagus R (2004) Characterization of mammalian eIF4E-family members. *Eur J Biochem* 271:2189-203.

40. Kong J, and Lasko P (2012) Translational control in cellular and developmental processes. *Nat Rev Genet* 13:383-94.

41. Cho PF, Gamberi C, Cho-Park YA, Cho-Park IB, Lasko P, and Sonenberg N (2006) Cap-dependent translational inhibition establishes two opposing morphogen gradients in Drosophila embryos. *Curr Biol* 16:2035-41.

42. Osborne MJ, Volpon L, Kornblatt JA, Culjkovic-Kraljacic B, Baguet A, and Borden KL (2013) eIF4E3 acts as a tumor suppressor by utilizing an atypical mode of methyl-7-guanosine cap recognition. *Proc Natl Acad Sci U S A* 110:3877-82.

43. Soares RJ, Cagnin S, Chemello F, Silvestrin M, Musaro A, De Pitta C, Lanfranchi G, and Sandri M (2014) Involvement of miRNAs in the Regulation of Muscle Wasting during Catabolic Conditions. *J Biol Chem.*

44. Joshi B, Cai AL, Keiper BD, Minich WB, Mendez R, Beach CM, Stepinski J, Stolarski R, Darzynkiewicz E, and Rhoads RE (1995) Phosphorylation of eukaryotic protein synthesis initiation factor 4E at Ser-209. *J Biol Chem* 270:14597-603.

45. Zuberek J, Wyslouch-Cieszynska A, Niedzwiecka A, Dadlez M, Stepinski J, Augustyniak W, Gingras AC, Zhang Z, Burley SK, Sonenberg N, Stolarski R, and Darzynkiewicz E (2003) Phosphorylation of eIF4E attenuates its interaction with mRNA 5' cap analogs by electrostatic repulsion: intein-mediated protein ligation strategy to obtain phosphorylated protein. *RNA* 9:52-61.

46. Herdy B, Jaramillo M, Svitkin YV, Rosenfeld AB, Kobayashi M, Walsh D, Alain T, Sean P, Robichaud N, Topisirovic I, Furic L, Dowling RJ, Sylvestre A, Rong L, Colina R, Costa-Mattioli M, Fritz JH, Olivier M, Brown E, Mohr I, and Sonenberg N (2012) Translational control of the activation of transcription factor NF-κB and production of type I interferon by phosphorylation of the translation factor eIF4E. *Nat Immunol* 13:543-50.

47. Furic L, Rong L, Larsson O, Koumakpayi IH, Yoshida K, Brueschke A, Petroulakis E, Robichaud N, Pollak M, Gaboury LA, Pandolfi PP, Saad F, and Sonenberg N (2010) eIF4E phosphorylation promotes tumorigenesis and is associated with prostate cancer progression. *Proc Natl Acad Sci U S A* 107:14134-9.

48. Lachance PE, Miron M, Raught B, Sonenberg N, and Lasko P (2002) Phosphorylation of eukaryotic translation initiation factor 4E is critical for growth. *Mol Cell Biol* 22:1656-63.

49. Xu X, Vatsyayan J, Gao C, Bakkenist CJ, and Hu J (2010) Sumoylation of eIF4E activates mRNA translation. *EMBO Rep* 11:299-304.

50. Xu X, Vatsyayan J, Gao C, Bakkenist CJ, and Hu J (2010) HDAC2 promotes eIF4E sumoylation and activates mRNA translation gene specifically. *J Biol Chem* 285:18139-43.

51. Gareau JR, and Lima CD (2010) The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. *Nat Rev Mol Cell Biol* 11:861-71.

52. Chen LZ, Li XY, Huang H, Xing W, Guo W, He J, Sun ZY, Luo AX, Liang HP, Hu J, Xu X, Xu YS, and Wang ZG (2014) SUMO-2 promotes mRNA translation by enhancing interaction between eIF4E and eIF4G. *PLoS One* 9:e100457.

53. Herrmann J, Lerman LO, and Lerman A (2007) Ubiquitin and ubiquitin-like proteins in protein regulation. *Circ Res* 100:1276-91.

54. Loeb KR, and Haas AL (1992) The interferon-inducible 15-kDa ubiquitin homolog conjugates to intracellular proteins. *J Biol Chem* 267:7806-13.

55. D'Cunha J, Ramanujam S, Wagner RJ, Witt PL, Knight E, and Borden EC (1996) *In vitro* and *in vivo* secretion of human ISG15, an IFN-induced immunomodulatory cytokine. *J Immunol* 157:4100-8.

56. Sloper-Mould KE, Jemc JC, Pickart CM, and Hicke L (2001) Distinct functional surface regions on ubiquitin. *J Biol Chem* 276:30483-9.

57. Zhao C, Denison C, Huibregtse JM, Gygi S, and Krug RM (2005) Human ISG15 conjugation targets both IFN-induced and constitutively expressed proteins functioning in diverse cellular pathways. *Proc Natl Acad Sci U S A* 102:10200-5.

58. Langevin C, van der Aa LM, Houel A, Torhy C, Briolat V, Lunazzi A, Harmache A, Bremont M, Levraud JP, and Boudinot P (2013) Zebrafish ISG15 exerts a strong antiviral activity against RNA and DNA viruses and regulates the interferon response. *J Virol* 87:10025-36.

59. Zhang D, and Zhang DE (2011) Interferon-stimulated gene 15 and the protein ISGylation system. *J Interferon Cytokine Res* 31:119-30.

60. Liu M, Reimschuessel R, and Hassel BA (2002) Molecular cloning of the fish interferon stimulated gene, 15 kDa (ISG15) orthologue: a ubiquitin-like gene induced by nephrotoxic damage. *Gene* 298:129-39.

61. Wang W, Zhang M, Xiao ZZ, and Sun L (2012) *Cynoglossus semilaevis* ISG15: a secreted cytokine-like protein that stimulates antiviral immune response in a LRGG motif-dependent manner. *PLoS One* 7:e44884.

62. Huang J, Zhao S, Zhu M, Wu Z, and Yu M (2009) Sequence and expression analyses of porcine ISG15 and ISG43 genes. *Comp Biochem Physiol B Biochem Mol Biol* 153:301-9.

63. Okumura F, Zou W, and Zhang DE (2007) ISG15 modification of the eIF4E cognate 4EHP enhances cap structure-binding activity of 4EHP. *Genes Dev* 21:255-60.

64. Paku KS, 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* 441:237-45.

65. Siddiqui N, Tempel W, Nedyalkova L, Volpon L, Wernimont AK, Osborne MJ, Park HW, and Borden KL (2012) Structural insights into the allosteric effects of 4EBP1 on the eukaryotic translation initiation factor eIF4E. *J Mol Biol* 415:781-92.

66. Lin TA, Kong X, Haystead TA, Pause A, Belsham G, Sonenberg N, and Lawrence JC (1994) PHAS-I as a link between mitogen-activated protein kinase and translation initiation. *Science* 266:653-6.

67. Pause A, Belsham GJ, Gingras AC, Donzé O, Lin TA, Lawrence JC, and Sonenberg N (1994) Insulin-dependent stimulation of protein synthesis by phosphorylation of a regulator of 5'-cap function. *Nature* 371:762-7.

68. Wang X, and Proud CG (2006) The mTOR pathway in the control of protein synthesis. *Physiology (Bethesda)* 21:362-9.

69. Topisirovic I, and Sonenberg N (2011) Translational control by the eukaryotic ribosome. *Cell* 145:333-4.

70. Rhoads RE, Dinkova TD, and Jagus R (2007) Approaches for analyzing the differential activities and functions of eIF4E family members. *Methods Enzymol* 429:261-97.

71. Richter JD, and Sonenberg N (2005) Regulation of cap-dependent translation by eIF4E inhibitory proteins. *Nature* 433:477-80.

72. Dostie J, Ferraiuolo M, Pause A, Adam SA, and Sonenberg N (2000) A novel shuttling protein, 4E-T, mediates the nuclear import of the mRNA 5' cap-binding protein, eIF4E. *EMBO J* 19:3142-56.

73. Sonenberg N (2008) eIF4E, the mRNA cap-binding protein: from basic discovery to translational research. *Biochem Cell Biol* 86:178-83.

74. Gosselin P, Martineau Y, Morales J, Czjzek M, Glippa V, Gauffeny I, Morin E, Le Corguillé G, Pyronnet S, Cormier P, and Cosson B (2013) Tracking a refined eIF4E-binding motif reveals Angel1 as a new partner of eIF4E. *Nucleic Acids Res* 41:7783-92.

75. Corey DR, and Abrams JM (2001) Morpholino antisense oligonucleotides: tools for investigating vertebrate development. *Genome Biol* 2:REVIEWS1015.

76. Hwang WY, Fu Y, Reyon D, Maeder ML, Tsai SQ, Sander JD, Peterson RT, Yeh JR, and Joung JK (2013) Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol* 31:227-9.

77. Hwang WY, Fu Y, Reyon D, Maeder ML, Kaini P, Sander JD, Joung JK, Peterson RT, and Yeh JR (2013) Heritable and precise zebrafish genome editing using a CRISPR-Cas system. *PLoS One* 8:e68708.

78. Auer TO, and Del Bene F (2014) CRISPR/Cas9 and TALEN-mediated knock-in approaches in zebrafish. *Methods* 69:142-50.

79. Bradbury J (2004) Small fish, big science. *PLoS Biol* 2:E148.

80. Dooley K, and Zon LI (2000) Zebrafish: a model system for the study of human disease. *Curr Opin Genet Dev* 10:252-6.

81. Amsterdam A, and Hopkins N (2006) Mutagenesis strategies in zebrafish for identifying genes involved in development and disease. *Trends Genet* 22:473-8.

82. Kari G, Rodeck U, and Dicker AP (2007) Zebrafish: an emerging model system for human disease and drug discovery. *Clin Pharmacol Ther* 82:70-80.

83. Ackermann GE, and Paw BH (2003) Zebrafish: a genetic model for vertebrate organogenesis and human disorders. *Front Biosci* 8:d1227-53.

84. Sémon M, and Wolfe KH (2007) Reciprocal gene loss between *Tetraodo*n and zebrafish after whole genome duplication in their ancestor. *Trends Genet* 23:108-12.

85. Sémon M, and Wolfe KH (2007) Rearrangement rate following the whole-genome duplication in teleosts. *Mol Biol Evol* 24:860-7.

86. Postlethwait JH, Yan YL, Gates MA, Horne S, Amores A, Brownlie A, Donovan A, Egan ES, Force A, Gong Z, Goutel C, Fritz A, Kelsh R, Knapik E, Liao E, Paw B, Ransom D, Singer A, Thomson M, Abduljabbar TS, Yelick P, Beier D, Joly JS, Larhammar D, Rosa F, Westerfield M, Zon LI, Johnson SL, and Talbot WS (1998) Vertebrate genome evolution and the zebrafish gene map. *Nat Genet* 18:345-9.

87. Brunet FG, Roest Crollius H, Paris M, Aury JM, Gibert P, Jaillon O, Laudet V, and Robinson-Rechavi M (2006) Gene loss and evolutionary rates following whole-genome duplication in teleost fishes. *Mol Biol Evol* 23:1808-16.

88. Joshi B, Robalino J, Schott EJ, and Jagus R (2002) Yeast "knockout-and-rescue" system for identification of eIF4E-family members possessing eIF4E-activity. *Biotechniques* 33:392-3, 395-6, 398 passim.

89. Maloof AC, Porter SM, Moore JL, Dudas FO, Bowring SA, Higgins JA, Fike DA, and Eddy MP (2010) The earliest Cambrian record of animals and ocean geochemical change. *Geol Soc Bull* 122:1731-1774.

90. Budd GE (2013) At the origin of animals: the revolutionary cambrian fossil record. *Curr Genomics* 14:344-54.

91. Erwin DH, Laflamme M, Tweedt SM, Sperling EA, Pisani D, and Peterson KJ (2011) The Cambrian conundrum: early divergence and later ecological success in the early history of animals. *Science* 334:1091-7.

92. Dornbos SQ, and Bottjer DJ (2000) Evolutionary paleoecology of the earliest echinoderms: Helicoplacoids and the Cambrian substrate revolution. *Geology* 28:839-842.

93. Marlétaz F, Holland LZ, Laudet V, and Schubert M (2006) Retinoic acid signaling and the evolution of chordates. *Int J Biol Sci* 2:38-47.

94. Blair JE, and Hedges SB (2005) Molecular phylogeny and divergence times of deuterostome animals. *Mol Biol Evol* 22:2275-84.

95. Zhang J (2003) Evolution by gene duplication: an update. *Trends in Ecology & Evolution* 18:292-298.

96. Meyer A, and Schartl M (1999) Gene and genome duplications in vertebrates: the one-to-four (-to-eight in fish) rule and the evolution of novel gene functions. *Curr Opin Cell Biol* 11:699-704.

97. Smith JJ, Kuraku S, Holt C, Sauka-Spengler T, Jiang N, Campbell MS, Yandell MD, Manousaki T, Meyer A, Bloom OE, Morgan JR, Buxbaum JD, Sachidanandam R, Sims C, Garruss AS, Cook M, Krumlauf R, Wiedemann LM, Sower SA, Decatur WA, Hall JA, Amemiya CT, Saha NR, Buckley KM, Rast JP, Das S, Hirano M, McCurley N, Guo P, Rohner N, Tabin CJ, Piccinelli P, Elgar G, Ruffier M, Aken BL, Searle SM, Muffato M, Pignatelli M, Herrero J, Jones M, Brown CT, Chung-Davidson YW, Nanlohy KG, Libants SV, Yeh CY, McCauley DW, Langeland JA, Pancer Z, Fritzsch B, de Jong PJ, Zhu B, Fulton LL, Theising B, Flicek P, Bronner ME, Warren WC, Clifton SW, Wilson RK, and Li W (2013) Sequencing of the sea lamprey (*Petromyzon marinus*) genome provides insights into vertebrate evolution. *Nat Genet* 45:415-21, 421e1-2.

98. Postlethwait J, Amores A, Cresko W, Singer A, and Yan YL (2004) Subfunction partitioning, the teleost radiation and the annotation of the human genome. *Trends Genet* 20:481-90.

99. Clark MS (2003) Genomics and mapping of teleostei (bony fish). *Comp Funct Genomics* 4:182-93.

100. Ulloa PE, Iturra P, and Neira C (2015) Zebrafish as a model organism for nutrition and growth: towards comparative studies of nutritional genomics applied to aquacultured fish. *Rev Fish Biol Fisheries.*

101. Betancur-R R, Broughton RE, Wiley EO, Carpenter K, López JA, Li C, Holcroft NI, Arcila D, Sanciangco M, Cureton Ii JC, Zhang F, Buser T, Campbell MA, Ballesteros JA, Roa-Varon A, Willis S, Borden WC, Rowley T, Reneau PC, Hough DJ, Lu G, Grande T, Arratia G, and Ortí G (2013) The tree of life and a new classification of bony fishes. *PLoS Curr* 5.

102. Vandepoele K, De Vos W, Taylor JS, Meyer A, and Van de Peer Y (2004) Major events in the genome evolution of vertebrates: paranome age and size differ considerably between ray-finned fishes and land vertebrates. *Proc Natl Acad Sci U S A* 101:1638-43.

103. Christoffels A, Brenner S, and Venkatesh B (2006) Tetraodon genome analysis provides further evidence for whole-genome duplication in the ray-finned fish lineage. *Comp Biochem Physiol Part D Genomics Proteomics* 1:13-9.

104. Volff JN (2005) Genome evolution and biodiversity in teleost fish. *Heredity (Edinb)* 94:280-94.

105. Glasauer SM, and Neuhauss SC (2014) Whole-genome duplication in teleost fishes and its evolutionary consequences. *Mol Genet Genomics* 289:1045-60.

106. Lauder G, and Liem K (1989) The evolution and interrelationships of the Actinopterygian fishes. *Bull Mus Comp Zool* 150:95-187.

107. Dahm R, and Geisler R (2006) Learning from small fry: the zebrafish as a genetic model organism for aquaculture fish species. *Mar Biotechnol (NY)* 8:329-45.

108. Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M, Claverie JM, and Gascuel O (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* 36:W465-9.

109. Blumenthal T (1998) Gene clusters and polycistronic transcription in eukaryotes. *Bioessays* 20:480-7.

110. Gherman A, Wang R, and Avramopoulos D (2009) Orientation, distance, regulation and function of neighbouring genes. *Hum Genomics* 3:143-56.

111. Ng YK, Wu W, and Zhang L (2009) Positive correlation between gene coexpression and positional clustering in the zebrafish genome. *BMC Genomics* 10:42.

112. Hershey JWB, and Merrick WC (2000) in *Translational control of gene expression* (Cold Spring Harbor, New York), p 33-88.

113. Aitken CE, and Lorsch JR (2012) A mechanistic overview of translation initiation in eukaryotes. *Nat Struct Mol Biol* 19:568-76.

114. Gingras AC, Raught B, and Sonenberg N (1999) eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. *Annu Rev Biochem* 68:913-63.

115. Marintchev A, Edmonds KA, Marintcheva B, Hendrickson E, Oberer M, Suzuki C, Herdy B, Sonenberg N, and Wagner G (2009) Topology and regulation of the human eIF4A/4G/4H helicase complex in translation initiation. *Cell* 136:447-60.

116. Hernández G, and Vazquez-Pianzola P (2005) Functional diversity of the eukaryotic translation initiation factors belonging to eIF4 families. *Mech Dev* 122:865-76.

117. Patrick RM, and Browning KS (2012) The eIF4F and eIFi*so*4F Complexes of Plants: An Evolutionary Perspective. *Comp Funct Genomics* 2012:287814.

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

119. Dehal P, and Boore JL (2005) Two rounds of whole genome duplication in the ancestral vertebrate. *PLoS Biol* 3:e314.

120. Lynch M, and Conery JS (2000) The evolutionary fate and consequences of duplicate genes. *Science* 290:1151-5.

121. Lynch M, and Conery JS (2003) The evolutionary demography of duplicate genes. *J Struct Funct Genomics* 3:35-44.

122. Sémon M, and Wolfe KH (2007) Consequences of genome duplication. *Curr Opin Genet Dev* 17:505-12.

123. Holland LZ (2013) Evolution of new characters after whole genome duplications: insights from *amphioxus*. *Semin Cell Dev Biol* 24:101-9.

124. Tee AR, Tee JA, and Blenis J (2004) Characterizing the interaction of the mammalian eIF4E-related protein 4EHP with 4E-BP1. *FEBS Lett* 564:58-62.

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

126. Lasko P (2011) Posttranscriptional regulation in *Drosophila* oocytes and early embryos. *Wiley Interdiscip Rev RNA* 2:408-16.

127. Westerfield M Embryonic and larval culture. *The Zebrafish Book, 5th edition.* Available at: https://zfin.org/zf_info/zfbook/zfbk.html [Accessed 2007].

128. Fling SP, and Gregerson DS (1986) Peptide and protein molecular weight determination by electrophoresis using a high-molarity tris buffer system without urea. *Anal Biochem* 155:83-8.

129. Mumberg D, Müller R, and Funk M (1994) Regulatable promoters of *Saccharomyces cerevisiae*: comparison of transcriptional activity and their use for heterologous expression. *Nucleic Acids Res* 22:5767-8.

