Greco Hernández · Rosemary Jagus
Editors

Evolution of the Protein Synthesis Machinery and Its Regulation

Springer

Evolution of the Protein Synthesis Machinery and Its Regulation

Greco Hernández · Rosemary Jagus
Editors

Evolution of the Protein Synthesis Machinery and Its Regulation

Springer

Editors
Greco Hernández
Division of Basic Research
National Institute of Cancer
Mexico City
Mexico

Rosemary Jagus
Institute of Marine and Environmental Technology
University of Maryland Center for Environmental Science
Baltimore, MD
USA

ISBN 978-3-319-39467-1 ISBN 978-3-319-39468-8 (eBook)
DOI 10.1007/978-3-319-39468-8

Library of Congress Control Number: 2016940795

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Preface

Proteins are one of the elementary molecules of the biosphere. They catalyze the majority of life-sustaining reactions and play structural, transport, and regulatory roles in all living organisms. Protein synthesis or "translation" is the process of decoding the genetic information of a messenger RNA by the ribosome along with translation factors to synthesize a protein. Translation regulation allows organisms to (1) rapidly respond to a variety of stresses, sudden environmental changes, and nutritional deficiencies, (2) produce proteins in tissues and developmental processes where transcription is strongly limited, and (3) elicit asymmetric localization of proteins when and where required. Thus, translation is a fundamental process for gene expression in all forms of life and should have evolved ever since the beginning of life.

The knowledge of basic processes and regulatory mechanisms of translation was established in the last five decades by the brilliant work of many scientists in different countries, mostly studying the bacteria *Escherichia coli*, human, mouse, rabbit, the budding yeast *Saccharomyces cerevisiae*, and the fruit fly *Drosophila melanogaster* as model organisms. In recent years, the advent of the powerful "omics" era (i.e., genomics, transcriptomics, and proteomics) has created a novel perspective in the study of biological processes at the genome-wide and thousands-of-species scales from many phyla never studied before. These studies have led to crucial findings on the origin and evolution of the process of translation.

Here, we have gathered experts in different aspects of translation to review the state of the art of their respective fields in the attempt to answer the question of how the protein synthesis machinery and its regulation might have originated and evolved. We wish to thank the authors for their excellent contributions. We also thank our editor team at Springer, especially Janet Slobodien and Eric Hardy, for producing this book.

Mexico City, Mexico — Greco Hernández
Baltimore, MD, USA — Rosemary Jagus

Contents

Evolution of tRNAs Was Driven by Entropic Forces

Marco V. José, Gabriel S. Zamudio and Sávio Torres de Farías

1 Introduction

Undoubtedly, the transfer RNA (tRNA) molecule played a central role in the origin of translation. The tRNA has the remarkable ability of decoding the language of nucleic acids to the stereochemical language of proteins. The molecule tRNA is the most important molecule in the origin of life, bridging the RNA and (RNA + proteins) worlds. Translation is the evolutionary transition from a nucleic acid-based world to the protein-based world of modern cells [20, 21]. Indeed, the ribosomal peptidyl transferase center (PTC) seems to have originated from concatemer-containing anticodons of ancestral sequences of tRNAs [5]. The PTC, which catalyzes peptide synthesis, is one of the most ancient enzymes that we know of [19].

The tRNA molecule itself displays two codes, the operational code and the anticodon code. Typically, two genetic codes are considered, to wit, the "classic" code represented in tRNA by an anticodon for reading codons in mRNA, and the other is the "second" [1] operational RNA code [7, 15–17] mapped mainly to the acceptor for appropriate aminoacylation at its 3′ terminus. As far as translation is concerned, it does not make sense to consider one code without the other. The present-day operational RNA code is intricately carved in the structure of tRNA acceptors and cognate aminoacyl-tRNA synthetases (aaRSs), whereas the anticodon code is reduced to codon-anticodon interactions. It has been observed that in pairs of consensus tRNAs with complementary anticodons, second bases in their acceptor stems prove to be complementarily related as well [13]. This relationship

M.V. José (✉) · G.S. Zamudio
Theoretical Biology Group, Instituto de Investigaciones Biomédicas, Universidad Nacional Autónoma de México, D.F. 04510 Mexico City, Mexico
e-mail: marcojose@biomedicas.unam.mx

S.T. de Farías
Centro de Ciencias Exactas y Naturales, Universidade Federal da Paraíba, João Pessoa, Brazil

G. Hernández and R. Jagus (eds.), *Evolution of the Protein Synthesis Machinery and Its Regulation*, DOI 10.1007/978-3-319-39468-8_1

is known as the dual complementarity [13]. Hence, the two codes appear to be originally one and the same.

The canonical aaRSs are composed by two distinct groups of protein subfamilies, class I and class II, with ten proteins each [2]. The present correspondence of the two codes is provided by 20 specific aaRSs divided into two strikingly dissimilar classes of 10 members each. There are only 20 aaRSs, one for each amino acid (and, respectively, for isoacceptor tRNAs); hence, the operational code is nondegenerate [1]. Such a nondegeneracy, inherent only to the acceptor code, may indicate the historically subsidiary role of anticodons in aminoacylation. Otherwise, more than 20 aaRSs could exist, one for each anticodon rather than one for each amino acid.

The modular structure of aaRS suggests that these proteins evolved from two class-defining, rather small modules, each having sites to bind ATP, cognate amino acids, and acceptors, but not anticodons [7, 15–17].

The two aaRSs recognize the acceptor helix from opposite sides: class I aaRS approaches the helix from the side of its minor groove and attaches the amino acid to the 2′OH group of the terminal adenosine ribose, while class II aaRS approaches from the side of major groove and attaches the amino acid to the 3′OH group [2].

Phylogenies proposed for each class separately show that these classes are monophyletic and have a pattern of diversification related to the chemical characteristics of the cognate amino acid [3, 11, 14].

Nowadays, we typically consider 64 triplets, 61 codons for 20 amino acids, 3 stop codons, 46 tRNAs, 2 classes of aaRSs, one classic genetic code, the operational code, and the anticodon code.

It has been proposed that the amino acid-accepting stem emerged before the anticodon loop of tRNAs, so that the first codification obeyed an operational code where amino acids were attached to their respective tRNA, without the need of anticodon loop recognition [7, 12, 17, 18].

However, it has also been shown that the specificities between tRNAs and aaRSs coevolved during the formation of the genetic code and were established by correlations of hydropathy of the anticodons between them [4].

We ignore how a given aaRS "knows" which amino acid has to be charged for the 46 types of tRNA. For example, given a hexacodonic amino acid, say Ser, there will be only one aaRS that will charge Ser to four different tRNAs (and its isoacceptors) with its corresponding anticodons.[1] In the genetic anticode, the wobbling of the third base of the codon and the fact that there are no anticodons that start with adenine must be taken into account.

There is no need for the aaRS to recognize the anticodon in order to properly aminoacylate the tRNA. This means that the two codes coevolved right at the origin of translation. This encoding system seems now lost in the dimness of the past.

[1]The codons for Ser are 6: UCC, UCU, UCA, UCG, AGC, and AGU, and their corresponding anticodons are 4: GGA, GGA, UGA, CGA, GCU, and GCU, respectively. tRNAS with AGA or ACU anticodons do not exist in the canonical standard or classic code.

After all, the second operational code does not make sense without the standard genetic code. However, the early relevance of the acceptor mini-helix in evolutionary development of the tRNA molecule cannot be understated [17]. The presumable antiquity of the operational code [17] is compatible with the logical primacy of anticodons [13].

It is known that the two complete tRNA-aaRS complexes look like mirror images of one another [2]. In this work, we calculate the entropy per site of each of the 20 tRNAs. We pose the following question: If the tRNAs are divided according to the two classes of aaRSs (class I and II), will their entropy profiles reflect this mirror symmetry?

1.1 The tRNA Code

Today we know that an amino acid is covalently bound at the 3′ end of a tRNA molecule and that a specific nucleotide triplet elsewhere in the tRNA interacts with a particular triplet codon in mRNA through hydrogen bonding of complementary bases. A striking feature of the genetic code is that an amino acid may be specified by more than one codon, so the code is described as degenerate. This does not suggest that the code is flawed: although an amino acid may have two or more codons, each codon specifies only one amino acid.

The degeneracy of the code is not uniform. Wobble allows some tRNAs to recognize more than one codon. Transfer RNAs base-pair with mRNA codons at a three-base sequence on the tRNA called the anticodon. The first base of the codon in mRNA (read in the 5′ → 3′ direction) pairs with the third base of the anticodon. tRNAs are grouped into families of isoacceptors, with each family recognized by a single cognate aminoacyl-tRNA synthetase.

All tRNAs conform to a secondary structure described as a “cloverleaf” and fold in three-dimensional space into an “L-shaped” molecule, in which the amino acid and the anticodon are at opposite ends of the molecule.

The 3′ end of all tRNAs have the sequence CCA, with the amino acid attached by the tRNA synthetase to the terminal adenosine residue. In eukaryotic cells, the 3′ terminal CCA is not encoded but is enzymatically added post-transcriptionally.

1.2 tRNA Genes

The tRNA gene sequences were obtained from the tRNA database (http://trnadb.bioinf.uni-leipzig.de), which correspond to 361 organisms distributed in the three domains of life. Due to the presence of non-canonical versions of the nucleosides in tRNAs sequences, the modified versions of these bases were considered as their

closest version that is canonical (adenosine, guanosine, cytidine, uridine). All sequences had the same length (76 characters), and the alignments were used for the calculations.

2 Entropy in tRNA Molecules

The equation for entropy per base of tRNA is: $H(X) = -\sum_{x\in\{\mathrm{A,G,U,C}\}} p(x)\log p(x)$, where the logarithm is taken to base 2 so that the outcome is measured in bits. This will yield a number in the interval $[0, 2]$ in each space of the sequence. Then the entropy per site just measures what the probability is of finding a given site in any of the 4 four possible states: A, U, G, or C.

In Fig. 1, the cloverleaf structure of two transcripts of tRNA is shown. The amount of entropy for each site ranges from the lowest (white) to the highest (black) values. Two gray values are for intermediate values.

The pairing in the stacks are Watson-Crick pairs (the ladder-like arrangement of bases that bind in pairs) of the secondary structure (see Fig. 1). Note that at these sites, the actual entropy diminishes, because two nucleotides that are bound together share their entropy. We included the anticodons because we are not calculating relative entropy. Note that the anticodons have high entropy.

The Watson-Crick binding in stacks is epistatic since their contributions to the fitness of the sequence are not independent, i.e., the probability to find a particular base at one position depends on the identity of a base at another position.

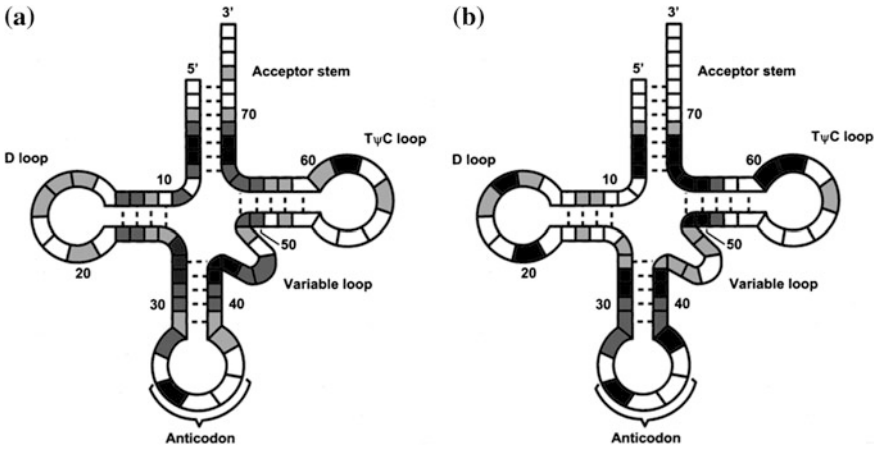


Fig. 1 Secondary cloverleaf structure of tRNA. In **a** $\mathrm{tRNA^{Gly}}$ and **b** $\mathrm{tRNA^{Ala}}$. The bases are colored *black* for high entropy (1.5–2 bits), *light gray* (0.5–1 bits) and *dark grey* (1–1.5) for intermediate values, and white (0–0.5 bits) for minimal entropy. There are 76 numbered sites

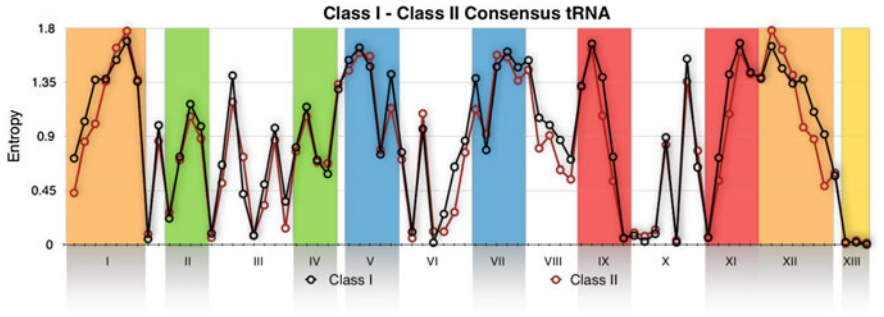


Fig. 2 Average entropy series for each class of tRNA according to the two classes of aaRS. The correlation coefficient between both series is $r^2 = 0.95$. (*I*) Acceptor stem; (*II*) D-stem; (*III*) D-loop; (*IV*) D-stem; (*V*) anticodon stem; (*VI*) anticodon loop; (*VII*) anticodon stem; (*VIII*) variable region; (*IX*) TΨC stem; (*X*) TΨC loop; (*XI*) TΨC stem; (*XII*) acceptor stem; (*XIII*) CCA

In Fig. 2, the average of entropy per site of the tRNAs is calculated according to their aaRSs class. Note that each profile is visually very similar. Indeed, they present a correlation coefficient of 0.95, so that on average their variabilities are statistically similar at each site. Assuming the null hypothesis that these two profiles are statistically similar, we cannot reject the hypothesis ($p < 0.0001$).

The colored bands correspond to regions in which Watson-Crick pairs are formed (regions I–XII; II–IV; V–VII; IX–XI). Note the mirror symmetry between the orange regions (acceptor stem), the green bands (before the D-loop), the blue region (before the anticodon loop), and the red bands (after the TΨC loop). The shape of the curve in each band is the mirror image of the same curve in the other band of the same color.

When we calculate other entropy profiles such as those amino acids encoded by RNY versus YNR triplets, the similarity of their entropy profiles is not as conspicuous as the one observed for the two classes of synthetases. Yet the correlation coefficient between the two series is $r^2 = 0.92$ (not shown).

3 Conclusions

It is widely accepted that the two aaRSs recognize the acceptor helix from the opposite side and that aaRS aminoacylates the tRNA without the need to recognize the anticodon. Yet, we have shown that there is mirror symmetry in the entropy profiles of the tRNAs when they are divided by their corresponding classes of aaRSs. This is a common property of the nondegenerate operational code. Two hairpins occur in tRNA, namely, the anticodon and acceptor arms, which participate in the translation mechanism in the ribosome. Despite differences in the D-loop and TΨC loops, we have found that the entropy profiles of both classes of aaRSs

display a remarkable mirror similarity: not only are their loops of different sizes, but they also show compositional differences. The RNY code, considered a primeval genetic code [6, 8, 9], does seem to reflect vestiges of these current entropy profiles. Information is not stored within a sequence, but rather in the correlations within the sequence. The analysis of information is an ongoing work of our group.

Acknowledgments MVJ was financially supported by PAPIIT-IN224015, UNAM, México.

References

1. De Duve C. The second genetic code. Nature. 1998;333:117–8.
2. Eriani G, Delarue M, Poch O, Gangloff J, Moras D. Partition of aminoacyl-tRNA synthetases into two classes based on mutually exclusive sets of conserved motifs. Nature. 1990;1990 (347):203–6.
3. Farias ST, Guimarães RC. Aminoacyl-tRNA synthetase classes and groups in prokaryotes. J Theor Biol. 2008;21:221–9.
4. Farias ST, Rêgo TG, José MV. Evolution of transfer RNA and the origin of the translation system. Front Genet. 2014;5:303. doi:10.3389/fgene.2014.00303.
5. Farias ST, Rêgo T, José MV. Origin and evolution of the Peptidyl transferase center from proto-tRNAs. FEBS Open Bio. 2014;2014(4):175–8.
6. Farías ST, Rêgo TG, José MV. A proposal of the proteome before the Last Universal Common Ancestor (LUCA). Int J Astrobiol. 2015;15:27–31. doi:10.1017/S1473550415000464.
7. Hou YM, Schimmel P. A simple structural feature is a major determinant of the identity of a transfer RNA. Nature. 1988;12:140–5.
8. José MV, Morgado ER, Guimarães RC, Zamudio GS, Farías ST, Bobadilla JR, Sosa D. Three-Dimensional algebraic models of the tRNA code and the 12 graphs for representing the amino acids. Life. 2014;4:341–73. doi:10.3390/life4030341.
9. José MV, Zamudio GS, Palacios-Pérez M, Bobadilla JR, Farías ST. Symmetrical and thermodynamic properties of phenotypic graphs of amino acids encoded by the primeval RNY code. Orig Life Evol Biosph. 2015;45:77–83. doi:10.1007/s11084-015-9427-4.
10. Moras D. Structural and functional relationships between aminoacyl-tRNA synthetases. Trends Biochem Sci. 1992;17:159–64.
11. Nagel GM, Doolittle RF. Phylogenetic analysis of the aminoacyl-tRNA synthetases. J Mol Evol. 1995;40:487–98.
12. Park SJ, Schimmel P. Evidence for interaction of an aminoacyl transfer RNA synthetase with a region important for the identity of its cognate transfer RNA. J Biol Chem. 1998;15:16527–30.
13. Rodin S, Rodin A, Ohno S. The presence of codon-anticodon pairs in the acceptor stem of tRNAs. Proc Natl Acad Sci USA. 1996;93:4537–42.
14. Rodin S, Ohno S. Four primordial modes of tRNA-synthetase recognition, determined by the (G, C) operational code. Proc Natl Acad Sci USA. 1997;94:5183–8.
15. Schimmel P. An operational RNA code for amino acids and variations in critical nucleotide sequences in evolution. J Mol Evol. 1995;40:531–6.
16. Schimmel P. In: Go M, Schimmel P, editors. Tracing biological evolution in protein and gene structures. Amsterdam: Elsevier; 1995. p. 1–10.
17. Schimmel P, Giegé R, Moras D, Yokoyama S. An operational RNA code for amino acids and possible relationship to genetic code. Proc Natl Acad Sci USA. 1993;90:8763–8.
18. Schwob E, Söll D. Selection of a 'minimal' glutaminyl-tRNA synthetase and the evolution of class I synthetases. EMBO J. 1993;15:5201–8.

19. Steitz TA, Moore PB. RNA, the first macromolecular catalyst: the ribosome is a ribozyme. Trends Biochem Sci. 2003;2003(28):411–8.
20. Woese CR. The emergence of genetic organization. In: Ponnamperuma C, editor. Exobiology. Amsterdam: North-Holland Publishing Co.; 1972. p. 301–41.
21. Yarus M. RNA–ligand chemistry: a testable source for the genetic code. RNA. 2000;6:475–84.

The Phylogenomic Roots of Translation

Derek Caetano-Anollés and Gustavo Caetano-Anollés

> In my version the history of life is counterpoint music, a two-part invention with two voices, the voice of the replicators attempting to impose their selfish purposes upon the whole network and the voice of homeostasis tending to maximize diversity of structure and flexibility of function. The tyranny of the replicators was always mitigated by the more ancient cooperative structure of homeostasis that was inherent in every organism. The rule of the genes was like the government of the old Hapsburg Empire: Despotismus gemildert durch Schlamperei, or 'despotism tempered by sloppiness'.
>
> —Freeman Dyson [1]

1 Introduction

The mechanisms behind translation and the specificities of the genetic code are well understood and are dependent on both nucleic acids and proteins [2]. In particular, transfer RNAs, or tRNAs for short, are central L-shaped nucleic acid molecules that are necessary for the transfer of genetic information from genomes and its interpretation during protein biosynthesis. They play fundamental roles during the entire translation process and during other processes of the cell as well. tRNAs recognize cognate aminoacyl-tRNA synthetase (aaRS) enzymes, which help them charge specific amino acids to the 3′ ends protruding from their acceptor stems. In turn, 'anticodon' sequences in their anticodon loops recognize complementary 'codon' sequences in messenger RNA (mRNA), translating genetic information that was

D. Caetano-Anollés (✉)
Department of Evolutionary Genetics, Max-Planck-Institut für Evolutionsbiologie, 24306 Plön, Germany
e-mail: caetano@evolbio.mpg.de

G. Caetano-Anollés (✉)
Evolutionary Bioinformatics Laboratory, Department of Crop Sciences, University of Illinois, Urbana, IL 61801, USA
e-mail: gca@illinois.edu

G. Hernández and R. Jagus (eds.), *Evolution of the Protein Synthesis Machinery and Its Regulation*, DOI 10.1007/978-3-319-39468-8_2

transcribed into RNA. The codon-anticodon recognition occurs within the confines of a complex ribonucleoprotein environment, the ribosome. tRNAs not only interact with mRNA but also with ribosomal RNA (rRNA) and proteins (r-proteins), as tRNAs are being ratcheted through the center of the biosynthetic complex and their amino acids unloaded during protein bond synthesis in the ribosomal peptidyl transferase center (PTC). The resulting polypeptides that are extruded through the ribosomal exit pore then fold according to hidden rules determined by interactions of tRNAs with all of its protein and nucleic acid partners. This 'structural code,' which holds deep historical information on the origin and evolution of proteins and life, differs from the 'genetic code.' It holds overarching specificities for the central molecular machinery that drives metabolism, translation, transcription and replication. Its vocabulary is currently unknown.

tRNAs are also very ancient molecules, a fact that is made evident by their universality and the many roles they play in translation and other biological processes [3]. For example, a recent study of the distribution of RNA molecules in a catalog of over a thousand RNA families revealed that tRNA molecules were part of only five families that were universally present in all biological organisms [4]. These families included rRNA and ribonuclease P (RNase P) RNA. The ubiquity and universality of the very central tRNA molecules have prompted their phylogenetic study using information in their sequences and structures [5–9]. Here we focus on the history of tRNA and its most fundamental interacting proteins and nucleic acid partners, aaRSs, elongation factors and ribosomal molecules, which are also part of a number of molecular complexes (e.g., ribosomes, multi-aaRS complexes). To unfold this history, we used phylogenomic information extracted from the sequence of millions of protein sequences and thousands of molecular structures to build a step-by-step timeline of accretion of their component parts, protein structural domains and RNA helical segments. We show that the gradual nucleation of these molecular modules, which behave as evolutionary units of proteins and nucleic acids, is ultimately responsible for the complexity of structures and molecular interactions unfolding in the biology of extant organisms.

2 A Structural Phylogenomic Method to Study the Evolution of Macromolecules

Phylogenetic analysis provides an objective criterion to study the natural history of biological entities of many kinds, beginning with the evolution of organisms, using information in specific features of those entities. The phylogenetic rationale of traveling back in time was made explicit by German entomologist Willi Hennig about half a century ago [10]. The systematization of evolutionary analyses gave rise to the fields of cladistics and systematic biology and provided background knowledge for the development of the field of molecular evolution and evolutionary genomics. It also resulted in the ongoing construction of a Tree of Life

(ToL) describing the evolution of organismal diversity at the planetary scale. Remarkably, no comparable community-driven effort is being pursued that would produce global views of the evolution of molecules of the kind advocated by Emile Zuckerkandl and Margaret Dayhoff in the early 1970s [11, 12]. Despite this shortcoming, the fields of structural biology and genomics have advanced considerably during past decades to provide wide-encompassing understanding of molecular diversity at atomic resolution [13]. As of 4 January 2016, there are 114,697 models of molecular structure deposited in the entries of the PROTEIN DATA BANK (PDB) [14], and their associated functions are encoded in the DNA of the 8,434 genomes and metagenomes that have been completely sequenced (GOLD DATABASE [15]). Genomic information has given rise to 0.55 million UNIPROTKB/SWISSPROT and ~50.4 million UNIPROTKB/TREMBL protein sequence entries and information in thousands of functional RNA molecules.

Phylogenetic analysis builds tree representations of genealogical relationships of the entities that are being studied, the *phylogenies*, by mining information in a number of biological features of interest, the phylogenetic *characters* [16]. Traditional characters that are useful include biochemical, morphological, physiological, developmental and molecular features with historical signal. The vast majority of molecular features that have been studied so far involve sequence information in alignments, i.e., sets of characters describing positions along a string of monomers that are homologous within groups of macromolecules. However, function impacts fitness and constrains evolution. Since molecular structures are the repositories of molecular functions, they are generally more resistant to change than sequences. They are therefore highly conserved at the evolutionary level and ideal candidates to study the history of life, from the very deep relationships to the most recent. For that reason, we have been studying the evolution of protein and nucleic acid structures for almost 2 decades using the wealth of information generated by the genomic revolution (first reviewed in [17]). We start by first summarizing the experimental strategies used to study molecular history (Fig. 1) and then describing some useful applications.

(1) *Evolution of proteins.* Advanced hidden Markov models (HMMs) of structural recognition assign fold structures to protein sequences with high accuracy and low error rates. These bioinformatic annotations permit the generation of a structural census of proteins, with structural domains defined at various levels of protein structural abstraction in the hierarchical classifications of SCOP [18] and CATH [19], the gold standards. We have computed the proteomic occurrence and abundance of each domain structure across a wide transect of organisms and used this proteomic census to construct data matrices (arrays) for phylogenetic analysis. Phylogenetic trees of domains (ToDs) and trees of proteomes (ToPs) were built from this census. The first study of this kind was published in 2003 and involved a proteomic analysis of only 32 organisms [20]. Recent analyses extended the approach to thousands of them and to viruses [21]. Since ToDs and ToPs can be rooted using direct methods of character polarization, the rooted trees describe the origin and evolution of parts and wholes, the structural domains (the evolutionary units of proteins) and the proteomes (the entire protein repertoire of an organism),

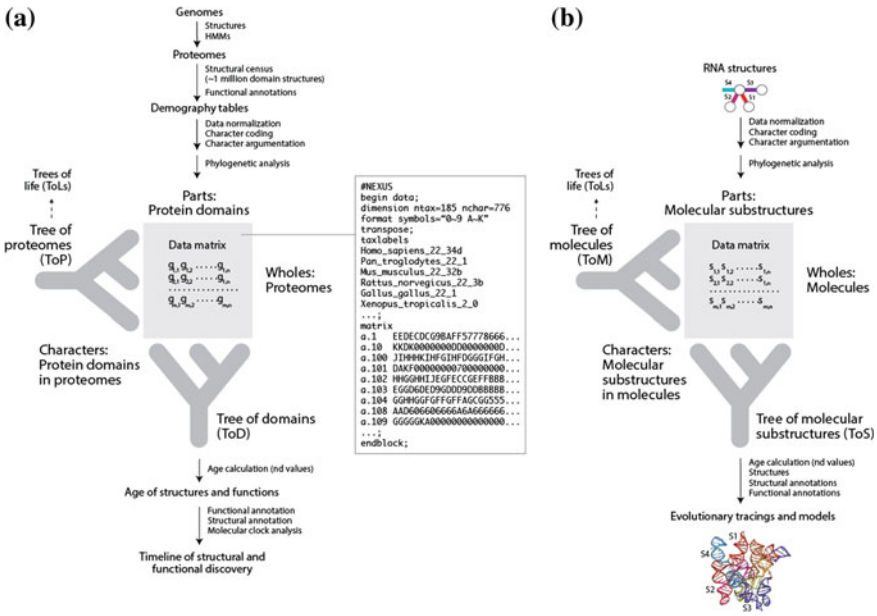
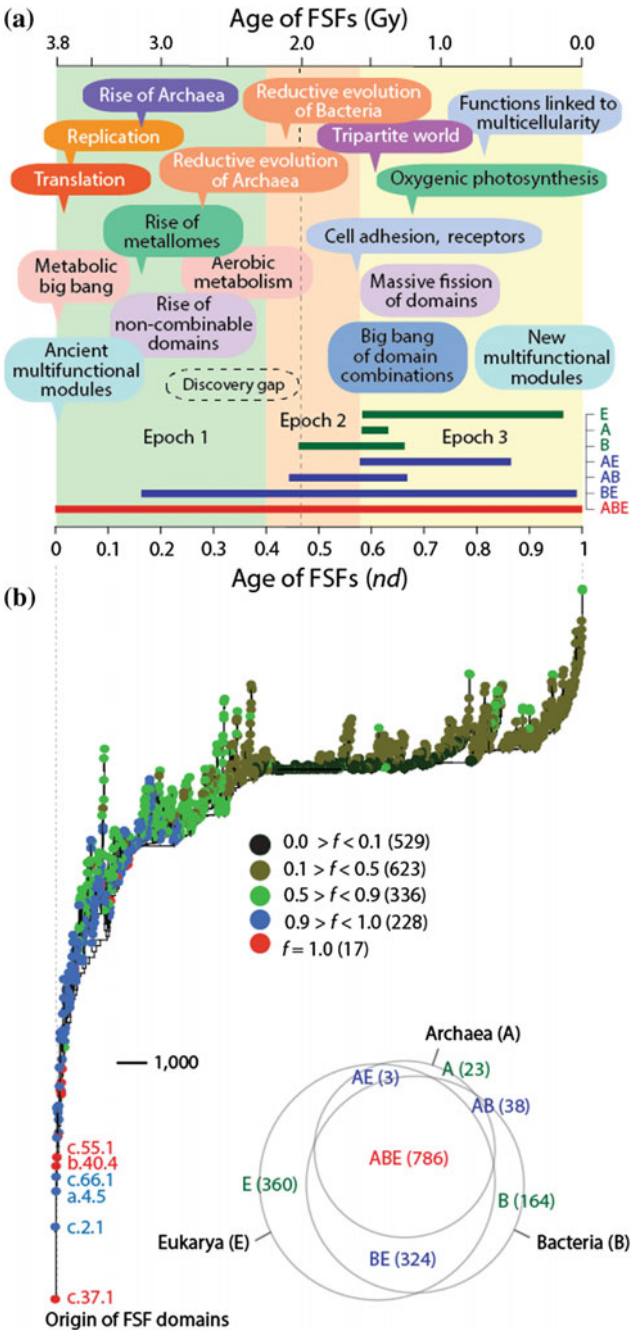


Fig. 1 Structural phylogenomic analyses of protein domains and RNA molecules. **a** The flow diagram describes the steps leading to the reconstruction of trees of domains (ToDs) and trees of proteomes (ToPs) and associated timelines of domain innovation. A census of domain structures in proteomes of thousands of completely sequenced organisms is used to compose data matrices (arrays with rows and columns corresponding to taxa/characters) for building phylogenomic trees. The trees describe the evolution of individual structural domains and proteomes, respectively. Elements of the array (g) represent genomic abundances of domains in proteomes, defined at different levels of classification of domain structure. The *inset* shows a very small segment of a NEXUS file holding readable information for tree computation. **b** The flow diagram to the right describes the phylogenetic reconstruction of trees of molecules (ToMs) and trees of substructures (ToSs) of RNAs. Molecular structures are first decomposed into substructures, including helical stem tracts and unpaired regions. Structural features of these substructures (e.g., length) are coded as phylogenetic characters and assigned character states according to an evolutionary model that polarizes character transformation toward an increase in conformational order (character argumentation). Coded characters (s) are arranged in data matrices. Phylogenetic analysis generates rooted phylogenetic trees. Embedded in ToDs and ToSs are timelines that assign age to molecular structures. These ages can be 'painted' onto 2D or 3D structural models of RNA or proteins, RNP complexes or protein complexes generating evolutionary heat maps

respectively. Furthermore, the fact that ToDs are comb-like enabled the construction of timelines of domain history by counting the number of nodes from the base of the tree to each taxon and expressing the value as a relative 'node distance' (*nd*). These *nd* values measure the relative age of each domain, which can be linked to the geological record through time calibration points. In fact, a remarkably linear relationship was observed between the *nd* and the age of biomarkers and geomarkers diagnostic of domain structures [22]. This relationship defined a molecular clock of domain structures, which effectively turned ToDs into 'timetrees' [23],

i.e., bona fide chronologies with time axes in billions of years (Gy). A ToD built from SCOP domain structures and its associated timeline is illustrated in Fig. 2. Note that the time of first appearance of a domain structure in a chronology records the time of origin of that structure and that the gradual evolutionary appearance of domains involves thousands of steps. Each step represents a domain structure with numerous and important annotations, including domain distribution in organisms and viruses, molecular functions, biochemical and biophysical properties, and biological network participation. Moreover, since the new discovery of domains has reached a plateau, the number of domains must be considered finite, and ToDs currently approach the highest level of universality that is possible in phylogenetic statements. The generation of timelines of structural innovation has already allowed exploration of a number of important questions. ToDs have been used to trace the origin and evolution of metabolic networks [27–29], study the rise of translation and the genetic code [30, 31], uncover the co-evolutionary history of the ribosome [32, 33], explore the evolution of metallomes and biological metal utilization [34], unfold the natural history of biocatalytic mechanisms [35] and protein flexibility [36], study the evolutionary dynamics of gain and loss of domains [37] and domain combinations [24], determine the makeup of the universal common ancestor of life [38], visualize a basal stem line of descent responsible for organismal biodiversity [25, 39] and generate a truly universal ToL that includes cellular organisms and viruses directly from the age of domains using multidimensional scaling approaches [21]. A recent encyclopedia entry summarizes some of the findings [40].