130. Gietz RD, and Woods RA (2002) Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol* 350:87-96.

131. Keogh MC, Kim JA, Downey M, Fillingham J, Chowdhury D, Harrison JC, Onishi M, Datta N, Galicia S, Emili A, Lieberman J, Shen X, Buratowski S, Haber JE, Durocher D, Greenblatt JF, and Krogan NJ (2006) A phosphatase complex that dephosphorylates gammaH2AX regulates DNA damage checkpoint recovery. *Nature* 439:497-501.

132. Schneider TD, and Stephens RM (1990) Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res* 18:6097-100.

133. Crooks GE, Hon G, Chandonia JM, and Brenner SE (2004) WebLogo: a sequence logo generator. *Genome Res* 14:1188-90.

134. Singh SK, Sundaram CS, Shanbhag S, and Idris MM (2010) Proteomic profile of zebrafish brain based on two-dimensional gel electrophoresis matrix-assisted laser desorption/ionization MS/MS analysis. *Zebrafish* 7:169-77.

135. Link V, Shevchenko A, and Heisenberg CP (2006) Proteomics of early zebrafish embryos. *BMC Dev Biol* 6:1.

136. Fletcher CM, McGuire AM, Gingras AC, Li H, Matsuo H, Sonenberg N, and Wagner G (1998) 4E binding proteins inhibit the translation factor eIF4E without folded structure. *Biochemistry* 37:9-15.

137. 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* 39:3496-503.

138. Siddiqui N, Tempel W, Nedyalkova L, Volpon L, Wernimont AK, Osborne MJ, Park HW, and Borden KL (2012) Structural insights into the allosteric effects of 4EBP1 on the eukaryotic translation initiation factor eIF4E. *J Mol Biol* 415:781-92.

139. Papadopoulos E, Jenni S, Kabha E, Takrouri KJ, Yi T, Salvi N, Luna RE, Gavathiotis E, Mahalingam P, Arthanari H, Rodriguez-Mias R, Yefidoff-Freedman R, Aktas BH, Chorev M, Halperin JA, and Wagner G (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.*

140. Gingras AC, Raught B, Gygi SP, Niedzwiecka A, Miron M, Burley SK, Polakiewicz RD, Wyslouch-Cieszynska A, Aebersold R, and Sonenberg N (2001) Hierarchical phosphorylation of the translation inhibitor 4E-BP1. *Genes Dev* 15:2852-64.

141. Oulhen N, Boulben S, Bidinosti M, Morales J, Cormier P, and Cosson B (2009) A variant mimicking hyperphosphorylated 4E-BP inhibits protein synthesis in a sea urchin cell-free, cap-dependent translation system. *PLoS One* 4:e5070.

142. Hidalgo M, Le Bouffant R, Bello V, Buisson N, Cormier P, Beaudry M, and Darribère T (2012) The translational repressor 4E-BP mediates hypoxia-induced defects in myotome cells. *J Cell Sci* 125:3989-4000.

143. Yogev O, Williams VC, Hinits Y, and Hughes SM (2013) eIF4EBP3L acts as a gatekeeper of TORC1 in activity-dependent muscle growth by specifically regulating Mef2ca translational initiation. *PLoS Biol* 11:e1001679.

144. Kuraku S (2008) Insights into cyclostome phylogenomics: pre-2R or post-2R. *Zoolog Sci* 25:960-8.

145. Kuraku S, Meyer A, and Kuratani S (2009) Timing of genome duplications relative to the origin of the vertebrates: did cyclostomes diverge before or after? *Mol Biol Evol* 26:47-59.

146. Smith JJ, Saha NR, and Amemiya CT (2010) Genome biology of the cyclostomes and insights into the evolutionary biology of vertebrate genomes. *Integr Comp Biol* 50:130-7.

147. Volpon L, Osborne MJ, Culjkovic-Kraljacic B, and Borden KL (2013) eIF4E3, a new actor in mRNA metabolism and tumor suppression. *Cell Cycle* 12:1159-60.

148. Rom E, Kim HC, Gingras A-, Marcotrigiano J, Favre D, Olsen H, Burley SK, and Sonenberg N (1998) Cloning and Characterization of 4EHP, a Novel Mammalian eIF4E-related Cap-binding Protein. *Journal of Biological Chemistry* 273:13104-13109.

149. Zuberek J, Kubacka D, Jablonowska A, Jemielity J, Stepinski J, Sonenberg N, and Darzynkiewicz E (2007) Weak binding affinity of human 4EHP for mRNA cap analogs. *RNA* 13:691-7.

150. Ferraiuolo MA, Basak S, Dostie J, Murray EL, Schoenberg DR, and Sonenberg N (2005) A role for the eIF4E-binding protein 4E-T in P-body formation and mRNA decay. *J Cell Biol* 170:913-24.

151. Andrei MA, Ingelfinger D, Heintzmann R, Achsel T, Rivera-Pomar R, and Lührmann R (2005) A role for eIF4E and eIF4E-transporter in targeting mRNPs to mammalian processing bodies. *RNA* 11:717-27.

152. Lee HC, Cho H, and Kim YK (2008) Ectopic expression of eIF4E-transporter triggers the movement of eIF4E into P-bodies, inhibiting steady-state translation but not the pioneer round of translation. *Biochem Biophys Res Commun* 369:1160-5.

153. 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.

154. Kulkarni M, Ozgur S, and Stoecklin G (2010) On track with P-bodies. *Biochem Soc Trans* 38:242-51.

155. Parker R, and Sheth U (2007) P bodies and the control of mRNA translation and degradation. *Mol Cell* 25:635-46.

156. Flicek P, Ahmed I, Amode MR, Barrell D, Beal K, Brent S, Carvalho-Silva D, Clapham P, Coates G, Fairley S, Fitzgerald S, Gil L, García-Girón C, Gordon L, Hourlier T, Hunt S, Juettemann T, Kähäri AK, Keenan S, Komorowska M, Kulesha E, Longden I, Maurel T, McLaren WM, Muffato M, Nag R, Overduin B, Pignatelli M, Pritchard B, Pritchard E, Riat HS, Ritchie GR, Ruffier M, Schuster M, Sheppard D, Sobral D, Taylor K, Thormann A, Trevanion S, White S, Wilder SP, Aken BL, Birney E, Cunningham F, Dunham I, Harrow J, Herrero J, Hubbard TJ, Johnson N, Kinsella R, Parker A, Spudich G, Yates A, Zadissa A, and Searle SM (2013) Ensembl 2013. *Nucleic Acids Res* 41:D48-55.

157. Lauermann V (1991) Ethanol improves the transformation efficiency of intact yeast cells. *Curr Genet* 20:1-3.

158. Hausmann S, Zheng S, Costanzo M, Brost RL, Garcin D, Boone C, Shuman S, and Schwer B (2008) Genetic and biochemical analysis of yeast and human cap trimethylguanosine synthase: functional overlap of 2,2,7-trimethylguanosine caps, small nuclear ribonucleoprotein components, pre-mRNA splicing factors, and RNA decay pathways. *J Biol Chem* 283:31706-18.

159. Yu YT, Shu MD, and Steitz JA (1998) Modifications of U2 snRNA are required for snRNP assembly and pre-mRNA splicing. *EMBO J* 17:5783-95.

160. Yanagiya A, Suyama E, Adachi H, Svitkin YV, Aza-Blanc P, Imataka H, Mikami S, Martineau Y, Ronai ZA, and Sonenberg N (2012) Translational homeostasis via the mRNA cap-binding protein, eIF4E. *Mol Cell* 46:847-58.

161. Thoreen CC, Chantranupong L, Keys HR, Wang T, Gray NS, and Sabatini DM (2012) A unifying model for mTORC1-mediated regulation of mRNA translation. *Nature* 485:109-13.

162. Fonseca BD, Smith EM, Yelle N, Alain T, Bushell M, and Pause A (2014) The ever-evolving role of mTOR in translation. *Semin Cell Dev Biol* 36:102-12.

163. Reijns MA, Alexander RD, Spiller MP, and Beggs JD (2008) A role for Q/N-rich aggregation-prone regions in P-body localization. *J Cell Sci* 121:2463-72.

164. Strzelecka M, Trowitzsch S, Weber G, Lührmann R, Oates AC, and Neugebauer KM (2010) Coilin-dependent snRNP assembly is essential for zebrafish embryogenesis. *Nat Struct Mol Biol* 17:403-9.

165. Osborne MJ, and Borden KL (2015) The eukaryotic translation initiation factor eIF4E in the nucleus: taking the road less traveled. *Immunol Rev* 263:210-23.

166. Bill BR, Petzold AM, Clark KJ, Schimmenti LA, and Ekker SC (2009) A primer for morpholino use in zebrafish. *Zebrafish* 6:69-77.

167. Sung YH, Kim JM, Kim HT, Lee J, Jeon J, Jin Y, Choi JH, Ban YH, Ha SJ, Kim CH, Lee HW, and Kim JS (2014) Highly efficient gene knockout in mice and zebrafish with RNA-guided endonucleases. *Genome Res* 24:125-31.

168. Ablain J, Durand EM, Yang S, Zhou Y, and Zon LI (2015) A CRISPR/Cas9 Vector System for Tissue-Specific Gene Disruption in Zebrafish. *Dev Cell* 32:756-64.

169. Kress TL, Yoon YJ, and Mowry KL (2004) Nuclear RNP complex assembly initiates cytoplasmic RNA localization. *J Cell Biol* 165:203-11.

170. Battle DJ, Kasim M, Yong J, Lotti F, Lau CK, Mouaikel J, Zhang Z, Han K, Wan L, and Dreyfuss G (2006) The SMN complex: an assembly machine for RNPs. *Cold Spring Harb Symp Quant Biol* 71:313-20.

171. Anderson P, and Kedersha N (2006) RNA granules. *J Cell Biol* 172:803-8.

172. Strzelecka M, Oates AC, and Neugebauer KM (2010) Dynamic control of Cajal body number during zebrafish embryogenesis. *Nucleus* 1:96-108.

173. Schisa JA (2012) New insights into the regulation of RNP granule assembly in oocytes. *Int Rev Cell Mol Biol* 295:233-89.

174. Buchan JR (2014) mRNP granules. *RNA Biol* 11:1019-30.

175. Li X, Lan Y, Xu J, Zhang W, and Wen Z (2012) SUMO1-activating enzyme subunit 1 is essential for the survival of hematopoietic stem/progenitor cells in zebrafish. *Development* 139:4321-9.

176. Rodriguez MS, Dargemont C, and Hay RT (2001) SUMO-1 conjugation in vivo requires both a consensus modification motif and nuclear targeting. *J Biol Chem* 276:12654-9.

177. Hsieh AC, and Ruggero D (2010) Targeting eukaryotic translation initiation factor 4E (eIF4E) in cancer. *Clin Cancer Res* 16:4914-20.

178. Avdulov S, Herrera J, Smith K, Peterson M, Gomez-Garcia JR, Beadnell TC, Schwertfeger KL, Benyumov AO, Manivel JC, Li S, Bielinsky AK, Yee D, Bitterman PB, and Polunovsky VA (2015) eIF4E threshold levels differ in governing normal and neoplastic expansion of mammary stem and luminal progenitor cells. *Cancer Res* 75:687-97.

179. Yaruin A, Harris RE, Mark PA, and Ashe HL (2011) Patterning of the *Drosophila* oocyte by a sequential translation repression program involving the d4EHP and Belle translational repressors. *RNA Biol* 8: 904-12.

180. Ma MX, Blenis J (2009) Molecular mechanisms of mTOR-mediated translational control. *Nat Rev Mol Cell Biol* 10: 307-18.

181. Lercher MJ, Urrutia AO, and Hurst LD (2002) Clustering of housekeeping genes provides a unified model of gene order in the human genome. *Nat Genet* 31: 180-3

APPENDICES

Table A2.1: Sequences of eIF4E family members

KEY
Joshi EST
NCBI
Embl
Hmmr
other database

organism name				eIF4E-1A		eIF4E-1B		eIF4E-1C
Tetrapod	**common name**	**abbrev**	**Accession**	**seq**	**Accession**	**seq**	**Accession**	**seq**
Homo sapiens	human	Hs	NP_001959.1, NM_001968.3	MATVEPETTPTPNPPTTEEEKTESNQEVANPEHYIKHPLQNRWALWFFKNDKSKTWQANLRLISKFDTVEDFWALYNHIQLSSNLMPGCDYSLFKDGIEPMWEDEKNKRGGRWLITLNKQQRRSDLDRFWLETLLCLIGESFDDYSDDVCGAVVNVRAKGDKIAWTTECENREAVTHIGRVYKERLGLPPKIVIGYQSHADTATKSGSTTKNRFVV	NP_001092878.1,NM_001099408.1	MLAVEVSEAEGGIREWEEEEKEEEAAERTPTGEKSPNSPRTLLSLRGKARTGGPMEVKLELHPLQNRWALWFFKNDRSRAWQDNLHLVTKVDTVEDFWALYSHIQLASKLSSGCDYALFKDGIQPMWEDSRNKRGGRWLVSLAKQQRHIELDRLWLETLLCLIGESFEEHSREVCGAVVNIRTKGDKIAWWTREAENQAGVLHVGRVYKERLGLSPKTIIGYQAHADTATKSNSLAKNKFVV	none found	
avian								
Taeniopygia guttata	zebrafinch	Tg	NP_001232122.1, NM_001245193.1.	MAAVEPETTPNPQPAEEEKTEPTPSQEVASPEQYIKHPLQNRWALWFFKNDKSKTWQANLRLISKFDTVEDFWALYNHIQLSSNLMPGCDYSLFKDGIEPMWEDEKNKRGGRWLITLNKQQRRSDLDRFWLETLLCLIGESFDDYSDDVCGAVVNVRTKGDKIAWTTECENRDAVTHIGRVYKERLGLPPKIVIGYQSHADTATKSGSTTKNRFVV	XM_002194165.1,XP_002194201.1	MATGEQRQQERRRQRARQQELLPAEILGKHPLQNRWALWFFKNDKSKMWQANLRLVTKFSTVEDFWALYSHIQLASKLTAGCDYSLFKDGIEPMWEDSQNKRGGRWLITLAKQQRHTELDRFWLDTLLCLIGEMFDEYSDEVCGAVINIRAKGDKIAIWTREADNQEGVTHIGRVYKEHLGLSQKVAIGYQAHADTATKSSSLAKTKFVM	none found	
Gallus gallus	chicken	Gg	XP_420655.2, XP_420655.4	NLLVNPFFKQETTPNPQPSEEEKTEPAPTQEVASPEQYIKHPLQNRWALWFFKNDKSKTWQANLRLISKFDTVEDFWALYNHIQLSSNLMPGCDYSLFKDGIEPMWEDEKNKRGGRWLITLTKQQRRSDLDRFWLETLLCLIGESFDDYSDDVCGAVVNVRTKGDKIAIWTTECENRDAVTHIGRVYKERLGLPPKIVIGYQSHADTATKSGSTTKNRFVV	BX931053.2,translated	PIWVAMLFLQRRQEERRRRRAQQQELQVAESLGKHPLHNRWALWFFKNDKSKMWQANLRLVTKFSTVEDFWALYTHIQLASKLTSGCDYSLFKDGIEPMWEDNQNKRGGRWLITLAKQQRHTELDRFWLETLLCLIGEMFDEYSDEVCGAVINIRAKGDKIAIWTREAENRGVTHIGRIYKEHLGLSQKVAIGYQAHADTATKSGSLTKNKFVVDWGGGSMGGAGPPLPDPTMNGFS	none found	
reptile								
Anolis carolinensis	lizard	Ac	XP_003225575.1	MATVEPETTSNPQTSEEEKTETPASQEVVSPEPYIKHPLQNRWALWFFKNDKSKTWQANLRLISKFDTVEDFWALYNHIQLSSNLMPGCDYSLFKDGIEPMWEDEKNKRGGRWLITLNKQQRRSDLDRFWLETLLCLIGESFDDYSDDVCGAVVNVRAKGDKIAIWTTECENRDAVTHIGRVYKERLGLPPKIVIGYQSHADTATKSGSTTKNRFVV	XP_003227829	MAAAMMVNTHEEQQKHKTERGEVVMEHITKHPLQNRWALWFFKNDKSKTWQDNLQLVTKFDTVEDFWALYSHIQLASKLTSGCDYSLFKDGIEPMWEDNWNKCGGRWLITLAKQQRHTELDRFWLETLLCLIGEMFSDYSDDVCGAVINIRTKGDKIAIWTREAENQDAVIHIGRIYKEHLGLSSKVVIGYQAHADTATKSGSLMKNKFVV	none found	
amphibian								
xenopus laevis	clawed frog	Xl	NP_001089212.1	MAAVEPENPNPQSTEEEKETGQEIVSPDQYIKHPLQNRWALWFFKNDKSKTWQANLRLISKFDTVEDFWALYNHIQLSSNLMSGCDYSLFKDGIEPMWEDEKNKRGGRWLITLNKQQRRNDLDRFWLETLMCLIGESFDEHSDDVCGAVVNIRAKGDKIAIWTTECDNRDAITHIGRVYKERLGLPAKVVIGFQSHADTATKSGSTTKNRFVV	**BQ398016,UniProtKB/Swiss-Prot: P48597.1**	MAAVEPENTNPQSTEEEKETGQEIVSPDQYIKHPLQNRWALWFFKNDKSKTWQANLRLISKFDTVEDFWALYNHIQLSSNLMSGCDYSLFKDGIEPMWEDEKNKRGGRWLITLNKQQRRNDLDRFWLETLMCLIGESFDEHSDDVCGAVVNVRAKGDKIAIWTTEFENKDAVTHIGRVYKERLGLPAKVVIGYQSHADTATKSGSTTKNRFVV	none found	
Xenopus tropicalis	clawed frog	Xt	CAJ83126.1,NP_001015909.1	MASVEPEGTNPQSTEEEKTETSQEIVSPDQYIKHPLQNRWALWFFKNDKSKTWQANLRLISKFDTVEDFWALYNHIQLSSNLMSGCDYSLFKDGIEPMWEDEKNKRGGRWLITLNKQQRRNDLDRFWLETLMCLIGESFDEYSDDVCGAVVNVRAKGDKIAIWTTEFENRDAVTHIGKVYKERLGLPAKVVIGYQSHADTATKSGSTTKNRFVV	XP_002936991.2,AAI54955.1	MAAAEAISIKELPREKLDNEKRRKKKESVILEKVIKHSLQSRWALWFFKNVKSQPWQCNLRLVTTFNTVEDFWSLYTHIQLASKLSSGCDYSLFKDGIEPMWEDSRNKRGGRWLITLSKQQRHSDLDALWLETLLCLIGEAFDEYSEEVCGAVINIRAKGDKIAIWTRETENREAVTHIGKVYKERLGLSSKVVIGYQAHADTATKSSSLSKNKFVV*	none found	

organism name				eIF4E-1A		eIF4E-1B		eIF4E-1C
urchin	**common name**	**abbrev**	**Accession**	**seq**	**Accession**	**seq**	**Accession**	**seq**
Strongylocentrotus purpuratus	sea urchin	Sp	CAM57099.1	MASVGVAKLPLHEDIVKDGKGIKDLQTQDEDGANQVVDPESLIKHPLQSRWSMWFFKNDKAKSWTENLRTVTAFDTVEDFWALYNHIQVASRITSGCDYSLFKEGVKPMWEDDKNKDGGRWLIGFDRKSKPQDIDRCWLETMLLMVGESFDDDSDLVNGAVVNIRSKGNKIAMWTGDWRKEDSITNIGRKFKERLGLPAKYSIGYEAHKDTMTKTGSMAKSLYTV	none found		none found	
tunicate								
ciona intestinalis	tunicate	Ci	XP_002126987.1	MSESDKSVKVSLDPGARPKSKESQEEKEVKTVSLKEDVIEEKPMGDPEDCIKHPLQNKWALWFFKNDRQKKWEDNLRLVTKFDTVEDFWALYNHIQLSSKLQSGCDYNLFKDGIQPMWEDKANKKGGKWMLQMSKQQRVNELDKVWLELLCLIGEGFGEDSDYVNGGVVQVRHKGDKVAIWTSDYKHREGIMNIGQICKTRLGLPKKAVLGFQAHEDTMSKSGSTVKTLWSV	none found		none found	
Amphioxus								
Branchiostoma floridea	lancelets	Bf	jgi|Brafl1|120045|	MASAEPAKQPQESPLEAAAVKTEEEKDEAAGAAAAKLDTDDLKYIKHPLQNRWALWFFKNDKTKTWAANLRCVSTFDTVEDFWALYNHIQVASRLQSGCDYSLFKDGVEPMWEDAFNKTGGRWLINIQKQQRHSDLDRFWLETLLCLIGEAFEEDSDEVCGAVINVRGKGDKIAIWTHDCKNSDAVIRIGRKFKERLNLPPKFVIGYQAHTDTMSKSGSTTKNRFSV*	none found		none found	

				eIF4E-1A		eIF4E-1B		eIF4E-1C
organism name	**common name**	**abbrev**	**Accession**	**seq**	**Accession**	**seq**	**Accession**	**seq**
shark								
Callorhinchus milii	elephant shark	Cm	V9KV53_CALMI		V9L709_CALMI		V9KN06_CALMI	
			XP_007897190.1. XM_007898999.1.	MATVESEATTQSQPPEEEGKTEAASPEIVKPEHYIKHPLQNRWALWFFKNDKNKTWQANLRLISKFDTVEDFWALYNHIQLCSNLISGCDYSLFKDGIEPMWEDEKNKRGGRWLLTLSKQQRKTDLDRIWLETLLCLIGEAFDEHSDDACGAVVNVRNKGDKIAIWTTDCENRESITYIGRIYKERLGLPPKVVIGYQSHTDTATKSGSTTKNKYVV	XP_007904436.1. XM_007906245.1.	MAMAHLVHLTGLQEPEEEEGFRPSSKASPTLEPYLKHPLQNRWTLWFFKNDKSKAWQANLRLVTKFDTVEDFWALYNHIQLSSKLMPGCDYSLFKDGIEPMWEDKQNKCGGRWLITLSKQQRLVELDQFWLETLLCLIGEAFDHYSDDVCGAVTNIRPKGDKIALWTRDTENRDAVLHIGKIYKEKLGLPVKVVIGYQAHADTATKSGLGIQNKFVV	XP_007897295.1. XM_007899104.1.	MAAAEPATGGGPEVTEVKEIDPVIAEVPSDAVTQDRYLKHPLQNRWALWYFKNDKTKSWTENLRLIAKFDTVEDFWALYNHIQQPSKLLFGCDYCLFKDGIKPMWEDDKNKKGGRWLMTLTKQQRHNDLDRYWLETLLCLIGEAFDEHSDEVCGAVVNVRPKGDKISIWTGNCQSREAVTSIGQSYKERLGLPMKALIGYQSHDDTSSKSGSTTKNMYTV
lobe fin fish								
Latimeria chalumnae	coelacanth	Lc	http://www.ensembl.org/id/ENSLACT00000010262	RTISHLQETTAKPPIKQEEKTEATGQEVVNPESYIKHPLQNRWALWFFKNDKSKTWQANLRLISKFDTVEDFWALYNHIQLSSNLMSGCDYSLFKDGIEPMWEDEKNKRGGRWLITLSKQQRRNDLDRFWLETLLCLIGEAFDDYSDDVCGAVVNVRAKGDKIAIWTTECENRDAITHIGRVYKERLGLPPKMVIGYQSHADTATKSGSTTKNRFVV	ENSLACT00000018309/ 218 AA	MATAKPQTCPPPPLQGQKEKSKIATQQDMSMPTRYMKHPLQSRWALWFYKNVKSKTWHANVRLITKFDTVEDFWALYNNIQIASKLMSGCDYSLFRDGIEPMWEDYRNKCGGRWLITFSKQQRHTELDRFWLETLLCLIGESFGDYSDDVCGAVINIRAKGDKIALWTTHAENRDAVLHIGRTLKERLGLPAKIVIGYQAHADTATKSGSIVNNKFVV	UCSC database: JH126614/JH128299/JH126722	MTLTGCRAAEESVSFWKMAACEQQRGTETQEINETKPVTSSPETVTPEQYIKHPLQNKWALWYFKNDKSKSWTENLRLIAKFDTVEDFWALYNHIQQPSKLQFGCDYCLFKDGIKPMWEDENNKRGGRWLMTLNKQQRHNDLDRYWLETLLCLIGESFGDYSDDVCGAVINIRAKGDKIALWTTHAENRDAVLHIGYKERLGLPLKALIGYQSHDDTSSKSGSTTKNMYTV*G
			AFYH01193633.1		UCSC database,JH126722		**XP_005993443.1,** was adjusted with use of UCSC database**	
basal ray finned fish								
Lepisosteus oculatus	Spotted Gar	Lo	XP_006629871.1	MATAEPESTSNPPNSEEEKTEATGQEVVSPEHYIKHPLQNRWALWFFKNDKSKTWQANLRLISKFDTVEDFWALYNHIQLSSNLMSGCDYSLFKDGIEPMWEDERNKRGGRWLITLSKQQRRSDLDRFWLETLLCLIGEAFDDYSDDVCGAVVNIRTKGDKIAIWTTDYENKDAVTHIGRVYKERLGVPQKVIIGYQSHADTATKSGSTTKNKFVV	XP_006632115.1	MATAALHPGVALPPRRDKGELRKTAKANGKEAALLLGKRLKHPLENRWALWFYKNDKSKTWQENLRLITKFDTVEDFWALYNNMKVASKLSSGCDYSVFKDGIEPMWEDRSNKCGGRWLITLSKQQRHTELDRFWLDTLLCIIGEGFDVYSADVCGAVINVRAKGDKIAIWTTNAENREAVTYIGRKFKEGLGLPAKIVIGYQAHADTATKSNSITKNKFVV	XP_006630507.1	MATSEPRGPEAEETSTETSSAVTSPDQYIKHPLQNRWALWYFKNDKSKSWTENLRLIAKFDTVEDFWALYNHIQQPSKLGFGCDYCLFKDGIKPMWEDDRNKLGGRWLMTLGKQQRHNDLDRYWMETLLCLIGESFDEASDDVCGAVVNVRPKGDKISIWTGNCQNKEAIMTIGQQYKERLNVPNKALLGYQSHDDTSSKSGSTTKNMYTV

Jawless fish				**eIF4E-1A**		**eIF4E-1B**		**eIF4E-1C**
organism name	**common name**	**abbrev**	**Accession**	**seq**	**Accession**	**seq**	**Accession**	**seq**
Petromyzon marinus	lamprey	Pm	Composite sequence	MAIRALTPEEPEKVEDLVTAETDVDPENYIKHPLQN RWALWFYKNDKSKSWQANLRLITKVDTVEDFWALYNHIQVASRLMPGCDYSLF KDGIEPMWEDERNKRGGRWLITLTKTQRHSDLDRYWLETLLCLIGEAFDDHSDDVCGAVV NVRPKADKIAVWTADCDNRESVVGIGRVYKDRLALPPRIIIGYQSHTDTATKSGSSTKNM FTV	**not found**		**not found**	
	EG336965-1	LIAPLMEQTQLNKHYLKNRSVTKEMAIRALTPEEPEKVEDLVTAETDVDPENYIKHPLQNRWALWFYKNDKSKSWQANLRLITKVDTVEDFWALYNHIQVASRLMPGCDYSLFKDGIEPMWEDERNKRGGRWLITLTKTQRHSDLDRYWL**ETLLCLIGEAFDDHSDDVCGAVVNVRPKADKIAVWTADCDNRESVVGIG**	**http://blast.imcb.a-star.edu.sg/cgi-bin/scripts/request_scaff.pl?db=jlamp_proteins&seqid=JL7239**	MNVCVCFRWALWFYKNDKSKSWQANLRLITKVDTVEDFWALYNHIQVASRLMPGCDYSLF KDGIEPMWEDERNKRGGRWLITLTKTQRHSDLDRYWLETLLCLIGEAFDDHSDDVCGAVV NVRPKADKIAVWTADCDNRESVVGIGRVYKDRLALPPRIIIGYQSHTDTATKSGSSTKNM FTV				

spiny fin fish lower teleost				eIF4E-1A		eIF4E-1B		eIF4E-1C
organism name	common name	abbrev	Accession	seq	Accession	seq	Accession	seq
Danio rerio	zebrafish	Dr	**NP_571808.1**	MATAEPETSTNPSNSEEKNEENEQQIVSLEDYIKHPLQNRWALWFFKNDKSKTWQANLRLISKFDTVEDFWALYNHIQLSSNLMSGCDYSLFKDGIEPMWEDERNKRGGRWLITLSKQQRRADLDRFWLETLLCLVGEAFDDHSDDVCGAVVNIRTKGDKIAIWTTDYENKDAIVHIGRVVYKERLGVPPKVIIGYQSHADTATKSGSTTKNKFVV	NP_571529.1	MASCAVQLIDKVPKKKVEKKKFEPNILKEPCMKHPLQNRWGLWFYKNDKSKMWQDNLRLITKFDTVEDFWGLYNNIQLPSKLSSGCDYSMFKDGIEPMWEDRSNKCGGRWLITLAKQHRHTELDHFWLETLLCLIGEGFSSFSRDICGSVINIRAKGDKIALWTSNAENCETVTYIGRKYKESLGLPQKLVIGYQAHADTATKSNSITKNKFVV	ADH59743.1	MATSEPRGTETEEVRADSPTAVVTTSPEQYIKHPLQNRWALWYFKNDKSKSWTENLRLISKFDTVEDFWALYNHIQQPSKLGFGCDYCLFKDGIKPMWEDDRNKLGGRWLMTLSKQQRHNDLDRYWMETLLCLIGESFDEASEDVCGAVVNVRPKGDKIAIWTGNCQNRDAIMTIGQQYKERLSLPSKTLIGYQSHDDTSSKSGSTTKNMYSV
Oncorhynchus mykiss	rainbow trout	Om	NP_001154092.1	MATAEPEININPPRHAEEEAETGQEIVSPESYIKHPLQNKWSLWFFKNDKTKTWQANLRLISKFDTVEDFWALYNHIQLSSNLISGCDYSLFKDGIEPMWEDERNKRGGRWLITLNKQQRRQDLDRFWLETLLCLVGEAFDDYSDEVCGAVVNIRTKGDKIAIWTADFDNREAITHIGRVYKERLGIPMKMTIGYQSHSDTATKSGSTTKNKFVV	CDQ69051.1	MACVAVRLDKGALKKIDGAKKKIDNKVARVVVGPHVKHPLQNRWALWFYKNDKSKMWQDNLRLITKFDTVEDFWALYNNIQLVSKLSSGCDYSVFKDGIEPMWEDRRNKCGGRWLITLSKQQRHTELDRFWKETLLCLIGEGFGSFSRDVCGAVINVRAKGDKIAIWTTDTENGEAVTYIGRKYKEGLGLPPKLVIGYQAHADTATKSNSITKNKFVV	CDQ63979.1	MATSEPVSEIENAPETEEPQPEVILTAPPVVAGSQQYIKHPLQNKWALWYFKNDKSKSWTENLRLIAKFDTVEDFWALYNHIQQPSKLGFGCDYCLFKDGVKPMWEDDKNKLGGRWLMTLSKQQRQIDLDRYWMETLLCLIGESFDEASEDVCGAVVNIRPKGDKISIWTGNCQNKEAIVTIGQQYKERLSIPIKLLIGYQSHDDTSSKSGSTTKNMYSV
Salmo Salar	salmon	Ss	SS2U046148 database http://salmondb.cmm.uchile.cl	MATAELESSLNPPNSEGKKSETTGQEVVNPEDYIKHPLQNRWALWFFKNDKTKTWQANLRLISKFDTVEDFWALYNHIQLSSNLISGCDYSLFKDGIEPMWEDERNKRGGRWLITLSKQQRRADLDRFWLETLLCLVGEAFDDHSDDVCGAVVNIRTKGDKLAIWTTDYENKDAITHIGQGWKERLGVPHKVIIGYQSHADTATKSGSTTKNKFVV	SSGWT00282 3 database http://salmondb.cmm.uchile.cl	VVGPHVKHPLQNRWALWFYKNDKSKMWQDNLRLITKFDTVEDFWALYNNIQLVSKLSSGCDYSVFKDGIEPMWEDRRNKCGGRWLITLSKQQRHTELDRFWKETLLCLIGEGFGSFSRDVCGAVINVRAKGDKIAIWTTDTENGEAVTYIGRKYKEGLGLPPKLVIGYQAHADTATKSNSITKNKFVV	NP_001140090.1,ACM09624.1, ACM09278.1	MATSEPVSEIENAPETEEPQPEVILTAPPVVDGSQQYIKHPLQNKWALWYFKNDKSKSWTENLRLIAKFDTVEDFWALYNHIQQPSKLGFGCDYCLFKDGVKPMWEDDKNKLGGRWLMTLSKQQRQIDLDRYWMETLLCLIGESFDEASEDVCGAVVNIRPKGDKISIWTGNCQNKEAIVTIGQQYKERLSIPIKLLIGYQSHDDTSSKSGSTTKNMYSV
Astyanax mexicanus	cavefish	Am	XP_007245597.1	MAAAELDINSKTLNSEEEKNCVSGQGQEVAIPEDYIKHPLQNRWALWFFKNDKSKTWQANLQLISKFDTVEDFWALYNHIQLSSNLISGCDYSLFKDGIEPMWEDERNKKGGRWLITLSKQQRRADLDRFWLETLLCLVGEAFDDHSDDVCGAVVNIRTKGDKIAIWTTDYENKDGIVHIGRVYKARLGVPPKVIIGYQSHADTATKSGSTTKNKFVV	**1B-like Xp_007236405**	MAACAVQLLLCMIGEGFGSYSRDVCGSVINVRAKGDKIAVWTTNTENAEAVTHIGRKYKEILGLPSKLVIGYQAHADTATKSNSITKNKFVV	XP_007240333.1	MRGTESEEASNDSPTPAVAEQYIKHPLQNRWALWYFKNDKSKSWTENLRLISKFDTVEDFWALYNHIQQPSKLGFGCDYCLFKDGIKPMWEDDRNKLGGRWLMTLSKQQRHNDLDRYWMETLLCLIGESFDEASEDVCGAVVNVRPKGDKIAIWTGNCQNRDAIMTIGLQYKERLNLPIKTLIGYQSHDDTSSKSGSTTKNMYSV

spiny ray fish higher Teleosts								
				eIF4E1A		eIF4E1B		eIF4E1C
organism name (genus, species)	common name	abbrev	accession	seq	accession	seq	accession	seq
Oreochromis niloticus	tilapia	On	ENSONIT00000001366	MRVVRTRRPNPSYFNTRRKTQSDFQKMATALVASNSVPANGETEKCETAVQKVVNPETYIKHPLQNRWALWFFKNDKSKTWQANLRLISKFDTVEDFWALYNHIQLSSNLMSGCDYSLFKDGIEPMWEDERNRRGGRWLITLSKQQRKSDLDRFWLETLLCLVGEAFDDYSDDVCGAVINVRAKGDKIAIWTTDYENKEAITHIGRVYKERLGVPPKVIIGYQSHADTATKSGSTTKNKFVA	**not found**		XP_003451978.1	MATSEPKTTETEDQQTDGQVVANPEQYIKHPLQNRWALWYFKNDKSKSWTENLRLISKFDTVEDFWALYNHIQQPSKLGFGCDYCLFKDGIKPMWEDDRNKLGGRWLMTLNKQQRHNDLDRYWMETLLCLVGESFDEASEDVCGAVVNVRPKGDKISIWTSNCQNRDAIMTIGQLYKERLNLPMKAMIGYQSHDDTSSKSGSTTKNMYSV
Oryzias latipes	Medaka	Ol	XP_004082704.1	MATAEPENSPSSPPPEEDGSEEVGQELVSPEAYVKHPLQNRWSLWFYKNDKSKTWQANLRLISKFDTVEDFWALYNHIQVSSNLISGCDYSLFKDGIEPMWEDERNKRGGRWLITLNKQLRRSDLDRFWLETLLCLVGEAFDDYSDDVCGAVVNVRNKGDKIAVWTSNYENREAVTHIGRVYKERLGLPMNMTIGYQSHADTATKSGSTTKNKFVV	**not found**		XP_004084917.1/BJ708113	MATSEPKAADTEDQQQTDGQVLANAEQYIKHPLQNRWALWYFKNDKTKSWTENLRLISKFDTVEDFWALYNHIQQPSKLVLGCDYCLFKDGIKPMWEDDKNKLGGRWLMTLNKQKHNDLDRYWMETLLCLVGESFDDASEEVCGAVVNVRHKGDKISIWTGNCQNKEAIMTIGQLYKERLNLPMKAIIGYQSHDDTSSKSGSTTKNMYSV
Takifugu rubripes	Fugu	Tr	XP_003978618.1	MATAEPEPSPNLSQPEEEETGETGQELVRPDPYIKHPLQNRWSLWFFKNDKSKTWQANLRLISKFDTVEDFWALYNHIQLSSNLMSGCDYSLFKDEIEPMWEDERNKRGGRWLITLNKQQRRFDLDRFWLETLLCLVGEAFDDYSDDVCGAVVNVRAKGDKIAIWTANYENRDAVTHIGRVYKERLGLPMKMTIGYQCHADTATKSGSTTKNKYVV	**not found**		ENSTRUT00000034803	MATSEPPQINCAFLQKTTETEDQQAESQVAAGSDSSIKHPLQNRWALWYFKNDKSKSWTENLRLISKFDTVEDFWALYNHIQQPSKLGFGCDYCLFKDGIKPMWEDDRNKLGGRWLMTLNKQQRHNDLDRFWMETLLCLVGESFDEASDDVCGAVVNVRPKGDKIAIWTSNCQNREAIMTIGQLYKERLNLPIKAMLGYQSHDDTSSKSGSTTKNMYSI
Gadus morhua	cod	Gm	ENSGMOT00000013641	VSLFVSAPISSNSKVETCEKTIESIVNALQYIKHPLQNRWALWFFKNDKSKTWQANLRLISKVDTVEDFWALYNHIQLSSNLVSGCDYSLFKDGIEPMWEDKLNRRGGRWLITLSKQQRRADLDRFWLETVLCLVGETFDEHSDDVCGAVINVRAKGDKIAIWTRDYENKEAITHIGRVYKERLGVPQKVIIGYQSHADTATKSGSSMKNKFVA	**not found**		ENSGMOT00000007368	KPADIEEPQPEIPVLPSSDEYIKHPLQNKWALWYFKNDKSKSWTENLRLISKFDTVEDFWALYNHIQQPSKLGFGCDYCLFKDGIKPMWEDDRNKLGGRWLMTLNKQQRHNDLDRYWMETLLCLVGESFDESSEDVCGAVVNVRPKGDKIAIWTSNCQNRDAIVTIGAGYKERLCLPSKPLISYQSHDDTSSKSGSTTKNMYSV
Gasterosteus aculeatus	stickleback	Ga	ENSGACT00000021787	LVFASQLSTSLPANPEKETCETIIQKIMSPESCIKHPLQNKWALWFFKNDKSKTWQANLRLISKFDTVEDFWALYNHIQLSSNLMSGCDYSLFKDGIEPMWEDARNRRGGRWLITLSKQQRRADLDRFWLETLLCLVGEAFDDHSDDVCGAVINVRAKGDKIAVWTTEYENKEAITHIGRVYKERLGVPQKVIIGYQSHADTATKSGSTTKNKFVA	**not found**		**ENSGACT00000010388**	MATSEPKAPEPEEPQAPDSQVVANPEQYIKHPLQNRWALWYFKNDKSKSWTENLRLISKFDTVEDFWALYNHIQQPSKLGFGCDYSLFKDGIKPMWEDDRNKLGGRWLMTLNKQQRHNDLDRYWMETLLCLVGESFDEASEDVCGAVVNVRPKGDKISIWTSQCQNREAIMTIGQNYKERLNIPTKAIIGYQSHDDTSSKSGSTTKNMYSV
Tetraodon nigroviridis	pufferfish	Tn	CAF94272.1	CR725540			CR706188	
				KHPLQNRWCLWFFKNDKSKTWQANLRLISKFDTVEDFWALYNHIQLSSNLMSGCDYSLFKDEIEPMWEDERNKRGGRWLITLNKQQRRFDLDRFWLETLLCLVGEAFDDYSDDVCGAVVNVRAKGDKIAIWTANYENRDAVTHIGKVYKERLGLPMKMTIGYQSHTDTATKSGSTTKNKYVV*	**not found**			MATSEPNQLKLKTNSAESQVAASPDKYIKHPLQNRWALWYFKNDKSKSWTDNLRLISKFATVEDFWALYNHIQQPSKLGFGCDYCLFKDGIKPMWEDDRNKLGGRWLITLNKQQRHNDLDRFWMETLLCLVGESFDEASDDVCGAVVNVRPKGDKLAIWTSNCQNRDAIMTIGQLYKERLNIPIKAMLGYQAHDDTSSMSGFTTKNMYSI*