(2) *Evolution of nucleic acids.* Since RNA molecules carry deep phylogenetic signal and the arrow of time in their structures, we have been able to derive historical accounts of molecular evolution directly from structural topology and thermodynamics [8, 41–44]. The evolutionary signal that we mine exists because the secondary structure is closely linked to structural conformation and dynamics [45]. RNA folding is negatively correlated with chain length, and the frustrated energetics and dynamics of folding allows only few conformations to reach stable states [46]. This forces structures to collapse by quickly reaching local folding solutions, which result in the formation of a number of helical structural modules compatible with the length and history of the molecules. Since the folding process is frustrated, numerous folding conformations are possible in molecules with randomized sequences. However, the number of possible conformations is actually culled by evolution to ensure that their average life is sufficiently long for the molecules to hold durable molecular functions [47]. This link between molecular evolution and the biophysics of RNA provides a rationale for our phylogenetic methodology: (1) characters describe features of helical stem and non-paired segments of RNA, and (2) minimization of conformations in RNA provides a definition of ‘evolution’s arrow’ for rooting of trees. Our methods make use of a census of geometrical features that measures the length and topology of RNA substructures, including stem and non-paired segments, or statistical features portraying stability and conformational diversity. The census produces data matrices with rows and columns representing molecules and molecular parts and phylogenetic characters describing molecular length or statistical features of structure


(a)
Age of FSFs (Gy)
3.8
3.0
2.0
1.0
0.0
Rise of Archaea
Reductive evolution of Bacteria
Functions linked to multicellularity
Replication
Tripartite world
Translation
Reductive evolution of Archaea
Oxygenic photosynthesis
Rise of metallomes
Cell adhesion, receptors
Metabolic big bang
Aerobic metabolism
Massive fission of domains
Rise of non-combinable domains
Big bang of domain combinations
New multifunctional modules
Ancient multifunctional modules
Discovery gap
Epoch 1
Epoch 2
Epoch 3
E
A
B
AE
AB
BE
ABE
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
(b)
Age of FSFs (nd)
0.0 > f < 0.1 (529)
0.1 > f < 0.5 (623)
0.5 > f < 0.9 (336)
0.9 > f < 1.0 (228)
f = 1.0 (17)
1,000
c.55.1
b.40.4
c.66.1
a.4.5
c.2.1
c.37.1
Origin of FSF domains
Archaea (A)
A (23)
AE (3)
AB (38)
ABE (786)
E (360)
B (164)
Eukarya (E)
BE (324)
Bacteria (B)

◀ **Fig. 2** The evolution of the protein world is visualized by studying its structural domain components. **a** Timeline of evolutionary appearance of fold superfamilies (FSFs) of structural domains describing the relative timing of important events in the history of life. Domain age was measured as a relative distance in the number of nodes from the base of the tree (*nd*) or was placed in a geological time scale of billions of years (Gy) using a molecular clock of domain structures [22]. Information in speech balloons without pointers was taken from trees of domain and domain combinations [24]. Their relative location is approximate. The three evolutionary epochs of the protein world are shaded in light green (Epoch 1, architectural diversification), salmon (Epoch 2, superkingdom specification) and light yellow (Epoch 3, organismal diversification) [25]. Boxplots display the FSF age distribution for the seven possible taxonomic groups. **b** Phylogenomic tree of domains (ToDs) describing the evolution of 1,733 FSFs reconstructed from structural domain abundance in the proteomes of 981 organisms. The tree was used to build the timeline of panel **a**. FSF taxa are colored according to FSF distribution (*f*) in the proteomes that were surveyed and used as characters to build the phylogenomic tree [26]. The most basal FSFs are labeled with SCOP concise classification strings (*ccs*; e.g., c.37.1 is the P-loop containing nucleoside triphosphate hydrolase FSF). The Venn diagram shows FSF distribution in superkingdoms

(e.g., branching, stability, diversity). Since the matrices can be transposed, the data can be used to build phylogenetic trees of molecules (ToMs) and trees of substructures (ToSs). ToMs and ToSs are data-driven models of the history of the molecular system or its component parts, respectively. The comb-like topologies of ToDs allow building timelines of the appearance of parts in molecules. These timelines define a 'natural history' of nucleic acids. The origin and evolution of the most ancient RNA molecules have been studied in this way, including tRNA [8, 48, 49], SINE elements [44], the large and small rRNA subunits [17, 32, 33, 42, 43], 5S rRNA [50] and RNase P RNA [51].

Note that the most parsimonious trees that describe the evolution of proteins and nucleic acids are retained after computational searches of tree space using the Wagner algorithm. Optimal trees are unrooted. They are only rooted a posteriori using phylogenetic process models that comply with Weston's generality criterion [52]. This criterion states that as long as ancestral characters are preponderantly retained in descendants, ancestral character states will always be more general than its derivatives given their nested hierarchical distribution in the rooted trees. Tracing the distribution of structural domains in proteomes (the *f* summary statistic) on the taxa of a ToD reveals compliance with the workings of Weston's rule (Fig. 2). When rooting a ToL, character change in domain abundance should be sequentially nested, with the most ancient structures being abundantly present in all or almost all of organismal lineages and more recent structures present at more moderate levels in increasingly more restricted groups of lineages. The ToD reflects that pattern; the most ancient domain structures (taxa) at the base of the tree are the most widely distributed in proteomes. A tracing of character state changes in the corresponding ToP (which is a ToL) shows that indeed these taxa (now characters) exhibit the widest distribution with change preponderantly restricted to the base of the tree. Weston's patterns also unfold by studying the distribution of domains across superkingdoms of life (Fig. 2). The Venn group of domains that are shared by all life (ABE) is the most ancient taxonomic group. Their domains span the entire time axis and are the most widely distributed in genomes. The evolutionary

appearance of the BE group shared by Bacteria and Eukarya occurs much later, coinciding with the first reductive loss of an FSF in Archaea. Domain structures specific to superkingdoms appear halfway in the timeline. These patterns also comply with the expected nesting of lineages.

Operationally, the direct character polarization method roots the trees of proteins by assuming domain structures accumulate in the evolution of the protein world and roots the trees of nucleic acids by assuming conformational stability increases in evolution as structures become canalized (reviewed in [17, 53, 54]). Biologically, domain structures spread by recruitment in evolution when genes duplicate and diversify, genomes rearrange, and genetic information is exchanged. Similarly, nucleic acid base pairs increase the stability and expand the size of RNA structures to match the increasing interactions with the expanding proteins and protein complexes that are responsible for cellular and functional makeup. This is a process of accumulation and retention of iterative homologies, such as serial homologs and paralogous genes, which is global, universal and largely unaffected by proteome or molecular size. The operational rooting (when made most parsimonious) complies with Weston's rule, and the axiomatic validity of character transformation can and has been tested using a number of approaches, including thermodynamics, phylogenetics and multidimensional scaling, proving its mettle.

3 The Early Emergence of Proteins and Metabolism

The structural domains are considered the evolutionary units of proteins. However, lower levels of structural granularity (abstraction) such as secondary structures (e.g., helix, strand, turns) or supersecondary structures (e.g., αα-hairpins, ββ-hairpins, βαβ-elements) could also hold evolutionary history. Remarkably, phylogenetic analyses, numerical approaches or machine learning techniques give no indication that these other levels hold strong phylogenetic signal or represent evolutionary modules (but see [55]). This may simply stem from our inability to suitably identify structural or non-structural lower level motifs that are responsible for molecular change. In contrast, domains have been carefully analyzed, unified into homologous groups and organized into a hierarchy in several classifications, including SCOP and CATH. For example, the SCOP classification groups domains into fold families (FFs), fold superfamilies (FSFs), folds and protein classes in a hierarchical classification system of decreasing granularity. Domains with pairwise amino acid sequence identities of more than 30 % are unified into FFs, and those FFs that share similar structures and functions are further unified into FSFs. FFs and FSFs have common evolutionary origins. FSFs sharing similar arrangements of secondary structures in three-dimensional space are further unified into folds, and those that share similar overall designs are further grouped into protein classes. The common evolutionary origin of FSFs in folds has not been systematically tested. In turn, classes unify large groups of folds that do not have a common evolutionary history.

Given these considerations, ToDs built at different hierarchical levels of protein classification should be considered phylogenetic statements solely related to structural domain history at those particular levels. Other possible structural modules at lower or higher levels of the hierarchical molecular system require separate exploration. The information gathered from ToDs has been however revealing since their inception [20]. The global emergent picture of molecular evolution derived from domain history is largely congruent regardless of the level of abstraction or the classification system. The global historical patterns obtained by tracing molecular functions annotated to domain structures in the timeline summarized almost a decade ago [17] still hold in updated timetrees and new studies. Here we highlight some of these patterns (summarized in Fig. 2):

1. The oldest domains are fully dependent on cellular membranes. It is therefore likely that the first proteins emerged enclosed in membrane containers forming primordial cells and evolved from there to form the wide diversity of globular proteins that today contribute to the complex make up of cellular organisms.
2. The very early proteins are first associated with organic cofactors but only later involve transition metals as ligands. This suggests an organismal response to increasing energy demands of the ancient world.
3. The very early, massive and then protracted appearance of domains with enzymatic functions indicates that the central metabolism played a primordial role in the early evolution of life.
4. The early but relatively late discovery of proteins involved in translation, including aaRSs, elongation factors and r-proteins, has a metabolic origin and is interrupted by a "discovery gap" that probably involves a historical revision of the translation apparatus.
5. The relatively early rise of metallomes (the Zn-metallome appearing first) and the late rise of oxygenic photosynthesis coincide with the late rise of aerobic metabolism. This explains the existence of the Great Oxygenation Event (GOE) $\sim$2.5 Gy ago, which is strongly supported by the geological record.
6. Domains involved in the synthesis of DNA precursors and replication complexes appear late. This indicates a late transition from storage of information in RNA genomes to storage in DNA genomes of cellular organisms.
7. Domains with functions that are typical of Eukarya, including cell adhesion, receptors, chromatin structure and functions linked to multicellularity, appear late and gradually and involve multidomain proteins. This suggests that modern Eukarya established as an organismal supergroup quite late in evolution.

Furthermore, a careful study of the origin and evolution of domains and domain combinations in multidomain proteins indicates the existence of a 'big bang' of protein discovery coinciding with the rise of eukaryotic organisms [24]. The conclusions of this study still hold and explain biphasic evolutionary patterns that exist in proteins [56]. The trees showed that the first proteins had single domains and were multifunctional, all of which produced fusion-driven combinations. These domain combinations arose early in the timeline (during Epoch 1), were

functionally specialized and later dominated the protein world. In contrast, fission processes developed late, notably during the big bang of domain combinations. These fissions produced many derived multifunctional single-domain proteins in Eukarya. The cyclic pattern of distribution of biological function along the architectural timeline is remarkable and reveals the emergence of a new class of protein module in evolution [17].

A major corollary from our phylogenomic studies is that the process of accretion of domains in proteomes occurs pervasively in nature and is a driving force for the evolution of macromolecules and life. Accretion is gradual, follows a molecular clock, and reconciles biology and planetary history. This finding crucially supports the *principle of spatiotemporal continuity*, the fundamental axiomatic necessity of evolution. We note that one could argue that the mere reconstruction of phylogenetic trees implies per se the gradual appearance of biological entities in evolution, i.e., that well-resolved tree topology cannot test spatiotemporal continuity. This is not so. The existence of a comb-like tree is an outcome of the existence of phylogenetic signal in the data and the existence of a timeline of natural structural discovery. Absence of such historical information would collapse branches into 'hard' polytomies, i.e., nodes supporting more than three branches with splits that arise from natural phenomena. These polytomies would distort the unbalanced tree structures toward a 'star' tree topology, making the construction of timelines impossible. The fact that we detect strong phylogenetic signal in the data diffuses such concerns. Furthermore, the molecular clock of folds extends the timeframe of domain diversification to the vast majority of the geological record. This supports the gradual spread of domain innovation in evolution. The recent mathematical modeling of the accretion process now makes the entire evolutionary process of protein domain accumulation explicit and prompts an exploration of how protein diversity extends through sequence space [57].

4 Insights into the Generation of the First Protein Structures

In a relatively recent study, we mapped the first evolutionary appearance of the oldest 54 FFs, tracing a number of properties of these domain structures, including their ability to bind cofactors, interact with RNA, and display broad molecular movements and flexibility [58]. These primordial FFs are important since they are responsible for jumpstarting metabolism and translation. Remarkably, their order of appearance provided detailed information about which central biological processes of the cell came first, metabolism, translation or replication, and what sub-processes were involved. The first four FFs were the ABC transporter ATPase domain-like family (c.37.1.12), the extended and tandem AAA-ATPase domain families (c.37.1.20 and c.37.1.19) and the tyrosine-dependent oxidoreductase domain family (c.2.1.2). All of these FFs currently unfold in membrane-structured cellular

environments. A detailed tracing of these structures in metabolic sub-networks defined by the KEGG database showed that these FFs provide hydrolase and transferase functions needed for nucleotide interconversion, storage and phosphate transfer-mediated recycling of chemical energy [29]. They are ultimately responsible for seeding the pathways of purine biosynthesis and establishing the chemical currency of energy storage in the biological world, the ATP and then GTP families of cofactors. Note that three of the four FFs hold the P-loop containing the nucleotide triphosphate (NTP) hydrolase fold (c.37), which is placed at the very base of each and every one of the ToDs we have ever generated. In the timeline, it appeared for the first time associated with a primordial bundle, the predominant structure of proteins associated with membranes. The archaic association of the "Rossmann-like" α/β/α-layered design that is typical of the c.37 and c.2 folds and the bundle structure was even made explicit in ToDs generated using CATH domain definitions, which split the SCOP FFs structure into finer grained modules [59]. Thus, the structural phylogenomic statements derived from structures appearing at the base of ToDs establish that the origin of proteins was unequivocally associated with metabolism and membranes. Thus, Dyson's "*more ancient cooperative structure of homeostasis*" typical of protein enzymes of metabolism indeed preceded the "*tyranny of the replicators*" underlying a nucleic acid-based genetic system [1], debunking the widely held belief of an ancient RNA world. The consequence of this finding is that first proteins had to unfold in the absence of genetic memory within cellular compartments.

An early appearance of peptide and protein molecules in cellular compartments is not an alien concept. Prebiotic chemistry supports the facile production of amino acids (even in artificial spark discharge experiments) and short peptide molecules (even in simple cycle desiccation experiments), which is much simpler than the synthesis of nucleic acid precursors. Amphiphilic molecules capable of forming vesicle containers are even present in meteorites. These emerging molecular systems are prone to hold molecular and cellular memory. Cellular compartments that are stabilized by addition of peptides could be more persistent [58]. Similarly, biases in self-catalyzed ligations of short peptides could result in longer and more stable emergent structures [60, 61]. These are hallmarks of 'homeostasis,' 'competitive optimization' and 'compositional selection.' Such forces could impart archaic memories about the expanding cellular and molecular systems.

If these conjectures are true, then we must invoke an ancestral 'origami' responsible for the generation of the first stable structural domains, which assembled from ancient peptides [62, 63]. Would this origami point toward the primordial α/β/α-layered structure present in the c.37 and c.2 Rossmann-like folds? We already have an answer! The use of advanced bioinformatics methods to survey and classify modular-like arrangements of helix, strand and turn segments $\sim$25–30 amino acid residues long identified the most conserved loop-forming building blocks [64]. Remarkably, the most popular of these structural motif prototypes (known as 'elementary functional loops') in archaeal proteomes and the most widely spread in fold superfamily domain structures preferentially involved superfamilies holding the c.37 and c.2 folds. A tracing of the bipartite network of elementary functional

loops and domain superfamilies in time showed patterns of emergence of modular scale-free behavior [65]. The ancient link between peptides and structural domains is therefore established and must be further studied.

5 Late Evolutionary Appearance of First Structural Domains Interacting with RNA

The structural domains that consistently appear at the base of the ToDs do not interact with nucleic acid macromolecules. Instead, the first nucleic acid-interacting domains made their debut relatively late, had metabolic origins and associated with tRNA [30, 58]. When studying the timeline of FFs, a number of FF domain structures appear after the rise of metabolism ~3.7–3.6 Gy ago (nd_{FF} = 0.02–0.045). The first four FFs of this group involve the class I aaRS catalytic domain (c.26.1.1), class II aaRS and biotin synthetases (d.104.1.1), G proteins (c.37.1.8) and actin-like ATPase domain (c.55.1.1) FFs [58]. These structural domains, which are also part of the catalytic makeup of enzymes important for fatty acid biosynthesis, appear before r-proteins in the timeline. All of them have the α/β/α-layered Rossmann-like design, and three of them define the catalytic domains of aaRSs and structures of elongation factors that are central for translation and the specificity of the genetic code. They catalyze crucial acylation and condensation reactions involved in the aminoacylation of tRNA bound to the aaRSs or phosphopantetheinyl arms of carrier proteins that are part of non-ribosomal peptide synthetase (NRPS) complexes.

6 The Co-evolutionary History of Emerging tRNA, Ribosomes and Proteins

Having established that translation started late by laying down a foundation of interactions among tRNA, aaRSs and factors, can we explore patterns of molecular growth indicative of the processes behind the rise of translation and the specificities of the genetic code? Phylogenomic analysis of thousands of RNA molecules and millions of protein structural domains supports three crucial historical patterns: (1) the co-evolution of tRNA and aaRS enzymes during the rise of genetic code specificities, (2) the co-evolution of ribosomal RNA and proteins, and (3) the co-evolution of tRNA and the emerging ribonucleoprotein structure of the ribosomes. We here define co-evolution as a coordinated succession of structural changes occurring within the emerging molecular environment. These changes should be considered mutually induced by the increasing interactions between and among protein and nucleic acid molecules that were being recruited to perform the very initial molecular functions. In all cases, co-evolution's goal was to fold

macromolecules into more stable and functionally efficient structures capable of extending the persistence of the molecules and the emergent primordial cells that would contain them. In these phylogenomic studies, the relative ages of structures of tRNA, rRNA, aaRS domains and r-protein domains were calculated from the phylogenetic trees (ToSs, ToMs and ToDs), indexed with structural, functional and molecular contact information and mapped onto three-dimensional models of molecules and molecular complexes.

The rise of the genetic code. The specificity of translation and the 'memory' of genetics is ultimately controlled by the specificities that define the genetic code. In vitro studies have shown that discrimination against non-cognate substrates is maximal in aminoacyl-tRNA synthesis, unknown but probably significant for EF binding and minimal for aaRS editing, aaRS resampling and ribosomal tRNA recognition and proofreading [2]. The rate of misincorporated amino acids in aaRSs is 1 in 200–10,000, at least an order of magnitude lower than other specificities, and the rate of misincorporated tRNA is 1 in more than 10,000. It is therefore clear that genetic code safekeeping has been entrusted to aaRSs and not the ribosomes. Reconstruction of phylogenies and evolutionary timelines showed that the history of catalytic, editing and anticodon-binding domains of aaRSs matched the history of tRNA charging and encoding [31]. The catalytic domains, which are the most ancient of the aaRSs molecules [30], interact with the acceptor arm of the tRNA that charges specific amino acids, which is the most ancient of the nucleic acid molecule [8] (Fig. 3). Similarly, the more recent anticodon-binding domain of aaRSs interacts with the more recent anticodon-binding arm of tRNAs. These co-evolutionary patterns that are derived from ToDs and ToSs can be complemented with more powerful tools that couple ToMs and phylogenetic constraint analysis to fine-grain the evolutionary history of the charging and encoding functions of translation [48, 49]. This allowed making historical inferences of the progression of specificities for both the 'operational' genetic code of the acceptor arm of tRNA [66] and the more derived 'standard' genetic code of the anticodon-binding stem of tRNA. The rise of the aminoacylation specificities of tRNA isoacceptors is described in the timelines of Fig. 4. The first specificities unfold by pre-transfer and post-transfer editing and trans-editing activities of aaRSs. These molecular activities are responsible for sieving amino acids by size in the active sites of the catalytic domains. They involve 11 of the 20 standard amino acids. Specificities are however split into two groups. Group 1 specificities associate with the older 'type II' tRNA structures holding a variable arm. Group 2 specificities associate with the standard 'type I' tRNA cloverleaf structures that lack the variable segment of the structure. These interactions, which unfolded $\sim$3.7–3.0 Gy ago, involve the acceptor stem of the tRNA molecule and probably defined the 'operational' genetic code in the absence of a fully functional ribosome and a full cloverleaf structure. In turn, codon specificities unfolded $\sim$3 Gy ago with the first anticodon (AC) binding domains, which interact with the more modern anticodon stem of tRNA. The evolution of this more modern 'standard' genetic code produced its own timeline of codon specificities that sometimes overlapped and enhanced the specificities of the 'operational' code (Fig. 4). Separate timelines of amino acid charging and codon

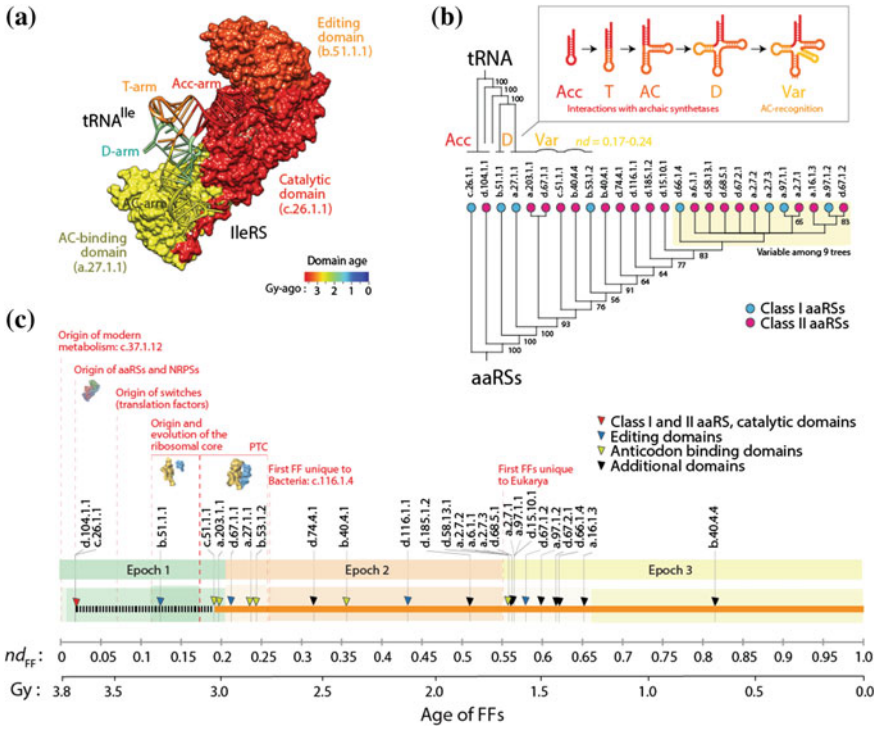


Fig. 3 The co-evolutionary history of tRNA and aaRSs. **a** The age of aaRS domains, exemplified by IleRS (PDB entry 1qu2), matches the age of the interacting arms of their tRNA isoacceptors. The oldest acceptor (*Acc*) arm interacts with the oldest catalytic domain and the more recent anticodon (*AC*) arm interacts with the more recent AC-binding domain. **b** One of nine most parsimonious phylogenomic tree reconstructions describing the history of aaRS domains [31]. Terminal leaves are colored according to aaRS class and indexed with domain *ccs* labels. The tree matches the corresponding subtree in the global tree of FFs described in the next panel. A tree of tRNA substructures describing the evolutionary growth of tRNA (made explicit in the *inset*; [8]) is mapped to the domains that interact with the unfolding tRNA substructures, showing tight co-evolution. **c** Timeline of FF domains directly obtained from a ToD reconstructed from information in the proteomes of 420 free-living organisms [31]. FFs (*indexed arrowheads*) are mapped along a timeline with landmarks derived from the domain history. The *dashed black segment* of the timeline indicates the aaRS history prior to the appearance of AC-binding domains and modern genetics. The three epochs of the protein world (described in Fig. 2) are *shaded*

recognition are therefore evident in the phylogenomic chronologies. Remarkably, a recent study shows that the acceptor and anticodon stem determinants encode the size and polarity of amino acid residues, respectively [67]. This matches the differential encoding of information in the top and bottom half of the tRNA molecule and the role of editing and anticodon binding recognition that differentiate these two sequential and apparently redundant codes [31]. This congruence supports the separate development of two genetic codes in evolution. A comparison of amino acid and dipeptide compositions of single-domain proteins appearing in the

Fig. 4 The history of the operational and standard genetic codes unfolds sequentially but the codes act redundantly. The operational code delimits amino acid charging, and the standard code delimits codon specificity. Phylogenomic analysis dissects their history [31, 49]

timelines before and after the first anticodon binding domains (i.e., the standard code) revealed enrichment of dipeptides with amino acids that are subject to aaRS editing (groups 1 and 2) [31]. Results uncover a hidden link between the emergence and expansion of the classic genetic code and protein flexibility [31].