Percomorpha Teleosts				eIF4E1A		eIF4E1A like		eIF4E1C
organism name (genus, species)	common name	abbrev	Accession	seq	Accession	seq	Accession	seq
Cynoglossus semilaevis	tongue sole	Cs	XP_008309132.1	MATAEPETSPGSPLPDDDAAEEAGLEVVSPEAYIKHPLQNRWSLWFFKNDKSKTWQANLRLISKFDTVEDFWALYNHIQLSSNLMSGCDYSLFKDDIEPMWEDERNKRGGRWLITLKQQRRLDLDRFWLETLLCLVGEAFDNYSDDVCGAVVNIRTKGDKIAVWTSDYENRDAVTHIGRVYKERLGLPMKMTIGYQSHADTATKSGSTTKNKFVV	XP_008324785.1	MATGLLVSTLAPSNPKKEECETSIQKVMNPELYIKHPLQNRWALWFFKNDKSKTWQANLRLISKFDTVEDFWALYNHIQVSSNLMSGCDYSLFKDGIEPMWEDERNRRGGRWLITLSKQQRKSDLDRFWLETLLCLVGEAFDDYSDDVCGAVINVRAKGDKIAIWTTDYENKDAIIHIGRVYKDRLGVPPKVIIGYQSHADTATKSGSTTKNKFVA	XP_008320130.1	MATSEPKTPETEDQQTEVVANPEQYIKHPLQNKWALWYFKNDKSKSWTENLRLISKFDTVEDFWALYNHIQQPSKLGFGCDYCLFKDGIKPMWEDDRNKLGGRWLMTLNKQQRHNDLDRYWMETLLCLVGESFDEASEDVCGAVVNVRPKGDKIAIWTSNCQNREAIMTIGQQYKERLNIPIKAMIGYQSHDDTSSKSGSTTKNMYSV
Pundamilia nyererei	cichlids	Pn	XP_005726528.1,XP_005750496.1	MATALVVSNSVPANGETEKCETAVQKVVNPETYIKHPLQNRWALWFFKNDKSKTWQANLRLISKFDTVEDFWALYNHIQLSSNLMSGCDYSLFKDGIEPMWEDERNRRGGRWLITLSKQQRKVDLDRFWLETLLCLVGEAFDDYSDDVCGAVINVRAKGDKMAIWTTDYENKEAITHIGRVYKERLGVPPKVIIGYQSHADTATKSGSTTKNKFVA	XP_005750496.1	MATAEPVSTPSPSQPDEDGAEQTGQEIVSPEAYIKHPLQNRWSLWFFKNDKSKTWQANLRLISKFDTVEDFWALYNHIQLSSNLMSGCDYSLFKDGIEPMWEDERNKRGGRWLVTLTKQQRRLDLDRYWLETLLCLVGEAFDDYSDDVCGAVVNVRTKGDKIAVWTSDYENREAVTHIGRVYKERLGLPVKMTIGYQSHADTATKSGSTTKNKFVV	XP_005743856.1	MATSEPKTTETEDQQTDGQVVANPEEYIKHPLQNRWALWYFKNDKSKSWTENLRLISKFDTVEDFWALYNHIQQPSKLGFGCDYCLFKDGIKPMWEDRNKLGGRWLMTLNKQQRHNDLDRYWMETLLCLVGESFDEASEDVCGAVVNVRPKGDKISIWTSNCQNRDAIMTIGQQLYKERLNLPMKAMIGYQSHDDTSSKSGSTTKNMYSV

				eIF4E1A		eIF4E1A like		eIF4E1B			eIF4E1C
organism name (genus, species)	common name	abbrev	Accession	seq	Accession	seq	Accession	seq		Accession	seq
Esox Lucius	northern pike	EI	XP_010897924.1	MATAEPESISNASNSEEEKSETTGQVVVNPEDYIKHPLQNRWALWFFKNDKTKTWQANLRLISKFDTVEDFWALYNHIQLSSNLISGCDYSLFKDGIEPMWEDERNKRGGRWLITLSKQQRRADLDRFWLETLLCLVGEAFDDHSDDVCGAVVNIRTKGDKLAIWTTDYDNKEAITHIGRGWKERLGVAPKLLIGYQSHADTATKSGSTTKNKFVV	XP_010890703.1	MWKMATAEPEINSNAPQHAEEGSEGTGQEIVSPESYIKHPLQNKWSLWFFKNDKSKTWQANLRLISKFDTVEDFWALYNHIQLSSNLISGCDYSLFKDGIEPMWEDERNKRGGRWLITLSKQQRRQDLDRFWLETLLCLVGEAFDDYSDEVCGAVVNIRTKGDKIAIWTADFDNREAITHIGRVVKERLGIPLKMTIGYQSHSDTATKSGSTTKNKFVV	XP_010866832.1	MMACVAVRVPDKRTIKKNDGDKKRLVTKVAYVVTGPHVKHPLQNRWGLWFYKNDKSKMWQDNLRLITKFDTVEDFWALYNNIQLASKLSSGCDYSVFKDGIEPMWEDRSNKCGGRWLITLSKQQRHTELDRFWKETLLCLIGEAFGSYSRDVCGAVINVRAKGDKIAIWTTNTENGEAVTYIGRKYKEGLGLPQKLIIGYQAHADTATKSNSIAKNKFVV		XP_010868333.1	MATSEPKAPETEESQPEVIPTAPPVVAGSEQYIKHPLQNKWALWYFKNDKSKSWTENLRLISKFDTVEDFWALYNHIQQPSKLGFGCDYCLFKDGVKPMWEDDKNKLGGRWLMTLSKQQRQVDLDRYWMETLLCLIGESFDEASDDVCGAVVNVRPKGDKISIWTGNCQNKEAIVTIGQQYKERLSLPIKLLIGYQSHDDTSSKSGSS

KEY						
Joshi EST						
Joshi seq						
NCBI						
Embl						
Hmmr						
other database						
				eIF4E2A		**eIF4E2B**
organism name (genus, species)	**common name**	**abbrev**	**eIF4E**	**seq**		**seq**
Tetrapod						
Homo sapiens	human	Hs				
bird			NP_004837	MNNKFDALKDDDSGDHDQNEE NSTQKDGEKEKTERDKNQSSS KRKAVVPGPAEHPLQYNYTFWY SRRTPGRPTSSQSYEQNIKQIG TFASVEQFWRFYSHMVRPGDLT GHSDFHLFKEGIKPMWEDDANK NGGKWIIRLRKGLASRCWENLIL AMLGEQFMVGEEICGAVVSVRF QEDIISIWNKTASDQATTARIRDT LRRVLNLPPNTIMEYKTHTDSIK MPGRLGPQRLLFQNLWKPRLN VP	none found	
Taeniopygia guttata	zebrafinch	Tg	ENSTGUT00000010390	FTRLKDDDSGDHDQNEENNTQ KDSEKEKNDREKPQSTTKRKAV VPGPAEHPLQYNYTFWYSRRTP GRPTSSQSYEQNIKQIGTFASVE QFWRFYSHMVRPGDLTGHSDF HLFKEGIKPMWEDDANKNGGK WIIRLRKGLASRCWENLILAMLG EQFMVGEEICGAVVSVRFQEDII SIWNKTASDQATTARIRDTLRRV LNLPPNTIMEYKTHTDSIKDNSS FRNTKITL	none found	
Gallus gallus	chicken	Gg	XP_422748.2, used by Joshi :isoform X2	MNNKFDALKDDDSGDHDQNEE NNTQKDSEKEKNDREKPQSTTK RKAVVPGPAEHPLQYNYTFWYS RRTPGRPTSSQSYEQNIKQIGTF ASVEQFWRFYSHMVRPGDLTG HSDFHLFKEGIKPMWEDDANKN GGKWIIRLRKGLASRCWENLILA MLGEQFMVGEEICGAVVSVRFQ EDIISIWNKTASDQATTARIRDTL RRVLNLPPNTIMEYKTHTDSIKD NSSFRNTKITL	not found	
		Gg	XP_004943442, used by Standart: isoform X10	MNNKFDALKDDDSGDHDQN EENNTQKDSEKEKNDREKPQ STTKRKAVVPGPAEHPLQYN YTFWYSRRTPGRPTSSQSYE QNIKQIGTFASVEQFWRFYSH MVRPGDLTGHSDFHLFKEGIK PMWEDDANKNGGKWIIRLRK GLASRCWENLILAMLGEQFM VGEEICGAVVSVRFQEDIISIW NKTASDQATTARIRDTLRRVL NLPPNTIMEYKTHTD**SIKAWE EFHGLVNSSGR**	not found	

				eIF4E2A		eIF4E2B
organism name (genus, species)	common name	abbrev	eIF4E	seq		seq
reptile						
Anolis carolinensis	lizard	Ac	XP_008116054.1	MEAEERKSGAGSKAAAGERPR RMNNKFDALKDDDSGDHEQNE ENNTQKDGEKEKNDRDKSQSSI KRKAVVPGPAEHPLQYNYTFWY SRRTPGRPTSSQSYEQNIKQIG TFASVEQFWRFYSHMIRPGDLT GHSDFHLFKEGIKPMWEDDANK NGGKWIIRLRKGLASRCWENLIL AMLGEQFMVGEEICGAVVSVRF QEDIISIWNKTASDQATTARIRDT LRRVLNLPPNTIMEYKTHTDSIK AWEEFHGLVNSGGR	none found	
amphibian						
xenopus laevis	frog	Xl	**AW643877**	MNNKFDALKDDDSGDHDQNEE NGTQKDGEKEKNDKEKNQGSS RKKSVVPGPAEHPLQYNYTFWY SRRTPGRPTSSQSYEQNIKQIGI FASVEQFWRFYSHMVRPGDLT GHSDFHLFKEGIKPMWEDDANK NGGKWIIRLRKGLASRCWENLIL	NP_001089549	MGLSGQEDLTTSEDEFTKXQKVKEVM VPPGEHPLQYKYTFWYSRRTPSRPAS THNYEQNIRPFGTVASVEQFWRIYSHI VRPGDLSGYSDFHLFKDGIKPMWEDE ANKNGGKWIIRLRKGLASRSWENIILA MLGEQFMVGEEICGVVVSIRFQEDILS IWNKTANDQFSTVRIRDTLRRVLNLPP
Xenopus tropicalis	frog	Xt	CAJ82134_/ NP_001005099.1	MNNKFDALKDDDSGDHDQNEE NGTQKDSEKEKNEKEKNQGSS RKKSVVPGPAEHPLQYNYTFWY SRRTPGRPTSSQSYEQNIKQIG TFASVEQFWRFYSHMVRPGDLT GHSDFHLFKEGIKPMWEDDANK NGGKWIIRLRKGLASRCWENLIL AMLGEQFMVGEEICGAVVSVRF QEDIISIWNKTASDQATTARIRDT LRRVLNLPPNTVMEYKTHTDSIK DKTSFRNTKIAL	NP_001016076.1;_AAI21525_	MGLSGQEDLTTAEDDFTKSQKVKEVM VPPGEHPLQYKYTFWYSRRTPSRPAS THNYEQNIRQFGTVASVEQFWRIYSHI VRPGDLTGYSDFHLFKDGIKPMWEDE ANKNGGKWIIRLRKGLASRFWENIILA MLGEQFMVGEEICGVVVSIRFQEDILS IWNKTANDQFSTVRIRDTLRRVLNLPP NTIMEYKTHTDSLKDNSSFRNTKITV

				eIF4E2A		eIF4E2B
organism name (genus, species)	common name	abbrev	eIF4E	seq		seq
Chondrichthyes			**XP_007901223.1**			
Callorhinchus milii	elephant shark	Cm	XP_007901376			
			>SINCAMP00000025009 len=250 SINCAMT00000025263 SINCAMG00000016215 1-250;AFP09199.1	MNNKFDALKDDDSGDQDQNED NHTQKESEKEKSDKEKSQNSIK RKAVVPGPGEHPFQYNYTFWY SRRTPGRPTSSQSYEQNVKQIG TFASVEQFWRFYSHLVRPGDLT GHSDFHLFKDGIKPMWEDEANK NGGKWIIRLRKGLASRCWENLIL AMLGEQFMVGEEVCGAVVSVR FQEDIISIWNKTASDQATTARIRD TLRRVLNLPPNTIMEYKTHTDSI KDNSSFRNTKLTL	none found	
lobe fin fish						
Latimeria chalumnae	coelacanth	Lc	ENSLACT00000018224	MNNNKFDALKDDDSGDHDQNE ENSAQKEGEKEKTEKDKPQSSA KRKTIVPGPGEHPLQYNYTFWY SRRTPGRPTSSQSYEQNIKQIG TFASVEQFWRFYSHMVRPGDLT GHSDFHLFKEGIKPMWEDEANK NGGKWIIRLRKGLASRCWENLIL AMLGEQFMVGEEICGAVVSVRF QEDIISIWNKTASDQATTARIRDT LRRVLNLPPNTIMEYKTHTDSIK DNSSFRNTKIAV	none found	
Basal ray finned fish						
Lepisosteus oculatus	Spotted Gar	Lo	ENSLOCT00000007011	VVPGAGEHPLQYNYTFWYSRR TPGRPTSSQSYEQNIKQIGSFAS VEQFWRFYSHMVRPGDLTGHS DFHLFKEGIKPMWEDDANKSGG KWIIRLRKGLASRCWENLILAML GEQFMVGEEICGAVVSVRFQED IISIWNKTASDQATTARIRDTLRR VLNLPPNTIMEYKTHTDSIKAWE EFHGLVNTSGGR	none found	

				eIF4E2A		eIF4E2B
organism name (genus, species)	common name	abbrev	eIF4E	seq		seq
Jawless fish						
Petromyzon marinus	lamprey	Pm	JL6818	FFRLKDDDSGDHDNNEENSTQ KNSEKEKSPNASRHKVTVPGPG EHPLQYNYSLWFSRRTPGRQA SKQNYEQNIKHIGTFASVEMFW RFYSHIVRPSDLTGHSDFHLFKE GIKPMWEDDANKSGGKWIVRLR KGLASRCWENLILAMLGEQFMV GEEICGAVVSLRFQEDIISIWNKT ASDQGTTSRIRDTLRRVLNLPPN TIMEYKTHTDSIKAWEDFHGLVN NR	none found	

				eIF4E2A		eIF4E2B
organism name (genus, species)	common name	abbrev	eIF4E	seq		seq
spiny fin fish						
Lower teleost						
Danio rerio	zebrafish	Dr	AG_W99949	MNNKFDALKDDDSGDHDQDNS SPKDGEKEKNDEEDKEANTTKR KAVVPGAGEHPLQYNYTFWYS RRTPGRPASTQSYEQNIKQIGSF ALVEQFWRFYSHMIRPGDLTGH SDFHLFKEGIKPMWEDDANKSG GKWIIRLRKGLASRCWENLILAM LGEQFMVGEEICGAVVSVRFQE DIISIWNKTASDQATTARIRDTLR RVLNLPPNTIMEYKTHTDSIKAW EDFHGLVNASGGR	AG_W99950	MNQFEHLKEEDCGDHEEMKDNNESD RASINNNNNNIRRKMVTPAAGEHPLQ YNYTFWYSRRTPSRPANTQSYEQNIR QMGTVASVEQFWKFYSHLVRPGDLT GHSDFHLFKEGIKPMWEDEANKNGG KWIIRLRKGLASRFWENIILAMLGEQF MVGEEICGVVVSIRFQEDILSIWNKTA NDQVTTSRIRDTLRRVLNLPPNTIMEY KTHNDSLKDNSSFRNTKITL
Oncorhynchus mykiss	rainbow trout	Om	CA3849333	MNNKFDALKDDDSGDHDQDQG SQKDCEKEKNDNDXDNDQNTA KKKIAVPGVGEHPLQYNYSFWY SRRTPGRPASTQSYESNIKQIGS FASVEQFWRFYSHMIRPGDLTG HSDFHLFKEGIKPMWEDDANKL GGKWIIRLRKGLASRCWENLILA MLGEQFMVGGEICGAVVSVRFQ EDIISIWNKTASDQATITRIRDTLX RVLNLPPNTIMEYKTHTDSIKAL GGLPWSGERCWWSLVSKAVSF CPXFVCSVFWMKSGGIYL	none found	
Astyanax mexicanus	cavefish	Am	ENSAMXT00000008126	SEMSSSLCVFQPVVPGAGEHPL QYNYTFWYSRRTPGRPASTQS YEQNIKQIGSFASVEQFWRFYS HMIRPGDLTGHSDFHLFKEGIKP MWEDDANKSGGKWIIRLRKGLA SRCWENLILAMLGEQFMVGEEI CGAVVSVRFQEDIISIWNKTASD QATTARIRDTLRRVLNLPPNTIM EYKTHTDSI	XP_007234185.1	MNQFEQLKEEEVDEREDNVCVREEK RDLDSSNRRKTITPGPGEHPLQYNYT FWYSRRTPSRPANTQSYEQNIRQIGT VASVEQFWKFYSHLVRPGDLTGHSDF HLFKEGIKPMWEDEANKNGGKWIIRL RKGLASRFWENIILAMLGEQFMVGEE VCGVVVSIRFQEDILSIWNKTASDQVT TSRIRDTLRRVLNLPPNTIMEYKTHND SLKDNSSFRNTKITL
Salmo Salar	salmon	Ss	ACN11073	MNNKFDALKDDDSGDHDQDQG SQKDCEKEKNDNDEDNDQNTA KKKVCVSPKYNATIAVPGVGEHP LQYNYSFWYSRRTPGRPASTQ SYESNIKQIGSFASVEQFWRFYS HMIRPGDLTGHSDFHLFKEGIKP MWEDDANKLGGKWIIRLRKGLA SRCWENLILAMLGEQFMVGEEI CGAVVSVRFQEDIISIWNKTASD QATITRIRDTLRRVLNLPPNTIME YKTHTDSIKAWEDFHGLVNAVG GR	ADM16298_	MNQFEHLKDDDPEDQDETVCNREED CSINNNRRKTVSPAAGEHPLQYNYTL WYSRRTPSRPANTQSYEQNIRQIGTV ASVEQFWKFYSHLVRPGDLTGHSDFH LFKEGIKPMWEDEANKNGGKWIIRLR KGLASRFWENIILAMLGEQFMVGEEIC GVVVSIRFQEDILSIWNKTASDQVTTS RIRDTLRRVLNLPPNTIMEYKTHNDSL KDNSSFRNTKITL

higher teleosts				eIF4E2A		eIF4E2B
organism name (genus, species)	common name	abbrev	eIF4E	seq	EST	seq
Gadus morhua	cod	Gm	ENSGMOT00000013210	AEGVVVDSGRIMNNKFDALKDDDSGDHDQDQGSPKNCEKEKNEDEDKEQNNAKKKMVVPGPGEHPLQYNYTFWYSRRTPGRPASTQSYEQNIKQIGGFASVEQFWRFYSHMIRPGDLTGHSDFHLFKEGIKPMWEDDANKMGGKWIIRLRKGLASRCWENLILAMLGEQFMVGEEICGAVVSVRFQEDIISIWNKTASDQATSGRIRDTLRRVLNLPPNTIMEYKTHTDSI	ENSGMOT00000001899	INLCVCVLQSVSPGPGEHPLQYNYSLWFSRRTPSRPASIQSYEQNIRQIATVASVEQFWKLYSHLIRPGDLTGHSDFHLFKEGIKPMWEDDANRSGGKWIIRLRKGLASRFWENIILAMVGEQFMVGEEVCGVVVSIRFQEDILSIWNRTSSDQTTTSRIRDTLRRVLNLPPNTIMEYKTHNDSIKDNSSFRNTKITL
Tetraodon nigroviridis	pufferfish	Tn	**ENSTNIT00000018580**2B like**	KDEKSQEESECHHDNSNGVNNSRRKTVCPAAGEHPLQYNYTFWYSRRTPSRPASSQSYEQNIRQIGTVASVEQFWRFYSHLIRPGDLSGHSDFHLFKEGIKPMWEDESNRSGGKWIIRRKGLASRFWENIILAMLGEQFMVGEEICGAVVSIRFQEDILSIWNRTSNDQMTTSRIRDTLRRVLNLPTNTIMEYKTHNDSLRDNSSFRNTKISL	**ENSTNIT00000018580**	MERPKDEKSQEESECHHDNSNGVNNSRRKTVCPAAGEHPLQYNYTFWYSRRTPSRPASSQSYEQNIRQIGTVASVDQFWRFYHLIRQGDLSGHSDHLFKEGIKLHVGDESNRSGGKWIIRLRKGLASRFWENIILAMLGEQFMVGEEICGAVVSIRFQEDILSIWNRTSNDQMTTSRIRDTLRRVLNLPTNTIMEYKTHNDSLRYRWLSTPERCRLQTSWEPGSDLD
Oreochromis niloticus	tilapia	On	ENSONIT00000012683	MNNKFDALKDDDSGDHDQDQGSPKDGEKEKNEDEEKEQNTSKKKMVVPGAGEHPLQYNYTFWYSRRTPGRPASTQSYEQNIKQIGSFASVEQFWRFYSHMIRPGDLTGHSDFHLFKEGIKPMWEDDANKMGGKWIIRLRKGLASRCWENLILAMLGEQFMVGEEICGAVVSVRFQEDIISIWNKTASDQATTARIRDTLRRVLNLPPNTIMEYKTHTDSIKAWEDFHGLVNASGGR	**XP_003459317.1/ ENSONIT00000022606**	MDQLERPQDDNAQETDCHVDNADGNNNNNRRKTVCPGAGEHPLQYNYTFWYSRRTPSRPASSLSYEQNIRQIGTVASVEQFWRFYSHLVRPGDLSGHSDFHLFKEGIKPMWEDESNRSGGKWIIRLRKGLASRFWENIILAMLGEQFMVGEEICGAVVSIRFQEDILSIWNRTSNDQTTTSRIRDTLRRVLNLPANTIMEYKTVNDTLKDNSSFRNTKISV
Pundamilia nyererei	cichlids	Pn	XP_005721275	MACASTKPLLCASPSLAAGFRRLWSDSEKTPIYKIMNNKFDALKDDDSGDHDQDQGSPKDGEKEKNEDEEKEQNTSKKKMVVPGAGEHPLQYNYTFWYSRRTPGRPASTQSYEQNIKQIGSFASVEQFWRFYSHMIRPGDLTGHSDFHLFKEGIKPMWEDDANKMGGKWIIRLRKGLASRCWENLILAMLGEQFMVGEEICGAVVSVRFQEDIISIWNKTASDQGTTARIRDTLRRVLNLPPNTIMEYKTHTDSIKAWEDFHGLVNASGGR	none found	
Stegastes partitus	damselfish	Spr	XP_008280487	MNNKFDALKDDDSGDHDQDQGSPKDGEKEKTEDEDKEQNTSKKKMVVPGAGEHPLQYNYTFWYSRRTPGRPASTQSYEQNIKQIGSFASVEQFWRFYSHMIRPGDLTGHSDFHLFKEGIKPMWEDDANKMGGKWIIRLRKGLASRCWENLILAMLGEQFMVGEEICGAVVSVRFQEDIISIWNKTASDQATTARIRDTLRRVLNLPPNTIMEYKTHTDSIKYVCLLDLLFSLQLFYLFHTCINVHAQKHFKCDVLLLLQRAQFYTTLHIPYMYSYSDVVLIN	none found	
Oryzias latipes	Medaka	Ol	**ENSORLT00000019703**	SLSQDRMNNKFDALKDDDSGDHDQDQSSPKDSETIKIEDDEKEQNTTKKKMVVPGAGEHPLQYNYTFWYSRRTPGRPASTQSYEQNIKQIGSFASVEQFWRFYSHMIRPGDLTGHSDFHLFKEGIKPMWEDDANKMGGKWIIRLRKGLASRCWENLILAMLGEQFMVGEEICGAVVSVRFQEDIISIWNKTASDQGTTARIRDTLRRVLNLPPNTIMEYKTHTDSIKYENSQYYTFKVTI	none found	

				eIF4E2A		eIF4E2B
organism name (genus, species)	common name	abbrev	eIF4E	seq	EST	seq
Takifugu rubripes	Fugu	Tr	ENSTRUT00000037357	MNNKFDALKDDDSGDHDQDQGSPKDGEKEKTEDEEKEQNVSKKKMVVPGAGEHPLQYNYTFWYSRRTPGRPASTQSYEQNIKQIGSFASVEQFWRFYSHMIRPGDLTGHSDFHLFKEGIKPMWEDDANKMGGKWIIRLRKGLASRCWENLILAMLGEQFMVGEEICGAVVSVRFQEDIISIWNKTASDQATTARIRDTLRRVLNLPPNTIMEYKTHTDSIKYSLGRLPRPVFAVRNDTGPRLRVC	ENSTRUT00000016942	MNQLERPKDERIQEESECHHDNSDGTNNNNRRKTVCPAVGEHPLQYNYTFWYSRRTPSRPASSQSYEQNIRQIGTVASVEQFWRFYSHLIRPGDLSGHSDFHLFKEGIKPMWEDESNRSGGKWIIRLRKGLASRFWENIILAMLGEQFMVGEEICGAVVSIRFQEDILSIWNRTSNDQMTTSRIRDTLRRVLNLPTNTIMEYKTHNDSLRLDYWHKLASVGSSSAP
Gasterosteus aculeatus	stickleback	Ga	**ENSGACT00000019950**	MNNKFDALKDDDSGDHDQDQGTPKDSETERTEDEDQSLSKKKMVVPAAGEHPLQYNYTFWYSRRTPGRPASTQSYEQNIKQIGSFASVEQFWRFYSHMIRPGDLTGHSDFHIFKEGIKPMWEDDANKMGGKWIIRLRKGLASRCWENLILAMLGEQFMVGEEICGAVVSVRFQEDIISIWNKTASDQSTTARIRDTLRRVLNLPPNTIMEYKTHTDSIKAWEDFHGLVNVSGGQ	**ENSGACT00000001546**	MNQLECTREEEDQEETECHHDNSDRTNNNNNNNRRKTVCPGVGEHPLQYNYTFWYSRRTPSRPASSQSYEQNIRQIGSVASVEQFWRFYSHLVRPGDLSGHSDFHLFKEGIKPMWEDDYNRSGGKWIIRLRKGLASRFWENIILAMLGEQFMVGEEICGAVVSIRFQEDILSIWNKTSNDQITTSRIRDTLRRVLNLPANTIMEYKTHNDSLRDNSSFRNTKISL
Xiphophorus maculatus	platyfish	Xm	ENSXMAT00000008617	MNNKFDALKDDDSGDHEQDQGSPKDSEKEKPEEDDKEQNIIKKKMVVPGPGEHPLQYNYTFWYSRRTPGRPASTQSYEQNIKQIGSFASVEQFWRFYSHMIRPGDLTGHSDFHLFKEGIKPMWEDDANKMGGKWIIRLRKGLASRCWENLILAMLGEQFMVGEEICGAVVSVRFQEDIISIWNKTASDQATTARIRDTLRRVLNLPPNTIMEYKTHTDSIKRSLEVHGNLEELQRSPAQSLGRLPWSGEC	none found	
Cynoglossus semilaevis	tongue sole	Cs	XP_008332066	MNNKFDALKDDDSGDHDQDQGSPKDSEKEKPEDEDKEQNTQRKKMVVPGAGEHPLQYNYTFWYSRRTPGRPASTQSYEQNIKQIGSFASVEQFWRFYSHMIRPGDLTGHSDFHLFKEGIKPMWEDDANKMGGKWIIRLRKGLASRCWENLILAMLGEQFMVGEEICGAVVSVRFQEDIISIWNKTASDQATTARIRDTLRRVLNLPPNTIMEYKTHTDSIKAWEDFHGLVNASGGR	none found	
Esox Lucius	northern pike	El	XP_010886656	NNKFDALKDDDSGDHDQDQGSQKDCEKEKNGNDDKDQNTAKKKIAVPGVGEHPLQYNYSFWYSRRTPGRPASTQSYESNIRQIGSFASVEQFWRFYSHMIRPGDLTGHSDFHLFKEGIKPMWEDDANKQGGKWIIRLRKGLASRCWENLILAMLGEQFMVGEEICGAVVSVRFQEDIISIWNKTASDQATTARIRDTLRRVLNLPPNTIMEYKTHTDSIKLIHNSPTPSWSTRHTPTALS	XP_010864634.1	MNQFDHLKDEDQEDLDETVCNREEDGSINNNPRKMVSPAAGEHPLQYNYTFWYSRRTPSRPANTLSYEQNIRQIGTVASVEQFWKFYSHLVRPGDLTGHSDFHLFKEGIKPMWEDEANKNGGKWIIRLRKGLASRFWENIILAMLGEQFMVGEEICGVVVSIRFQEDILSIWNKTSSDQVTTSRIRDTLRRVLNLPPNTIMEYKTHNDSLKDNSSFRNTKITL

KEY

Joshi EST
Joshi seq
NCBI
Embl
Hmmr
other database

organism name	common name	abbrev	Accession	eIF4E-3 seq
Tetrapod				
Homo sapiens	human	Hs	NP_001128123	MALPPAAAPPAGAREPP GSRAAAAAAAPEPPLGL QQLSALQPEPGGVPLHS SWTFWLDRSLPGATAAE CASNLKKIYTVQTVQIFW SVYNNIPPVTSLPLRCSY HLMRGERRPLWEEESNA KGGVWKMKVPKDSTSTV WKELLLATIGEQFTDCAA ADDEVIGVSVSVRDRED VVQVWNVNASLVGEATV LEKIYELLPHITFKAVFYK PHEEHHAFEGGRGKH
avian				
Taeniopygia guttata	zebrafinch	Tg	H0ZI21 - H0ZI21_TAEGU partial fragment ABQF01038076	LPLHSPWTFWLDKSLPG TTAAECALNLKKIYTVQT VQDFWSVYNNIPPVTNL PLRCSYHLMRGERRPLW EEGSNAKGGIWKMKVPK ESTAAVWKELLLATIGEQ FTDCCAADDEVIGVSVSV RDREDVVQVWNMNSSS ASEAKVLEKIHKLLPHTSF KVIFYKSHREHHAFEG
Gallus gallus	chicken	Gg	F7BFT4_CHICK	RALPRPAPSLGGGTAGP AGQDMELSAPGPEPPRS RGGDEEGAEAPGLPLHS AWTFWLDKSLPGTTAAE CASNLKKIYRVQTVQDF WSVYNNIPPVTSLPLRCS YHLMRGERRPLWEEESN AKGGIWKMKVAKESTAA VWKELLLATVGEQFTDC CAADDEVIGVSISVRDRE DVVQVWNGNASLASEAK VLEKIHKLLPHTSFRAVFY KPHREHHAFEGRRGRH
reptile				
Anolis carolinensis	lizard	Ac	XP_003217870	MALSGPERQPEPPQSSS GAAEEADLDPQRLRALE PEDGGAPGIPLHSPWTF WLDKSLPGTTAAECASN LKKIYTVQTVQIFWSVYN NIPPVTNLPLRCSYHLMR GERRPLWEEESNAKGGV WKMKIPKDNTASVWKEL LLATIGEQFADRCAGGDE VIGVSVSVRDREDVVQV WNVNASLASEATVLEKIY ELLPHMSFKAVFYKPHKE HHAFEGGRGRH
amphibian				
xenopus laevis	frog	Xl	AAI70573	MALPAAPADRRLQPEPD EQLHLNHQDIGELGLPQE PDTEGIPLHSPWTFWLD RSLPGTTAAECESNLKKI YTVHTIQSFWSVYNNIPL VTNLPVRWSYHLMRGER KPLWEEESNAKGGVWK MKVPKEASSLVWKELLLA TIGEQFTDRCAPEDEVIG VSVSVRDREDIVQVWNG NASVVAEATVLEKIYELLP NTSFKAVFYKPHEEHHAF EGGRSRH
Xenopus tropicalis	frog	Xt	NP_001016049.1	MSGAELQRRVRRKMALP AAPADRRLQPEPDEQLH LNHRELGELALPQEPDTE GIPLHSPWTFWLDRSLP GTTAAECESNLKKIYTVH TIQSFWSVYNNIPQVTNL PLRWSYHLMRGERKPLW EEESNAKGGVWKMKVP KEASSLVWKELLLATIGE QFTDRCAPEDEVIGVSVS VRDREDVVQVWNGNAS VVGEATVLEKIYELLPNTS FKAVFYKPHEEHHAFEG GRSRH

urchin				eIF4E-3
organism name	**common name**	**abbrev**	**Accession**	**seq**
Strongylocentrotus purpuratus	sea urchin	Sp	NP_001091926	MATSEFVSSAIQRQQVQ RAISFEGLADRVVAQEEG VPLNTPWTFWLERSMPI ATAAEVEANLEEIYTVKTV ENFWGVYNNIPDASDLP LRFSYHLMRGNVKPLWE DPCNAQGGDWKFKVQK QNTTKLWKEVLLATVGE QFATSISPDDEICGVSVSI RNTNDVIQVWNRYSRFS EQASIVEKVQTLTPDIDFR ATFYKPHHQHDSFEKNR P
tunicate				
ciona intestinalis	tunicate	Ci	XP_002122741. 1	MAKALFAGGVDVEHIFG NKFNKASRDRSSSPVPM VNGESPKLSRKAISNLGD TEKTGVPLNSAWTLWLD RSVPNLTASEYEANLRKI YTVSTIESFWGVYNHIPP PSRLIPRYSYHLMRNNIR PVWEDEVNANGGMWKL RCHKSVTDNVWNELLLS CIGEQFTGYVNKGDDIIG LSVSIRKNDDLVQVWNIN SSNIEQCKVLEKIKEVLPH VIFETSFYKPHQLHRAFE GRKM
Amphioxus				
Branchiostoma floridea	lancelets	Bf	XP_002613647. 1composite	**MAASTVDSLQSPEPPVP GSSPKLPRAA**IDGIQRNE KTGIPLNTAWTFWLDKSV RGATAAEYEANLRKIYTV NTVESFWGVFNNIPDVS EIQDRYGYHLMREERRPI WEDECNMRGGYWKMK CFKKDTSVVWKELLLAVI GEQFTDHTAEGDEVVGL SVSVRERDDIIQIWNQNA EAAEKATVVSKFRELLPN TNFPTLFYKPHQAHHAFE KDRTNFYRK
basal ray finned fish				
Lepisosteus oculatus	Spotted Gar	Lo	XP_006630669. 1	MAVPAAPDLQMNAETQN SPANNPENNIHIDETELE NITNNDEDGTALPLHSPW TFWLDKSLPGTTAAECE SNLKKIYTVQTVQSFWSV YNNIPPVTNLPLRCSYHL MRGERRPLWEEESNAK GGVWKMKVPKESTAAV WKELLLATIGEQFADYCA SEDEVVGVSVSVRERED VVQVWNVNASFASEANIL GKIHELLPHISFKAVFYKP HEEHHAFEGGRSRH

jawless fish				eIF4E-3
organism name	common name	abbrev	Accession	seq
Petromyzon marinus	lamprey	Pm	CO549413	MDEAAAATTTTAVAAERP KQQQQQHGDADDMIVLD GIVEEGVPLNSPWTFWL DRSVRGITVTEYESNLKK IYTVRSVEGFWSVYNNIP SVESLPVRCTYHLMRGE RQPLWEDPSNCGGGIW KMKCTKEQTITVWKELLL ATIGEQFSDSCEKDDEVV GVSVSIREREDVIQVWNK NARLADRATVLPKLFSLL PSVGFKGVFYKEHEEHD AFERGRTQRHHGGGSG GGGGFYRNHE*
EG336965-1	LIAPLMEQTQLNKHYLK NRSVTKEMAIRALTPE EPEKVEDLVTAETDVD PENYIKHPLQNRWALW FYKNDKSKSWQANLR LITKVDTVEDFWALYNH IQVASRLMPGCDYSLF KDGIEPMWEDERNKR GGRWLITLTKTQRHSD LDRYWL**ETLLCLIGEA** **FDDHSDDVCGAVVNV** **RPKADKIAVWTADCDN** **RESVVGIG**	**http://blast.imcb.a-star.edu.sg/cgi-bin/scripts/request_scaff.pl?db=jlamp_proteins&seqid=JL7239**		