The rise of the ribosome. Domain history indicates that r-proteins appeared 3.3–3.4 Gy ago, later than aaRSs and factors but earlier than anticodon binding specificities. The ribosome was therefore present while the 'operational code' was being developed. Since the small (SSU) and large (LSU) subunits of the ribosome contain 30–40 and 30–45 proteins, respectively, r-protein history unfolds considerable detail about the origin and evolution of the ribosome. Similarly, SSU and LSU hold about 50 and 100 universal helical segments, respectively, which can also provide details about the evolutionary growth of the RNA molecules. Indeed, ToDs and ToSs enabled construction of detailed timelines of the history of r-proteins and nucleic acids, respectively [17, 32, 33, 43, 50]. More importantly, the structural interactions present in models of the atomic structure of the ribosome permitted mapping interactions in both timelines, effectively linking the two. Remarkably, the exercise showed strong co-evolutionary relationships between the age of r-proteins and the age of interacting rRNA helices in the universal ribosomal core [32, 50], which were expressed as a significant correlation (Fig. 5). The oldest proteins (S12, S17, S9, L3) appeared together with the oldest rRNA substructures responsible for decoding and ribosomal dynamics. These structures include the ratchets and two hinges of SSU rRNA and the L1 and L7/L12 stalks important for ribosomal movement of tRNA in the complex. As the ribosome continued to unfold in evolution, the age of rRNA helical regions in both subunits (see Fig. 5) and interacting domains of r-proteins co-evolved simultaneously to form a fully functional ribosomal core. Importantly, the appearance of RNA substructures at first occurred in orderly fashion until the formation of five-way LSU and ten-way SSU junctions in SSU and LSU, respectively, at which point a 'major transition' in ribosomal evolution occurred 2.8–3.1 Gy ago (Fig. 6). This transition, which coincided with the start of planet oxygenation [28], brought ribosomal subunits together through inter-subunit bridge contacts [32]. It also stabilized loosely evolving ribosomal

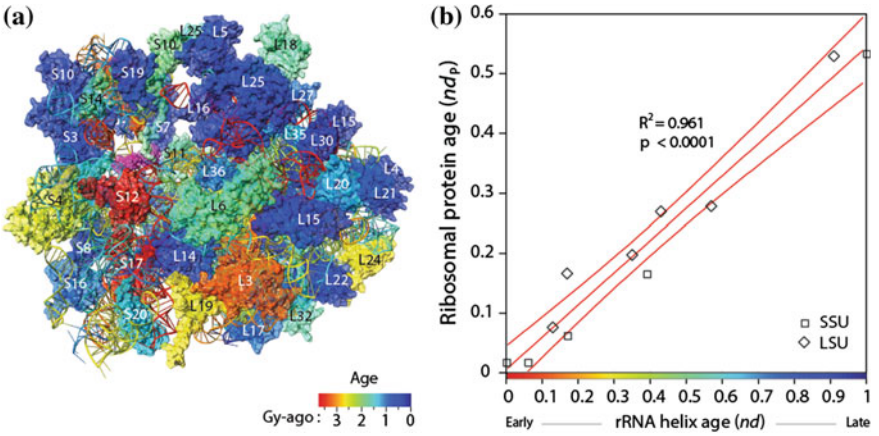


Fig. 5 The molecular evolution of the ribosome. **a** Secondary structure models of the small (*SSU*) and large (*LSU*) subunits of the *Escherichia coli* ribosome with rRNA helical segments and proteins colored according to their age (in Gy). Note the very ancient and central translocation core of helix 44 and r-proteins S12 and S17 (colored red) develops into a complex patchwork of molecular ages. **b** Tight co-evolution between r-proteins and rRNA helical segments. The relative ages are expressed as node distances (*nd*) derived from ToDs and ToSs. Figure modified from [32]

components and developed tRNA-interacting structures and a fully-fledged PTC with exit pores capable of protein biosynthesis. The implications of these co-evolutionary patterns of ribosomal history are profound. They debunk the idea of an origin of the ribosome in an ancient 'RNA world' since the growth of RNA and protein structure occurred in close interaction.

tRNA is at the center of ribosomal evolution. The timelines of ribosomal history showed that tRNAs were the centerpiece of important structures that were being accreted [32]. The gradual development of tRNA-rRNA molecular interfaces revealed that known interactions occurring before the major transition involved contacts between ancient SSU helices and the anticodon arm of tRNA. After the transition, most contacts involved newer LSU helices and the older half of tRNA. Contacts with the T-arm of tRNA formed soon after the transition. The T-arm is the only tRNA substructure that interacts with the two major subunits of the ribosome. Importantly, all tRNA contacts with the PTC unfolded abruptly during the major transition. Coupling the evolutionary timelines of tRNA and rRNA structure with annotations of their interactions with protein domains revealed that the tRNA cloverleaf structure was already fully formed when the PTC made its appearance [68]. Thus, fully formed tRNA molecules played other roles before being recruited for protein biosynthesis, perhaps both as cofactors of peptide-producing dipeptidases and ligases [31, 58] and as primordial genomes [69].

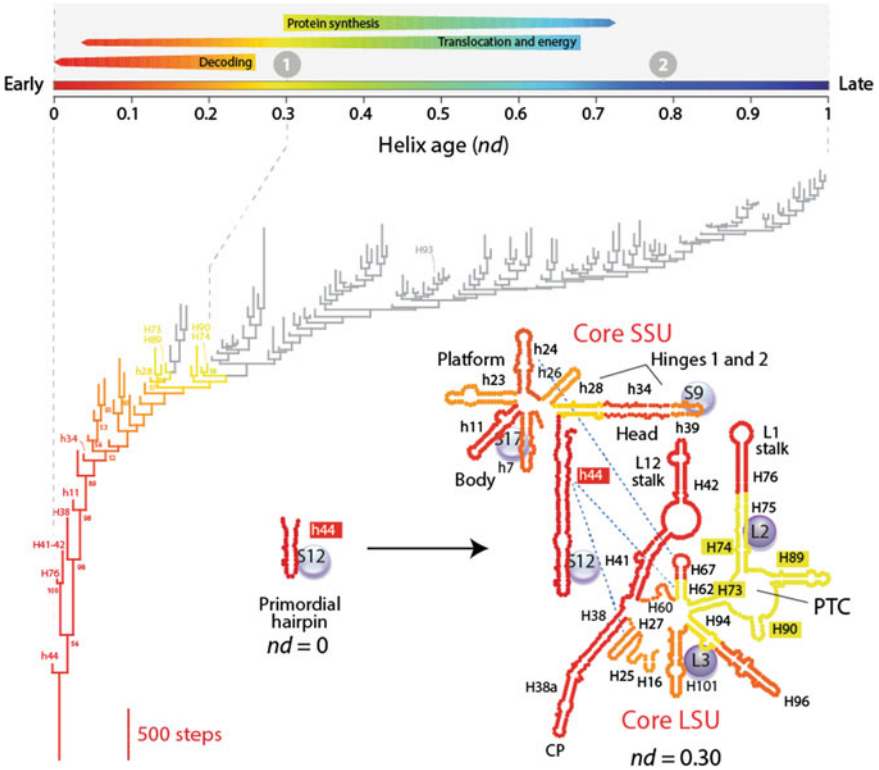


Fig. 6 Timeline of rRNA history. The first (major) and second transitions are indicated with *encircled numbers* in the timeline of rRNA substructures, which unfolds in time from *left* to *right* and is indexed with molecular functions. The timeline was inferred from a ToS, which is shown below. The branches of the ToS leading to the major transition are colored according to the age of evolving substructures. The major transition occurred once the decoding apparatus was in place, the H73, H74, H89 and H90 of the LSU formed the PTC responsible for protein synthesis (*yellow region* of the tree), and inter-subunit bridges (*dashed lines* in the model) were brought together and stabilized the SSU and LSU subunits. A model of the ribosome at the time of the major transition (*nd* = 0.3) is shown below the ToS with secondary structures colored according to their age. r-Proteins are indicated with *labeled buttons*. The growth of helical segments was modeled with growth rates of 100 base pairs/nd (~26 base pairs/Gy) and an average start length of 15.9 ± 11 (SD) bp to assume recruitment

7 tRNAs Are Evolutionary Building Blocks of Ribosomes and Genomes

A recent study generated lists of non-overlapping pairwise global alignments between tRNA and rRNA molecules that identified a number of remote homologies, which were often overlapping [70]. Similarly, sequential and overlapping remote homologies were detected between reconstructed tRNA and the PTC core of

LSU rRNA [71, 72]. These results suggest that both subunits of the ribosome were built piecemeal from primordial tRNA molecules. They also support an early proposal that the PTC originated from two tRNA halves by ancestral duplication [73] and even an earlier proposal supported by early bioinformatics analyses that tRNA and rRNA shared a common history [74]. Remarkably, we recently explored how the putative tRNA accretion process gave rise to functional rRNA by tracing the age of rRNA regions associated with the isoacceptor tRNA relics [75]. The ages of rRNA were taken directly from the work of Harish and Caetano-Anollés [32]. Remarkably, tRNA relics were enriched in older regions of the rRNA molecules, and these older regions harbored isoacceptor tRNA homologies that were also enriched in the oldest group 1 and 2 editing specificities for amino acid charging [31]. Thus, it appears tRNA relics preserve information about charging functions developed during the rise of the 'operational' code. What is even more remarkable is the existence of remote homologies to genes encoding very old proteins of metabolism, translation and replication that are also hidden in rRNA [70]. Thus, ancient rRNA had dual roles. It acted as a macromolecular machine or as a genome capable of encoding the information that the machine translated into proteins.

8 Conclusions

Translation is a biological process of interpretation of genetic information for the biosynthesis of proteins. Structural phylogenomic analysis suggests translation is ancient but developed later than the most primordial enzymatic functions of metabolism. Interactions with tRNA involve domains that were not at the base of the phylogenomic trees. Even the most ancient translation-related domains had metabolic functions (e.g., amino acylation of tRNA in catalytic domains), which preceded ribosomal-mediated protein biosynthesis. This has profound consequences for our understanding of how the molecular machinery of the cell originated. In current efforts to jumpstart a cellular system in vitro with the tools of synthetic biology, the "*cooperative structure of homeostasis*," which is embedded in proteins and cellular structure, must be established first, before ever attempting to impose a "*tyranny of replicators*" on the emerging system. Bioengineering should interface with knowledge from evolutionary history.

We note that the historical explorations we here describe started almost two decades ago. Their premise is that phylogenetic history exists in the structure of extant molecules. Its approach is grounded in cladistic methodology widely applied to the systematic survey of organismal biodiversity. Inferences about molecular structure are made with state-of-the-art HMM methods taking advantage of genomic information that is increasingly available. Phylogenetic trees are built using algorithmic implementations that extract deep phylogenetic signal from protein and nucleic acid molecules. Our studies have been followed by a handful of explorations from other laboratories, including building trees of life [76, 77], tracing domain changes in their branches [78] or constructing databases of structures

present in the last universal common ancestor of life [79]. Some explorations have been misguided by the use of unrealistic evolutionary models [80]. Since cladistics offers an objective criterion to reconstructing history, explorations follow the hypothetico-deductive method for overthrowing theories that supports scientific growth [81]. The strength of relationships of homology is tested at every stage of the exploration. The goal is to enhance the breath and scope, universality and degree of precision of the evidence that supports the historical conjectures. The effort increases explanatory power, empirical content and degree of corroboration. In the process, phylogenomics has repeatedly falsified the ancient "RNA world" theory in favor of other alternatives. The exercise attempts avoidance of recently highlighted fallacies that exist in the origin-of-life research field [82]. The experimental research of this field, which is predicated on deductive logic, appears largely immature, lacks "*patterns of progress*," and cannot integrate empirical evidence and theory from many domains of inquiry. Uncertainties in origin-of-life research are a "*breeding ground for a proclivity to combine wild speculation with dogmatic defense*" [82]. This explains a number of pernicious tendencies, including the adoption of extreme skepticism, collapse into metaphysics, and retreat to aprioristic narration and mythology. Phylogenomics provides one avenue out of the impasse. This avenue can systematize knowledge about the natural history of biological molecules and life.

Acknowledgments Computational biology is supported by grants from NSF (OISE-1172791 and DBI-1041233) and USDA (ILLU-802-909) to GCA. DCA is the recipient of NSF postdoctoral fellowship award 1523549.

References

1. Dyson F. Origins of life. Cambridge: Cambridge University Press; 1999.
2. Reynolds NM, Lazazzera BA, Ibba M. Cellular mechanisms that control mistranslation. Nature Rev Microbiol. 2010;8:849–56.
3. Francklyn CS, Minajigi A. tRNA as active chemical scaffold for diverse chemical transformations. FEBS Lett. 2010;584:366–75.
4. Hoeppner MP, Gardner PP, Poole AM. Comparative analysis of RNA families reveals distinct repertoires for each domain of life. PLoS Comput Biol. 2012;8:e1002752.
5. Fitch WM, Upper K. The phylogeny of tRNA sequences provides evidence for ambiguity reduction in the origin of the genetic code. Cold Spring Harb Symp Quant Biol. 1987;52:759–67.
6. Eigen M, Lindemann BF, Tietze M, Winkler-Oswatitsch R, Dress A, von Haeseler A. How old is the genetic code? Science. 1989;244:673–9.
7. Di Giulio M. The phylogeny of tRNA molecules and the origin of the genetic code. Orig Life Evol Biosph. 1994;24:425–34.
8. Sun F-J, Caetano-Anollés G. The origin and evolution of tRNA inferred from phylogenetic analysis of structure. J Mol Evol. 2008;66:21–35.
9. Farias ST. Suggested phylogeny of tRNAs based on the construction of ancestral sequences. J Theor Biol. 2013;335:245–8.
10. Hennig W. Phylogenetic systematics. Urbana: University of Illinois Press; 1966.

11. Zuckerkandl E. The appearance of new structures and functions in proteins during evolution. J Mol Evol. 1975;7:1–57.
12. Dayhoff MO. The origin and evolution of protein superfamilies. Fed Proc. 1976;35:2132–8.
13. Almo SC, Garforth SJ, Hillerich BS, Love JD, Seidel RD, Burley SK. Protein production from the structural genomics perspective: achievements and future needs. Curr Opin Struct Biol. 2013;23:335–44.
14. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The protein data bank. Nucleic Acids Res. 2000;28:235–42.
15. Reddy TBK, Thomas A, Stamatis D, Bertsch J, Isbandi M, Jansson J, Mallajosyula J, Pagani I, Lobos E, Kyrpides N. The Genomes OnLine Database (GOLD) v. 5: a metadata management system based on a four level (meta)genome project classification. Nucleic Acids Res. 2014;. doi:10.1093/nar/gku950.
16. Wheeler WC. Systematics: a course of lectures. Hoboken: John Wiley & Sons; 2012.
17. Caetano-Anollés G, Sun F-J, Wang M, Yafremava LS, Harish A, Kim HS, Knudsen V, Caetano-Anollés D, Mittenthal JE. Origin and evolution of modern biochemistry: insights from genomes and molecular structure. Front Biosci. 2008;13:5212–40.
18. Murzin AG, Brenner SE, Hubbard T, Chothia C. SCOP: a structural classification of proteins database for the investigation of sequences and structures. J Mol Biol. 1995;247:536–40.
19. Orengo CA, Michie A, Jones S, Jones DT, Swindells M, Thornton JM. CATH–a hierarchic classification of protein domain structures. Structure. 1997;5:1093–109.
20. Caetano-Anollés G, Caetano-Anollés D. An evolutionarily structured universe of protein architecture. Genome Res. 2003;13:1563–71.
21. Nasir A, Caetano-Anollés G. A phylogenomic data-driven exploration of viral origins and evolution. Science Adv. 2015;1:e1500527.
22. Wang M, Jiang Y-Y, Kim KM, Qu G, Ji H-F, Zhang H-Y, Caetano-Anollés G. A molecular clock of protein folds and its power in tracing the early history of aerobic metabolism and planet oxygenation. Mol Biol Evol. 2011;28:567–82.
23. Laurin M. Recent progress in paleontological methods for dating the Tree of Life. Front Genet. 2012;3:130.
24. Wang M, Caetano-Anollés G. The evolutionary mechanics of domain organization in proteomes and the rise of modularity in the protein world. Structure. 2009;17:66–78.
25. Wang M, Yafremava LS, Caetano-Anollés D, Mittenthal LE, Caetano-Anollés G. Reductive evolution of architectural repertoires in proteomes and the birth of the tripartite world. Genome Res. 2007;17:1572–85.
26. Nasir A, Caetano-Anollés G. Comparative analysis of proteomes and functionomes provides insights into origins of cellular diversification. Archaea. 2013;2013:648746.
27. Caetano-Anollés G, Kim HS, Mittenthal JE. The origin of modern metabolic networks inferred from phylogenomic analysis of protein architecture. Proc Natl Acad Sci USA. 2007;104:9358–63.
28. Kim KM, Qin T, Jiang Y-Y, Chen L-L, Xiong M, Caetano-Anollés D, Zhang H-Y, Caetano-Anollés G. Protein domain structure uncovers the origin of aerobic metabolism and the rise of planetary oxygen. Structure. 2012;20:67–76.
29. Caetano-Anollés K, Caetano-Anollés G. Structural phylogenomics reveals gradual evolutionary replacement of abiotic chemistries by protein enzymes in purine metabolism. PLoS ONE. 2013;8:e59300.
30. Caetano-Anollés D, Kim KM, Mittenthal JE, Caetano-Anollés G. Proteome evolution and metabolic origins of translation and cellular life. J Mol Evol. 2011;72:14–33.
31. Caetano-Anollés G, Wang M, Caetano-Anollés D. Structural phylogenomics retrodicts the origin of the genetic code and uncovers the evolutionary impact of protein flexibility. PLoS ONE. 2013;8:e72225.
32. Harish A, Caetano-Anollés G. Ribosomal history reveals origins of modern protein synthesis. PLoS ONE. 2012;7:e32776.

33. Caetano-Anollés G, Caetano-Anollés D. Computing the origin and evolution of the ribosome from its structure—uncovering processes of macromolecular accretion benefiting synthetic biology. Comp Struct Biotech J. 2015;13:427–47.
34. Dupont CL, Butcher A, Valas RE, Bourne PE, Caetano-Anollés G. History of biological metal utilization inferred through phylogenomic analysis of protein structures. Proc Natl Acad Sci USA. 2010;107:10567–72.
35. Nath N, Mitchel JOB, Caetano-Anollés G. The natural history of biocatalytic mechanisms. PLoS Comput Biol. 2014;10:e1003642.
36. Debès C, Wang M, Caetano-Anollés G, Gratër F. Evolutionary optimization of protein folding. PLoS Comput Biol. 2013;9:e1002861.
37. Nasir A, Kim KM, Caetano-Anollés G. Global patterns of domain gain and loss in superkingdoms. PLoS Comput Biol. 2014;10:e1003452.
38. Kim KM, Caetano-Anollés G. The proteomic complexity and rise of the primordial ancestor of diversified life. BMC Evol Biol. 2011;11:140.
39. Caetano-Anollés G, Mittenthal JE, Caetano-Anollés D, Kim KM. A calibrated chronology of biochemistry reveals a stem line of descent responsible for planetary biodiversity. Front Genet. 2014;5:306.
40. Vandergon TL. Protein domain structure evolution. Molecular Life Sciences. New York: Springer; 2014. doi:10.1007/978-1-4614-6436-5_19-2 .
41. Caetano-Anollés G. Novel strategies to study the role of mutation and nucleic acid structure in evolution. Plant Cell Tissue Org Cult. 2001;67:115–32.
42. Caetano-Anollés G. Evolved RNA secondary structure and the rooting of the universal tree of life. J Mol Evol. 2002;4:333–45.
43. Caetano-Anollés G. Tracing the evolution of RNA structure in ribosomes. Nucleic Acids Res. 2002;30:2575–87.
44. Sun F-J, Fleurdépine S, Bousquet-Antonelli C, Caetano-Anollés G, Deragon J-M. Common evolutionary trends for SINE RNA structures. Trends Genet. 2007;23:26–33.
45. Bailor MH, Sun X, Al-Hashimi HM. Topology links RNA secondary structure with global conformation, dynamics, and adaptation. Science. 2010;327:202–6.
46. Hyeon C, Thirumalai D. Chain length determines the folding rates of RNA. Biopys J. 2012;102:L11–3.
47. Fontana W. Modeling 'evo-devo' with RNA. BioEssays. 2002;24:1164–77.
48. Sun F-J, Caetano-Anollés G. Evolutionary patterns in the sequence and structure of transfer RNA: Early origins of Archaea and viruses. PLoS Comput Biol. 2008;4:e1000018.
49. Sun F-J, Caetano-Anollés G. Evolutionary patterns in the sequence and structure of transfer RNA: A window into early translation and the genetic code. PLoS ONE. 2008;3:e2799.
50. Sun F-J, Caetano-Anollés G. The evolutionary history of the structure of 5S ribosomal RNA. J Mol Evol. 2009;69:430–43.
51. Sun F-J, Caetano-Anollés G. The ancient history of the structure of ribonuclease P and the early origins of Archaea. BMC Bioinformatics. 2010;11:153.
52. Weston PH. Indirect and direct methods in systematics. In: Humphries CJ, editor. Ontogeny and Systematics. New York: Columbia University Press; 1988. p. 27–56.
53. Caetano-Anollés G, Wang M, Caetano-Anollés D, Mittenthal JE. The origin, evolution and structure of the protein world. Biochem J. 2009;417:621–37.
54. Sun F-J, Harish A, Caetano-Anolles G. Phylogenetic utility of RNA structure: evolution's arrow and emergence of modern biochemistry and diversified life. In: Caetano-Anollés G, editor. Evolutionary bioinformatics and systems biology. Hoboken: Wiley-Blackwell; 2010. p. 329–60.
55. Przytycka T, Aurora R, Rose GD. A protein taxonomy based on secondary structure. Nature Struct Biol. 1999;6:672–82.
56. Mittenthal JE, Caetano-Anollés D, Caetano-Anollés G. Biphasic patterns of diversification and the emergence of modules. Front Genet. 2012;3:147.
57. Tal G, Boca SM, Mittenthal JE, Caetano-Anollés G. A dynamic model for evolution of protein structure. J Mol Evol. 2016;82:230–243.

58. Caetano-Anollés G, Kim KM, Caetano-Anollés D. The phylogenomic roots of modern biochemistry: Origins of proteins, cofactors and protein biosynthesis. J Mol Evol. 2012;74:1–34.
59. Bukhari SA, Caetano-Anollés G. Origin and evolution of protein fold designs inferred from phylogenomic analysis of CATH domain structures in proteomes. PLoS Comput Biol. 2013;9: e1003009.
60. Ikehara K. Possible steps to the emergence of life: The [GADV]-protein world hypothesis. Chem Rec. 2005;5:107–18.
61. Jakschitz T, Rode BM. Evolution from simple in- organic compounds to chiral peptides. Chem Soc Rev. 2012;41:5484–9.
62. Söding J, Lupas AN. More than the sum of their parts: On the evolution of proteins from peptides. BioEssays. 2003;25:837–46.
63. Trifonov EN, Frenkel ZM. Evolution of protein modularity. Curr Op Struct Biol. 2009;18:335–40.
64. Goncearenco A, Berezovsky IN. Protein function from its emergence to diversity in contemporary proteins. Phys Biol. 2015;12:045002.
65. Aziz MF, Caetano-Anollés G. The early history and emergence of molecular functions and modular scale-free behavior. Sci Rep. 2016;6:25058.
66. Schimmel P, Giege R, Moras D, Yokoyama S. An operational RNA code for amino acids and possible relation to the genetic code. Proc Natl Acad Sci USA. 1993;90:8763–8.
67. Carter CW Jr, Wolfenden R. tRNA acceptor stem and anticodon bases form independent codes related to protein folding. Proc Natl Acad Sci USA. 2015;112:7489–94.
68. Caetano-Anollés G, Sun F-J. The natural history of transfer RNA and its interactions with the ribosome. Front Genet. 2014;5:127.
69. Rodin SN, Rodin AS. On the origin of the genetic code: signatures of its primordial complementarity in tRNAs and aminoacyl-tRNA synthetases. Heredity. 2008;100:341–55.
70. Root-Bernstein M, Root-Bernstein R. The ribosome as a missing link in the evolution of life. J Theor Biol. 2015;367:130–58.
71. Farias ST, Rêgo TG, José MV. Origin and evolution of the peptidyl transferase center from proto-tRNAs. FEBS Open Bio. 2014;4:175–8.
72. Farias ST, Rêgo TG, José MV. tRNA core hypothesis for the transition between the RNA world to the ribonucleoprotein world. 2016 (submitted).
73. Agmon I, Bashan A, Yonath A. On ribosome conservation and evolution. Israel J Ecol Evol. 2006;52:359–74.
74. Bloch D, McArthur B, Widdowson R, Spector D, Guimarães RC, Smith J. tRNA-rRNA sequence homologies: a model for the origin of a common ancestral molecule, and prospects for its reconstruction. Orig Life. 1984;14:571–8.
75. Caetano-Anollés G, Root-Bernstein R, Caetano-Anollés G. tRNA: building blocks of ribosomes and genomes. 2016 (submitted).
76. Yang S, Doolittle RF, Bourne PE. Phylogeny determined by protein domain content. Proc Natl Acad Sci USA. 2005;102:373–8.
77. Fang H, Oates ME, Pethica RB, Greenwood JM, Sardar AJ, Rackham OJ, Donoghue PC, Stamatakis A, de Lima Morais DA, Gough J. A daily-updated tree of (sequenced) life as a reference for genome research. Sci Rep. 2013;3:2015.
78. Edwards H, Abeln S, Deane CM. Exploring fold preferences of new-born and ancient protein superfamilies. PLoS Comput Biol. 2013;9:e1003325.
79. Goldman AD, Bernhard TM, Dolzhenko E, Landweber LF. LUCApedia: a database for the study of ancient life. Nucleic Acids Res. 2013;41:D1079–82.
80. Kim KM, Nasir A, Caetano-Anollés G. The importance of using realistic evolutionary models for retrodicting proteomes. Biochimie. 2014;99:129–37.
81. Farris JS. Parsimony and explanatory power. Cladistics. 2008;24:1–23.
82. Wächtershäuser G. In praise of error. J Mol Evol. 2016. doi:10.1007/s00239-015-9727-3.

Origins and Early Evolution of the Ribosome

George E. Fox

1 The Beginnings: Origins of the PTC and Possible Existence of an RNA World

1.1 Background

A central question in the origin of the modern translation machinery is the source of the extremely conserved PTC where the peptide bond is formed. Indeed, the formation of the PTC is by definition at the beginning of ribosome history even if some of the components actually predate it. For many years, investigators sought to discover which ribosomal protein or combinations thereof were responsible for catalyzing peptide bond formation. Failure to solve this riddle gradually led to the hypothesis that the ribosomal RNA (rRNA) had a significant role as reviewed by Noller [1]. When a high-resolution crystal structure of the 50S ribosomal subunit was obtained [2], it quickly became clear that since no protein appeared to be in proximity to the PTC region, the RNA itself must be catalytic [3]. Later it was found, in the case of the *Thermus thermophilus* ribosome, that ribosomal protein L27 did in fact closely approach the PTC [4, 5]. However, elimination of multiple residues that most closely approach the PTC does not completely prevent peptide bond formation [6]. Most recently it was shown that the rate of peptide bond formation was independent of the presence or absence of L27 [7]. Thus, the ribosome is widely and likely correctly regarded as an RNA machine.

G.E. Fox (✉)
Department of Biology and Biochemistry, University of Houston,
3455 Cullen Blvd. Suite 342, Houston, TX 77204-5001, USA
e-mail: Fox@uh.edu

G. Hernández and R. Jagus (eds.), *Evolution of the Protein Synthesis Machinery and Its Regulation*, DOI 10.1007/978-3-319-39468-8_3

1.2 RNA World

This brings the question of the RNA World to the fore. The evidence that is most indicative of the existence of a primitive RNA World that predated protein synthesis is the potential involvement of nucleotide coenzymes in prebiotic reactions as originally proposed by White in 1976 [8] and recently reviewed and pursued further by Yarus [9]. In many cases, the core structure of the coenzyme is an adenine diphosphate with a side chain that sometimes contains amide bonds attached to the 5′ diphosphate. This is seen, for example, in nicotinamide adenine dinucleotide and coenzyme A. This same feature is also seen in modern aminoacyl tRNA synthetases that initially attach the cognate amino in the same position creating the aminoacyl adenylate.

However, to obtain an extensive RNA World in which RNA catalyzes biochemical pathways that many have envisioned, it is thought necessary to obtain an RNA-based RNA replicase. This would need to be accomplished before peptide synthesis became established as the latter could quickly terminate the RNA World in favor of an RNA/peptide World. Proteins that can replicate RNA clearly exist, whereas effective RNA-catalyzed RNA replication of RNAs has proven difficult to demonstrate. Although template-directed synthesis by ribozymes is feasible [10], enzymes that synthesize RNAs approaching their own size had been difficult to obtain [11]. Using selection in ice, this hurdle was recently overcome in a specialized laboratory setting [12].

More recent efforts have highlighted the value of freeze-thaw cycles in the possible generation of self-replicating RNAs [13]. This progress notwithstanding, the complexity of these synthetic self-replicating RNAs is not substantially less than that of the peptidyl transferase center (PTC) found in extant ribosomes. So perhaps the quest to find a self-replicating RNA should be augmented/replaced with a quest to find a minimal peptide that can replicate an RNA. At the least, such a peptide RNA polymerase must emerge at some early stage as protein polymerases, and other complex proteins such as aminoacyl tRNA synthetases are likely present at the time of the LUCA [14]. Depending on its size, evidence of such a peptide could strongly support an abbreviated RNA World where the discovery of peptide synthesis predated the discovery of an RNA replicase. Indeed, RNAs that can synthesize peptides have been obtained in selection experiments [15]. As noted by Lilley [16], “Ultimately the finest achievement of the RNA World was probably the creation of proteins. These then took over most of the catalytic functions, leaving the ribosome as the most permanent monument to a heroic era.” However, even this may be an overstatement, as the RNA World may simply have never existed, as is suggested by the primitive nature of the mechanism of peptide synthesis in ribosomes and limited evidence, if any, of proteins replacing ribozymes [17].

A second problem with even an abbreviated RNA World, which is shared with ribosomal origins, is of course the source of the RNA. Although recent progress has been made [18, 19], the prebiotic synthesis of chiral RNA or a precursor nucleic acid remains uncertain. One alternative view is a prebiotic world in which peptides

of perhaps 3–8 residues could be produced without ribosomes [20]. The synthesis of even longer polymers of amino acids on illite and hydroxyapatite has been demonstrated [21]. Non-coded peptide synthesis in fact occurs in extant organisms using coded synthetases [22]. In the absence of the ribosomal machinery, various atypical features can be utilized including the incorporation of D-amino acids and/or non-standard amino acids [23]. Irrespective of how or when RNA or a precursor nucleic acid first became available, it would clearly be an important step in origins. This is because complexity is readily increased even in the absence of a replicase by hybridization and/or ligation. These processes are obtainable in random RNA pools [10]. It has been shown, for example, that a ligation of a simple RNA stem loop structure to itself could in principal produce a tRNA with the full complexity of the modern molecule [24].

1.3 Peptidyl Transferase Center

With regard to the modern PTC, Ilana Agmon in Yonath's research group found evidence of structural similarity between the A-site and P-site portions of the modern PTC [25]. This suggests that the modern PTC actually arose from an ancient duplication or hybridization event [26, 27]. Consistent with this, in the LSU rRNAs of *Chlamydomonas reinhaardtii* mitochondria [28] and the *Euglena gracilis* cytoplasm [29], the PTC is actually formed by noncovalent base pairing between two separate RNA fragments. These fragments are, however, very large and likely more indicative of possibility rather than actual history.

Regardless of its origin, a popular question regarding the PTC is the mechanism of the two reactions it is involved in [30]. The more studied of these is the entropic peptidyl transferase reaction in which the ester bond that links the nascent peptide to the 3′ hydroxyl of the 3′ terminal ribose of the P-site tRNA is subject to aminolysis by the alpha-amino group of the incoming tRNA. The reaction is fast with 15–50 peptide bonds produced per second [31]. The PTC is thought to be an entropy trap, which means it is the positioning of the substrates that is central rather than conventional chemical catalysis [32, 33]. The chemistry has been the subject of considerable discussion [30, 34–39] with no consensus resolution. Unlike a typical enzyme, the reaction is also not specific. The PTC can synthesize other products including esters and thioesters [40–44] and utilize non-standard amino acids [45–47]. The second reaction is the subsequent release of the peptide from the P-site tRNA by ester bond hydrolysis, which occurs when a stop codon is reached. This reaction is much slower [48].