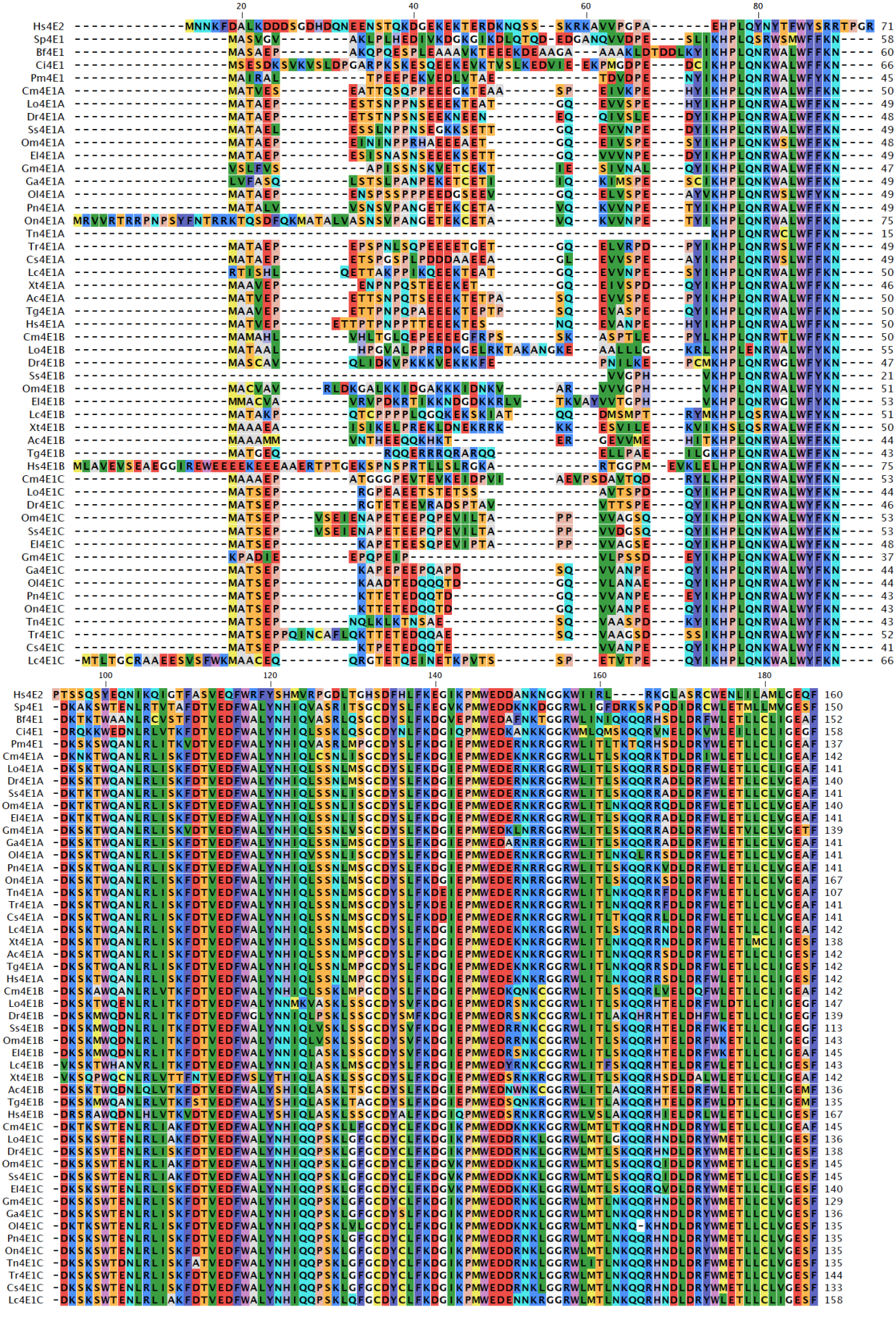
spiny fin fish lower teleost				eIF4E-3			eIF4E3 Like
organism name	common name	abbrev	Accession	seq		accession	seq
Danio rerio	zebrafish	Dr	NP_001004589.1	MAVPAAPNLQLNTARQSSPVNSTENDIHIDERELENITNHVEDGTSLPLHSPWTFWLDRSLPGTTAAECESNLKKIYTVHTVQSFWSVYNNIPPVSCLPLRCSYHLMRGERRPLWEEESNAKGGVWKMKVPKESTLAVWKELLLATIGEQFTDYCASEDEVVGVSVSVRERED VVQVWNGNASFANEANVLGRIYELLPQISFKAVFYKPHEEHHAFEGGRSRH		none found	
Oncorhynchus mykiss	rainbow trout	Om	CDQ75850.1	MAVPAVPNLQLNKTVSQNSPERNIHIDERELENITNNDGNGTLPLHSPWTFWLDRSLPGTTAAECESGLKKIYTVQTVQSFWSVYNNIPGVSSLPLRCSYHLMRGERRPLWEEESNAKGGVWKMKVPKESTPAVWKELLLATIGEQFTDYCASEDEVVGVSVSVRDREDVVQVWNGNAFFANDANILGRIYELLPQITFKAVFYKPHEEHHAFEGGRPRH		none found	
Salmo Salar	salmon	Ss	SS2U053917 unnamed protein product	MAVPAVPTLQLNNTVSQNSAERNIHIDERELGNITNNDGNGTLPLHSPWTFWLDRSLLGTTAAECESGLKKIYTVQTVQSFWSVYNNIPGVSSLPLRCSYHLMRGERRPLWEEESNAKGGVWKMKVPKESTPAVWKELLLATIGEQFTDYCASEDEVVGVSVSIRDREDVVQIWNGIASFANEANVLGRIYELLPQITFKAVFYKPHEEHHAFEGGRPRH		SS2U053917 unnamed protein product	MAVPAVPNLQLNKTVSQNSPERNIHIDERELENITNNDGNDTLPLHSPWTFWLDRSLPGTTAAECESGLKKIYTVQTVQSFWSVYNNIPGVSSLPLRCSYHLMRGERRPLWEEESNAKGGVWKMKVPKESTPAVWKELLLATIGEQFTDYCASEDEVVGVSVSVRDREDVVQIWNGNAFFANEANILGRIYELLPQITFKAVFYKPHEEHHAFEGGRPRH
Astyanax mexicanus	cavefish	Am	ENSAMXP00000017959	MAVPATPNLQLNAGRQSGSPVSSSENNIHIDEKELENLTNHAEDGTSLPLHSPWTFWLDRSLPGTTAAECESNLKKIYTVQTVQSFWSVYNNIPTVSYLPLRCSYHLMRGERRPLWEEESNAKGGVWKMKVPKESTPAVWKELLLATIGEQFADYCASEDEVVGVSVSVRDREDVVQVWNGNASFANEANILGRIYELLPQITFKAVFYKPHEEHHAFEGGRSRH		XP_007256411	MAVPATPNLQLNAGRQSGSPVSSSENNIHIDEKELENLTNHAEDGTSLPLHSPWTFWLDRSLPGTTAAECESNLKKIYTVQTVQSFWSVYNNIPTVSYLPLRCSYHLMRGERRPLWEEESNAKGGVWKMKVPKESTPAVWKELLLATIGEQFADYCASEDEVVGVSVSVRDREDVVQVWNGNASFANEANILGRIYELLPQISFKAVFYKPHEEHHAFEGGRSRH

spiny fin fish higher teleosts				eIF4E3			eIF4E3 Like
organism name (genus, species)	common name	abbrev	accession	seq		accession	seq
Oreochromis niloticus	tilapia	On	ENSONIT00000025557	MAVPAGQTDVQMDRGALSGQTVSSENNIDIDEKELENITKKHREEDTATLPLHSPWTFWLDRSLPGTTAAECESNLKKIYTVQTVQMFWSVYNNIPPVTALPLRCSYHLMRGERRPLWEEESNAKGGVWKMKIPKENTSAVWKELLLATIGEQFADYCASDDEVVGVSVSVRDREDVVQVWNSDASLANEANILGKVYELLPYISFKAVFYKSHMEHHAFEGGRSRH		none found	
Oryzias latipes	Medaka	Ol	ENSORLT00000013758/XP_004068975	MALPAGQTGLQPNRGAPSGHSVPCENNIDIDEKELENITKKHREDDATALPLHSPWTFWLDRSLPGTTAAECESNLKKIYTVQTVQTFWSVYNNIPPVTALPLRCSYHLMRGERRPLWEEESNARGGVWKMKIPKESSAAVWKELLLATIGEQFADYCAIDDEVVGVSVSVRDREDVVQIWNSEASLANEANILGKVYELLPSISFKAVFYKSHMEHHAFEGGRSRH		XP_004070877	MAVPVAALQLSSQHNLRDLGRVPHDSESDTLPLHSSWTFWLDRSLPGTTAAECESNLKKIYTVETVQSFWRVYNNIPGVSSLPLRCSYHLMRGERKPLWEEESNAKGGVWKMKVPKECTPSVWKELLLATIGEQFSDFCASEDEVVGVSVSVRDREDVVQVWNENAFCSNESNVLERIYQLLPQISFKAVFYKPHKEHHAFEGGRSRH
Takifugu rubripes	Fugu	Tr	ENSTRUT00000008253/XP_003973808	MAVPLGQADPQMDRAALSGPAGSSQHDIDIDEQELENITKKHRDDGASQTLPLHSPWTFWLDRSLPGTTAAQCESNLKKIYTVQSVQMFWSVYNNIPLVTALPLRCSYHLMRGERRPLWEEDGNARGGVWKMKVPKDGTSDVWKELLLATIGEQFADYCASDDEVVGVSVSVRDREDVVQIWNRDASVASEANVLGKVHELLPFVSFRAVFYKPHMDHHAFEGGRSRH		XP_003963542	MAALQLSTSSSPVLLEHNMHVSDRDLGKTSTNEDDILPLHSPWTFWLDRSLPGTTAAQCESNLKKIYTVETVQNFWRVYNNIPSVSSLPLRCSYHLMRGERKPLWEEESNAKGGVWKMKVPKEYTSVVWKELLLATIGEQFSDYCASEDEVVGVSVSIRDREDVIQVWNGNASCANKSNILGRIHELLPHTPFKAVFYKPHEEHHAFEGGRSRR
Gadus morhua	cod	Gm	ENSGMOP00000007894	TTSSENNIDIDEKELEKITKNHREDPTALTLHSPWTFWLDRSLPGTTAAECASNLKKIYTVKTVQTFWSVYNNIPQVTALPLRCSYHLMRGERRPLWEEESNAKGGVWKMKVPKDSTAVVWKELLLATIGEQFADYCASADEVIGVSVSVRDREDVVQVWNGNAPLAQEASILAKIYELLPQISFKAVFYKSHQEHHAFEGGRSRH		none found	
Gasterosteus aculeatus	stickleback	Ga	ENSGACT00000015686	AAAMSVPAGRTVTPERGIDIDEEELENIARSQRGGEAAQLPLHSPWTFWLDRSLPGTTAAQCESGLKKIYTVQSVQLFWSVYNNIPAATALPLRCSYHLMRGERRPLWEEESNAKGGVWKMKIPKESTSAVWKELLLATIGEQFADYCSSDDEVVGVSVSVRDREDVVQIWNSDASLAGEANILGKVYELLPHMSFKAVFYKSHMEHRAFEGGRSRH		none found	
Tetraodon nigroviridis	pufferfish	Tn		MAALQLSTSSSPVLLEHNIHISDRDLGKISSNENDILPLHSPWTFWLDRSLPGTTAAECESNLKKIYTVETVQNFWRVYNNIPSVSSLPLRCSYHLMRGERKPLWEEESNAKGGVWKMKVPKEYTSVVWKELLLATIGEQFSDYCALEDEVVGVSVSIRDREDVFQVWNGNACCANKSDILGRIHELLPHTPFKAVFYKPHEEHHAFEGGRSRR*		none found	

Percomorpha Teleosts				eIF4E-3			eIF4E-3 Like
organism name (genus, species)	common name	abbrev	Accession	seq		Accession	seq
Cynoglossus semilaevis	tongue sole	Cs	XP_008318656.1	MAVPVGQADVQMDRAALSGQTVSSESDIDIDEKELENITKKHSEEPSLPLHSPWTFWLDRSLPGTTAACESNLKKIYTVQTVQMFWSVYNNIPPVTALPLRCSYHLMRGERRPLWEEESNANGGVWKMKTPKESTAVVWKELLLATIGEQFADYCSSDDEVVGVSVSVRDREDVVQIWNKDASLANEANILGKVYELLPYISFKAVFYKSHMEHHAFEGGRSRH		XP_008317405.1	MAVPVAVFPPSTSSSSPALLEHNIHMNDRDLTRIPTNTDIDKLPLHSPWTFWLDRSLPGTTAAECESNLKKIYTVETVQNFWRVYNNIPGISSLPLRCSYHLMRGERKPLWEEESNAKGGVWKMKVPKECTSAVWKELLLATIGEQFSDYCAEEDEVVGVSVSVRDREDVVQVWNGNASCVNDSKVLERINELLPQTPFKAVFYKPHEEHHAFERGRARY
Pundamilia nyererei	cichlids	Pn	XP_005726669.1	MAVPAGQTDVQMDRGALSGQTVSSENNIDIDEKELENITKKHREEDTATLPLHSPWTFWLDRSLPGTTAAECESNLKKIYTVQTVQMFWSVYNNIPPVTALPLRCSYHLMRGERRPLWEEESNAKGGVWKMKIPKENTSAVWKELLLATIGEQFADYCASDDEVVGVSVSVRDREDVVQVWNSDASLANEANILGKVYELLPYISFKAVFYKSHMEHHAFEGGRSRH		XP_005731770.1	MAVPVAALQLSSSSSPSLLEHNIHISDRNLGGISNNNNNGNDTLPLHSPWTFWLDRSLPGTTAAECESNLKKIYTVETVQNFWRVYNNIPGVSSLPLRWSYHLMRGERKPLWEEESNAKGGVWKMKVPKECTSAVWKELLLAIIGEQFSDYCAAEDEVVGVSVSIRDREDVVQVWNGNASCVSESNIIGRINELLPQVPFKAVFYKPHEEHHAFEGGRSRH

				eIF4E-3			eIF4E-3 Like
organism name (genus, species)	common name	abbrev	Accession	seq		Accession	seq
Esox Lucius	northern pike	El	XP_010880364.1	MAVPAGTNPQTNTANPSENDIHIDEKELENLTKHHEDGSTFPLHSPWTFWLDRSLPGTTAAECESNLKKIYTVQNVQSFWSVYNNIPPVIALPLRCSYHLMRGERRPLWEEESNSKGGVWKMKIPKESTLAVWKELLLATIGEQFADYCASEDEVVGVSVSIRDRDDVVQVWNGNASLASEANILGKVYELLPHITFKAVFYKSHREHHAF		XP_010873758	MSQQPANEALCTSKGETKSFEKTLKSFIFTNEEMAVPAVPNLQLNNTASKNSPERIIHIDESELENITNEDGNGTFPLHSPWTFWLDRSLPGTTAAECESGLKKIYTVQTVQSFWSVYNNIPGVSSLSLRCSYHLMRGERRPLWEEESNAKGGVWKMKVPKESTSAVWKELLLATIGEQFTDYCASDDEVVGVSVSVRDREDVVQIWNGNASLANEANILGRIYELLPQITFKAVFYKPHEEHHAFEG

Figure A2.1 Multiple alignments Class I eIF4E family members



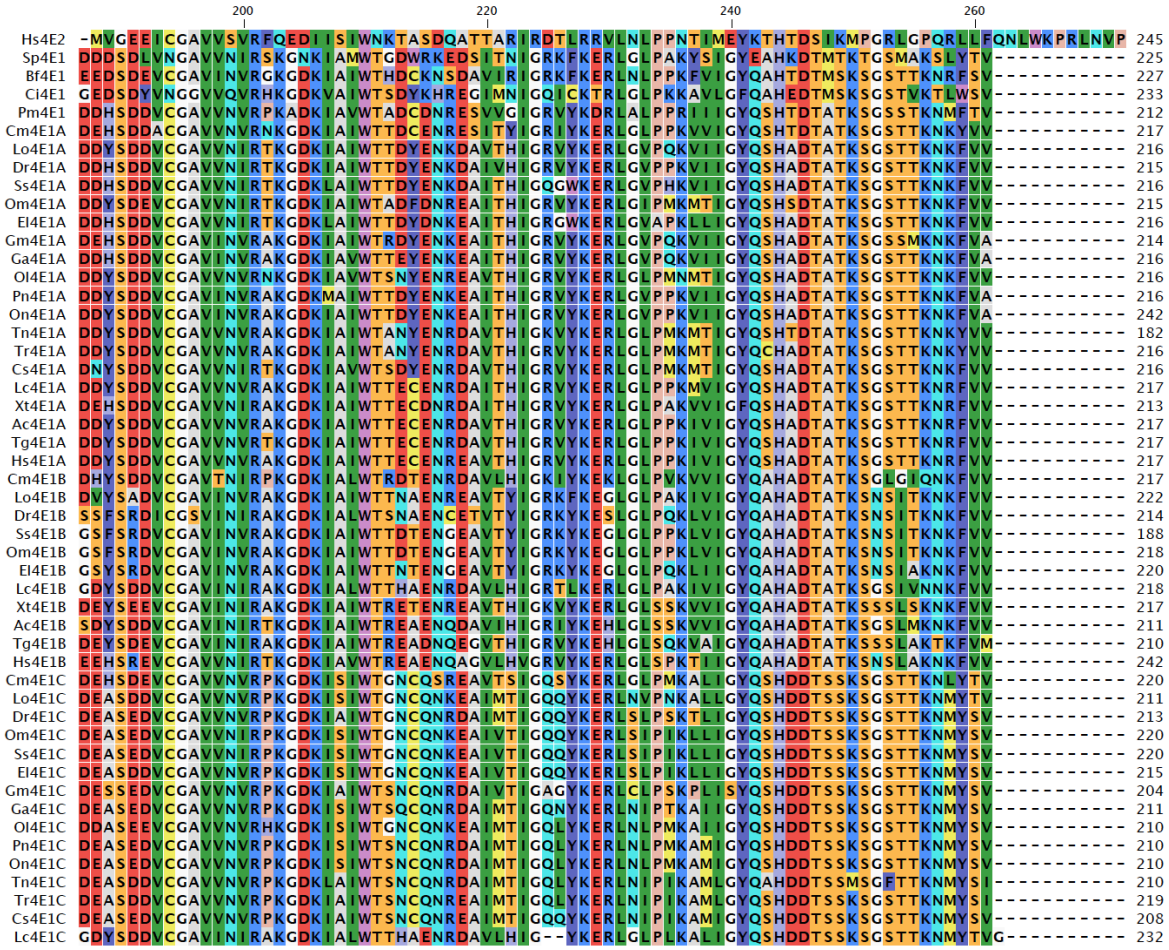
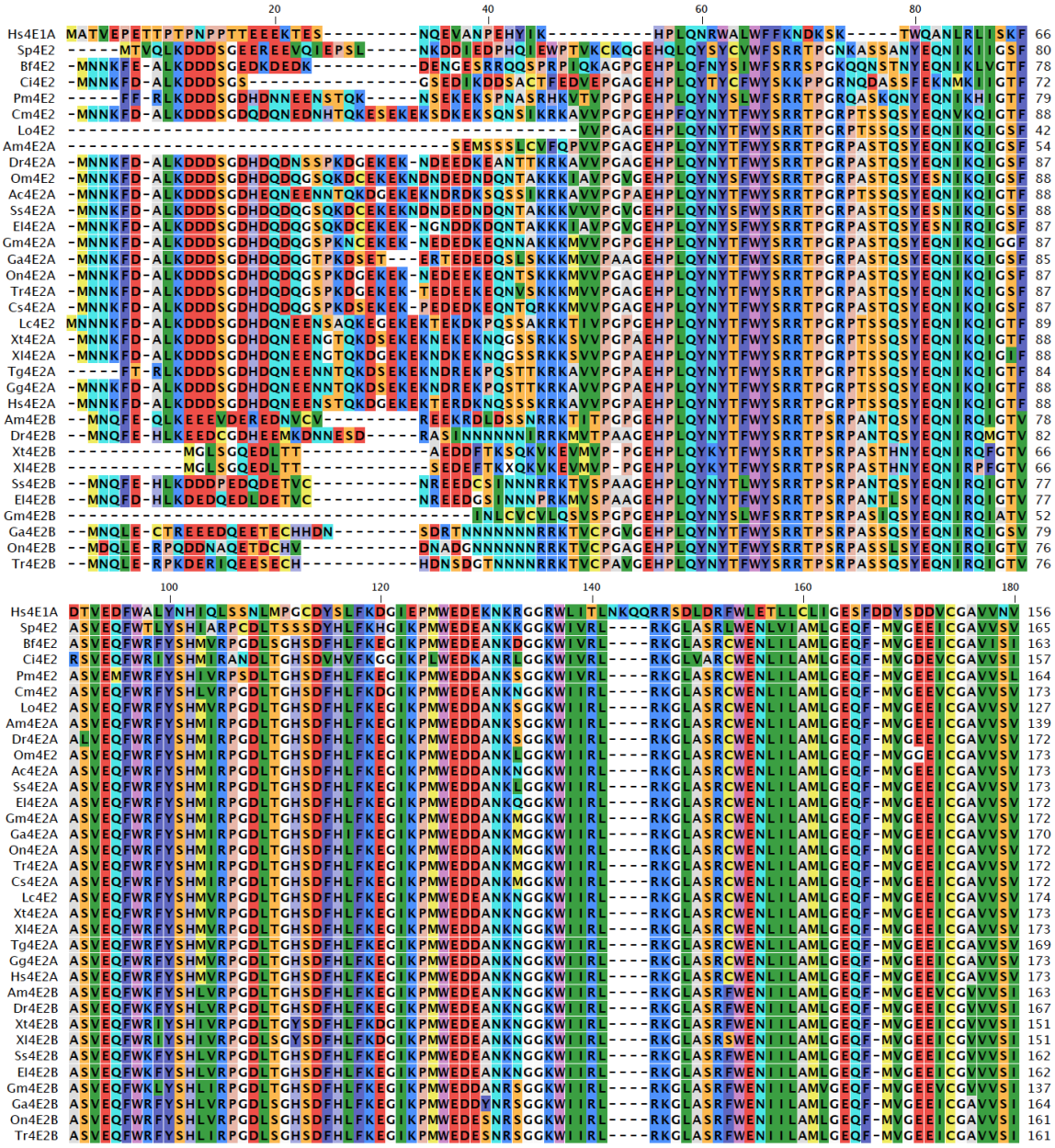