However, a focus on the specifics of the chemistry obscures the larger picture of the utility of a proto-ribosome in a prebiotic world. It is not just the ability to synthesize peptide or perhaps ester bonds that matters, but rather the ability of the process to be done over and over again in order to create a polymer [49]. The chemistry itself is inherently favorable in that the synthesis of a dipeptide is relatively easy, but its subsequent release is slow. Thus, further extension is

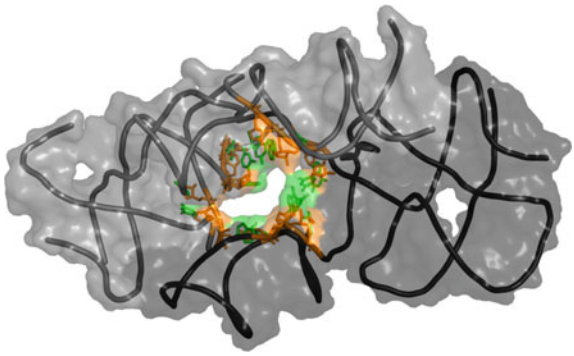


Fig. 1 Space filling model of the PTC region highlighting the residues that surround the exit pore. Backbone atoms are colored in *brown*, and the bases are in *green*. The *black line* traces the RNA backbone throughout the region. Unlike other known RNA pores that are lined almost exclusively with backbone atoms, the PTC pore provides access to the bases

possible if additional activated amino acids approach the growing chain before it is released. The modern ribosome exploits this by holding the growing chain in place in the P site while the next charged tRNA is recruited to the A site. When the PTC region of the ribosome is closely examined, one sees that the PTC actually forms a nanopore (Fig. 1).

The PTC pore serves as the entrance to the exit tunnel that provides a path for the growing peptide to pass through the ribosome and ultimately emerge as a mature protein. To accomplish this, various bases in the PTC are accessible at the pore surface. In the modern ribosome, any blockage of the exit tunnel rapidly inhibits protein synthesis, in essence because the active site becomes clogged. This is the mechanism employed by many macrolide antibiotics [50–56]. Many other antibiotics simply prevent the correct positioning of acceptor or donor substrates at the PTC [30].

With these facts in mind, it was previously proposed [49] that the entrance to the exit tunnel was part of the PTC from the very beginning. In fact, in combination with the favorable chemistry, it was likely this feature that provided a critical advantage over alternative peptide synthesis systems by facilitating the polymerization process as well as the synthesis of individual peptide bonds. For example, if pairs of activated amino acids bound to very small RNAs [57] are brought to the surface of a proto-ribosome they could, if positioned properly, form a peptide bond such that one of the RNAs now carries a dipeptide while the other is vacant. This is then the critical moment. If complexity is to continue to increase, the small RNA that contains the dipeptide must now be more likely to stay associated with the proto-ribosome. In contrast, the now naked small RNA must be more likely to leave the reaction center thereby making the creation of a trimer possible when a new RNA carrying an amino acid reaches the surface of the proto-ribosome. It is argued that the presence of the exit pore will facilitate precisely this by providing a means

to slow the departure of the small RNA carrying the dipeptide and preventing it from interfering with the next synthesis step.

What amino acid would the small RNAs carry? This question has been addressed not from the PTC perspective, but rather from the origins of coding. The basic idea is that the proposed small aminoacylated RNAs would be the ancestors of modern tRNAs. These RNAs may have delivered different amino acid to the proto-PTC depending on their individual structure and sequence. For example, the CCA stem of early proto-tRNAs may have favored some amino acids over others. Thus, as recently revisited [58], it has been speculated that there was originally an operational genetic code [59, 60] that would have predated the usual genetic code. In support of this idea, some modern aminoacyl tRNA synthetases do not rely on the anticodon to determine tRNA identity. For example, an alanine minihelix RNA and an even smaller microhelix consisting of only the acceptor stem are readily charged [61, 62].

What peptides would be made? It is not envisioned that any magic peptide would have emerged. In fact, as stated explicitly by Noller [1], "it is proposed that translation originally arose not to synthesize functional proteins, but to provide simple (perhaps random) peptides that bound to RNA, increasing its available structure space, and therefore its functional capabilities." Being random, noncoded, and likely lacking complex structures, the earliest peptides would provide one obvious benefit: stabilization of the growing proto-ribosome [49]. For example, simply increasing the lifetime of the machinery would facilitate further developments in the pre-LUCA World. However, stabilization of the very core of the emerging PTC was likely initially facilitated by metal ion interactions, not peptides. The modern PTC region is largely devoid of protein interactions. Instead, inner sphere interactions of magnesium with phosphate oxygens play a key role [63, 64]. However, at the time the ribosome was first developing, the Earth was anaerobic. As a result, ferrous ions may have been used rather than magnesium [65]. If a ribosome structure from an anaerobic organism becomes available, it will be of interest to see if ferrous iron is associated with the ribosomes.

Given the proposed importance of the PTC in terminating the RNA World one would expect that a suitable PTC analog would be stable, able to bind the substrates, and easy to form. Although it remains to be proven experimentally, initial quantum kernel energy studies indicate that Yonath's proto-ribosome would in fact form a stable structure to which small RNA substrates could stably attach [66]. But, how easily could a PTC-like structure be formed in a prebiotic world? A search of known RNA structures was performed, and in fact 11 additional nanometer size pores were found in the large rRNAs, and more examples were found in other RNAs [67]. These additional pores are made by folding of the RNA and in several cases encompass less than 100 nucleotides. However, unlike the PTC case, all of these additional pores are formed primarily by the backbone with the bases facing away from the pore surface. One might argue that a second pore in the rRNAs with a structure akin to the PTC pore would be detrimental because of possible competition. However, pores were also found in other RNAs, and although smaller, these too typically hide the bases from the pore surface. A second issue is

conservation. The residues that comprise the modern PTC region are found to be extremely conserved when a representative group of organisms is examined [68]. This reflects on one hand the importance of the PTC, but if only a few sequences can perform the function, then its discovery in the prebiotic world would be unlikely. Of course, the modern PTC is central to all extant organisms, and it likely was highly optimized by selective forces well before the emergence of the LUCA. It would be helpful to further study whether an artificial RNA capable of performing the PTC reaction is available. However, to date such an RNA-alone system has not yet been obtained.

2 Toward a Timeline for the Subsequent Evolution of the Translation Machinery

2.1 Initial Models of Ribosomal RNA Age

Compared to the PTC, the modern ribosome is incredibly complex. The large subunit RNA in bacteria is alone approximately 2900 residues in length, and clearly some parts of it must be older than other parts [69]. This raises the obvious challenge to develop the history and function of individual sections of the RNA and possibly ribosomal proteins, (r-proteins). Thus, structural insights have been used to deduce the relative age of various ribosomal components. The large subunit RNA in particular has been targeted by multiple approaches. An initial attempt examined sequence and secondary structure conservation to identify functionally important regions of the RNA [68]. One might also infer relative age from this comparison as the most conserved regions might be expected to be the oldest.

Further progress was spawned by the availability of atomic resolution structures. Initially, Hury et al. [70] examined tertiary interactions that provide connectivity between distant regions of the 23S rRNA. It was argued that connectivity likely increased over time with the most connected regions being the oldest. It was concluded that the oldest regions were likely domain 5, e.g. the PTC, followed essentially simultaneously by a portion of domain 2 (helices 31–35), which encompasses part of the exit tunnel, and domain IV, which is a major site of bridges to the 30S subunit. Somewhat later by this criterion would be the addition of parts of domain I and domain VI. Domain 3 and the GTPase center would be the most recent additions. A comparison with the LSU secondary structure analysis revealed that with the exception of the GTPase Center, essentially these same regions were universally conserved [70]. The GTPase exception reminds us that universality of structure is primarily about maintaining function.

An alternative approach was developed by Hsiao et al. [64]. They aligned three-dimensional structures of the 23S $rRNA^{HM}$ (23S rRNA of *Haloarcula marismortui*) and 23S $rRNA^{TT}$ (23S rRNA of *Thermus thermophilus*) and obtained objective local and global superimpositions of the two LSU rRNAs. They then

sectioned the superimposed LSUs into concentric shells, like an onion, using the site of peptidyl transfer as the origin. Next they approximated ribosomal evolution by accretion of spherical layers. Thus, RNA regions near the PTC are regarded as the oldest and increasingly recent as one approaches the surface.

This approximation appears to capture significant information along the evolutionary timeline revealing, for example, that the sequence and conformational similarity of these 23S rRNAs are greatest near the PTC origin and diverge smoothly with distance from it (i.e., with increasing spherical shell radius). Unlike the Hury model [70], the onion model [64] provided a relative age for individual helices in the RNA rather than just local regions. It was found that characteristics such as (1) rRNA conformation, (2) rRNA base pairing interactions, (3) rRNA interactions with Mg^{2+} ions, and (4) ribosomal protein conformation and interactions vary with distance from the PTC origin. The results suggest that the conformation, environment, and interactions of both RNA and protein can be described as changing in an observable manner over evolutionary time. An examination of the exit tunnel similarly provides insight into the age of ribosome regions as it was present from the beginning and necessarily maintained as the ribosome grew larger. Thus, it progresses from its beginning at the PTC in domain V to domains IV, II, I, and III respectively.

Subsequently, the Hury approach [70] was substantially refined [71]. In particular, it was recognized that A-minor tertiary interactions [72, 73] were potentially directional in time. These interactions connect adenosine stacks with helical regions that are distant in the primary sequence. The adenosine stack is not stable by itself, whereas the helical stack is thereby implying the latter is older. This idea coupled with a dismantling of the ribosome by systematically eliminating regions whose absence does not compromise the integrity of the remaining structure allowed construction of the first helix-by-helix model of rRNA history [71]. It was found that essentially all of the A-minor interactions associated with the PTC region of domain 5 involved a helix that interacted with an A-stack elsewhere in the RNA. Thus, the PTC was concluded to be the oldest portion of the LSU rRNA. Other old regions again included helices H31–H35 of domain II and helices H61 and H64–H67 of domain IV. In contrast, regions identified as very recent included helices H42–H44, which comprise much of the modern GTPase center, and helix H38, the A-site finger.

2.2 Accretion Model for Ribosomal RNA History

All of these initial studies focused on the bacterial and/or archaeal ribosomes. Eukaryotic ribosomes typically have larger RNAs. However, there is a common structural core shared by archaea, bacteria, and eukaryotes. The increased size of the eukaryotic rRNAs results from local blocks of additional residues, which are often referred to as expansion segments [74, 75]. In actuality, the first expansion

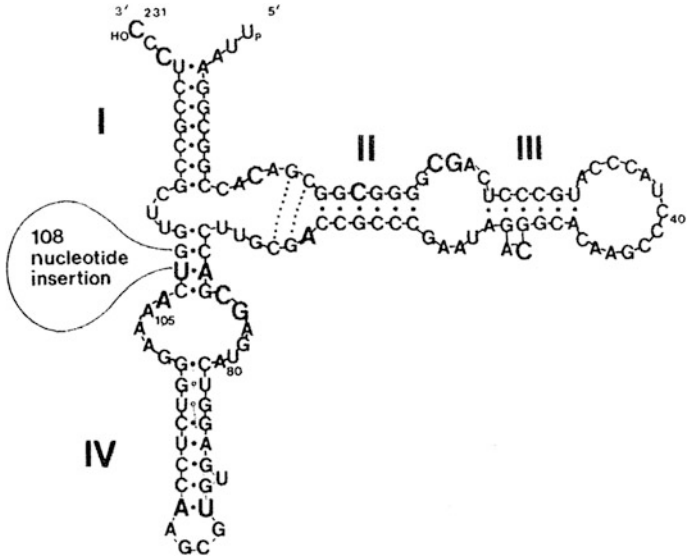


Fig. 2 The secondary structure of *H. morrhuae* 5S rRNA is shown with the location of the 108-nucleotide expansion sequence indicated [76]. The four primary helical regions in the 5S rRNA secondary structure are labeled by *bold Roman numerals*. In the presence of the *insert*, the backbone connection between residues 108 and 109 is broken, but the base pairs needed to make the usual fifth helical region appear to be present

segment that was observed was found in the 5S rRNA of *Halococcus morrhuae* (Fig. 2) [76].

In this case, the 108-nucleotide insertion nearly doubles the size of the 5S rRNA. The insertion emerges from what would otherwise be a helical segment of the RNA. By simply removing the insertion, one can reconnect the bases at the beginning and end of the insert and create a perfectly normal 5S rRNA. Because the insert is only present in one genus, one can readily infer that it was a recent addition. The crucial insight that had previously been overlooked is that the same processes were likely also in effect before the LUCA. With this in mind, one can basically work backwards looking for examples in the common core structure where a helix can be removed without disrupting the rest of the core. These frequently appear to have arisen as the result of insertion of a branch helix into a preexisting trunk helix as in the case of *H. morrhuae* 5S rRNA [77]. Such a helix, which is now universal, is then envisioned as having been inserted at a pre-LUCA time. A second example of insertion fingerprints arises by elongation of an existing helix [77]. In addition, local rearrangements of a stem may occur [78]. In multiple cases, one can see that additional structure has have been added over time while preserving more widely found inserts that presumably occurred at earlier times. Thus, by looking for different types of insertion fingerprints in combination with A-minor interactions and continuity, Petrov sought to infer the relative age of local regions in the LSU [77].

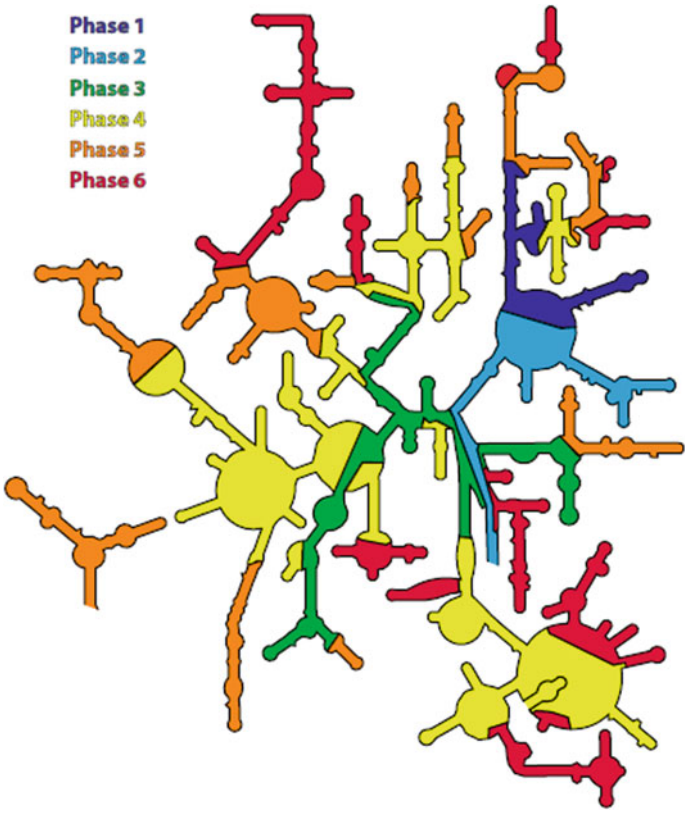


Fig. 3 The LSU common core is shown [77] using the updated LSU secondary structure [79]. Each region is assigned to one of six phases. The phase assignments are indicated by differing *colors*. The figure is reprinted with permission from Anton et al. [77]

A total of 58 accretion elements were identified and numbered (not shown) according to their possible order of addition to the growing LSU. The various insertion elements were then grouped into one of six phases (Fig. 3). Eukaryotes frequently have post-LUCA insertions and are envisioned to have two additional phases.

The resulting grouping is in many respects essentially a phylogenetic tree of the accretion events. As in a phylogenetic tree, timing along different branches can differ once they have diverged. Thus, in particular the exact temporal order of many peripheral additions is uncertain as they were acquired independently of one another [77]. It should also be appreciated that individual events in ribosome evolution such as the introduction of coding are likely occurring both over extended time periods and in parallel with other events. Thus, a linear timeline, although useful, is not realistic. Instead, the idea of phases is introduced to group events that are likely occurring in a similar time period. The accretion model addresses the

order of additions along various branches in the structure, but phase assignment must take into account other information. In sectors where additions have occurred in all the proposed phases, one can correctly infer their likely order, for example, from dark blue to light blue to green, to yellow, to orange, to red. However, if accretion was very fast in some periods, from a larger perspective two accretions may actually have occurred in the same phase. In many instances, not all phases are represented in a lineage that progresses, for example, directly from green to red. In this case one does not immediately know from the accretion data alone whether the final accretion actually occurred in phase 4, 5, or 6. The phase assignment usually takes into account other information, too. Thus, phase assignment is sometimes subjective, so future refinement of the model can be expected.

Another complication is relating traditional helix numbering and definitions to the accretion model and various ribosome studies. This reflects the fact that the traditional helix numbering scheme is rather random. Diagrams indicating the numbering system are available at (http://rna.ucsc.edu/rnacenter/images/figs/thermus_23s_2ndry.jpg) and for the SSU at (http://rna.ucsc.edu/rnacenter/images/figs/thermus_16s_2ndry.jpg). For example, both helix H34 and helix H42 actually consist of two stems separated by an interior loop. Using the recently proposed naming system for ribosomal proteins [80], it is seen for the case of H34 that ribosomal protein uL4 interacts in part with one of these helices but not the other. So saying uL4 interacts with helix H34 is misleading. In the accretion model, this numbering system is discarded in favor of a system in which each accretion element is given a number. So now the portion of helix H34 that interacts with uL4 is seen to be part of accretion element 10. However, accretion element 10 also contains what in the traditional helix numbering system is helix H35. Because the traditional numbering system has been widely used in numerous papers, the best solution might be to break offending helices in the traditional nomenclature up such that each stem has a specific name that relates to the original naming system. Then we would see that uL4 binds to H34a but not H34b.

Most recently, the accretion analysis was extended to the small subunit [81]. Again, it was possible to assign a relative age to various helices, and phases could be assigned. The crucial issue at this point is the time of initial collaboration between the two subunits. The small subunit might be envisioned to initially start as a small growing RNA with its own history that begins to interact with the emerging large subunit. In the most extreme case, the small subunit RNA is envisioned to have absolutely no prior history and thus is associated perhaps by hybridization with the large subunit as a small RNA. The key to correlating the small and large subunit timelines is the bridge elements and the extended two-domain tRNAs, which ultimately connect the two subunits. The bridges can only form when both partners exist. The timing relationship between the two subunits was based on bridge elements with directionality supported by A-minor elements in which the younger A-minor donors were found to be in the SSU. In the accretion model, the earliest bridge elements are B3 and B2a. Once the bridge elements were in place, it was possible to correlate subsequent events in the two subunits. However, prior to the amalgamation of the two subunits, each would have a separate timeline. The

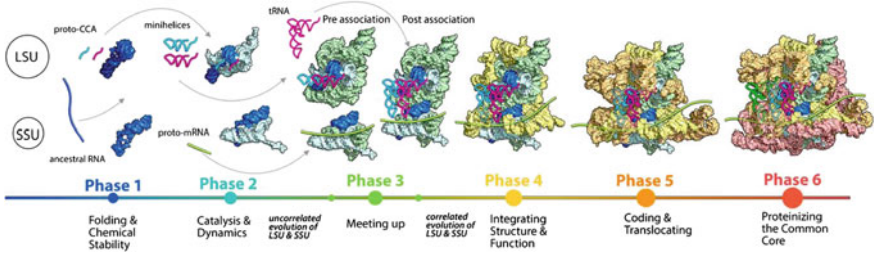


Fig. 4 A summary timeline representing major events in the evolution of the full ribosome is reprinted with permission [81]. In this version, the LSU and SSU are envisioned to have initially evolved independently and subsequently are joined together in phase 3

combined model provides an approximate time of appearance for each helical region and thereby provides an approximate timeline (Fig. 4) against which various ribosomal features and activities can be mapped. This timeline is organized into six phases for bacterial ribosomes, and events associated with each phase are described in detail [81].

2.3 Is the PTC the Oldest Portion of the Ribosome?

Separately an alternative model of ribosome origins that putatively relies primarily on phylogenetic methods applied to the ribosomal RNA structure has been proposed [82, 83]. This model and Petrov's accretion model [80] have generated recent controversy [84–86]. A primary issue is whether or not the origin of protein synthesis began with the creation of the PTC region of the large subunit. One strong indicator of this is that the PTC can be made from a single self-folding RNA, whereas the decoding center is not [87]. From a purely semantic point of view, it would appear the origin should be the structure that was first able to make peptide bonds. However, there is a possible complication in that it has been suggested on several occasions that the translocation process that moves the mRNA may have originated in the RNA World, possibly as the original replicase [88–90].

In fact, both Petrov and Caetano-Anollés recognize that a portion of the ancestral small subunit rRNA may have existed as an independent body prior to any association with a precursor large subunit. Harish and Caetano-Anollés [83] clearly state, "Intersubunit bridge history indicates early independent evolution of subunits." In this view, both subunits already consist of multiple helices at the time they first associate. In the case of the accretion model, the large subunit precursor is envisioned to contain the major components of the PTC and thought to already be making non-coded peptides at the time it acquires (or is acquired by) the proto-SSU in phase 3. Indeed, the modern 50S particle can synthesize peptides by itself. It therefore is appropriate to simply consider the addition of the small subunit as another step toward the modern ribosome thereby indicating that translation

evolved from the PTC even if a portion of the 30S subunit had an earlier prior history. However, this argument does not apply to the Caetano-Anollés model because at the time of the envisioned merger the PTC region was not yet included in the emerging large subunit.

The calculations that support this alternative view produce a tree of relationships between various helical elements. In this tree, helices H76, H41–42, H38, H67, H96, H60, H55, H101, H27, H25, and H16 all predate the helices that comprise the PTC. Several of these non-PTC putative aboriginal helices are not part of the universal core seen by sequence comparisons [68]. In addition, they are not actually backbone interconnected in the modern structure. This raises the question of how they were held together to create the proto-LSU structure. It might be argued that this could be done by hybridization. Consistent with this, in some cases modern rRNAs have been found to be produced by multiple smaller fragments that hybridize together to make the full rRNA [69]. In fact, it has been posited that the PTC itself may have emerged from hybridization of two fragments [69] and/or duplication of a smaller fragment [25]. From this writer's perspective, the Caetano-Anollés model needs to show that the proposed early fragments could have interacted with one another by hybridization without disrupting their essential structure. In addition, an explanation for how additional RNA was subsequently inserted between the putative early helices would be helpful. More detailed discussion of the Caetano-Anollés model will be found elsewhere in this volume.

3 Major Events Along the Timeline

In the accretion model, events are described at a level in which local structural elements of the RNAs are added at specific relative times. This allows one to infer the likely addition of other features. For example, it is reasonable to speculate that a particular ribosomal protein is not present until its binding site has been added. There are many inferences of this type that remain to be fully appreciated. In order to begin organizing this very detailed information, the emergence of various ribosomal subsystems and components have been broken up into approximately six phases. Major events associated with various phases in ribosome history will be discussed in detail in this section.

3.1 Homochirality and the Ribosome

A long-standing issue in the origin of life is when and how homochirality was introduced into emerging biological systems [91]. Homochirality is generally perceived to be the result of chemical properties and therefore thought to likely precede the emergence of ribosomes. That said, as reviewed in some detail previously [92], the modern ribosome can in fact be convinced to accept tRNAs carrying D-amino

acids [93, 94]. In addition, mutations in the PTC can improve the tolerance for D-amino acids [45, 95–97] and even β-amino acids [98]. If mutations can decrease the specificity of the modern ribosome, then it is clear that mutations in the past may have improved the specificity. Separately, a variety of processes are utilized to insure that D-amino acids are not attached to tRNAs by the amino acyl tRNA synthetases [99, 100], thereby indicating that this also may have been a problem in the past.

Thus, there is strong evidence that the early ribosome likely would have incorporated D-amino acids with some likelihood. However, if both processes have the same chiral preference then working together with both having perhaps 70 % accuracy they quickly produce a product that is over 85 % homochiral. So even if L-amino acids were not predominant in the pre-ribosome period, the early ribosomes would soon be producing nearly homochiral peptides. It should also be noted that incorporation of a D-amino acids into a modern protein is not necessarily destructive to the formation of a peptide, but instead it depends on where they are located [96, 97]. Alpha helices in particular are intolerant of D-amino acids. Thus, the first peptides made by a primitive ribosome would be more likely to contain beta sheets than alpha helices. Amino acids that do not present a chirality problem, e.g., glycine, might also be favored.

3.2 tRNA and the Timeline

The tRNAs span the two subunits and are therefore central to the modern ribosome. The universal CCA sequence at their 3′ end carries either the incoming amino acid or the partially synthesized growing peptide chain and is thus associated with the PTC. At the opposite extreme, approximately 70 Angstroms away, the anticodon regions of two tRNAs interact with the mRNA at the decoding center of the small subunit.

The origin of the typical modern tRNA of 76 nucleotides has been the subject of considerable discussion [101]. It has in particular been observed that the tRNA consists of essentially two domains with the presumably older top half carrying the amino acid and the younger lower half interacting with the mRNA [60, 102, 103]. The top portion includes the T-stem, the acceptor stem, and the universal, but typically not coded, terminal CCA to which the amino acid is attached to the modern tRNA. It has been suggested that this top portion may have originated as an even simpler single hairpin to which was later added the T-stem and loop [104]. A subsequent ligation of two copies of a stem loop structure could in principle produce the essence of the tRNA's 3D structure including long-range interactions [24].

As attractive as this general idea is, it is important to appreciate that other facts raise concern and tRNA origins may be consistent with other hypotheses. For example, in the case of some mitochondria, the D arm or T arm is deleted from the tRNA [105]. This feature normally interacts with the L1 stalk in the large subunit

rRNA, which is also missing. The L1 stalk is a phase 5/phase 6 addition to the LSU RNA, thereby raising the possibility that the D-stem is a recent addition. Likewise, the discovery of split tRNA genes in *Nanoarchaeum equitans* [106] suggests models in which two half structures formed the tRNA structure by hybridization rather than duplication [101].

However, the idea of a bottom-to-top addition of the lower portion of the tRNA stem has been greatly strengthened by the discovery of an ancient insertion fingerprint that supports the addition of the D arm and anticodon arm into an earlier structure consisting of the CCA stem and T arm [81]. The new addition to the tRNA would thus make interaction with a proto-mRNA possible, thereby representing the origins of true coding within the context of ribosome history. When did this occur? Although it is currently not obvious how to directly time the origin of the two-domain tRNA, one can infer its presence by the relative time of appearance of the small subunit helices associated with the decoding center. This places it as most likely in late phase 3 or early phase 4.

3.3 *Ribosome History Has Implications for the Origin of the Genetic Code*

The holy grail of ribosome evolution, which will not be addressed here, is an explanation of the origin of the modern codon assignments. This relates at least in part to the aminoacyl tRNA synthetases and has been recently examined in several reviews [107, 108]. The genetic code, as we know it, is often thought to arise in two stages. As small RNAs brought possible substrates to the proto-PTC, the sequence of the RNA carrier would favor the attachment of some amino acids (or perhaps esters) over others, resulting in an operational code that did not rely on any genomic information. The peptides made would be largely random depending on the availability and affinity of particular amino acids. Over time, as the carrier RNA became larger, additional residues would be added to the CCA stem thereby expanding the initial operational code. True coding as is found in modern organisms would require a two-domain tRNA and a proto-ribosome that includes the basic components of both subunits as well as an mRNA. Thus, in the context of the accretion model, coded synthesis would not begin until late in phase 3 or phase 4. The likely initial driving force was the dynamic h44–h28 combination, which could spontaneously switch between two or more configurations. This motion might have been coupled to a noncoding proto-mRNA that hybridized to h44 much as the Shine-Dalgarno sequence [109] of modern bacterial mRNAs does. The addition of the lower portion of the tRNA would then allow the tRNAs to interact indirectly with the h44–h28 driver by hybridization to the proto-mRNA. Alternatively the tRNA may have been extended first and then serendipitously interacted with a small RNA associated with the h44–h28 complex [110]. This would in either case greatly improve the speed of peptide synthesis by increasing the rate of motion from the A

to P site. The ability to more rapidly make random peptides of upwards of even 10–15 residues likely would have provided a new level of complexity to the prebiotic world. The sequence of the proto-mRNA would govern which tRNAs would best interact with it. This would create an environment in which, in a manner perhaps similar to the operational code, a genetic code as we know it could evolve. Thus, the proto-mRNA that may have initially simply facilitated movement of the tRNAs would now be available to carry information. Understanding the early history of the ribosome has not yet provided a specific explanation for any codon assignments. However, it does provide potentially useful insights into the environment in which the code was established.

Although the modern code is largely universal, there are numerous minor exceptions found in various organisms and most especially mitochondria. When a tRNA is mutated such that it recognizes two anticodons, e.g., UGA and UGG, as Trp, for example, this eliminates a stop codon. Alternatively, a tRNA mutation may result in a methionine tRNA that now reads both AUA and AUG. Thus, changes in codon assignments in modern organisms may really be about changes in the mRNA and tRNA populations. Especially interesting was the observation that in human mitochondria the arginine codons AGA and AGG were extremely rare and apparently the cell did not produce a tRNA to recognize them [111]. Hence, they were thought to have become de facto stop codons. However, it was subsequently found that they instead provoked frame-shifting events that then restored a normal stop codon [112].

These sorts of observation may have implications for the manner in which we envision coding prior to the LUCA. Initial genomes, especially RNA genomes, would have initially been small. As a result, many codons would potentially be rare and have to be dealt with in some manner. If they were treated as stop codons this would be undesirable as it would limit the size of peptides that could be made. In this context, tRNAs that could recognize multiple codons might have been preferred. The role of stop codons and small genomes has been discussed before [113] and likely should be revisited.

3.4 Ribosomal Proteins Line the Path to Increasing Complexity

Before the sequences and structures of the ribosomal proteins were known, it was thought that they might be descended from some small number of ancient proteins by gene duplication. In fact, there are only a few examples of such relationships, and lateral transfer events involving the ribosomal proteins are at best rare [114, 115]. In our original attempt to establish a timeline for ribosome evolution [116], the ribosomal proteins played a key role. It was observed that the non-universal proteins were typically late additions in ribosome assembly. This led to the hypothesis that the assembly map recapitulated to some significant extent the

evolutionary history of the ribosome [116]. It was thus argued that the oldest ribosomal proteins were likely uL2, uL3, uL4, and uL24, as these were at the core of the assembly process. Experimental studies showed that 50S particles alone containing 5S rRNA and eight proteins including uL2 and uL3 were active [117]. Subsequently, the relative ordering of both the SSU and LSU chronologies were optimized with respect to differences in amino acid usage bias [114]. The results, in conflict with the cladistics model [82, 83], strongly supported a scenario in which the LSU predates the SSU. Ribosomal proteins uL2, uL3, and uL4 in particular again appeared to be the oldest.

Modern genomics has shown that gene order is not typically conserved over vast phylogenetic distances [118]. In contrast with this observation, analysis of the ribosomal proteins revealed that in bacteria and archaea there are six clusters/operons consisting almost exclusively of ribosomal proteins that are in fact conserved. Within this group, the two largest clusters are the S10 operon, which includes uL2, uL3, uL4, and the spc operon, which includes uL24. In both cases, the order of the genes, not just the gene content, is typically preserved. In *E. coli* where experimental studies have been made, these operons are all regulated at the translational (e.g., RNA) level rather than the transcriptional level. This has been cited in support of these clusters having been part of a primitive RNA genome [118, 119].

An examination of the ribosomal proteins in the SCOP database release 4.75 [120] reveals that many of the ribosomal proteins have structures that are either all alpha or all beta. However, the subunits differ in that proteins with all beta structures, such as uL2 and uL3, and proteins comprised mostly of parallel beta sheets, such as uL4, occur more frequently in the large subunit than in the SSU. The universal proteins uL2, uL3, uL4, and uL22 all have extensions that reach the PTC region. It has been pointed out [121] that the regions that comprise these extensions are peptides with no secondary structure, loops with beta turns, or beta hairpins. The exit tunnel, which starts at the PTC, by its role in providing a path to the exterior of the ribosome, must have been maintained as the ribosome grew larger. Moving along this path is then basically a timeline. In doing so, one sees initial beta structures such as an extension of uL22 followed only later by an alpha helix. This absence of alpha helical structures in the oldest parts of the PTC may reflect the difficulty in obtaining alpha helices in peptides that initially may have been comprised of mixed chirality amino acids.

When equivalent proteins from the three domains of life were aligned, they were frequently found to exhibit a block structure in which some regions are universal and other segments are specific to either the Bacteria or the Archaea [122]. In a few cases, e.g., uL6, uL11, uL14, uS9, and uS11, there is essentially only a single universal block [122]. The existence of domain-specific blocks strongly supports the notion that, like the ribosomal RNAs, portions of the proteins may have arisen at different times. One clear example of a protein with interesting history is uL2. It has two distinct domains. The N-terminal domain forms an OB fold, while the C-terminal domain forms an SH3 fold [120, 123]. What is especially noteworthy is that these two folds are in fact very similar [123]. By simply moving the location of

one beta segment, it is possibly to effectively convert the SH3 fold to an OB fold. Thus, it is likely that one of the domains may have arisen from the other by duplication in conjunction with a partial rearrangement.

The question then is which domain of uL2 is the oldest. The accretion model may provide the answer by combining RNA-protein interaction sites with the likely age of the interaction sites. In examining a preliminary protein-RNA interaction map available at the CRW web site (http://www.rna.icmb.utexas.edu/SIM/4A/CRWStructure/rpi/rpi.23S.hb.pdf) [124], one finds that the primary site of interaction of the more universal proteins typically interacts with the predicted older regions of the RNA. At the very oldest regions of both subunits, protein interactions are minimal to nonexistent. The issue of the primary binding site where a protein interacts is complicated by the fact that there are typically multiple interaction sites for each protein. However, in most cases there is a cluster of contacts representing the primary interaction site. In the case of uL2, the SH3 domain interacts with accretion elements 12 and 12a, while the OB domain interacts primarily with accretion elements 9, 13, and 14. In both cases, these regions are assigned to phase 3, and although both domains are clearly essential to the early history of the ribosome, it remains unclear which domain is older. Interestingly, there is a small segment of uL2 that is between the two major domains that forms a small loop that interacts with the loop of helix 93 in what is regarded as phase 2 in the accretion model. Rather than either domain, this small element may in fact be the very oldest part of uL2.

Consistent with their early evolution, both the SH3 and OB domains are devoid of an alpha helix, and each is found in other ribosome-relevant proteins. The SH3 domain is found in bL21 and uL24 as well as a variety of intracellular or membrane-associated proteins. The OB fold is frequently associated with oligonucleotide binding and is found in several ribosomal proteins including uS12, uS17, and bS1. The bS1 protein is associated with initiation and actually contains six copies of the OB fold that form what is known as an S1 domain.

Useful insight can also be obtained by examining the non-universal proteins. Although these proteins likely have a post-LUCA origin, the way they have evolved may provide insight into the pre-LUCA acquisition of proteins. In a seminal paper, Klein et al. [125] compared the structures and locations of ribosomal proteins from the archaeon *Haloarcula marismortui* and the bacterium *Deinococcus radiodurans*. It was observed in several cases that one or more of the unique *H. marismortui* proteins in fact have non-homologous analogs in the *Deinococcus* structure. For example, eL21, eL24, and eL37 have analogs bL27, bL19, and bL34, respectively. In each case, the analogous proteins are distinctly different in sequence and structure. However, at the level of secondary structure at least, one sees essentially no significant change in the LSU RNA in these cases. When differences do occur, they are relatively minor. There are also examples of proteins that are in one structure with no analog in the other. Thus, for example bL25 and bL36 are unique to *Deinococcus*, and eL18, eL19, and eL39 are unique to *Haloarcula.* In some cases there are again no obvious RNA differences. This raises the evolutionary question in each case of whether the unique protein is a post-LUCA gain where it occurs or a loss where it is not found.

However the RNA sometimes does noticeably change in association with a protein. One interesting case is eL18, which is unique to *Haloarcula* and has partial sequence homology with uL15. It turns out that its binding site partially overlaps with uL15 on helix H27. In addition, there is a small helix insert (H30) between H29 and H31, which is the primary binding site for eL18. This insert and the protein are both absent in *Deinococcus*.

3.5 Origins of the Dynamic Ribosome

The modern ribosome is a dynamic molecular machine that executes protein synthesis one residue at a time. An incoming tRNA carrying the next amino acid in association with elongation factor EF-Tu enters at the A site following GTP cleavage. Peptide bond formation is spontaneous, and the tRNA now carrying the growing peptide chain moves to the P site with the help of elongation factor EF-G and a second GTP cleavage. A new tRNA now enters at the A site, and the process is repeated. As the process proceeds, the tRNA is now deacylated and moves to the E site from which it exits the ribosome. During the process the small subunit exhibits head swivel and rotates relative to the large subunit. Although facilitated by EF-Tu and EF-G, the core motions are actually inherent to the ribosome itself. In the absence of the factors, protein synthesis can still proceed [126, 127]. Thus, the ribosome is thought to be inherently a processive Brownian motor [128].

Well before any structures were available, Woese proposed that the tRNA itself was dynamic, ratcheting between two configurations of the anticodon loop [129]. This proved to be the correct general idea, but not the correct specific structure. The first atomic resolution tRNA structure revealed what appeared to be a hinge-like structure in the lower portion of the tRNA [130, 131]. It is now appreciated that as the tRNA enters and ultimately leaves the ribosome, various conformations are produced as a result of motions facilitated by the pivot point as indicated in Fig. 5, which was adapted from Dunckle et al. [132].

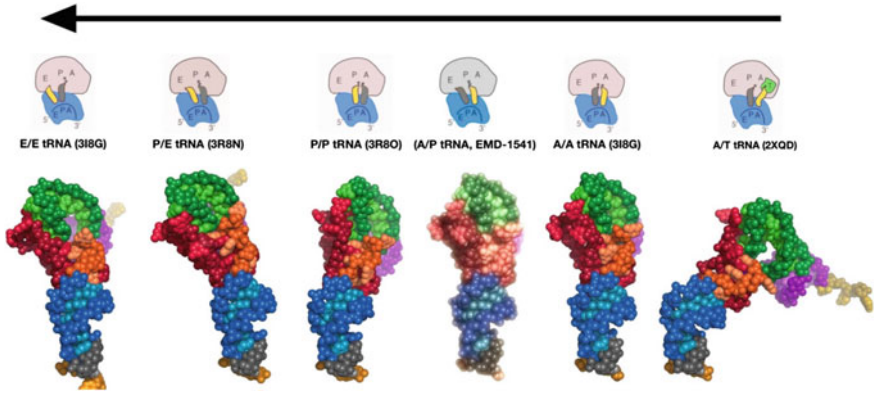


Fig. 5 Alternative tRNA orientations seen during a single translation cycle

The existence of this key center of motion inspired a search for additional pivots within the large RNAs. Initially, atomic resolution structures of the ribosome before and after EF-G associated GTP cleavage were compared. A total of 23 pivots were found [133]. Of these, 15 were in the small subunit and 8 in the large subunit. A similar analysis of pivots associated with EF-Tu associated GTP cleavage found 16 pivots, 4 of which are uniquely associated with EF-Tu. [134]. That many shared pivots are found is consistent with the long-standing observation that the EF-G and EF-Tu sites of interaction partially overlap [135, 136]. Both EF-G and EF-Tu interact with the factor-binding site, which includes helices H43 and H44.The pivoting positions are consistently associated with weak spots in the RNA structure such as non-standard base pairs and bulge loops.

In the case of EF-G, pivots in small subunit helices h28 and h44 are especially interesting. Motions associated with these two pivots strongly affected other pivots, Helix h28 appeared to control h31, h33, h36, h37, h39, h40, h41, h42, and h43, while pivot h32 appears to control motions at pivots h33, h36, h37, h39, and h40. These pivots are associated with the head swivel and head rotation. Overall, there appears to be a network of interacting pivots as outlined in Fig. 6, which is reprinted with permission [134].

The dynamic aspects of the modern ribosome clearly make it faster and likely more accurate as well. Placing the emergence of this key event on the timeline of ribosome evolution is an important goal. To begin to answer this question, one can examine when the various helices containing pivots were incorporated according to the accretion model [77]. In the case of the LSU, the ten pivots associated with either EF-Tu or EF-G or both are listed as phase 5 or 6. The one exception is the helix H89 pivot, which is part of the original aes1 element that initiated the P site

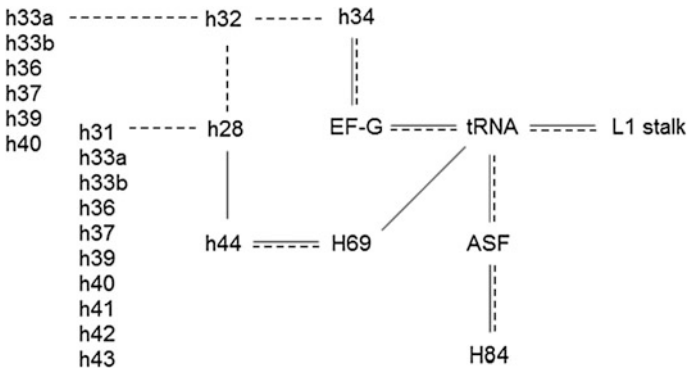


Fig. 6 Proposed partial network of EF-G-associated pivoting elements found in the ribosomal RNAs [134, 137]. *Black lines* indicate direct physical contact between moving helixes. *Dashed lines* indicate the motion that results from an upstream pivoting motion. Helixes h28, h32, and h34 form the head domain of the SSU and lie in sequence. Helices h28 and h32 influence the motions of a number of more external pivots, which are listed vertically in the figure, as well as helix h34, which contacts EF-G. Thus, a cascade of motion originating with EF-G-GTP hydrolysis is plausible in either direction—forward toward the tRNA or in the reverse direction toward h34

portion of the PTC. This helix connects with helix H91 and thereby indirectly with the alpha sarcin/ricin loop. As has been pointed out [77], helix elongation events do not leave insertion fingerprints, and therefore in the absence of clues from A-minor interactions, such events cannot be readily established. Therefore, it is not inconsistent with the accretion model to infer that helix H89 arose at a later time as a simple insertion in the pivoting element. It therefore seems likely that all the mobile aspects of the LSU RNA were added rather late in the process but nevertheless well before the LUCA.

The small subunit is clearly more dynamic than the large subunit. It is actively involved in both the head swivel and head rotation. In the accretion model, the vast majority of the pivots are either phase 5 (helices h6, h41, h42, and h44) or phase 6 (helices h8, h21, h26, h36/37, h38, h39, and h40). Pivots in helices h31, h32, and 43 are assigned to phase 4. There are small subunit pivots in helices h6, h32, h 41, and h44, which are also assigned to phase 5. Pivots in helices h8, h21, h26, h32, h36/37, h39, and h40 are designated as phase 6 and hence more recent. Pivots in helices h31, h32, and h43 are assigned to phase 4. The remaining two pivots h28 and h44 are regarded as being among the first elements to be incorporated into the growing small subunit. In the accretion model, they are assigned to phase 2 and phase 1, respectively. This is consistent with the fact that motions at these pivots strongly influence the motion seen at multiple other pivots. Helix h44 includes the Shine-Dalgarno sequence [109], which facilitates the initial interaction of the mRNA with the small ribosomal subunit during initiation and is involved in multiple bridges to the large subunit.

Overall these results indicate that the dynamic ribosome arose in essentially two separate stages. It is likely that initially the H28/H44 pair were part of an inherently dynamic element that independent of the ribosome would spontaneously alternate between two or more conformations. When incorporated into the emerging ribosome during phase 3, this inherent motion would have driven both the mRNA and tRNA movement. Subsequently, as the ribosome grew, other mobile elements were added, primarily in phase 5 or 6, resulting in head swivel and the ratcheting motion.

3.6 Recent Aspects of Ribosome Evolution

The ribosome clearly continued to evolve as complexity reached and extended beyond that of the LUCA. For example, as discussed elsewhere in this volume, initiation is distinctly different in eukaryotes. Likewise, the various factors have evolved over time. In bacteria, the addition of the tmRNA system for recovering stalled ribosomes is a noteworthy achievement. Another key development is the introduction of post-transcriptional modifications. In this section, three important additions that occurred very near the time of the LUCA are discussed.

3.6.1 Trigger Factor and Factor Binding Site

The involvement of ribosomes in co-translational folding in conjunction with the signal recognition particle and trigger factor was an important development in the integration of the ribosome with other cellular systems [138]. In this case, the key docking protein for the trigger factor is uL23, which interacts near helices H51 and H53, which are added in phase 5. A second protein that interacts with the trigger factor, uL29, is associated with phase 6. It thus appears that the trigger factor and co-translational folding predate the LUCA.

Although the ribosome is at its core a Brownian motor, the modern version is driven by GTP hydrolysis events associated with the initiation factor IF-2, elongation factors EF-Tu and EF-G, and release factors such as RF-2. The addition of these factors to the emerging ribosome would have greatly increased the rate of peptide synthesis and may have been as a result in part responsible for the emergence of the LUCA [70, 139]. In bacteria, the factor-binding site contains multiple copies of ribosomal protein bL12 attached to an underlying stalk formed by helix H42, which is thought to be added in phase 5, and ribosomal protein uL10. However, bL12 is replaced by P1 and P2 in archaea and eukaryotes [140], thereby associating the continuing formation of the factor-binding site with the post-LUCA time period. The structures of these various factors have been inter-compared in some detail, and it was concluded that they likely originated from a fusion of an OB fold to a Ras-like GTPase [121].

3.6.2 5S rRNA

5S rRNA is a small, independent RNA that largely forms the central protuberance of the large ribosomal subunit [141, 142]. In combination with several ribosomal proteins, its incorporation is essentially the last step in ribosome assembly [143]. 5S rRNA is universally found in all three domains of life and therefore was presumably present at the time of the LUCA. Reconstituted ribosomes lacking 5S rRNA have minimal but not completely eliminated activity [144]. In particular, factor-dependent tRNA binding at the A site is interrupted [142]. It was originally proposed that the sequence CGAA that frequently occurs in loop C of 5S rRNAs interacted with the T-loop of the tRNA [145]. As reviewed in detail [142, 146, 147], it is now thought that 5S rRNA may facilitate communication between different ribosome regions. This is a key ribosome function, as the movement of the mRNA likely needs to be coordinated with the synthesis of the peptide bond for the machine to function smoothly. Among 5S rRNAs' multiple contacts, its interaction with ribosomal protein uL5 is of special interest. This is because uL5 in turn interacts with uS13 to form bridge B1b/c. This bridge is the only protein-protein bridge between the ribosomal subunits, and its time of appearance likely coincides with that of 5S rRNA. Ribosomal protein uS13 also has a long C-terminal extension that reaches the coding site [148, 149]. Hence, a signal may be passed between the

LSU and the SSU decoding site. However, this is not essential as a deletion of the uS13 tail only minimally reduced ribosome activity [150].

Although 5S rRNA is regarded as "universal," this is subject to some interpretation as its structure is somewhat different in various organisms. In the initial *E. coli* structure obtained by comparative analysis, 5S rRNA was envisioned to have four extended helical regions [151, 152]. In contrast, the equivalent eukaryotic and archaeal structures were soon shown to have a fifth helix [153, 154] in a region that was originally envisioned as a bulge in the prokaryotic cases. The length of this helix varies and can be quite extended in organisms such as *Thermoplasma acidophilum* that are subjected to harsh environments [155]. Subsequent X-ray studies of this bulge region in *E. coli* revealed that it was actually a non-standard duplex formed by what is now known as a bulge E motif [156]. In addition, a few characteristic differences are seen in many 5S rRNA such as a single deletion in the loop C region in many eukaryotes. The appearance of equivalent structures is, however, not fully supported by experimental data. Reconstitution assays were conducted in which 5S rRNAs from various sources were incorporated into the large subunit of *Bacillus stearothermophilus* ribosomes [157]. It was found that 5S rRNAs from prokaryotic organisms typically produced active ribosomes, whereas those from eukaryotes did not [158].

The issue of 5S rRNA universality has long been questioned because it appears to be missing in many mitochondria and some chloroplasts. Recently, a detailed search for 5S rRNA homologs in organelle genomes was conducted [159]. Numerous new examples of 5S rRNA-like RNAs with additional structural variations were found in multiple cases. In the case of the mammal Porcine (*S. scrofa*), the structure of the large ribosomal subunit was examined by cryoelectron microscopy. It appears that a greatly reduced form of 5S rRNA is in fact present [160]. Just as the universality of 5S rRNA was on the brink of establishment, it was definitively shown this was not the case. The crystal structures of human [161] and yeast [162] were determined, and in both cases 5S rRNA was completely absent. In the case of the human mitochondrial large subunit, the 5S rRNA was replaced by a mitochondrial valine tRNA. This tRNA occupies essentially the same region in the structure as the abbreviated RNA found in the Porcine example. In summary, 5S rRNA was likely a late addition to the ribosome shortly before the LUCA.

4 Summary and Future Studies

It is suggested here that the matter of ribosome origins is best regarded as two distinct problems: the origins of the core machinery and its subsequent evolution. Much of ribosome history predates the LUCA, and thus its study takes us back in time some considerable distance into the pre-LUCA world. It is unclear how far back the origin of the PTC itself or its inherently dynamic SSU counterpart actually extends into the pre-LUCA time period. Clearly RNA is crucial, and the issue of where it comes from has significant implications. If an early path to a true RNA

World of any complexity existed, then the core aspects of the ribosome would be a direct and possibly quite early product. If, however, the path to RNA was arduous, perhaps involving catalytic prebiotic peptides, for example, then in the context of the pre-LUCA time frame, the proto-ribosome may not have been so early.

At this juncture considerable attention has been focused on the mechanism of peptide synthesis. In addition, much of what can be inferred from ribosome structures and primary sequences of various components about the very beginnings has likely been uncovered. One especially interesting reminder is that one should perhaps pay attention to the nature of the pre-LUCA earth and thereby realize that ferrous iron may have preceded magnesium in stabilizing RNA [65]. The immediate goal that several groups are pursuing is to uncover a minimal RNA that incorporates the PTC and makes peptides. This model PTC would provide final proof that it is the RNA that is catalytic. Such an experimental system would also provide a starting point for studies to understand the origins and evolution of the modern PTC. It would be possible, for example, to examine how specificity for L or D amino acids is affected by changes in the sequence or structure of the model PTC. Less obvious but likewise important would be to synthesize a minimal segment based on helices h28, h44, and h45 in the small subunit in order to look for an inherent strong alternation between two structures. Again, if such an RNA were available, one could explore its properties with mutational studies.

With regard to the post PTC period, multiple attempts to establish a timeline of major events have been described [64, 70, 71, 92], culminating in the accretion model, which is based primarily on RNA insertion fingerprints and A-minor interactions [81]. As a result, events in the growth of the rRNAs can now be discussed at the level of individual helices or groups of helices.

This timeline of the relative age of various regions of the RNA can be correlated with existing knowledge to form an integrated understanding of ribosome history. This will be especially useful in refining the order of accretion of peripheral elements in the RNAs. Ongoing, for example, are efforts to correlate protein-binding sites as provided by atomic resolution studies with helix age. In the case of proteins with multiple domains, it may be possible to assign a relative age of each domain. Especially noteworthy is the obvious increasing complexity that is encountered in the later phases. Simply put, there is rapidly increasing parallelism as one reaches phases 5 and 6. Much of what is occurring in these and later phases will likely be found to relate to other cellular processes. This may ultimately allow a unified history of the cell.

References

1. Noller HF. Evolution of protein synthesis from an RNA World. Cold Spring Harb Perspect Biol. 2010;4:a003681.
2. Ban N, Nissen P, Hansen J, Moore PB, Steitz TA. The complete atomic structure of the large ribosomal subunit at 2.4 A resolution. Science. 2000;2000(289):905–20.

3. Nissen P, Hansen J, Ban N, Moore PB, Steitz TA. The structural basis of ribosome activity in peptide bond synthesis. Science. 2000;289:920–30.
4. Maguire BA, Beniaminov AD, Ramu H, Mankin AS, Zimmermann RA. A protein component at the heart of an RNA machine: the importance of protein l27 for the function of the bacterial ribosome. Mol Cell. 2005;20:427–35.
5. Selmer M, Dunham CM, Murphy FV 4th, Weixlbaumer A, Petry S, Kelley AC, Weir JR, Ramakrishnan V. Structure of the 70S ribosome complexed with mRNA and tRNA. Science. 2006;313:1935–42.
6. Wower IK, Wower J, Zimmermann RA. Ribosomal protein L27 participates in both 50 S subunit assembly and the peptidyl transferase reaction. J Biol Chem. 1998;273:19847–52.
7. Maracci C, Wohlgemuth I, Rodnina MV. Activities of the peptidtl transferase center of ribosomes lacking protein L27. RNA. 2015;21:1–6.
8. White HB III. Coenzymes as fossils of an earlier metabolic state. J Mol Evol. 1976;1976 (7):101–4.
9. Yarus M. Getting past the RNA World: The initial Darwinian Ancestor. Cold Spring Harb Perspect Biol. 2010;3:a003590.
10. Bartel DP, Szostak JW. Isolation of new ribozymes from a large pool of random sequences. Science. 1993;261:1411–8.
11. Zaher HS, Unrau PJ. Selection of an improved RNA polymerase ribozyme with superior extension and fidelity. RNA. 2007;13:1017–26.
12. Attwater J, Wochner A, Holliger P. In-ice evolution of RNA polymerase ribozyme activity. Nat Chem. 2013;5:1011–8.
13. Mutswchler H, Wochner A, Hollinger P. Freeze-thaw cycles as drivers of complex ribozyme assembly. Nat Chem. 2015;7:502–8.
14. Wolf YI, Koonin EV. On the origin of the translation system and the genetic code in the RNA world by means of natural selection, exaptation, and subfunctionalization. Biol Direct. 2007;2:14.
15. Zhang B, Cech TR. Peptide bond formation by *in vitro* selected ribozymes. Nature. 1997;1997(390):96–100.
16. Lilley DM. The origins of RNA catalysis in ribozymes. Trends Biochem Sci. 2003;28: 495–501.
17. Bowman JC, Hud NV, Williams LD. The ribosome challenge to the RNA World. J Mol Evol. 2015;80:143–61.
18. Cafferty BJ, Hud NV. Abiotic synthesis of RNA in water: a common goal of prebiotic chemistry and bottom-up synthetic biology. Curr Opin Chem Biol. 2014;22:146–57.
19. Sczepanski JT, Joyce GF. A cross-chiral RNA polymerase ribozyme. Nature. 2014;515: 440–2.
20. Van der Gulik P, Massar S, Gilis D, Buhrman H, Rooman M. The first peptides: the evolutionary transition between prebiotic amino acids and early proteins. J Theor Biol. 2009;261:531–9.
21. Ferris JP, Hill WR Jr, Rihe L, Orgel LE. Synthesis of long prebiotic oligomers on mineral surfaces. Nature. 1996;381:59–61.
22. Stachelhaus T, Marahiel MA. Modular structure of genes encoding multifunctional peptide synthetases required for non-ribosomal peptide synthesis. FEMS Microbiol Lett. 1995;125:3–14.
23. Marahiel MA. Working outside the protein-synthesis rules: insights into non-ribosomal peptide synthesis. J Peptide Sci. 2009;15:799–807.
24. Nagaswamy U, Fox GE. RNA ligation and the origin of tRNA. Orig Life Evol Biosph. 2003;33:199–209.
25. Agmon I. The dimeric proto-ribosome: Structural details and possible implications on the origin of life. Int J Mol Sci. 2009;2009(10):2921–34.
26. Davidovich C, Belousoff M, Wekselman I, Shapira T, Krupkin M, Zimmerman E, Bashan A, Yonath A. The proto-ribosome: an ancient nano-machine for peptide bond formation. Isr J Chem. 2010;50:29–35.

27. Krupkin M, Matzov D, Tang H, Metz M, Kalaora R, Belousoff MJ, Zimmerman E, Bashan A, Yonath A. A vestige of a prebiotic bonding machine is functioning within the contemporary ribosome. Philos Tran R Soc Lond B Biol Sci. 2011;366:2972–8.
28. Boer PH, Gray MW. Scambled ribosomal RNA gene pieces in *Chlamydomonas reinhardtii* mitochondrial DNA. Cell. 1988;55:399–411.
29. Schnare MN, Gray MW. Sixteen discrete RNA components in the cytoplasmic ribosome of *Euglena gracilis*. J Mol Biol. 1990;215:73–83.
30. Polacek N, Mankin AS. The ribosomal peptidyl transferase center: Structure, function, evolution, inhibition. Crit Rev Biochem Mol Biol. 2005;40:285–311.
31. Katunin VI, Muth GW, Strobel SA, Wintermeyer W, Rodnina MV. Important contribution to catalysis of peptide bond formation by a single ionizing group within the ribosome. Mol Cell. 2002;10:339–46.
32. Sievers A, Beringer M, Rodnina MV, Wolfenden R. The ribosome as an entropy trap. Proc Natl Acad Sci USA. 2004;101:7897–901.
33. Schroeder GK, Wolfenden R. The rate enhancement produced by the ribosome: an improved model. Biochemistry. 2007;46:4037–44.
34. Moore PB, Steitz TA. After the ribosome structures: How does peptidyl transferase work? RNA. 2001;9:155–9.
35. Wohlgemuth I, Beringer M, Rodnina MV. Rapid peptide bond formation on isolated 50S ribosomal subunits. EMBO reports. 2006;7:699–703.
36. Rodina MV, Beringer M, Wintermeyer W. How ribosomes make peptide bonds. Trends Biochem Sci. 2007;32:20–6.
37. Zaher HS, Shaw JJ, Strobel SA, Green R. The 2'-OH group of the peptidyl-tRNA stabilizes an active conformation of the ribosomal PTC. EMBO J. 2011;30:2445–53.
38. Carrasco N, Hiller DA, Strobel SA. Minimal transition state charge stabilization of the oxyanion during peptide bond formation by the ribosome. Biochemistry. 2011;50:10491–8.
39. Świderek K, Marti S, Tuñón I, Moliner V, Bertran J. Peptide bond formation mechanism catalyzed by ribosome. J Am Chem Soc. 2015;137:12024–34.
40. Fahnestock S, Rich A. Ribosome-catalyzed polyester formation Science. 1971;173:340–3.
41. Fahnestock S, Neumann H, Shashoua V, Rich A. Ribosome catalyzed ester formation. Biochemistry. 1970;9:2477–83.
42. Victorova LS, Kotusov VV, Azhaev AV, Krayevsky AA, Kukhanova MK, Gottikh BP. Synthesis of thioamide bond catalyzed by *E. coli* ribosomes. FEBS Lett. 1976;68:215–8.
43. Ohta A, Murakami H, Suga H. Polymerization of alpha-hydroxy acids by ribosomes. Chem Bio Chem. 2008;9:2773–8.
44. Subtelny AO, Hartman MC, Szostak JW. Ribosomal synthesis of n-methyl peptides. J Am Chem Soc. 2008;130:6131–6.
45. Tan Z, Forster AC, Blacklow SC, Cornish VW. Amino acid backbone specificity of the *Escherichia coli* translation machinery. J Am Chem Soc. 2004;126:12752–3.
46. Hartman MC, Josephson K, Lin CW, Szostak JW. An expanded set of amino acids analogs for the ribosomal translation of unnatural peptides. PLoS ONE. 2007;2:e972.
47. Kang TJ, Suga H. Ribosomal synthesis of nonstandard peptides. Biochem Cell Biol. 2008;86:92–9.
48. Zavialov AV, Mora L, Buckingham RH, Ehrenberg M. Release of peptide promoted by the GGQ motif of class 1 release factors regulates the GTPase activity of RF3. Mol Cell. 2002;10:789–98.
49. Fox GE, Tran Q. Yonath. An exit cavity was crucial to the polymerase activity of the early ribosome. Astrobiology. 2012;12:57–60.
50. Nakatogawa H, Ito K. The ribosomal exit tunnel functions as a discriminating gate. Cell. 2002;108:629–36.
51. Lu J, Deutsch C. Folding zones inside the ribosomal exit tunnel. Nat Struct Mol Biol. 2005;12:1123–9.