200
220
240
260
Hs4E2 -MVGEEICGAVVSVRFQEDIISIWNKTASDQATTARIRDTLRRVLNLPPNTIMEYKTHTDSIKMPGRLGPQRLLFQNLWKPRLNVP 245
Sp4E1 DDDSDLVNGAVVNIRSKGNKIAMWTGDWRKEDSITNIGRKFKERLGLPAKYSIGYEAHKDTMTKTGSMAKSLYTV---------- 225
Bf4E1 EEDSDEVCGAVINVRGKGDKIAIWTHDCKNSDAVIRIGRKFKERLNLPPKFVIGYQAHTDTMSKSGSTTKNRFSV---------- 227
Ci4E1 GEDSDYVNGGVVQVRHKGDKVAIWTSDYKHREGIMNIGQICKTRLGLPKKAVLGFQAHEDTMSKSGSTVKTLWSV---------- 233
Pm4E1 DDHSDDVCGAVVNVRPKADKIAVWTADCDNRESVVGIGRVYKDRLALPPRIIIGYQSHTDTATKSGSSTKNMFTV---------- 212
Cm4E1A DEHSDDACGAVVNVRNKGDKIAIWTTDCENRESITYIGRIYKERLGLPPKVVIGYQSHTDTATKSGSTTKNKYVV---------- 217
Lo4E1A DDYSDDVCGAVVNIRTKGDKIAIWTTDYENKDAVTHIGRVYKERLGVPQKVIIGYQSHADTATKSGSTTKNKFVV---------- 216
Dr4E1A DDHSDDVCGAVVNIRTKGDKIAIWTTDYENKDAIVHIGRVYKERLGVPPKVIIGYQSHADTATKSGSTTKNKFVV---------- 215
Ss4E1A DDHSDDVCGAVVNIRTKGDKLAIWTTDYENKDAITHIGQGWKERLGVPHKVIIGYQSHADTATKSGSTTKNKFVV---------- 216
Om4E1A DDYSDEVCGAVVNIRTKGDKIAIWTADFDNREAITHIGRVYKERLGIPMKMTIGYQSHSDTATKSGSTTKNKFVV---------- 215
El4E1A DDHSDDVCGAVVNIRTKGDKLAIWTTDYDNKEAITHIGRGWKERLGVAPKLLIGYQSHADTATKSGSTTKNKFVV---------- 216
Gm4E1A DEHSDDVCGAVINVRAKGDKIAIWTRDYENKEAITHIGRVYKERLGVPQKVIIGYQSHADTATKSGSSMKNKFVA---------- 214
Ga4E1A DDHSDDVCGAVINVRAKGDKIAVWTTEYENKEAITHIGRVYKERLGVPQKVIIGYQSHADTATKSGSTTKNKFVA---------- 216
Ol4E1A DDYSDDVCGAVVNVRNKGDKIAVWTSNYENREAVTHIGRVYKERLGLPMNMTIGYQSHADTATKSGSTTKNKFVV---------- 216
Pn4E1A DDYSDDVCGAVINVRAKGDKMAIWTTDYENKEAITHIGRVYKERLGVPPKVIIGYQSHADTATKSGSTTKNKFVA---------- 216
On4E1A DDYSDDVCGAVINVRAKGDKIAIWTTDYENKEAITHIGRVYKERLGVPPKVIIGYQSHADTATKSGSTTKNKFVA---------- 242
Tn4E1A DDYSDDVCGAVVNVRAKGDKIAIWTANYENRDAVTHIGKVYKERLGLPMKMTIGYQSHTDTATKSGSTTKNKYVV---------- 182
Tr4E1A DDYSDDVCGAVVNVRAKGDKIAIWTANYENRDAVTHIGRVYKERLGLPMKMTIGYQCHADTATKSGSTTKNKYVV---------- 216
Cs4E1A DNYSDDVCGAVVNIRTKGDKIAVWTSDYENRDAVTHIGRVYKERLGLPMKMTIGYQSHADTATKSGSTTKNKFVV---------- 216
Lc4E1A DDYSDDVCGAVVNVRAKGDKIAIWTTECENRDAITHIGRVYKERLGLPPKMVIGYQSHADTATKSGSTTKNRFVV---------- 217
Xt4E1A DEHSDDVCGAVVNIRAKGDKIAIWTTECDNRDAITHIGRVYKERLGLPAKVVIGFQSHADTATKSGSTTKNRFVV---------- 213
Ac4E1A DDYSDDVCGAVVNVRAKGDKIAIWTTECENRDAVTHIGRVYKERLGLPPKIVIGYQSHADTATKSGSTTKNRFVV---------- 217
Tg4E1A DDYSDDVCGAVVNVRTKGDKIAIWTTECENRDAVTHIGRVYKERLGLPPKIVIGYQSHADTATKSGSTTKNRFVV---------- 217
Hs4E1A DDYSDDVCGAVVNVRAKGDKIAIWTTECENREAVTHIGRVYKERLGLPPKIVIGYQSHADTATKSGSTTKNRFVV---------- 217
Cm4E1B DHYSDDVCGAVTNIRPKGDKIALWTRDTENRDAVLHIGKIYKEKLGLPVKVVIGYQAHADTATKSGLGIQNKFVV---------- 217
Lo4E1B DVYSADVCGAVINVRAKGDKIAIWTTNAENREAVTYIGRKFKEGLGLPAKIVIGYQAHADTATKSNSITKNKFVV---------- 222
Dr4E1B SSFSRDICGSVINIRAKGDKIALWTSNAENCETVTYIGRKYKESLGLPQKLVIGYQAHADTATKSNSITKNKFVV---------- 214
Ss4E1B GSFSRDVCGAVINVRAKGDKIAIWTTDTENGEAVTYIGRKYKEGLGLPPKLVIGYQAHADTATKSNSITKNKFVV---------- 188
Om4E1B GSFSRDVCGAVINVRAKGDKIAIWTTDTENGEAVTYIGRKYKEGLGLPPKLVIGYQAHADTATKSNSITKNKFVV---------- 218
El4E1B GSYSRDVCGAVINVRAKGDKIAIWTTNTENGEAVTYIGRKYKEGLGLPQKLIIGYQAHADTATKSNSIAKNKFVV---------- 220
Lc4E1B GDYSDDVCGAVINIRAKGDKIALWTTHAENRDAVLHIGRTLKERLGLPAKIVIGYQAHADTATKSGSIVNNKFVV---------- 218
Xt4E1B DEYSEEVCGAVINIRAKGDKIAIWTRETENREAVTHIGKVYKERLGLSSKVVIGYQAHADTATKSSSLSKNKFVV---------- 217
Ac4E1B SDYSDDVCGAVINIRTKGDKIAIWTREAENQDAVIHIGRIYKEHLGLSSKVVIGYQAHADTATKSGSLMKNKFVV---------- 211
Tg4E1B DEYSDEVCGAVINIRAKGDKIAIWTREADNQEGVTHIGRVYKEHLGLSQKVAIGYQAHADTATKSSSLAKTKFVM---------- 210
Hs4E1B EEHSREVCGAVVNIRTKGDKIAVWTREAENQAGVLHVGRVYKERLGLSPKTIIGYQAHADTATKSNSLAKNKFVV---------- 242
Cm4E1C DEHSDEVCGAVVNVRPKGDKISIWTGNCQSREAVTSIGQSYKERLGLPMKALIGYQSHDDTSSKSGSTTKNLYTV---------- 220
Lo4E1C DEASDDVCGAVVNVRPKGDKISIWTGNCQNKEAIMTIGQQYKERLNVPNKALLGYQSHDDTSSKSGSTTKNMYTV---------- 211
Dr4E1C DEASEDVCGAVVNVRPKGDKIAIWTGNCQNRDAIMTIGQQYKERLSLPSKTLIGYQSHDDTSSKSGSTTKNMYSV---------- 213
Om4E1C DEASEDVCGAVVNIRPKGDKISIWTGNCQNKEAIVTIGQQYKERLSIPIKLLIGYQSHDDTSSKSGSTTKNMYSV---------- 220
Ss4E1C DEASEDVCGAVVNIRPKGDKISIWTGNCQNKEAIVTIGQQYKERLSIPIKLLIGYQSHDDTSSKSGSTTKNMYSV---------- 220
El4E1C DEASDDVCGAVVNVRPKGDKISIWTGNCQNKEAIVTIGQQYKERLSLPIKLLIGYQSHDDTSSKSGSTTKNMYSV---------- 215
Gm4E1C DESSEDVCGAVVNVRPKGDKIAIWTSNCQNRDAIVTIGAGYKERLCLPSKPLISYQSHDDTSSKSGSTTKNMYSV---------- 204
Ga4E1C DEASEDVCGAVVNVRPKGDKISIWTSQCQNRDAIMTIGQNYKERLNIPTKAIIGYQSHDDTSSKSGSTTKNMYSV---------- 211
Ol4E1C DDASEEVCGAVVNVRHKGDKISIWTGNCQNKEAIMTIGQLYKERLNLPMKAIIGYQSHDDTSSKSGSTTKNMYSV---------- 210
Pn4E1C DEASEDVCGAVVNVRPKGDKISIWTSNCQNRDAIMTIGQLYKERLNLPMKAMIGYQSHDDTSSKSGSTTKNMYSV---------- 210
On4E1C DEASEDVCGAVVNVRPKGDKISIWTSNCQNRDAIMTIGQLYKERLNLPMKAMIGYQSHDDTSSKSGSTTKNMYSV---------- 210
Tn4E1C DEASDDVCGAVVNVRPKGDKLAIWTSNCQNRDAIMTIGQLYKERLNIPIKAMLGYQAHDDTSSMSGFTTKNMYSI---------- 210
Tr4E1C DEASDDVCGAVVNVRPKGDKIAIWTSNCQNREAIMTIGQLYKERLNIPIKAMLGYQSHDDTSSKSGSTTKNMYSI---------- 219
Cs4E1C DEASEDVCGAVVNVRPKGDKIAIWTSNCQNREAIMTIGQQYKERLNIPIKAMIGYQSHDDTSSKSGSTTKNMYSV---------- 208
Lc4E1C GDYSDDVCGAVINIRAKGDKIALWTTHAENRDAVLHIG--YKERLGLPLKALIGYQSHDDTSSKSGSTTKNMYTVG---------- 232

Figure A2.2 Multiple alignments Class II eIF4E family members



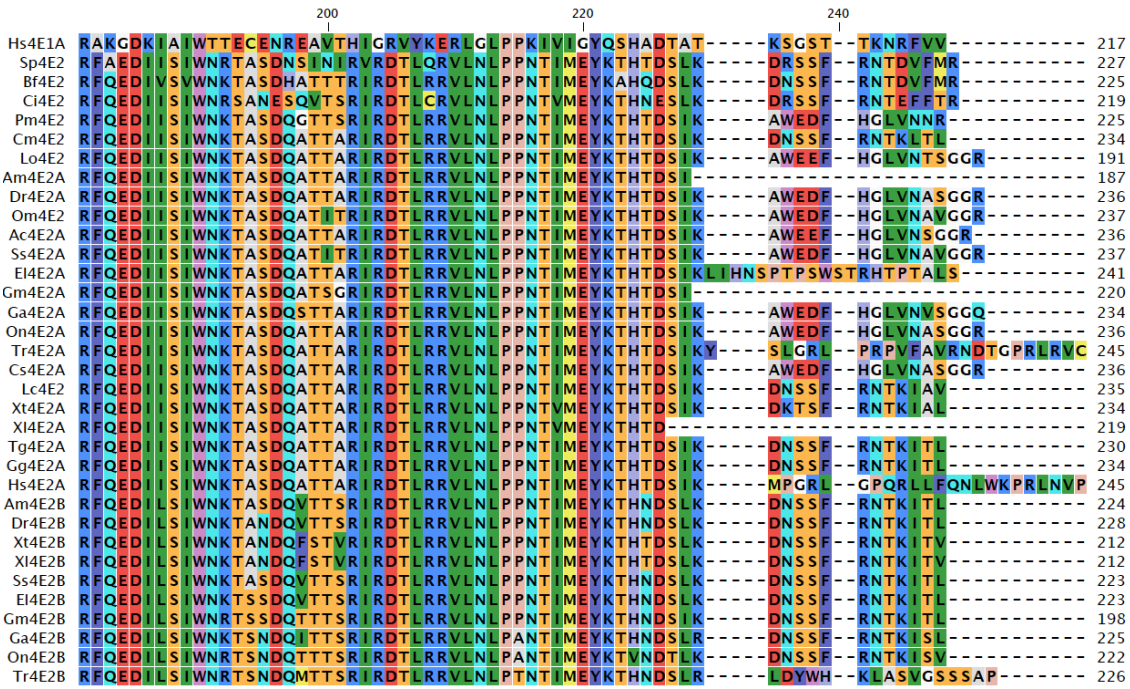
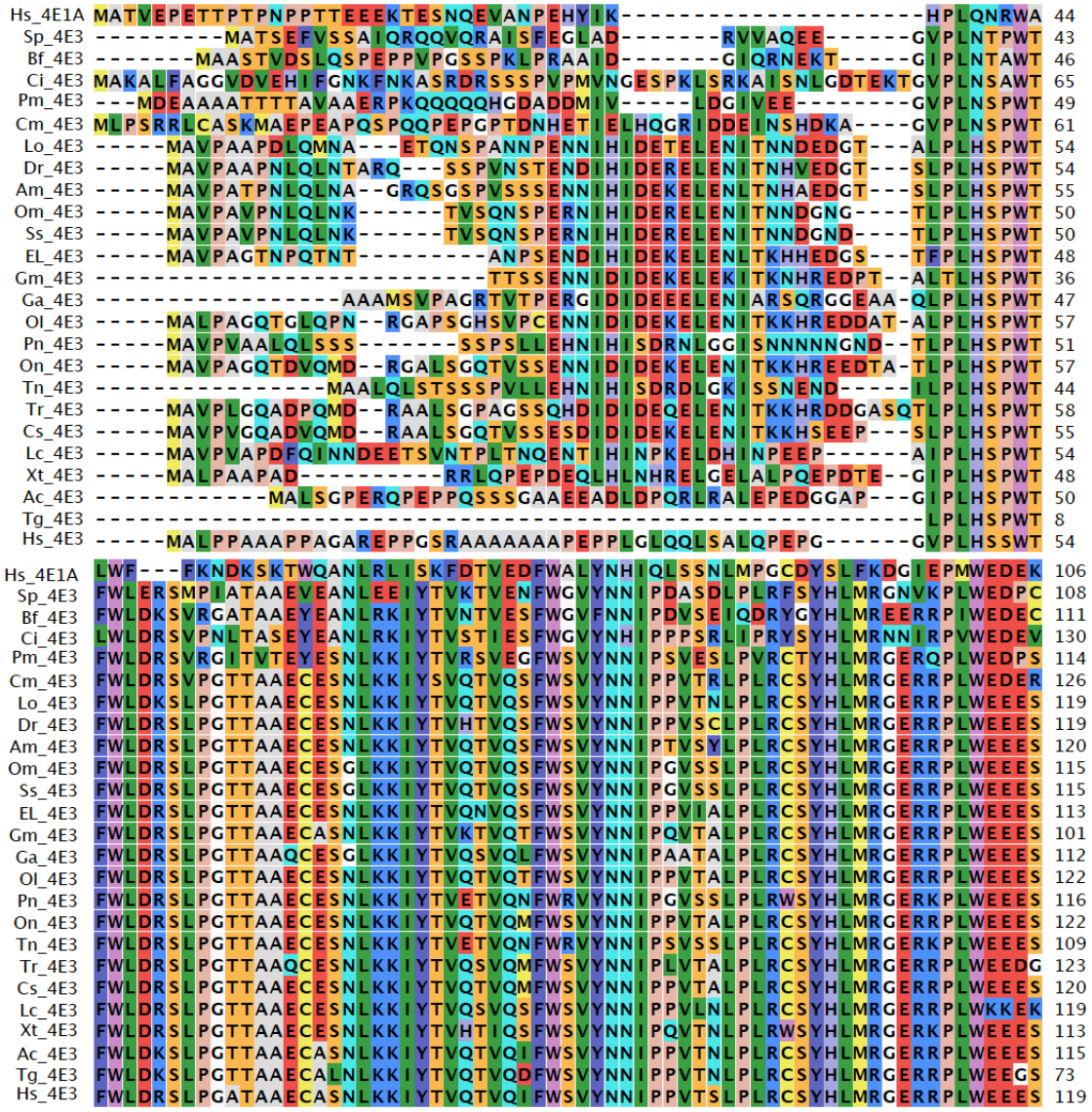