52. Yonath A. Antibiotics targeting ribosomes: resistance, selectivity, synergism and cellular regulation. Annu. Rev. Biochem. 2005;74:649–79.
53. Ito K, Chiba S, Pogliano K. Divergent stalling sequences sense and control cellular physiology. Biochem Biophys Res Commun. 2010;393:1–5.
54. Ranu H, Mankin A, Vazquez-Laslop N. Programmed drug-dependent ribosome stalling. Mol Microbiol. 2009;71:811–24.
55. Ramu H, Vazquez-Laslop N, Klepacki D, Dai Q, Piccrilli J, Micura R, Mankin AS. Nascent peptide in the ribosome exit tunnel affects functional properties of the A-site of the peptidyl transferase center. Mol Cell. 2011;41:321–30.
56. Chiba S, Kanamori T, Ueda T, Akiyama Y, Pogliano K, Ito K. Recruitment of a species-specific translational arrest module to monitor different cellular processes. Proc Natl Acad Sci USA. 2011;108:6073–8.
57. Turk RM, Illangasekare M, Yarus M. Catalyzed and spontaneous reactions on ribozyme ribose. J Am Chem Soc. 2011;133:6044–50.
58. Schaul S, Berel D, Benjamini Y, Graur D. Revisiting the operational code for amino acids: Ensemble attributes and their implications. RNA. 2010;16:141–53.
59. Schimmel P, Giege R, Moras D, Yokoyama S. An operational RNA code for amino acids and possible relationship to genetic code. Proc Natl Acad Sci USA 1993;90:8763–8.
60. Schimmel P. Ribas de Pouplana L. Transfer RNA: From Minihelix to Genetic Code. Cell. 1995;81:983–6.
61. Francklyn C, Schimmel P. Aminoacylation of RNA minihelices with alanine. Nature. 1989;337:478–81.
62. Francklyn C, Schimmel P. Enzymatic aminoacylation of an eight base pair mivrohelix with histidine. Proc Natl Acad Sci USA. 1990;87:8655–9.
63. Kleine DJ, Moore PB, Steitz TA. The contribution of metal ions to the structural stability of the large ribosomal subunit. RNA. 2004;10:1366–79.
64. Hsiao C, Mohan S, Kalahar BK, Williams LD. Peeling the Onion: ribosomes are ancient molecular fossils. Mol Biol Evol. 2009;26:2415–25.
65. Hsiao C, Chou IC, Okafor CD, Bowman JC, O'Neill EB, Athavale SS, Petrov AS, Hud NV, Wartell RM, Harvey SC, Williams LD. RNA with iron (II) as a cofactor catalyzes electron transfer. Nat Chem. 2013;5:525–8.
66. Huang L, Krupkin M, Bashan A, Yonath A, Massa L. Protoribosome by quantum kernel energy method. Proc Natl Acad Sci USA. 2013;110:14900–5.
67. Rivas M, Tran Q, Fox GE. Nanometer scale pores similar in size to the entrance of the ribosomal exit cavity are a common feature of large RNAs. RNA. 2013;19:1349–54.
68. Mears JA, Cannone JJ, Stagg SM, Gutell RR, Agrawal RK, Harvey SC. Modeling a minimal ribosome based on comparative sequence analysis. J Mol Biol. 2002;321:215–34.
69. Gray MW, Schnare MN. Evolution of RNA gene organization. In: Zimmermann RA, Dahlberg AE, editors. *Ribosomal RNA—Structure, evolution, processing, and function in protein synthesis*. Boca Raton, Florida: CRC Press; 1996. p. 49–69.
70. Hury J, Nagaswamy U, Larios-Sanz M, Fox GE. Ribosome origins: The relative age of 23S rRNA domains. Orig Life Evol Biosphere. 2006;36:421–9.
71. Bokov K, Steinberg SV. A hierarchical model for evolution of 23S ribosomal RNA. Nature. 2009;457:977–80.
72. Nissen P, Lppolito JA, Ban N, Moore PB, Steitz TA. RNA tertiary interactions in the large ribosomal subunit: the A-minor motif. Proc Natl Acad Sci USA. 2001;98:4899–903.
73. Réblová K, Šponer JE, Špačková N, Beššeová I, Šponer J. A-minor tertiary interactions in RNA kink-turns. Molecular dynamics and quantum chemical analysis. J Phys Chem B. 2011;115:13897–910.
74. Gerbi SA. Expansion segments: Regions of variable size that interrupt the universal core secondary structure of ribosomal RNA. In: Zimmermann RA, Dahlberg AE, editors. *Ribosomal RNA—Structure, evolution, processing, and function in protein synthesis*. Boca Raton, Florida: CRC Press; 1996. p. 71–87.

75. Gerbi SA. Evolution of ribosomal DNA. In: MacIntyee RJ, editor. Molecular evolutionary genetics. New York: Plenum Publishing Corporation; 1985. p. 419–51774.
76. Luehrsen KR, Nicholson DE, Eubanks DC, Fox GE. An archaebacterial 5S rRNA contains a long insertion sequence. Nature. 1981;293:755–7.
77. Petrov AS, Bernier CR, Hsiao C, Norris AM, Kovacs NA, Waterbury CC, Stepanov VG, Harvey SC, Fox GE, Wartell RM, Hud NV, Williams LD. Evolution of the ribosome at atomic resolution. Proc Natl Acad Sci USA. 2014;111:10251–6.
78. Caisova L, Melkonian M. Evolution of helix formation in the ribosome internal transcribed spacer 2 (ITS2) and its significance for RNA secondary structures. J Mol Evol. 2014;78: 324–7.
79. Petrov AS, Bernier CR, Gulen B, Waterbury CC, Hershkovits E, Hsiao C, Harvey SC, Hud NV, Fox GE, Wartell RM, Williams LD. Secondary structures of tRNAs from all three domains of life. PLoS One. 2014;9:e88222.
80. Ban N, Beckmann R, Cate JH, Dinman JD, Dragon F, Ellis SR, Lafontaine DL, Lindahl L, Liljas A, Lipton JM, McAlear MA, Moore PB, Noller HF, Ortega J, Panse VG, Ramakrishnan V, Spahn CM, Steitz TA, Tchorzewski M, Tollervey D, Warren AJ, Williamson JR, Wilson D, Yonath A, Yusupov M. A new system for naming ribosomal proteins. Curr Opin Struct Biol. 2014;24:165–9.
81. Petrov AS, Gulen B, Norris A, Kovacs NA, Bernier CR, Lanier KA, Fox GE, Harvey SC, Wartell RM, Hud NV, Williams LD. The history of the ribosome and the origin of translation. Proc. Natl. Acad Sci USA. 2015;112:15396–401.
82. Caetano-Anollés G. Tracing the evolution of RNA structure in ribosomes. Nucl Acids Res. 2002;30:2575–87.
83. Harish A, Caetano-Anollés G. Ribosomal history reveals origins of modern protein synthesis. PLoS One. 2012;7:e32776.
84. Petrov AS, Williams LD. The ancient heart of the ribosomal large subunit: a response to Caetano-Anollés. J Mol Evol. 2015;80:166–70.
85. Caetano-Anollés D, Caetano-Anollés G. Ribosomal accretion aprioism and the phylogenetic method: A response to Petrov and Williams. Front Genet. 2015;6:194.
86. Caetano-Anollés G. Ancestral insertions and expansions of rRNA do not support an origin of the ribosome in its peptidyl transferase center. J Mol Evol. 2015;80:162–5.
87. Smith TF, Gutell R, Lee J, Hartmann H. The origin and evolution of the ribosome. Biol Direct. 2008;3:16.
88. Campbell J. An RNA repisome as the ancestor of the ribosome. J Mol Evol. 1991;32:3–5.
89. Gordon K. Were RNA replication and translation directly coupled in the RNA (+protein) World? J Theor Biol. 1995;173:179–93.
90. Poole AM, Jeffares DC, Penny D. The path from the RNA world. J Mol Evol. 1998;46:1–17.
91. Blackmond DG. The origin of biological homochirality. Cold Spring Harb Perspect Biol. 2010;2:a002147.
92. Fox GE. Origin and evolution of the ribosome. Cold Spring Harb Perspect Biol. 2010;3: a003483.
93. Yamane T, Miller DL, Hopfield JJ. Discrimination between D and L-tyrosyl transfer ribonucleic acids in peptide chain elongation. Biochemistry. 1981;20:7059–68.
94. Heckler TG, Roesser JR, Xu C, Chang PI, Hecht SM. Ribosomal binding and dipeptide formation by misacylated $tRNA^{Phe's}$. Biochemistry. 1988;27:7254–62.
95. Starck SR, Qi X, Olsen BN, Roberts RW. The puromycin route to assess stero- and regiochemical constraints on peptide bond formation in eukaryotic ribosomes. J Am Chem Soc. 2003;125:8090–1.
96. Dedkova LM, Fahmi NE, Golovine SY, Hecht SM. Enhanced D-amino acid incorporation into proteins by modified ribosomes. J Am Chem Soc. 2003;125:6616–7.
97. Dedkova LM, Fahmi NE, Golovine SY, Hecht SM. Construction of modified ribosomes for incorporation of D-amino acids into proteins. Biochemistry. 2006;45:15541–51.

98. Dedkova LM, Fahmi NE, Paul R, del Rosario M, Zhang L, Chen S, Feder G, Hecht SM. β-Puromycin selection of modified ribosomes for *in vitro* incorporation of β-amino acids. Biochemistry. 2012;51:401–15.
99. Soutourina J, Plateau P, Blanquet S. Metabolism of D-aminoacyl-tRNAs in *Eschrichia coli* and *Saccharomyces cerevisiae* cells. J Biol Chem. 2000;275:32535–42.
100. Yang H, Zheng G, Peng X, Qiang B, Yuan J. D-amino acids and D-Tyr-tRNA^{tyr} deacylase: Stereospecificity of the translation machine revisited. FEBS Lett. 2008;552:95–8.
101. Di Giulio M. A comparison among the models proposed to explain the origin of the tRNA molecule: A synthesis. J Mol Evol. 2009;69:1–9.
102. Maizels N, Weiner AM. The genomic tag hypothesis: modern viruses as molecular fossils of ancient strategies for genomic replication. In: Gesteland RF. Atkins JF editors. The RNA world, Plainview. New York: Cold Springs Harbor Laboratory Press; 1993. p. 577–602.
103. Maizels N, Weiner AM. Phylogeny from function: evidence from the molecular fossil record that tRNA originated in replication, not translation. Proc Natl Acad Sci USA. 1994;91: 6729–34.
104. Di Giulio M. On the origin of the transfer RNA molecule. J Theor Biol. 1992;159:199–214.
105. Watanabe Y, Suematsu T, Ohtsuki T. Losing the stem-loop structure From metazoan mitochondrial tRNAs and co-evolution of interacting factors. Front Genet. 2014;5:109.
106. Randau L, Munch R, Hohn M, Jahn D, Soll D. *Nanoarchaeum equitans* creates functional tRNAs from separate genes for their50- and 30-halves. Nature. 2005;433:537–41.
107. Carter CW. What RNA world? Why a peptide/RNA partnership merits renewed experimental attention. Life (Basel). 2015;5:294–320.
108. Smith TF, Hartman H. The evolution of class II aminoacyl-tRNA synthetases and the first code. FEBS Lett. 2015;589:3499–507.
109. Shine J, Dalgarno L. The 3'-terminal sequence of Escherichia coli 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc Natl Acad Sci USA. 1974;71:1342–6.
110. Bernhardt HS, Tate WP. The transition from noncoded to coded protein synthesis: did coding mRNAs arise from stability-enhancing binding partners to tRNA? Biol Direct. 2012;7:4.
111. Barrell BG, Anderson S, Bankier AT, de Bruijn MH, Chen E, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJ, Staden R, Young IG. Different pattern of codon recognition by mammalian mitochondrial tRNAs. Proc Natl Acad Sci USA. 1980;77:3164–6.
112. Temperley R. Richter R, DEnnerlein S, Lightowlers, Chrzanowska-Lightowlers Z. Hungry codons promote frameshifting in human mitcchondrial ribosomes. Science. 2010;327:301.
113. Lehman N, Jukes TH. Genetic code development by stop codon takeover. J Theor Biol. 1988;135:203–14.
114. Fournier GP, Neumann JE, Gogarten JP. Inferring the ancient history of the translation machinery and genetic code via recapitulation of ribosomal subunit assembly orders. PLoS One. 2010;27:1792–801.
115. Wang J, Dasgupta I, Fox GE. Many nonuniversal archaeal ribosomal proteins are found in conserved gene clusters. Archaea. 2009;2:241–51.
116. Fox GE, Naik AK. The evolutionary history of the translation machinery. In Ribas, de Pouplana L, editor), The genetic code and the origin of genetic code. New York, New York: Springer. US. 2004;2007:92–105.
117. Khaitovich P, Mankin AS, Green R, Lancaster L, Noller HF. Characterization of functionally active subribosomal particles from Thermus aquaticus. Proc Natl Acad Sci USA. 1999;96:85–90.
118. Siefert JL, Fox GE. Conserved gene clusters in bacterial genomes provide further support for the primacy of RNA. J Mol Evol. 1997;45:467–72.
119. Joyce GF. RNA evolution and the origins of life. Nature. 1989;338:217–24.
120. Murzin AG, Brenner SE, Hubbard T, Chothia C. SCOP: a structural classification of proteins database for the investigation of sequences and structures. J Mol Biol. 1995;247:536–40.

121. Hartman H, Smith TF. The evolution of the ribosome and the genetic code. Life (Basel). 2014;4:227–49.
122. Vishwanath P, Favaretto P, Hartman H, Mohr SC, Smith TF. Ribosomal protein-sequence block structure suggests complex prokaryotic evolution with implications for the origin of eukaryotes. Mol Phylogenet Evol. 2004;33:615–25.
123. Agrawal V, Kishan RK. Functional evolution of two subtly different (similar) folds. BMC Struct Biol. 2001;1:5.
124. Cannone JJ, Subramanian S, Schnare MN, Collett JR, D'Souza LM, Du Y, Feng B, Lin N, Madabusi LV, Muller KM, Pande N, Shang Z, Yu N, Gutell RR. The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. BMC Bioinformatics. 2002;3:2.
125. Klein DJ, Moore PB, Steitz TA. The roles of ribosomal proteins in the structure assembly, and evolution of the large ribosomal subunit. J Mol Biol. 2004;340:141–77.
126. Spirin AS. Ribosome as a molecular machine. FEBS Lett. 2002;514:2–10.
127. Gavrilova LP, Kostiashkina VE, Koteliansky NM, Rutkevich NM, Spirin AS. Factor free ("Non-enzymatic") and factor-dependent systems of translation of polyuridylic acid by *Escherichia coli* ribosomes. Mol Biol. 1976;101:537–52.
128. Frank J, Gonzalez RL Jr. Structure and Dynamics of a processive Brownian motor: the translating ribosome. Annu Rev Biochem. 2010;2010(79):381–412.
129. Woese C. Molecular mechanics of translation: a reciprocating ratchet mechanism. Nature. 1970;226:817–20.
130. Robertus JD, Ladner JE, Finch JT, Rhodes D, Brown RS, Clark BF, Klug A. Structure of yeast phenylalanine tRNA at 3S resolution. Nature. 1974;250:546–51.
131. Harvey SC, McCammon JA. Intramolecular flexibility in phenylalanine transfer RNA. Nature. 1981;294:286–7.
132. Dunkle JA, Wang L, Feldman MB, Pulk A, Chen VB, Kapral GJ, Noeske J, Richardson JS, Blanchard SC. Cate JH. Structures of the bacterial ribosome in classical and hybrid states of tRNA binding. Science. 2011;332:981–4.
133. Paci M, Fox GE. Major centers of motion in the large ribosomal RNAs. Nucl Acids Res. 2015;43:4640–9.
134. Paci M, Fox GE. Centers of motion associated with EF-Tu binding to the ribosome. RNA Biol. 2016;Jan 19.0 [Epub ahead of print].
135. Moazed D, Robertson JM, Noller HF. Interaction of elongation factors EF-G and EF-Tu with a conserved loop in 23S rRNA. Nature. 1998;334:362–4.
136. Sergiev PV, Bogdanov AA, Dontsova OA. How can elongation factors EF-G and EF-Tu discriminate the functional state of the ribosome using the same binding site? FEBS Lett. 2005;579:5439–42.
137. Fox GE, Paci M, Tran Q, Petrov, Williams LD. Ribosome dynamics and the Evolutionary history of ribosomes. Proceedings SPIE Conference 9606 Instruments, Methods and Missions for Astrobiology XVII 2015: 96060G1-96060G-5.
138. Hoffmann A, Bukau B, Kramer G. Structure and function of the molecular chaperone trigger factor. Biochim Biophys Acta. 2010;1803:650–61.
139. Frank J, Gao H, Sengupta J, Gao N, Taylor DJ. The process of mRNA-tRNA translocation. Proc Natl Acad Sci USA. 2007;104:19671–8.
140. Grela P, Bernadó P, Svergun D, Kwiatowski J, Abramczyk D, Grankowski N, Tchórzewski M. Structural relationships among the ribosomal stalk proteins from the three domains of life. J Mol Evol. 2008;67:154–67.
141. Szymanski M, Barciszewska MZ, Erdmann VA, Barciszewski J. 5 S rRNA: structure and interactions. Biochem J. 2003;371:641–51.
142. Gongadz GM. 5S rRNA and Ribosome. Biochemistry (Moscow). 2011;76:1450–64.
143. Dohme F, Nierhaus KH. Role of 5S RNA in assembly and function of the 50S subunit from Escherichia coli. Proc Natl Acad Sci USA. 1976;73:2221–5.
144. Erdmann VA, Fahnestock S, Higo K, Nomura M. Role of 5S RNA in the function of 50S ribosomal subunits. Proc Natl Acad Sci USA. 1971;68:2932–6.

145. Forget BG, Weissman SM. Nucleotide sequence of KB cell 5S RNA. Science. 1967;158:1695–9.
146. Dontsova OA, Dinman JD. 5S rRNA: Structure and Function from Head to Toe. Int J Biomed Sci. 2005;1:1–7.
147. Bogdanov AA, Dontsova OA, Dokudovskaya SS, Lavrik IN. Structure and function of 5S rRNA in the ribosome. Biochem Cell Biol. 1995;73:869–76.
148. Dokudovskaya S, Dontsova O, Shpanchenko O, Bogdanov A, Brimacombe R. Loop IV of 5S ribosomal RNA has contacts both to domain II and to domain V of the 23S RNA. RNA. 1996;2:146–52.
149. Stark H, Rodnina MV, Wieden HJ, Zemlin F, Wintermeyer W, van Heel M. Ribosome interactions of aminoacyl-tRNA and elongation factor Tu in the codon-recognition complex. Nat. Struct Biol. 2002;9:849–54.
150. Hoang L, Fredrick K, Noller HF. Creating ribosomes with an all-RNA 30S subunit P site. Proc Natl Acad Sci USA. 2004;101:12439–43.
151. Fox GE, Woese CR. 5S rRNA secondary structure. Nature. 1975;256:505–7.
152. Fox GE, Woese CR. The architecture of 5S rRNA and its relation to function. J Mol Evol. 1975;6:61–76.
153. Nishikawa K, Takemura S. Nucleotide sequence of 5 S RNA from *Torulopsis utilis*. FEBS Lett. 1974;40:106–9.
154. Luehrsen KR, Fox GE. The secondary structure of eucaryotic cytoplasmic 5S ribosomal RNA. Proc Natl Acad Sci USA. 1981;78:2150–4.
155. Fox GE. The structure and evolution of archaebacterial RNA. In: Gunzalus IC, editor. The Bacteria, vol 8. New York: Academic Press; 1985. p. 267–310.
156. Correll CC, Freeborn B, Moore PB, Steitz TA. Metals, motif, and recognition in the crystal structure of a 5S rRNA domain. Cell. 1997;91:705–12.
157. Wrede P, Erdmann VA. Activities of *B. stearothermophilus* 50S ribosomes reconstituted with prokaryotic and eukaryotic 5S RNA. FEBS Lett. 1975;33:315–9.
158. Bellemare G, Vigne R, Jordan B. Interaction between *Escherichia coli* ribosomal proteins and 5S RNA molecules: recognition of prokaryotic 5S RNAs and rejection of eukaryotic 5S RNAs. Biochimie. 1973;55:29–35.
159. Valach M, Burger G, Gray MW, Lang BF. Widespread occurrence of organelle genome-encoded 5S rRNAs including permuted molecules. Nucleic Acids Res. 2014;42: 13764–77.
160. Greber BJ, Boehringer D, Leitner A, Bieri P, Voigts-Hoffmann F, Erzberger JP, Leibundgut M, Aebersold R, Ban N. Architecture of the large subunit of the mammalian mitochondrial ribosome. Nature. 2014;515:283–6.
161. Brown A, Amunts A, Bai XC, Sugimoto Y, Edwards PC, Murshudov G, Scheres SH, Ramakrishnan V. Structure of the large ribosomal subunit from human mitochondria. Science. 2014;346:718–22.
162. Amunts A, Brown A, Bai XC, Llácer JL, Hussain T, Emsley P, Long F, Murshudov G, Sjors Scheres SHW, Ramakrishnan V. Structure of the yeast mitochondrial large ribosomal subunit. Science. 2014;2014(43):1485–9.

Evolution of Translational Initiation: From Archaea to Eukarya

Dario Benelli, Anna La Teana and Paola Londei

1 Translation Initiation: An Evolutionary Overview

Because of its obvious importance for cell survival, translation is perhaps the most conserved of cellular mechanisms in evolution. For this reason, translation has long been regarded as a rather dull, if essential, process having only secondary importance in regulating gene expression. However, in relatively recent years, and especially since the discovery of miRNAs, translational control has been recognized as being of central importance in a wide range of cellular processes, from differentiation to tumoral transformation. It has become evident that decoding of an mRNA is a complex and finely tuned task, whose unbalance may have very far-reaching consequences for the fate of both the individual cell and the organism as a whole [30, 39, 45, 47, 49].

Albeit translation may be subjected to regulation at any stage, the main target of the principal control mechanisms is the initiation phase, during which the ribosome interacts with the mRNA and recognizes the correct starting point for decoding. In all organism, this task requires the small ribosomal subunit, a special initiator tRNA (carrying methionine in Eukarya and Archaea and formylmethionine in Bacteria), and must be aided by a set of specific proteins called translation initiation factors (TIF). Despite these common themes, initiation is the least conserved in evolution of all steps of translation. Indeed, the mechanism and molecular machinery for initiation have diverged extensively in the three primary domains of life.

D. Benelli · P. Londei (✉)
Istituto Pasteur-Fondazione Cenci- Bolognetti and Dpt. Biotecnologie Cellulari ed Ematologia, Policlinico Umberto I, University of Rome Sapienza,
Viale Regina Elena 324, 00161 Rome, Italy
e-mail: londei@bce.uniroma1.it

A. La Teana
Dpt. Scienze Della Vita e dell'Ambiente, Università Politecnica delle Marche,
Via Brecce Bianche, 60131 Ancona, Italy

G. Hernández and R. Jagus (eds.), *Evolution of the Protein Synthesis Machinery and Its Regulation*, DOI 10.1007/978-3-319-39468-8_4

In the Bacteria, the initiation mechanism has been deciphered in fine detail and is known to be relatively simple, economic but very efficient. The small ribosomal subunit interacts directly with the mRNA, often using the Shine-Dalgarno (SD) interaction with the 3′ end of the 16S rRNA, although this is not an obligatory requirement. A minimal set of only three protein factors, IF1, IF2 and IF3, assists the initiation process [21], performing the really essential tasks: ensuring that the right initiation codon is identified and that the fmet-tRNAi is correctly located in the P site as well as preventing the premature joining of the 50S subunit to the initiation complex.

In the Eukarya, the translation initiation process is also known in fairly good detail and is far more complicated than in Bacteria. There is a general consensus that the eukaryotic small ribosomal subunit cannot interact directly with the mRNA, except for some special cases. Rather, mRNA/ribosome interaction and the selection of the correct initiation codon are believed to take place by a "scanning" mechanism, whereby the 40S subunit binds at, or near, the capped 5′ end of the mRNA and moves in a 3′ direction until the initiator AUG—usually the first available one—is found [28]. This process requires the assistance of a host of protein factors, several of which are multisubunit proteins, and is subjected to a complex regulation, which may entail the modification of certain IFs by phosphorylation [25, 26].

As to the Archaea, in silico studies on complete genomic sequences have revealed that genes are generally organized in operons. The initiation codons may be preceded by SD motifs, although this is by no means a general rule [29]; indeed, in some archaeal species, only a minority of cistrons is preceded by SD motifs. Characteristically, the Archaea possess a significant proportion of mRNAs lacking 5′ untranslated leader sequences [46, 51]. In certain archaeal species, a majority of mRNAs has this kind of "leaderless structure." Translation initiation on leaderless mRNAs dispenses with an SD/anti-SD interaction and takes place with a different mechanism, somewhat reminiscent of the eukaryotic one [4, 11, 51].

In silico genomic analysis has also revealed that the archaea are endowed with a set of putative translation initiation factors more complex than the bacterial one and mostly homologous to eukaryotic proteins. This discovery disproved the simplistic assumption that translation initiation is intrinsically simpler in prokaryotic cells and raised new questions about the evolution of this step of protein synthesis. While a few initiation factors are shared by the three domains of life, some of the complexity observed in present-day eukaryotic cells originated already during the common evolutionary path shared by the Eukarya and the Archaea [3]. This is reflected in the larger-than-bacterial number of archaeal initiation factors, as well in the presence of some factors common to the Archaea and the Eukarya but not found in Bacteria. The archaeal/eukaryal (a/e) factors are the proteins known as eIF2-a/eIF2 and eIF6-aIF6. They are especially interesting, since it is known that they are at the core of important translational control pathways in eukaryotes, and they will be described in detail later in this chapter.

2 mRNA Features in the Three Domains of Life

Translation initiation is about finding the right place on the mRNA to begin to decode the genetic message. Albeit obvious, this is far from being trivial, and the problem must have been dealt with at the very start of the evolution of translation. There are reasons to believe that primitive mRNAs were read simply beginning from the first nucleotide at the 5′ end, dispensing with the untranslated leader sequences (UTRs) prevalent in modern messages. Such "leaderless" mRNAs are abundant in Archaea, not uncommon in bacteria and rare only in eukaryotes. Nevertheless, it has been demonstrated that a leaderless mRNA, regardless of source, is universally readable by all types of modern ribosomes [19], lending support to the idea that these were the ancient, universal kind of messenger RNAs.

Modern mRNAs, of course, have acquired different kinds of specializations. One major difference between the eukaryotes and both prokaryotic domains is the presence of the nucleus, wherein eukaryotic mRNAs are synthesized and maturated. This creates a temporal and spatial gap between transcription and translation, which does not exist in prokaryotic organisms. The reason for the creation of a specific compartment for RNA transcription and maturation probably derives from the particular structure of eukaryotic genes, namely the customary presence of the introns that make the primary products of transcription unreadable as such.

The major structural differences between mature eukaryotic and prokaryotic mRNAs are, as is well known, the presence in the latter of several cistrons in a row (polycistronic mRNAs) and of a specific ribosome recognition sequence (the Shine-Dalgarno sequence). Both features may be absent in prokaryotic mRNAs, but are nevertheless quite distinctive features of this class of messages. Characteristically, archaea often harbor leaderless polycistronic mRNAs in which the first gene is leaderless and the subsequent ones may be preceded by a SD sequence for ribosome recognition [11, 40].

It is not obvious why the Eukarya should not have SD sequences or polycistronic mRNAs, but the reason is probably to be found, once again, in the process of mRNA maturation unique to this domain of life. In prokaryotes, transcription and translation are contemporary, and the ribosomes begin translation as soon as the 5′ terminus of the mRNA is available. In this scenario, a ribosome may continue translation decoding several cistrons in a row, or new ribosomes may read the internal cistrons binding to their translation initiation regions as soon as they appear on the transcript.

On the contrary, in the Eukarya, mRNAs can only be translated after their maturation is complete and they are shipped from the nucleus to the cytoplasm, tightly packed with protein. The ribosomes are then confronted with the task to find the place to begin translation; in order to do this, the mRNA must be unpacked and its sequence inspected. The uncovering of the translation initiation region is performed by a set of translation factors unique to eukaryotes. To unpack the mRNA, these factors take advantage of a modification unique to eukaryotic mRNAs, the 5′ cap structure, which is recognized and bound by a protein complex called eIF4F. In

addition to the cap-binding function, this complex also includes a helicase activity for unwinding the mRNA and a docking site for the small ribosomal subunit. It is believed that the 40S subunit "rides" the unwinding complex scanning the sequence in search of the initiation codon. To this end, it carries along the initiator tRNA, whose anticodon must pair with the AUG initiator codon. Once this happens, scanning is arrested and translation may begin [26]. This is the basic mechanism for translation initiation in eukaryotes, although many variations on the theme exist.

A crucial point in initiation is the way in which the interaction between the ribosome and the mRNA is established. There are several solutions to this problem, some direct and some indirect. Except for the case of leaderless mRNAs, the ribosome-binding region (also called translation initiation region, or TIR) is more or less removed from the 5′ end of the message. In prokaryotic mRNAs, the TIR often contains the SD sequence, to which the 30S ribosomal subunits can bind directly, thereby framing the initiation codon in the P site. Eukaryotic mRNAs, as said above, mark the mRNA 5′ end with the cap, which allows the binding of the protein complex and ribosome scanning. The interaction between the ribosome and the initiation codon is achieved through the mediation of met-tRNAi.

Besides its function in protein synthesis initiation, it is well known that the cap has an important function in protecting the mRNAs against 5′ end degradation. This protection is particularly important in eukaryotes, where mRNAs undergo a long and complex maturation process before being released in the cytoplasm for translation. Indeed, it is very likely that the protective function of the cap is the ancestral one, while its role as a facilitator of translation initiation evolved later.

Since prokaryotic mRNAs are generally shorter-lived than eukaryotic ones, one would think that the prokaryotic domains of life may dispense with evolving special devices for mRNA protection. In fact, in Bacteria mRNA capping does not exist. A true capping is also absent in Archaea. Interestingly, however, the Archaea seem to have evolved a device similar to eukaryotic capping. They possess a protein factor that is capable of recognizing specifically the triphosphate 5′ end of an mRNA, thereby protecting it from degradation until translation begins [24]. The archaeal mRNA-binding factor is quite unlike the eukaryal cap-binding complex, being instead the homolog of another eukaryotic factor, the so-called eIF2 (see below).

As most Archaea live in extreme environments, protecting the mRNAs against 5′-end degradation may be useful under stress conditions, when translation is more or less shut off. Ensuring a reservoir of mRNAs ready to resume translation when conditions revert to normal may give a selective advantage to archaeal cells.

The presence of phosphate on the mRNA 5′ end may have a cap-like function also in the sense that it is recognized by the ribosomes binding to a leaderless mRNA [17]. The mechanism of leaderless initiation is imperfectly understood. In Archaea, the terminal AUG initiator codon may be recognized by a 30S subunit carrying a tRNAi [4], while in bacteria a mechanisms implicating void, preformed 70S ribosomes has been described [34].

In summary, as regards mRNA structure, Archaea and Bacteria have several common features, absent in Eukarya, notably a polycistronic organization and specific ribosome-binding sequences ahead of the initiation codons. However, it

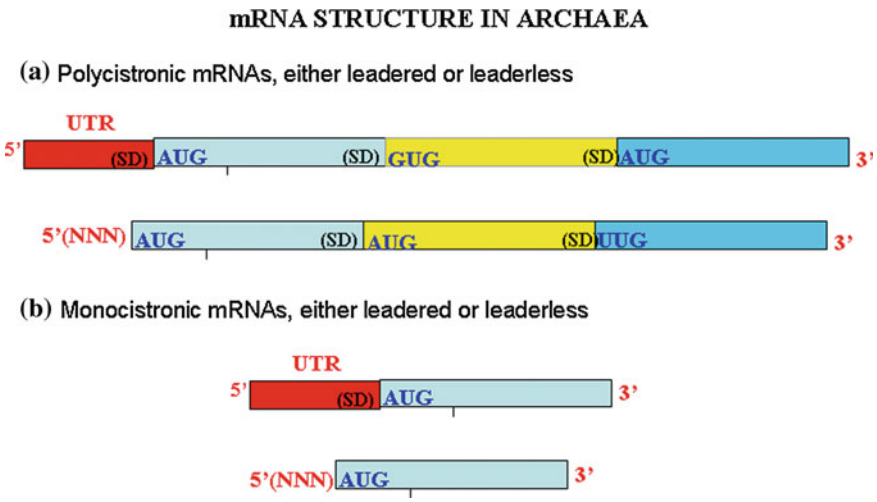


Fig. 1 mRNA structure in Archaea. **a** Leadered and leaderless polycistronic mRNAs; **b** leadered and leaderless monocistronic mRNAs

would be reductive to state that archaeal mRNAs are bacteria-like, since they very often are leaderless and make scarcer use of the SD sequence for ribosome-mRNA interaction. The presence of a system for 5′-end protection of their mRNAs, which is reminiscent of eukaryotic capping and may be its evolutionary precursor, is another notable unique characteristic of archaeal mRNAs.

A summary of the mRNA structures present in Archaea is presented in Fig. 1.

3 Translation Initiation Factors: Conservation and Divergence in the Three Domains of Life

Protein synthesis initiation, as every other complex cellular process, requires the assistance of numerous protein factors that are not structural components of the main translating "enzyme," the ribosome. The principal functions performed by these factors, in all domains of life, include ribosome/mRNA recognition, delivery of initiator tRNA (tRNAi) to the ribosome and proofreading of the initiation complex. However, the number and structure of the translation initiation factors (TIF) vary widely in the three domains of life. Eukarya have the largest number of TIFs, which are quite often also large, multimeric proteins. Bacteria are the most streamlined, having just three TIFs, all of them monomeric, relatively small proteins. Archaea are somewhere in between: they have a larger number of TIFs than bacteria, but not so many as the eukarya. Interestingly, the Archaea and the Eukarya share a number of TIFs not possessed by the Bacteria (Fig. 2). Some of these archaeal-eukaryal proteins have still ill-defined functions [31].

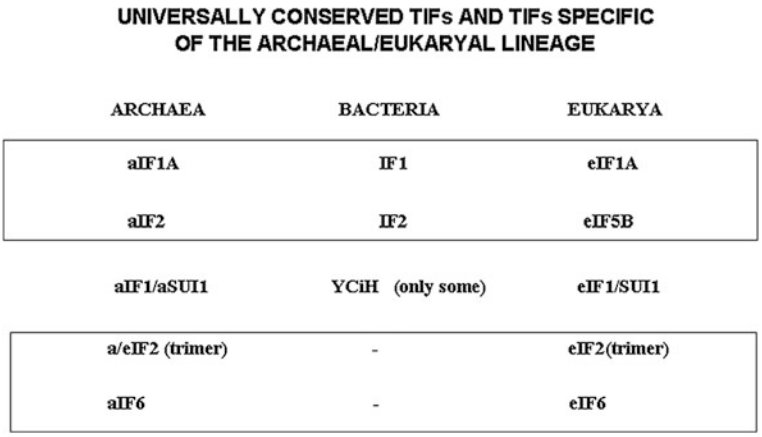


Fig. 2 Universally conserved and archaea/eukarya specific translation initiation factors

3.1 Universally Conserved TIFs

There is a small set of two "core" TIFs that are universally conserved in the three domains of life (Fig. 2). Depending on whether they belong to the bacterial or to the archaeal/eukaryal lineage, these proteins are termed IF1/IF1A and IF2/IF5B. The fact that the archaeal and eukaryal factors have the same name is not casual, but means that they resemble each other in primary sequence while the bacterial protein stands apart. A third universal factor, EFP/IF5A, has recently turned out to be actually an elongation factor and will not be taken into consideration here.

Remarkably, the universally conserved factors are not so conserved when it comes to their function. At least, IF/IF1As are all thought to have the main role of occupying the ribosomal A site, thereby helping the unambiguous placement of the tRNAi in the P site [7, 9]. But IF2/IF5Bs tell another story. All versions of this protein have a similar, curious shape, resembling a calyx with a stem and a basement (Fig. 3; [38]), but there most of the resemblance ends. Bacterial IF2 is a truly central player in initiation: it interacts specifically with fmet-tRNAi, directing it to the ribosomal P site [6]. Apparently, the presence of the formyl group on the methionine is instrumental for the recognition of the tRNAi by IF2 [22], which may be the reason why this specific modification is only found in Bacteria.

By contrast, archaeal/eukaryal IF5Bs do not bind tRNAi, albeit they may stimulate its binding in the P site [10, 32]. Apparently, the main function of IF5B is to promote the joining of the ribosomal subunits to form the monomeric ribosome in a late stage of initiation [36]. This function has been clearly demonstrated in eukaryotes, while some doubts remain as to whether it also exists in archaea. In both archaea and eukarya, on the other hand, the tRNAi-binding function has been allotted to another factor, a/eIF2, which probably arose after the splitting of the bacterial branch from the universal tree of life.

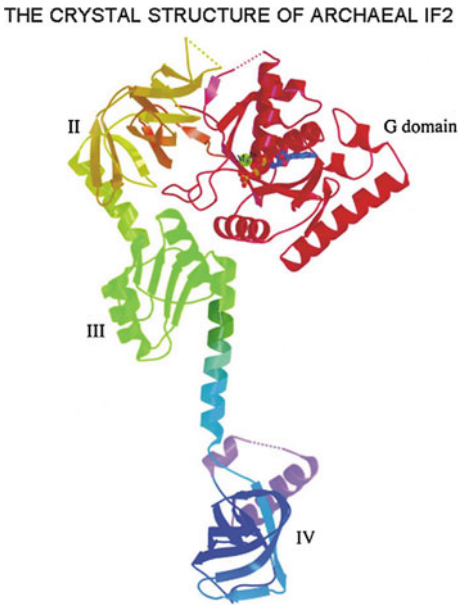


Fig. 3 The crystal structure of the universal factor IF2/5B. The protein is composed of four domains, one of which contains the guanine-nucleotide-binding pocket

In spite of its functional divergence, there is little doubt that IF2/IF5B is a very old protein, probably existing already, along with IF1/IF1A, in the last universal common ancestor of extant cells. Why it came to play different roles in bacteria and in eukarya/archaea is not clear, although a number of hypotheses can be advanced, as we shall see later in this chapter.

Finally, there is a third semi-universal factor, in the sense that it is found in all Archaea and Eukarya (where it is called IF1 or SUI1), but only in certain bacterial species (where it is called YciH). In Archaea and Eukarya, this is an essential protein that helps the correct identification of the initiation codon, while in Bacteria its function is unknown.

3.2 The Archaeal/Eukaryal Factors

Perhaps the most interesting among the TIFs are those specifically shared by the Archaea and the Eukarya (Fig. 2). They are interesting also because their relative functions are still not entirely understood.

If a factor is found specifically in the archaeal/eukaryal lineage, this means that it arose after this lineage separated from the common stem of the evolutionary tree and that it had an important and specific function in the common ancestor of the Archaea and Eukarya. Deciphering such functions, therefore, is very important for understanding the evolution of the modern eukaryal translational apparatus and how much it is indebted to its archaeal ancestor for its specific functions.

It is noteworthy that the archaeal/eukaryal factors, where their function is at least partially understood, seem to have an important role in translational regulation.

3.2.1 a/eIF2

A central step in translation initiation is the delivery of tRNAi to the ribosomal P site. This task is assisted in Bacteria by the monomeric protein IF2 and is performed in Eukarya by the trimeric complex eIF2, consisting of the α, β and γ subunits, none of which is homologous to bacterial IF2. The textbook rationale for the usage of a different tRNAi-binding factor in Eukarya and Bacteria was that the former, because of their multicellularity and greater functional complexity, needed to evolve a more elaborated tRNAi-binding protein to achieve a more sophisticated control of translation. In fact, eIF2 is a central target for translational regulation in eukaryotes. It is a G-protein, which is only active in the GTP-bound form. After delivering met-tRNAi to the ribosomal P site, eIF2 hydrolyzes its GTP and is ejected from the ribosome in the inactive, GDP-bound form. To participate in another round of initiation, eIF2 must be reactivated by GTP/GDP exchange, which requires the intervention of an exchange factor called eIF2B, a pentameric protein composed of two catalytic and three regulatory subunits. In most conditions (such as stress) when a rapid shut-off of protein synthesis is desirable, eIF2 can be inactivated by the phosphorylation of its α subunit, carried out by certain stress-activated kinases [15]. This modification converts the α protein in a competitive inhibitor of eIF2B, thereby inhibiting GTP/GDP exchange and blocking the recycling of eIF2. This mechanism of translational control is essential and widespread in eukaryotic cells, but does not exist in Bacteria and also, most probably, in Archaea.

Because of this, it was very surprising to discover that all archaeal genomes contain homologs of the three subunits of eIF2, thereby showing that the presence of a trimeric tRNAi-binding factor is not a specific adaptation to "complex" eukaryotic lifestyle. The three archaeal IF2 subunits were cloned and purified, and it was demonstrated that they do assemble to form a trimeric factor that binds met-tRNAi selectively and adapts it in the ribosomal P site [23, 35].

Archaeal IF2 is smaller in size than its eukaryal counterpart, due to the greatly reduced length of its β subunit (15 kDa instead of ∼50 kDa) (Fig. 4). In all archaea, the γ subunit is the largest protein (∼45 kDa) followed by the α subunit (∼30 kDa). The archaeal β polypeptide is smaller than its eukaryal counterpart due to the lack of the domains that, in the eukaryal subunit, interact with the guanine nucleotide exchange factor, eIF2B, and with the GTPase activator, eIF5.

Three-dimensional structures are available for all three subunits of aIF2 (Fig. 4) as well as for the αγ dimer and the complete trimeric protein [48, 55]. The largest subunit, γ, has a striking resemblance to the EF-1A (EF-Tu in bacteria) ([41]; Fig. 5), consistent with the fact that it contains the guanine-nucleotide-binding domain and is principally involved in the interaction with met-tRNAi. The small β

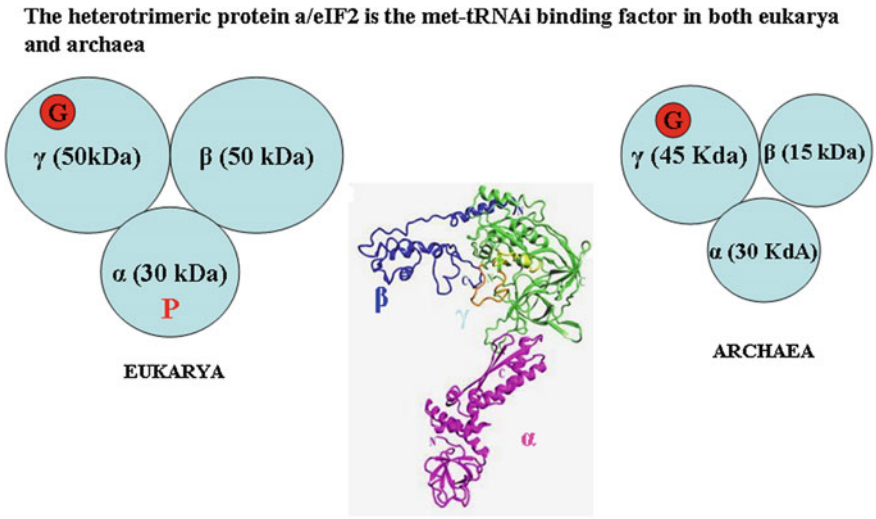


Fig. 4 The tRNAi-binding factor in Archaea and Eukarya. The drawings on the *left* and *right* show schemes of the subunit structure of, respectively, the eukaryal and the archaeal IF2. The size of the subunits is indicated in parentheses; the G in the *red circle* indicated the G domain of the γ subunit. The three-dimensional structure shown in the middle refers to archaeal IF2 and shows the relative positions of the three subunits, with the γ subunit in the middle that bridges together the α and the β subunits

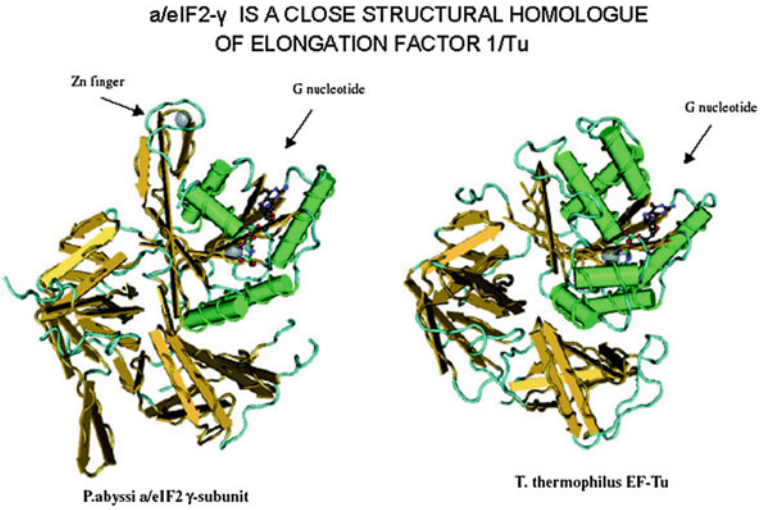


Fig. 5 Structural homology between the a/eIF2 γ-subunit and the translation elongation factor 1/Tu. The structures shown are those of *Pyrococcus abyssi*, aIF2 γ subunit, and of *Thermus thermophilus* EF-Tu. The positions of the guanine nucleotide-binding pocket and of a zinc-finger motif on the archaeal protein are indicated

subunit contains a zinc-finger motif. The α subunit is composed of three distinct domains, two of which have RNA-binding properties.

In spite of their functional and phylogenetic homology, eIF2 and aIF2 display several notable functional differences. One of these regards the nature of the tRNA-binding site. The met-tRNAi-binding site of the archaeal factor is all contained in the αγ dimer, to the exclusion of the β subunit. By contrast, met-tRNAi binding by eIF2 seems to involve mainly the γ and β subunits, while, as explained above, the eukaryal α-polypeptide is the main target of regulatory kinases that control the factor's activity.

A report that Archaeal (*Pyrococcus abyssi*) IF2-α is also phosphorylated [50] has not been confirmed to date. However, translational control by IF2-α phosphorylation is unlikely to happen in Archaea, because it has been shown that aIF2 has a similar affinity for GDP and GTP and therefore does not require a guanine nucleotide exchange factor to be reactivated [35]. This is consistent with the observation that all Archaea lack a homolog of eIF5, as well as three of the five subunits of eukaryal eIF2B. Archaeal genomes do include homologs of the α, β and δ subunits of eIF2B, but lack counterparts of the γ and ε subunits that catalyze guanine nucleotide exchange on eIF2. It has been shown that archaeal IF2B is able to interact with the aIF2 α subunit, but the significance of this interaction is unknown [14].

Another difference between eIF2 and aIF2 is the probable lack of a GTPase activator protein for the latter. GTP hydrolysis of eukaryal eIF2-phosphate is triggered by the helper factor eIF5, and no recognizable homolog of this protein is present in the genome sequences of Archaea. It is therefore likely that aIF2 has an intrinsic, ribosome-triggered GTPase activity, although this has not yet been demonstrated experimentally. Alternatively, GTP hydrolysis on aIF2 may be facilitated by an unidentified GTPase activator.

Apart from the problem of tRNAi binding, archaeal IF2 has an intriguing and unexpected property. It is capable of interacting with the 5′ terminus of an RNA, but only when this carries three phosphate groups [1, 24]. Most mRNAs fall in this category; therefore, it seemed likely that aIF2 is an mRNA-binding factor. Indeed, it was demonstrated that mRNAs carrying a 5′ triphosphate are protected against 5′ end degradation by bound aIF2, while degradation from the 3′ end of from the central parts of the molecule remained unaffected by aIF2 binding. Moreover, detachment of one or more of the terminal phosphates drastically lowers the affinity of the factor for the mRNA. This situation is evidently reminiscent of eukaryotic capping, which is a molecular device originally evolved just to protect the mRNAs against 5′-end degradation. In Archaea there is no "capping" modification of the mRNA 5′ end (i.e., no addition of a 7-methylguanosine by means of a 5′-5′ phosphate bond), but the 5′ terminal phosphates act themselves as a sort of primitive "cap," in the sense that they, just as the eukaryotic cap, bind specifically a protein complex aimed at protecting the RNA molecule against nucleolytic degradation.

Interestingly, aIF2 can bind mRNA in vitro both as an intact trimer as well as an isolated γ subunit; neither of the other two subunits binds RNA independently of the others. It is not known whether this is true also in vivo. Some indication that this might be the case stems from the fact that the a/eIF2 γ subunit seems to be more abundant than the other two in exponentially growing *Sulfolobus solfataricus* cells (Londei, unpublished work), but more accurate measurements of the aIF2 subunit's stoichiometry as well as other evidence are needed before this can be confirmed.

In any event, the mRNA-binding capacity of aIF2 seems to be at odds with its other function, that of binding met-tRNAi and promoting its interaction with the ribosome.

However, measurements of the relative affinity of free and ribosome-bound aIF2 for met-tRNAi and for the mRNA 5′ end yielded an explanation of this apparent paradox [24]. It turned out that aIF2 in a free state has a higher affinity for mRNA than for met-tRNAi, while the opposite is true when the factor sits on the ribosome. It must be pointed out that, unlike what happens in the eukaryotes, the interaction between aIF2 and met tRNAi most probably takes place when the factor is already ribosome-bound. Thus, in a physiological situation where there is an abundance of ribosomes, such as during rapid growth, most aIF2 is in a ribosome-bound state, thereby ensuring efficient translation and a rapid mRNA turnover. By contrast, when growth is arrested, such as under certain stress conditions, ribosomes become scarcer and there is more free aIF2 in the cell cytosol. In these conditions, translation is lagging, and it becomes important to protect the untranslated mRNAs from degradation to ensure a faster recovery of the cell once the environmental conditions revert to normal. In support of this, it has recently emerged that the removal of aIF2 from the 5′ end of a leaderless mRNA is stimulated by a new protein factor, Trf, and that *S. solfataricus* Trf deletion mutants have an impaired protein synthesis during outgrowth from the stationary phase [33].

Thus, the Archaea seem indeed to possess a rudimentary cap-like system for mRNA protection. Its relationship with the more sophisticated eukaryotic system is uncertain, but the similarity is nevertheless intriguing.

Although more data are evidently needed to fully understand the dual role of aIF2, its mRNA-binding properties may suggest ideas to unravel the main, still unanswered question: why did the Archaea adopt a trimeric tRNAi-binding factor (later inherited by the Eukarya) instead of recruiting the universal factor IF2 for this function, as the Bacteria did?

To gain insight into this problem, it is useful to bear in mind that, in the last universal common ancestor, translation most probably dispensed with initiation factors but did make use of the two elongation factors still present today, with very few modifications, in all life forms [3]. Elongation factor 1 (called Tu in Bacteria) delivers aminoacyl-tRNA to the ribosomal A site and checks the accuracy of codon-anticodon interaction, while elongation factor 2 (called EFG in Bacteria) promotes translocation, i.e., the reciprocal movement of ribosomes and mRNA that permits the positioning of the next unread codon in the A site. Both EF1/Tu and

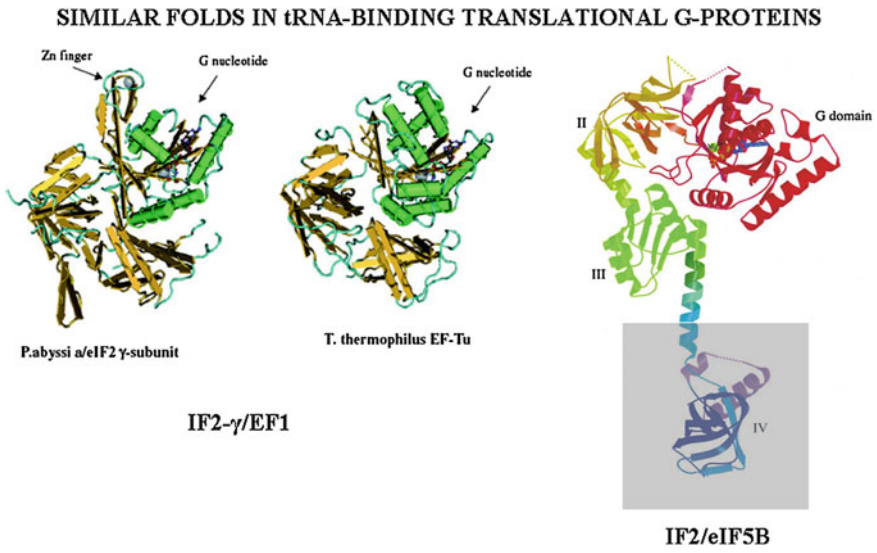


Fig. 6 Structural homologies between translational G-proteins. The structures compared are those of archaeal IF2-γ, bacterial EF-Tu (see Fig. 5) and archaeal IF5B. It can be seen that the "head" of the latter has a marked resemblance with the other two factors

EF2/G are G proteins, endowed with a guanine-nucleotide-binding site whose regulated interaction with GDP or GTP controls the activity of the factor.

As shown in Fig. 6, the three-dimensional structure of EF1/Tu closely resembles that of the a/eIF2 γ subunit and also the globular "head" of bacterial IF2 [41]. This suggests that the specialized tRNAi-binding proteins are all derived by gene duplication from the ancestral EF1/Tu, adapting themselves to interact with a specific initiator tRNA instead of with any aminoacyl-tRNA.

The aIF2 γ subunit was the specialized derivative of EF1/Tu that acquired met-tRNAi-binding properties in the common ancestor of Archaea and Eukarya. However, aIF2-γ also interacts with the triphosphate 5′ end of the mRNAs by virtue of its G domain. It may be supposed that this dual function was important in ancestral times, as it is important in modern Archaea. Before the branching of the eukaryotic lineage the protein evolved further, adding two other subunits, probably to create a factor that had an optimal tRNA-binding capacity, while maintaining the mRNA-protecting function. Once inherited by the eukaryotes, the trimeric eIF2 retained its original met-tRNA-binding capacity, while a novel and more reliable "capping" system, which combined mRNA protection with control of translational initiation, was evolved de novo.

A model depicting the possible evolution of the tRNAi-binding proteins in the three domains of life is shown in Fig. 7.

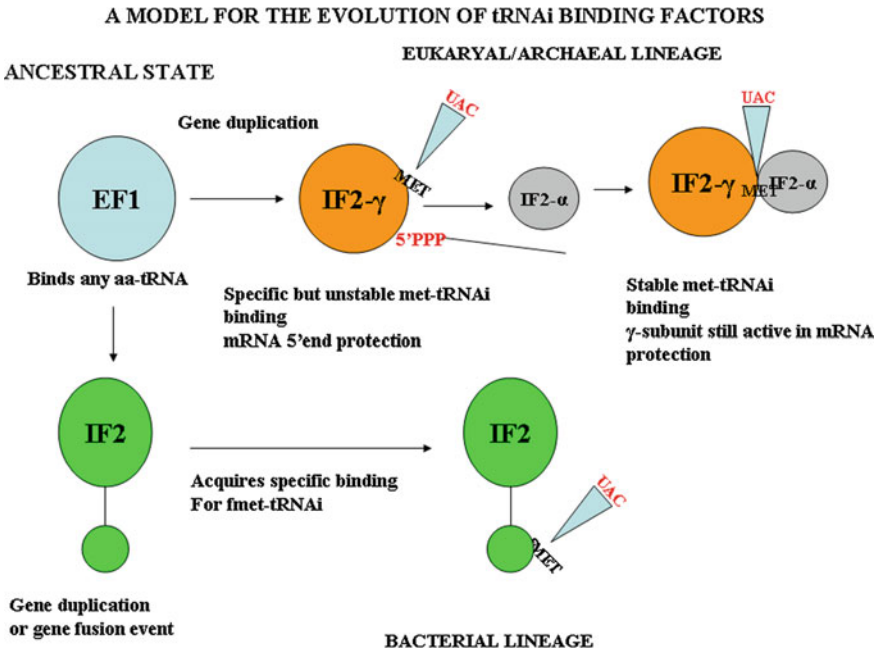


Fig. 7 A model of the evolution of tRNAi-binding factors. On the *left* of the figure the LUCA stage is shown, including the ancestral factors EF1/Tu and IF2, the latter probably derived from the former by a gene duplication or fusion event. On the *right* the separate evolution of the bacterial and of the archaeal/eukaryal lineages is shown. In the bacterial lineage, IF2 acquires the ability of binding f-met-tRNAi specifically, while the archaeal/eukaryal lineage evolves a new tRNAi-binding factor that also protects mRNAs from 5′-end degradation

3.2.2 a/eIF6

In both Archaea and Eukarya, the translation factor termed IF6 is a small (27 kDa) monomeric protein that binds specifically to the large ribosomal subunit. In eukaryotes, eIF6 was initially identified as an anti-association factor for its ability to bind the 60S ribosomal subunits preventing their association with the 40S ribosomal subunits to form the 80S initiation complex [52]. Later, however, experiments performed in yeast revealed that eIF6 may have a role in ribosome biogenesis, since it is also found in the nucleolus and its loss produces a decrease of 60S particles [44].

aIF6, the archaeal homolog, is a few amino acids shorter than its eukaryal counterpart, but otherwise shares a high degree of homology with it. IF6's three-dimensional structure was first determined for the archaeal protein [20] serving as a guide to model the structure of the eukaryotic homolog. As shown in Fig. 8, a/eIF6 has a characteristic flower-like shape with a hole in the middle. The particular fold of IF6 has been called a "pentein" because it is composed by a repetition of five very similar domains.

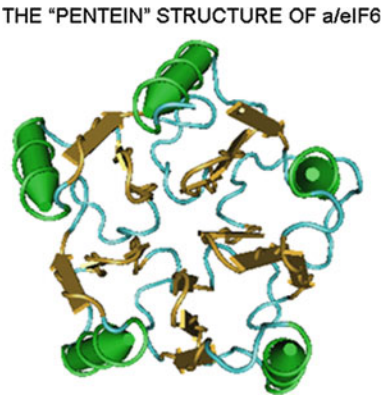


Fig. 8 The "pentein" structure of a/eIF6

Both the archaeal and the eukaryal proteins bind with high affinity to the large ribosomal subunit. The binding site, first determined for the archaeal factor [5, 18], and later also for the eukaryal one [27], lies on the side of the large subunit engaged in pairing with the smaller one, in close vicinity to proteins L14p and L24e. The position of eIF6 justifies its role as an anti-association factor since it masks the RNA and protein sites on the large subunit involved in contacting the small subunit and in bridging the two particles together.

Despite the detailed structural information, the role in translation of both archaeal and eukaryal IF6 remains puzzling. Only about one in ten large ribosomal subunits carry a/eIF6 in the cytoplasm, so one might hypothesize that the factor undergoes a series of fast binding and release cycles, regulating subunit joining in the course of translational initiation. However, the interaction of a/eIF6 with the ribosome is quite strong, and it appears that, in eukaryotes at least, specific factors are required for its release. Two different mechanisms have been proposed for eIF6 release from eukaryotic ribosomes, and it is not yet clear whether they co-exist or operate in different circumstances or in different cells. On the one hand, eIF6 detachment from the 60S subunits has been shown to be promoted by a GTPase, called Efl1, which acts in concert with the ribosome-binding factor Sdo1 (also called SBDS) to couple GTP hydrolysis with IF6 release [53]. On the other hand, eIF6 release has also been described to depend on the phosphorylation of the factor, which is in turn triggered by translation-promoting signaling transduced by the ribosome-bound kinase RACK1 [8].

Thus far, the prevalent model for the function of eukaryal eIF6 posits that the factor is important for the late maturation of the large ribosomal subunit. Immature 60S ribosomes would emerge from the nucleus with bound eIF6, being thus still unable to participate in translation. EIF6 release (by whichever mechanism) would permit the particles to take part in translation. In this model, the main role of eIF6 would be that of fine-tuning translation by regulating the number of active 60S subunits. However, many doubts remain, and many data are still needed to confirm (or refute) the model.

In Archaea, aIF6 has been shown to bind tightly and specifically to the 50S ribosomal subunit and thereby to inhibit subunit association [5]. However, the mechanism for aIF6 release from the 50S subunit is still unknown. The factor does not appear to undergo phosphorylation. Interestingly, the Archaea possess a ribosome-binding protein homologous to Sdo1/SBDS, but so far no clear proof of its involvement in aIF6 release has been obtained, nor has it been possible to identify a GTPase promoting aIF6 detachment from the ribosome (Benelli, La Teana and Londei, unpublished work). One study suggests that aIF6, rather than in initiation, may be involved in ribosome recycling [2], and eIF6 may well have a similar role.

The confusion about the role of a/eIF6 in translation is all the more frustrating since this factor, in eukaryotes at least, clearly has a very important role in regulating certain crucial cellular processes. Remarkably, eIF6 overexpression is observed in many natural cancers, while, conversely, eIF6 haplo-insufficiency protects against certain types of tumors [16]. Moreover, the overexpression of eIF6 has been described to produce developmental defects in Xenopus [12, 13].