200
220
240
Hs4E1A RAKGDKIAIWTTECENREAVTHIGRVYKERLGLPPKIVIGYQSHADTAT-----KSGST--TKNRFVV---------- 217
Sp4E2 RFAEDIISIWNRTASDNSINIRVRDTLQRVLNLPPNTIMEYKTHTDSLK-----DRSSF--RNTDVFMR---------- 227
Bf4E2 RFQEDIVSVWNKTASDHATTTRIRDTLRRVLNLPPNTIMEYKAHQDSLK-----DNSSF--RNTDVFMR---------- 225
Ci4E2 RFQEDIISIWNRSANESQVTSRIRDTLCRVLNLPPNTVMEYKTHNESLK-----DRSSF--RNTEFFTR---------- 219
Pm4E2 RFQEDIISIWNKTASDQGTTSRIRDTLRRVLNLPPNTIMEYKTHTDSIK-----AWEDF--HGLVNNR---------- 225
Cm4E2 RFQEDIISIWNKTASDQATTARIRDTLRRVLNLPPNTIMEYKTHTDSIK-----DNSSF--RNTKLTL---------- 234
Lo4E2 RFQEDIISIWNKTASDQATTARIRDTLRRVLNLPPNTIMEYKTHTDSIK-----AWEEF--HGLVNTSGGR-------- 191
Am4E2A RFQEDIISIWNKTASDQATTARIRDTLRRVLNLPPNTIMEYKTHTDSI---------- 187
Dr4E2A RFQEDIISIWNKTASDQATTARIRDTLRRVLNLPPNTIMEYKTHTDSIK-----AWEDF--HGLVNASGGR-------- 236
Om4E2 RFQEDIISIWNKTASDQATIITRIRDTLRRVLNLPPNTIMEYKTHTDSIK-----AWEDF--HGLVNAVGGR-------- 237
Ac4E2A RFQEDIISIWNKTASDQATTARIRDTLRRVLNLPPNTIMEYKTHTDSIK-----AWEEF--HGLVNSGGR--------- 236
Ss4E2A RFQEDIISIWNKTASDQATITRIRDTLRRVLNLPPNTIMEYKTHTDSIK-----AWEDF--HGLVNAVGGR-------- 237
El4E2A RFQEDIISIWNKTASDQATTARIRDTLRRVLNLPPNTIMEYKTHTDSIKLIHNSPTPSWSTRHTPTALS---------- 241
Gm4E2A RFQEDIISIWNKTASDQATSGRIRDTLRRVLNLPPNTIMEYKTHTDSI---------- 220
Ga4E2A RFQEDIISIWNKTASDQSTTARIRDTLRRVLNLPPNTIMEYKTHTDSIK-----AWEDF--HGLVNVSGGQ-------- 234
On4E2A RFQEDIISIWNKTASDQATTARIRDTLRRVLNLPPNTIMEYKTHTDSIK-----AWEDF--HGLVNASGGR-------- 236
Tr4E2A RFQEDIISIWNKTASDQATTARIRDTLRRVLNLPPNTIMEYKTHTDSIKY----SLGRL--PRPVFAVRNDTGPRLRVC 245
Cs4E2A RFQEDIISIWNKTASDQATTARIRDTLRRVLNLPPNTIMEYKTHTDSIK-----AWEDF--HGLVNASGGR-------- 236
Lc4E2 RFQEDIISIWNKTASDQATTARIRDTLRRVLNLPPNTIMEYKTHTDSIK-----DNSSF--RNTKIAV---------- 235
Xt4E2A RFQEDIISIWNKTASDQATTARIRDTLRRVLNLPPNTVMEYKTHTDSIK-----DKTSF--RNTKIAL---------- 234
Xl4E2A RFQEDIISIWNKTASDQATTARIRDTLRRVLNLPPNTVMEYKTHTD---------- 219
Tg4E2A RFQEDIISIWNKTASDQATTARIRDTLRRVLNLPPNTIMEYKTHTDSIK-----DNSSF--RNTKITL---------- 230
Gg4E2A RFQEDIISIWNKTASDQATTARIRDTLRRVLNLPPNTIMEYKTHTDSIK-----DNSSF--RNTKITL---------- 234
Hs4E2A RFQEDIISIWNKTASDQATTARIRDTLRRVLNLPPNTIMEYKTHTDSIK-----MPGRL--GPQRLLFQNLWKPRLNVP 245
Am4E2B RFQEDILSIWNKTASDQVTTSRIRDTLRRVLNLPPNTIMEYKTHNDSLK-----DNSSF--RNTKITL---------- 224
Dr4E2B RFQEDILSIWNKTANDQVTTSRIRDTLRRVLNLPPNTIMEYKTHNDSLK-----DNSSF--RNTKITL---------- 228
Xt4E2B RFQEDILSIWNKTANDQFSTVRIRDTLRRVLNLPPNTIMEYKTHTDSLK-----DNSSF--RNTKITV---------- 212
Xl4E2B RFQEDILSIWNKTANDQFSTVRIRDTLRRVLNLPPNTIMEYKTHTDSLK-----DNSSF--RNTKITV---------- 212
Ss4E2B RFQEDILSIWNKTASDQVTTSRIRDTLRRVLNLPPNTIMEYKTHNDSLK-----DNSSF--RNTKITL---------- 223
El4E2B RFQEDILSIWNKTSSDQVTTSRIRDTLRRVLNLPPNTIMEYKTHNDSLK-----DNSSF--RNTKITL---------- 223
Gm4E2B RFQEDILSIWNRTSSDQTTTSRIRDTLRRVLNLPPNTIMEYKTHNDSIK-----DNSSF--RNTKITL---------- 198
Ga4E2B RFQEDILSIWNKTSNDQITTSRIRDTLRRVLNLPANTIMEYKTHNDSLR-----DNSSF--RNTKISL---------- 225
On4E2B RFQEDILSIWNRTSNDQTTTSRIRDTLRRVLNLPANTIMEYKTVNDTLK-----DNSSF--RNTKISV---------- 222
Tr4E2B RFQEDILSIWNRTSNDQMTTSRIRDTLRRVLNLPTNTIMEYKTHNDSLR-----LDYWH--KLASVGSSSAP-------- 226

Figure A2.3 Multiple alignments Class III eIF4E family members



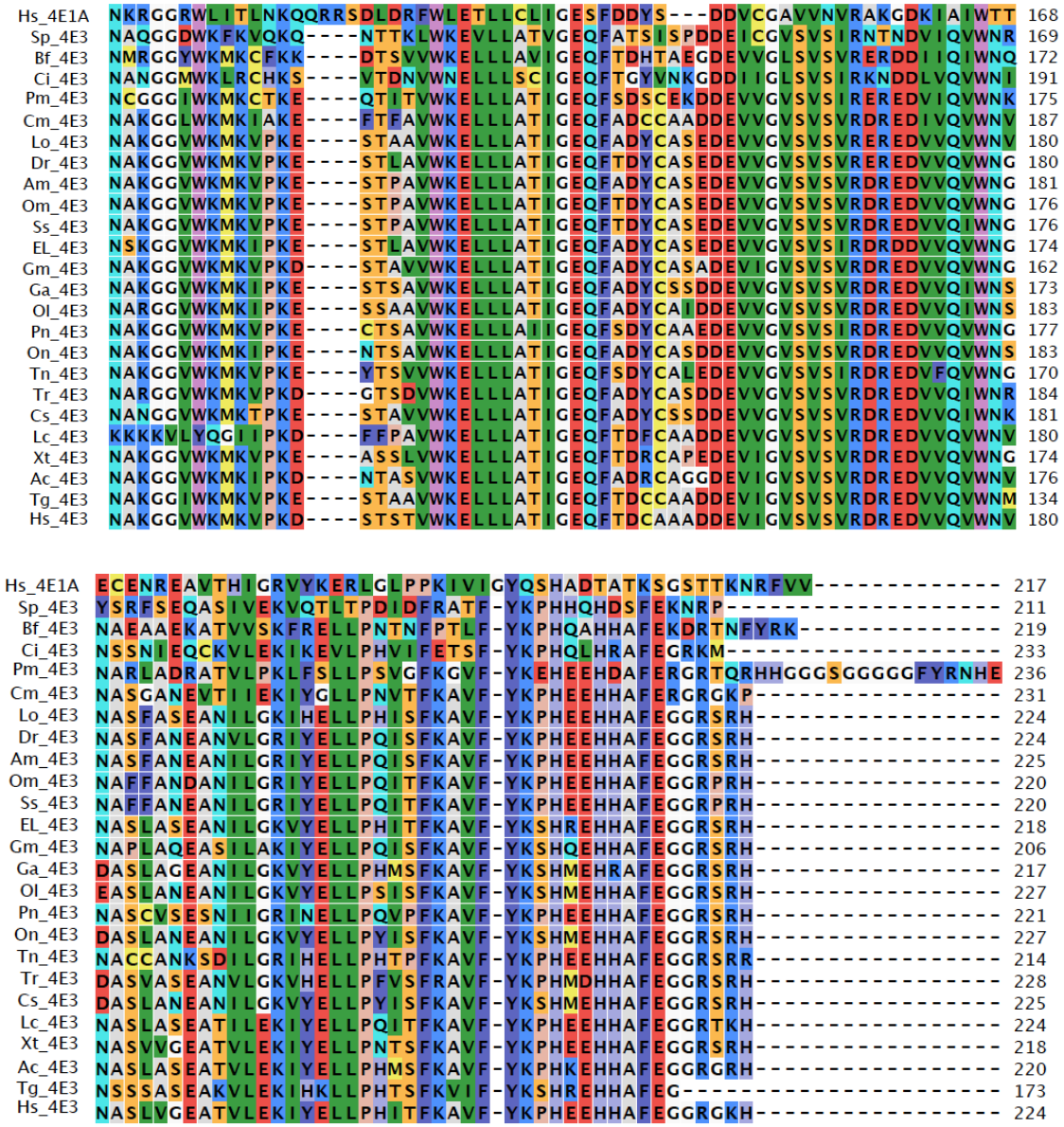

Hs_4E1A NKRGGRWLITLNKQQRRSDLDRFWLETLLCLIGESFDDYS---DDVCGAVVNVRAKGDKIAIWTT 168
Sp_4E3 NAQGGDWKFKVQKQ----NTTKLWKEVLLATVGEQFATSISPDDEICGVSVSIRNTNDVIQVWNR 169
Bf_4E3 NMRGGYWKMKCFKK----DTSVVWKELLLAVIGEQFTDHTAEGDEVVGLSVSVRERDDIIQIWNQ 172
Ci_4E3 NANGGMWKLRCHKS----VTDNVWNELLLSCIGEQFTGYVNKGDDIIGLSVSIRKNDDLVQVWNI 191
Pm_4E3 NCGGGIWKMKCTKE----QTITVWKELLLATIGEQFSDSCEKDDEVVGVSVSIRERDEDVIQVWNK 175
Cm_4E3 NAKGGLWKMKIAKE----FTFAVWKELLLATIGEQFADCCAADDEVVGVSVSVRDREDIVQVWNV 187
Lo_4E3 NAKGGVWKMKVPKE----STAAVWKELLLATIGEQFADYCASEDEVVGVSVSVRERDEDVVQVWNV 180
Dr_4E3 NAKGGVWKMKVPKE----STLAVWKELLLATIGEQFTDYCASEDEVVGVSVSVRERDEDVVQVWNG 180
Am_4E3 NAKGGVWKMKVPKE----STPAVWKELLLATIGEQFADYCASEDEVVGVSVSVRDREDVVQVWNG 181
Om_4E3 NAKGGVWKMKVPKE----STPAVWKELLLATIGEQFTDYCASEDEVVGVSVSVRDREDVVQVWNG 176
Ss_4E3 NAKGGVWKMKVPKE----STPAVWKELLLATIGEQFTDYCASEDEVVGVSVSVRDREDVVQIWNG 176
EL_4E3 NSKGGVWKMKIPKE----STLAVWKELLLATIGEQFADYCASEDEVVGVSVSIRDRDDVVQVWNG 174
Gm_4E3 NAKGGVWKMKVPKD----STAVVWKELLLATIGEQFADYCASADEVIGVSVSVRDREDVVQVWNG 162
Ga_4E3 NAKGGVWKMKIPKE----STSAVWKELLLATIGEQFADYCSSDDEVVGVSVSVRDREDVVQIWNS 173
Ol_4E3 NARGGVWKMKIPKE----SSAAVWKELLLATIGEQFADYCAIDDEVVGVSVSVRDREDVVQIWNS 183
Pn_4E3 NAKGGVWKMKVPKE----CTSAVWKELLLAIIGEQFSDYCAAEDEVVGVSVSIRDREDVVQVWNG 177
On_4E3 NAKGGVWKMKIPKE----NTSAVWKELLLATIGEQFADYCASDDEVVGVSVSVRDREDVVQVWNS 183
Tn_4E3 NAKGGVWKMKVPKE----YTSVVWKELLLATIGEQFSDYCALEDEVVGVSVSIRDREDVFQVWNG 170
Tr_4E3 NARGGVWKMKVPKD----GTSDVWKELLLATIGEQFADYCASDDEVVGVSVSVRDREDVVQIWNR 184
Cs_4E3 NANGGVWKMKTPKE----STAVVWKELLLATIGEQFADYCSSDDEVVGVSVSVRDREDVVQIWNK 181
Lc_4E3 KKKKVLYQGIIPKD----FFPAVWKELLLATIGEQFTDFCAADDEVVGVSVSVRDREDVVQVWNV 180
Xt_4E3 NAKGGVWKMKVPKE----ASSLVWKELLLATIGEQFTDRCAPEDEVIGVSVSVRDREDVVQVWNG 174
Ac_4E3 NAKGGVWKMKIPKD----NTASVWKELLLATIGEQFADRCAGGDEVIGVSVSVRDREDVVQVWNV 176
Tg_4E3 NAKGGIWKMKVPKE----STAAVWKELLLATIGEQFTDCCAADDEVIGVSVSVRDREDVVQVWNM 134
Hs_4E3 NAKGGVWKMKVPKD----STSTVWKELLLATIGEQFTDCAAADDEVIGVSVSVRDREDVVQVWNV 180
Hs_4E1A ECENREAVTHIGRVYKERLGLPPKIVIGYQSHADTATKSGSTTKNRFVV---------- 217
Sp_4E3 YSRFSEQASIVEKVQTLTPDIDFRATF-YKPHHQHDSFEKNRP---------- 211
Bf_4E3 NAEAAEKATVVSKFRELLPNTNFPTLF-YKPHQAHHAFEKDRTNFYRK---------- 219
Ci_4E3 NSSNIEQCKVLEKIKEVLPHVIFETSF-YKPHQLHRAFEGRKM---------- 233
Pm_4E3 NARLADRATVLPKLFSLLPSVGFKGVF-YKEHEEHDAFERGRTQRHHGGGSGGGGGFYRNHE 236
Cm_4E3 NASGANEVTIIEKIYGLLPNVTFKAVF-YKPHEEHHAFERGRGKP---------- 231
Lo_4E3 NASFASEANILGKIHELLPHISFKAVF-YKPHEEHHAFEGGRSRH---------- 224
Dr_4E3 NASFANEANVLGRIYELLPQISFKAVF-YKPHEEHHAFEGGRSRH---------- 224
Am_4E3 NASFANEANILGRIYELLPQISFKAVF-YKPHEEHHAFEGGRSRH---------- 225
Om_4E3 NAFFANDANILGRIYELLPQITFKAVF-YKPHEEHHAFEGGRPRH---------- 220
Ss_4E3 NAFFANEANILGRIYELLPQITFKAVF-YKPHEEHHAFEGGRPRH---------- 220
EL_4E3 NASLASEANILGKVYELLPHITFKAVF-YKSHREHHAFEGGRSRH---------- 218
Gm_4E3 NAPLAQEASILAKIYELLPQISFKAVF-YKSHQEHHAFEGGRSRH---------- 206
Ga_4E3 DASLAGEANILGKVYELLPHMSFKAVF-YKSHMEHRAFEGGRSRH---------- 217
Ol_4E3 EASLANEANILGKVYELLPSISFKAVF-YKSHMEHHAFEGGRSRH---------- 227
Pn_4E3 NASCVSESNIIGRINELLPQVPFKAVF-YKPHEEHHAFEGGRSRH---------- 221
On_4E3 DASLANEANILGKVYELLPYISFKAVF-YKSHMEHHAFEGGRSRH---------- 227
Tn_4E3 NACCANKSDILGRIHELLPHTPFKAVF-YKPHEEHHAFEGGRSRR---------- 214
Tr_4E3 DASVASEANVLGKVHELLPFVSFRAVF-YKPHMDHHAFEGGRSRH---------- 228
Cs_4E3 DASLANEANILGKVYELLPYISFKAVF-YKSHMEHHAFEGGRSRH---------- 225
Lc_4E3 NASLASEATILEKIYELLPQITFKAVF-YKPHEEHHAFEGGRTKH---------- 224
Xt_4E3 NASVVGEATVLEKIYELLPNTSFKAVF-YKPHEEHHAFEGGRSRH---------- 218
Ac_4E3 NASLASEATVLEKIYELLPHMSFKAVF-YKPHKEHHAFEGGRGRH---------- 220
Tg_4E3 NSSSASEAKVLEKIHKLLPHTSFKVIF-YKSHREHHAFEG---------- 173
Hs_4E3 NASLVGEATVLEKIYELLPHITFKAVF-YKPHEEHHAFEGGRGKH---------- 224

Table A3.1: Species used for eIF4E-1A and eIF4E-1B logos		
organism name	**common name**	**abbrev**
mammal		
Homo sapiens	human	Hs
bird		
Taeniopygia guttata	zebrafinch	Tg
reptile		
Anolis carolinensis	lizard	Ac
amphibian		
Xenopus tropicalis	frog	Xt
shark		
Callorhinchus milii	elephant shark	Cm
spiny fin fish		
Danio rerio	zebrafish	Dr
Oncorhynchus mykiss	rainbow trout	Om
Astyanax mexicanus	cavefish	Am
Lepisosteus oculatus	spotted gar	Lo
lobe fin fish		
Latimeria chalumnae	coelacanth	Lc
Additional species used for eIF4E-1C logo		
Spiny Fin Fish	**common name**	**abbrev**
Salmo Salar	salmon	Ss
Oreochromis niloticus	tilapia	On
Oryzias latipes	medaka	Ol
Takifugu rubripes	fugu	Tr
Gadus morhua	cod	Gm
Gasterosteus aculeatus	stickleback	Ga