Recent studies have provided some hint into the mechanism of action of eIF6 as a tumor-promoting factor. It has been shown that overexpression of the factor in an ovarian cancer cell line increases the motility and invasiveness of the cells, thus enhancing their metastatic potential [37]. Proteomic analysis has revealed that eIF6 overexpression perturbs the cellular proteome, upregulating a set of proteins involved in cell motility and invasiveness and downregulating another set of proteins, mainly ribosomal ones. On the whole, the data suggest that eIF6 may influence the cell's behavior indirectly, by modifying the translational landscape and altering the relative amounts of certain functionally crucial proteins. However, how it does so it is still a matter of conjecture. A possibility is that eIF6 overabundance may affect ribosome biosynthesis, thereby enhancing the production of specialized ribosomes that would preferentially translate certain classes of mRNAs. That such specialized ribosomes indeed exist has been shown in a number of reports [43, 54] but it remains to be seen whether eIF6 is actually involved in their biosynthesis.

Unfortunately, there are no data in Archaea to show whether aIF6 imbalances have any kind of physiological effects. It is only known that aIF6 is overexpressed under stress conditions, a fact that may suggest a role in controlling cellular behavior roughly similar to that observed for its eukaryotic counterpart. However, we are still a long way from understanding all the functional facets of this fascinating factor, let alone the motive for its evolutionary conservation in the archaeal/eukaryal line.

4 Conclusions: The Early Evolution of the Translation Initiation Machinery

As we have seen, present-day mechanisms of translational initiation are widely divergent in the primary domains of life. This is not very surprising: as initiation is the main target for translational regulation, each cell lineage has adapted its details

to its specific physiological needs. Nevertheless, it is still possible to discern conserved functional traits and common components, which allow tracing a model, if fragmentary and tentative, of the early evolution of the translation initiation machinery.

The story of the evolution of translational initiation, as we currently understand it, begins with a primordial cell, the last universal common ancestor (LUCA), whose ribosomes decoded leaderless mRNAs interacting with their 5′ extremity and probably using any codon as the start. The two elongation factors were in all probability already present. One of them, EF1/Tu, was (and still is) a tRNA-binding protein that interacted with any aminoacyl-tRNA adapting it in the A site. From EF1/Tu, a first gene duplication event generated the universal factor IF2/5B. This, in collaboration with the other universal factor IF1/1A, may have had the function of facilitating the entrance of the first tRNA in the P site, thus speeding up the translation initiation process [3].

The bacterial lineage was the first to branch off the common tree of life and to evolve its own system for translational initiation. According to their generally parsimonious lifestyle, the Bacteria built up a very streamlined mechanism, keeping the two initiation factors already present in the common ancestor and adding only a newly evolved one, IF3. IF2/5B specialized as the initiator tRNA binding factor IF2, acquiring a specificity for fmet-tRNAfmet. In parallel, the Bacteria modified mRNA structure to include leader sequences and SD motifs for ribosome binding.

The archaeal/eukaryal lineage found different solutions. As shown in Fig. 7, another derivative of EF1/Tu, the a/eIF2 γ subunit, was created by gene duplication, assuming the dual function of mRNA 5′ end protection and of specific met-tRNAi binding. The latter was achieved by the addition of the α subunit, while the β subunit probably came later. IF2/5B, by contrast, maintained only a generic capacity to stabilize met-tRNAi binding in the P site while at the same time promoting the joining of the ribosomal subunits in a late phase of initiation.

A further, and essential, component of the common archaeal/eukaryal translation initiation machinery was a/eIF6, but as we have seen very little can be conjectured on the ancestral function of this protein. Perhaps the primordial IF6 had a main role as a ribosome anti-association factor, being the functional counterpart of bacterial IF3, or perhaps it assisted ribosome synthesis. Further research on the function of modern archaeal and eukaryal IF6 is needed to answer this question.

After the branching off of the Eukaryal lineage, the Archaea made only a few improvements on the original translation initiation machinery, which may be the reason why they retain to date a large proportion of leaderless mRNAs. The Eukarya, on the other hand, modified the system, extensively evolving new mRNA features, a new system to locate the initiation codon and new TIFs to deal with these novelties.

References

1. Arkhipova V, Stolboushkina E, Kravchenko O, Kljashtorny V, Gabdulkhakov A, Garber M, Nikonov S, Märtens B, Bläsi U, Nikonov O. Binding of the 5′-Triphosphate end of mRNA to the γ-subunit of translation initiation factor 2 of the Crenarchaeon Sulfolobus solfataricus. J Mol Biol. 2015 Sep 25;427(19):3086–95.
2. Barthelme D, Dinkelaker S, Albers SV, Londei P, Ermler U, Tampé R. Ribosome recycling depends on a mechanistic link between the FeS cluster domain and a conformational switch of the twin-ATPase ABCE1. Proc Natl Acad Sci USA. 2011 Feb 22;108(8):3228–33.
3. Benelli D, Londei P. Translation initiation in Archaea: conserved and domain-specific features. Biochem Soc Trans. 2011;39(1):89–93.
4. Benelli D, Maone E, Londei P. Two different mechanisms for ribosome/mRNA interaction in archaeal translation initiation. Mol Microbiol. 2003;50:635–43.
5. Benelli D, Marzi S, Mancone C, Alonzi T, La Teana A, Londei P. Function and ribosomal localization of aIF6, a translational regulator shared by archaea and eukarya. Nucleic Acids Res. 2009 Jan;37(1):256–67.
6. Boelens R, Gualerzi CO. Structure and function of bacterial initiation factors. Curr Protein Pept Sci. 2002 Feb;3(1):107–19.
7. Carter AP, Clemons WM Jr, Brodersen DE, Morgan-Warren RJ, Hartsch T, Wimberly BT, Ramakrishnan V. Crystal structure of an initiation factor bound to the 30S ribosomal subunit. Science. 2001;291:498–501.
8. Ceci M, Gaviraghi C, Gorrini C, Sala LA, Offenhauser N, Marchisio PC, Biffo S. Release of eIF6 (p27BBP) from the 60S subunit allows 80S ribosome assembly. Nature. 2003;426:579–84.
9. Chaudhuri J, Si K, Maitra U. Function of eukaryotic translation initiation factor 1A (eIF1A) in initiation of protein synthesis. J Biol Chem. 1997;272:7883–91.
10. Choi SK, Lee JH, Zoll WL, Merrick WC, Dever TE. Promotion of met-tRNAi Met binding to ribosomes by yIF2, a bacterial IF2 homolog in yeast. Science. 1998;280:1757–60.
11. Condò I, Ciammaruconi A, Benelli D, Ruggero D, Londei P. Cis-acting signals controlling translational initiation in the thermophilic archaeon Sulfolobus solfataricus. Mol Microbiol. 1999;19:5233–40.
12. De Marco N, Iannone L, Carotenuto R, Biffo S, Vitale A, Campanella C. p27(BBP)/eIF6 acts as an anti-apoptotic factor upstream of Bcl-2 during Xenopus laevis development. Cell Death Differ. 2010 Feb;17(2):360–72.
13. De Marco N, Tussellino M, Vitale A, Campanella C. Eukaryotic initiation factor 6 (eif6) overexpression affects eye development in Xenopus laevis. DifferentiationSep. 2011;82 (2):108–15.
14. Dev K, Santangelo TJ, Rothenburg S, Neculai D, Dey M, Sicheri F, Dever TE, Reeve JN, Hinnebusch AG. Archaeal aIF2B interacts with eukaryotic translation initiation factors eIF2alpha and eIF2Balpha: implications for aIF2B function and eIF2B regulation. J Mol Biol. 2009 Sep 25;392(3):701–22.
15. Donnelly N, Gorman AM, Gupta S, Samali A. The eIF2α kinases: their structures and functions. Cell Mol Life Sci. 2013 Oct;70(19):3493–511.
16. Gandin V, Miluzio A, Barbieri AM, Beugnet A, Kiyokawa H, Marchisio PC, Biffo S. Eukaryotic initiation factor 6 is rate-limiting in translation, growth and transformation. Nature. 2008;455:684–8.
17. Giliberti J, O'Donnell S, Etten WJ, Janssen GR. A 5′-terminal phosphate is required for stable ternary complex formation and translation of leaderless mRNA in Escherichia coli. RNA. 2012 Mar;18(3):508–18.
18. Greber BJ, Boehringer D, Godinic-Mikulcic V, Crnkovic A, Ibba M, Weygand-Durasevic I, Ban N. Cryo-EM structure of the archaeal 50S ribosomal subunit in complex with initiation factor 6 and implications for ribosome evolution. J Mol Biol. 2012 May 4;418(3–4):145–60.

19. Grill S, Gualerzi CO, Londei P, Bläsi U. Selective stimulation of translation of leaderless mRNA by initiation factor 2: evolutionary implications for translation. EMBO J. 2000 Aug 1;19(15):4101–10.
20. Groft CM, Beckmann R, Sali A, Burley SK. Crystal structures of ribosome anti-association factor IF6. Nat Struct Biol. 2000 Dec;7(12):1156–64.
21. Gualerzi CO, Pon CL. Initiation of mRNA translation in prokaryotes. Biochemistry. 1990 Jun 26;29(25):5881–9.
22. Guenneugues M, Caserta E, Brandi L, Spurio R, Meunier S, Pon CL, Boelens R, Gualerzi CO. Mapping the fMet-tRNA f-met binding site of initiation factor IF2. EMBO J. 2000.
23. Guillon L, Schmitt E, Blanquet S, Mechulam Y. Initiator tRNA binding by e/aIF5B, the eukaryotic/archaeal homologue of bacterial initiation factor IF2. Biochemistry. 2005 Nov 29;44(47):15594–601.
24. Hasenöhrl D, Lombo T, Kaberdin V, Londei P, Bläsi U. Translation initiation factor a/eIF2(-gamma) counteracts 5′ to 3′ mRNA decay in the archaeon Sulfolobus solfataricus. Proc Natl Acad Sci USA. 2008 Feb 12;105(6):2146–50.
25. Hinnebusch AG. The scanning mechanism of eukaryotic translation initiation. Annu Rev Biochem. 2014;83:779–812.
26. Hinnebusch AG, Lorsch JR. The mechanism of eukaryotic translation initiation: new insights and challenges. Cold Spring Harb Perspect Biol. 2012 Oct 1;4(10).
27. Klinge S, Voigts-Hoffmann F, Leibundgut M, Arpagaus S, Ban N. Crystal structure of the eukaryotic 60S ribosomal subunit in complex with initiation factor 6. Science. 2011 Nov 18;334(6058):941–8.
28. Kozak M. The scanning model for translation: an update. J Cell Biol. 1989 Feb;108 (2):229–41.
29. Kramer P, Gäbel K, Pfeiffer F, Soppa J. Haloferax volcanii, a prokaryotic species that does not use the Shine Dalgarno mechanism for translation initiation at 5′-UTRs. PLoS One. 2014 Apr 14;9(4):e94979. doi:10.1371/journal.pone.0094979.
30. Le Quesne JP, Spriggs KA, Bushell M, Willis AE. Dysregulation of protein synthesis and disease. J Pathol. 2010 Jan;220(2):140–51.
31. Londei P. Evolution of translational initiation: new insights from the archaea. FEMS Microbiol Rev. 2005 Apr;29(2):185–200.
32. Maone E, Di Stefano M, Berardi A, Benelli D, Marzi S, La Teana A, Londei P. Functional analysis of the translation factor aIF2/5B in the thermophilic archaeon Sulfolobus solfataricus. Mol Microbiol. 2007 Aug;65(3):700–13.
33. Märtens B, Manoharadas S, Hasenöhrl D, Zeichen L, Bläsi U. Back to translation: removal of aIF2 from the 5′-end of mRNAs by translation recovery factor in the crenarchaeon Sulfolobus solfataricus. Nucleic Acids Res. 2014 Feb;42(4):2505–11.
34. Moll I, Hirokawa G, Kiel MC, Kaji A, Blasi U. Translation initiation with 70S ribosomes: an alternative pathway for leaderless mRNAs. Nucleic Acids Res. 2004;32:3354–63.
35. Pedullà N, Palermo R, Hasenöhrl D, Bläsi U, Cammarano P, Londei P. The archaeal eIF2 homologue: functional properties of an ancient translation initiation factor. Nucleic Acids Res. 2005 Mar 23;33(6):1804–12.
36. Pestova TV, Lomakin IB, Lee JH, Choi SK, Dever TE, Hellen CU. The joining of ribosomal subunits in eukaryotes requires eIF5B. Nature. 2000;403:332–5.
37. Pinzaglia M, Montaldo C, Polinari D, Simone M, La Teana A, Tripodi M, Mancone C, Londei P, Benelli D. EIF6 over-expression increases the motility and invasiveness of cancer cells by modulating the expression of a critical subset of membrane-bound proteins. BMC Cancer. 2015 Mar 15;15:131.
38. Roll-Mecak A, Cao C, Dever TE, Burley SK. X-ray structures of the universal translation initiation factor IF2/eIF5B: conformational changes on GDP and GTP binding. Cell. 2000 Nov 22;103(5):781–92.
39. Ruggero D. Translational control in cancer etiology. Cold Spring Harb Perspect Biol. 2013 Feb 1;5(2).

40. Sartorius-Neef S, Pfeifer F. In vivo studies on putative Shine-Dalgarno sequences of the halophilic archaeon Halobacterium salinarum. Mol Microbiol. 2004 Jan;51(2):579–88.
41. Schmitt E, Blanquet S, Mechulam Y. The large subunit of initiation factor aIF2 is a close structural homologue of elongation factor Tu. EMBO J. 2002;21:1821–32.
42. Schmitt E, Naveau M, Mechulam Y. Eukaryotic and archaeal translation initiation factor 2: a heterotrimeric tRNA carrier. FEBS Lett. 2010 Jan 21;584(2):405–12.
43. Shi Z, Barna M. Translating the genome in time and space: specialized ribosomes, RNA regulons, and RNA-binding proteins. Annu Rev Cell Dev Biol. 2015;31:31–54.
44. Si K, Maitra U. The *Saccharomyces cerevisiae* homologue of mammalian translation initiation factor 6 does not function as a translation initiation factor. Mol Cell Biol. 1999;19:1416–26.
45. Siddiqui N, Sonenberg N. Signalling to eIF4E in cancer. Biochem Soc Trans. 2015 Oct;43 (5):763–72.
46. Slupska MM, King AG, Fitz-Gibbon S, Besemer J, Borodovsky M, Miller JH. Leaderless transcripts of the crenarchaeal hyperthermophile Pyrobaculum aerophilum. J Mol Biol. 2001 Jun 1;309(2):347–60.
47. Spriggs KA, Bushell M, Willis AE. Translational regulation of gene expression during conditions of cell stress. Mol Cell. 2010 Oct 22;40(2):228–37.
48. Stolboushkina E, Nikonov S, Nikulin A, Bläsi U, Manstein DJ, Fedorov R, Garber M, Nikonov O. Crystal structure of the intact archaeal translation initiation factor 2 demonstrates very high conformational flexibility in the alpha- and beta-subunits. J Mol Biol. 2008 Oct 10;382(3):680–91.
49. Stumpf CR, Ruggero D. The cancerous translation apparatus. Curr Opin Genet Dev. 2011 Aug;21(4):474–83.
50. Tahara M, Ohsawa A, Saito S, Kimura M. In vitro phosphorylation of initiation factor 2 alpha (aIF2 alpha) from hyperthermophilic archaeon Pyrococcus horikoshii OT3. J Biochem (Tokyo). 2004;135:479–85.
51. Tolstrup N, Sensen CW, Garrett RA, Clausen IG. Two different and highly organized mechanisms of translation initiation in the archaeon Sulfolobus solfataricus. Extremophiles. 2000 Jun;4(3):175–9.
52. Valenzuela DM, Chaudhuri A, Maitra U. Eukaryotic ribosomal subunit anti-association activity of calf liver is contained in a single polypeptide chain protein of *M*r = 25,500 (eukaryotic initiation factor 6). J Biol Chem. 1982;257:7712–9.
53. Weis F, Giudice E, Churcher M, Jin L, Hilcenko C, Wong CC, Traynor D, Kay RR, Warren AJ. Mechanism of eIF6 release from the nascent 60S ribosomal subunit. Nat Struct Mol Biol. 2015 Nov;22(11):914–9.
54. Xue S, Barna M. Specialized ribosomes: a new frontier in gene regulation and organismal biology. Nat Rev Mol Cell Biol. 2012 May 23;13(6):355–69.
55. Yatime L, Mechulam Y, Blanquet S, Schmitt E. Structural switch of the gamma subunit in an archaeal aIF2 alpha gamma heterodimer. Structure. 2006 Jan;14(1):119–28.

On the Origin and Early Evolution of Translation in Eukaryotes

Greco Hernández, Vincent G. Osnaya, Alejandra García and Mitzli X. Velasco

1 Introduction

Proteins are one of the elementary molecules of the biosphere, catalyzing the majority of reactions sustaining life, as well as playing structural, transport and regulatory roles in all living organisms. Hence, protein synthesis or "translation" is a fundamental process for all forms of life [1, 2], and translational control plays a crucial role in gene expression during many cellular and developmental processes. Accordingly, the process of translation as well as different regulatory mechanisms should have evolved ever since the beginning of life. Later in evolution, the emergence of eukaryotes represented a profound hallmark in the history of life on our planet, leading to crucial changes at the ecological, morphological, biochemical and molecular levels in living organisms. How translation ever originated and what changes the process of translation underwent during the arousal and radiation of eukaryotes is still the subject of intense debate.

The knowledge of the mechanism and regulation of translation has been established in the last 5 decades by the work of brilliant scientists from many laboratories across different countries. In recent years, the advent of the powerful "omics" era has created a huge data set regarding the molecular composition of cells from hundreds of species from many phyla never studied before, giving rise to an innovative perspective in the study of biological processes. This approach has led to the surprising discovery that a number of components of the translation apparatus have undergone diversification across eukaryotes and that distinct regulatory mechanisms have evolved in different phyla at different times [3–7]. This also has allowed performing phylogenomic analyses across the three domains of life, namely Bacteria, Archaea and Eukarya, to gain insight into how the translation

G. Hernández (✉) · V.G. Osnaya · A. García · M.X. Velasco
Division of Basic Research, National Institute of Cancer (INCan),
22 San Fernando Avenue, Tlalpan, 14080 Mexico City, Mexico
e-mail: ghernandezr@incan.edu.mx

G. Hernández and R. Jagus (eds.), *Evolution of the Protein Synthesis Machinery and Its Regulation*, DOI 10.1007/978-3-319-39468-8_5

machinery might have evolved during the emergence of eukaryotes. Yet, despite the impressive advances in the field of translation, many questions regarding the emergence and early evolution of translation in eukaryotes still remain open.

In this chapter, we will review recent research shedding light on the evolution of the translation apparatus during the onset of eukaryotes and how it might have evolved right afterwards. Since the elongation and termination steps, or translation, are very well conserved among all kingdoms of life, and the initiation step has undergone substantial modification in eukaryotes as compared to both Archaea and Bacteria, we will focus on the evolution of the initiation step.

2 Translation Initiation in the Prokaryotic World

In prokaryotes, translation happens simultaneously in time and space with transcription, which always synthesizes polycistronic mRNAs. Translation initiation in bacteria consists of the recruitment of the 5′ end of an mRNA by the 30S ribosomal subunit, i.e., the formation of the complex among mRNA, fMet-tRNA$_i^{fMet}$ (initiator formylmethionyl-tRNA) and the ribosomal subunit. It is assisted by the translation initiation factors (IFs) IF1, IF2 and IF3. IF2 binds tRNA$_i^{fMet}$ and delivers it to the P site of the ribosomal subunit 30S, and its activity is stimulated by IF1; IF3 controls the accuracy of codon-anticodon recognition. In Archaea, initiation is more complex since it possesses at least five archaeal initiation factors (aIFs): aIF1, aIF1A, eIF2, aIF5B and eIF6. aIF1 drives mRNA binding to the ribosome and also confers fidelity of the start codon selection; aIF2 binds tRNA$_i^{fMet}$; aIF2 along with aIF5B delivers tRNA$_i$ to the P site; aIF6 keeps ribosomal subunits dissociated. So far, no role has been found for aIF1A [8–10].

John Shine and Lynn Dalgarno discovered in the early 1970s that mRNA recruitment to the ribosome occurs by a direct base pairing between a purine-rich region located ~7 nucleotides upstream the mRNA start codon, the so-called Shine-Dalgarno (SD) sequence with the consensus AGGAGG, and a complementary sequence at the 3′ end of the 16S rRNA on the 30S small ribosomal subunit (referred to as anti-Shine-Dalgarno sequence, anti-SD) [11, 12]. The critical role of the SD sequence in translation initiation was further experimentally corroborated in a variety of species, both eubacterial and archaeal [10, 13–16]. This, together with the large number of genes possessing SD sequences in many bacteria, led to the general idea that for prokaryotic mRNAs the SD sequence is the essential (although not necessarily the sole) element for ribosome recruitment and for selection of the correct initiation codon. It was then assumed that the SD/anti-SD interaction during initiation is conserved in all prokaryotes [9, 10, 17].

Besides the SD motif, it was later found that ribosomal protein (RP) S1 interacts with the 3′ end of the 16S rRNA, particularly with helices h26 and h45, which contains the anti-SD sequence, as well as with 11 nucleotides of the mRNA 5′-UTR located immediately upstream of the SD sequence. Thus, it is thought that the major function of RPS1 is to bring the mRNA onto the 30S subunit during translation

initiation, thereby assisting the interactions between the SD motif in the mRNA and the anti-SD sequence of 16S rRNA. This is consistent with the observation that translation of leaderless mRNAs does not require RPS1, since it does not depend on SD interaction [18–23].

The recent advent of genome- and transcriptome-wide studies of thousands of species led to the discovery in the past few years of a significant number of naturally occurring mRNAs lacking an SD sequence spread across a wide variety of eubacterial and archaeal phyla, being more abundant in Archaea. These include thousands of mRNAs devoid of 5′-UTR (and hence referred to as "leaderless" mRNAs) produced from single genes and from the first genes of operons, as well as mRNAs that possess a 5′-UTR but lack an SD sequence [8, 9, 24–43]. These findings were further confirmed in a more recent study comprising 2,458 bacterial complete genomes [44]. Indeed, several studies have shown that the major pathway to initiate translation initiation in Archaea might involve mostly leaderless mRNAs [29, 32–35].

Thus, in addition to the aforementioned SD/anti-SD-dependent initiation, two other major mechanisms for prokaryotic translation initiation have been described. (1) For leaderless mRNAs the AUG start codon itself was found to serve as the most important signal for ribosome recruitment and translation initiation. Here, the initiator tRNA and IF2 are critical for complex formation between the start codon and the ribosome. It is noteworthy that translation initiation of leaderless mRNAs involves the undissociated ribosome 70S instead of the 30S ribosomal subunit [8, 9, 24, 31, 40, 41, 45–52]. (2) For mRNAs possessing a 5′-UTR but lacking an SD motif, mRNA recruitment into the ribosome can be mediated exclusively by RPS1 [18–23, 30, 33, 53]. These mRNAs exhibiting a pronounced minimum in secondary structure and AUG start codon reside in single-stranded regions of the mRNAs, and ribosome binding to the mRNA is a sequence-independent event but is strictly dependent on the local absence of secondary RNA [54].

Intriguingly, neither archaebacteria nor eukaryotes contain an *RPS1* gene, raising the question of how leadered mRNAs devoid of an SD motif are translated in Archaea [8, 17, 18, 30, 51, 55]. Finally, in different species alternate non-SD sequences have been reported to mediate 16S rRNA-mRNA interaction in a variety of prokaryotic mRNAs, including domain #17 of *E. coli* 16S rRNA [56], the "translation initiation region" sequence of the *Mycoplasma genitalium tuf* gene [57] and a region of *Thermus thermophilus thrS* gene mRNA [58]. Moreover, genomic studies of several archaeal species found a strong conservation of a GGTG atypical putative ribosome binding site within 15 nucleotides upstream of the start codon of hundreds of genes [26, 36]. Yet, whether or not these sequences undergo base-pairing with the anti-SD sequence of 16S rRNA is unknown. The biological relevance of these sequences, if any, remains poorly understood.

Overall, the emerging view is that in the prokaryotic world, both SD-dependent and -independent translation mechanisms are present in all major lineages, showing that the 16S rRNA recruitment by prokaryotic mRNAs is a variable process. Since elongation and termination steps of translation are highly conserved in prokaryotes, these findings support the hypothesis that the *l*ast *u*niversal *c*ommon *a*ncestor (hereafter termed LUCA) of extant life already possessed an established

fundamental translational apparatus, but the mechanisms of initiation further evolved in the bacterial and archaeal lineages and afterwards changed in eukaryotes even more [3, 4, 30, 33, 43, 59–61]. Indeed, evidence suggests that a highly developed translation system was a necessary condition for the emergence of cells on earth [62]. However, the aforementioned variety of prokaryotic mechanisms raises the question of how LUCA might have initiated translation.

2.1 mRNA Recruitment in the Last Universal Common Ancestor of Extant Organisms

As mentioned above, ribosome recruitment by mRNAs is a variable process in prokaryotes. Thus, a crucial question is whether or not all prokaryotes possess an anti-SD sequence at the 3′ end of 16S rRNA on the 30S ribosomal subunit. Complete genome analyses of 277 [30] and 162 [29] prokaryote species (both bacterial and archaeal) and 18 archaeal species [26] surprisingly found that the anti-SD sequence is highly conserved among all species analyzed. Since thousands of prokaryotic mRNAs have been found to lack either an SD motif or a leader sequence, this paradox could be explained by three alternative evolutionary scenarios, i.e., (1) LUCA mRNAs possessed SD sequences at the 5′-UTR of mRNAs, but were lost multiple and independent times in different prokaryotic lineage [29, 30]. For them, RPS1-mediated or leaderless mRNA mechanisms of translation initiation work to a great extent [29, 30]. In this case, the evolutionary pressures that led to the loss of SD sequences, if any, are completely unknown. (2) Only some organisms have possessed SD motifs ever since the beginning of life, opening the possibility that the anti-SD motif present in 16S rRNA from many species might play more, yet-unidentified roles in translation or even in different process, such as ribosomal RNA biogenesis, export or stability. (3) One or more hypothetical sequences, other than the SD motif, might have driven translation initiation in LUCA mRNAs and are currently present in different prokaryotic mRNAs but have not been identified. This idea is supported by the finding that a variety of alternate sequences support base-pairing between the 16S rRNA and mRNA to drive translation initiation [56, 58, 63]. We may conclude that the current knowledge does not shed light on what the mechanism of ribosomal recruitment by LUCA mRNAs might have been.

3 Translation Initiation in Modern Eukaryotes

In modern eukaryotes, the vast majority of mRNAs initiate translation by the so-called cap-dependent mechanism, which is mediated by the eukaryotic initiation factors (eIFs) and consists of the recruitment of the mRNA to the 40S ribosomal subunit upon recognition of the cap structure (m^7GpppN, where N is the nucleotide located at the very 5′ end of the mRNA) by the cap-binding protein eIF4E.

It begins with the dissociation of the ribosomal subunits 60S and 40S by IF6. Afterward, the free 40S ribosomal subunit, which is stabilized by eIF3, eIF1 and eIF1A, binds to a ternary complex—consisting of eIF2 bound to an initiator Met-tRNA$_i^{Met}$ and GTP (eIF2-GTP/Met-tRNA$_i^{Met}$)—to form a 43S pre-initiation complex. eIF5 interacts with eIF2 and eIF3 and is probably also recruited to the 40S ribosomal subunit. On the other hand, and most likely simultaneously, the cap structure of an mRNA is recognized by eIF4E in complex with the scaffold eIF4G. Then, recruitment of the mRNA 5′-UTR by the 43S pre-initiation complex happens, a process that is coordinated by eIF4G via its interaction with eIF4E, the ATPase/RNA helicase eIF4A, the poly (A)-binding protein (PABP) and the 40S ribosomal subunit-associated eIF3. eIF4G-PABP interaction causes a crosstalk between both mRNA ends, hence prompting mRNA circularization, a spatial conformation that stimulates translation and that is known as the closed-loop model. This complex scans in a 5′ ➜ 3′ direction along the 5′-UTR to reach the start codon, usually an AUG. During the scanning (a process that requires ATP), eIF4B stimulates the activity of eIF4A, which unwinds secondary RNA structures in mRNA. eIF1, eIF1A and eIF5 assist in the positioning and fidelity of the 40S ribosomal subunit at the correct start codon so that eIF2 can deliver the anti-codon of the initiator Met-tRNA$_i^{Met}$ as the cognate partner for the start codon directly to the peptidyl-site of the 40S ribosomal subunit. Once the ribosomal subunit is placed on the correct start codon a 48S pre-initiation complex is formed. Then eIF5 promotes GTP hydrolysis by eIF2 and the release of the initiation factors. eIF2B and GTP afterward recycle the dissociated eIF2-GDP complex so that it can associate with a new Met-tRNA$_i^{Met}$ and take part in a new round of initiation. Finally, the GTPase eIF5B is required for the assembling of the 60S ribosomal subunit to the 48S complex to form an 80S initiation complex. Thereafter, the polypeptide elongation begins [64–66].

In the late 1980s, the groups of Nahum Sonenberg and Eckard Wimmer independently discovered that there is an internal sequence in the 5′-UTR of picornaviral mRNAs located in the proximity of the start codon that allows the 40S ribosomal subunit to land directly on the mRNA in a cap-independent manner and without involvement of eIF4E [67, 68]. This sequence is termed an *i*nternal *r*ibosome *e*ntry *s*ite (IRES). A few years later, the first cellular IRES was discovered in the mRNA of the immunoglobulin heavy-chain binding protein [69], an mRNA that is translated upon poliovirus infection. Since that time a large number of cellular and viral mRNAs have been found translated via different IRESs elements [17, 70–73]. In the following, we will analyze the evolutionary phenomena that might have spurred the emergence of present-day translation in eukaryotes.

4 The Emergence of Eukaryotic Translation

About 1.8 billon years ago, the endosymbiotic association of respiratory, alpha-proteobacterium-like prokaryotes (the ancestors of the mitochondria) with host organisms that possessed an archaeal genetic identity led to the emergence of

eukaryotes. The posterior association of this cellular consortium with photosynthetic, cyanobacterium-like endosymbionts led to the evolution of plastids [74–77]. The onset of eukaryotes caused the emergence of novel, much more sophisticated levels of cellular architecture than their prokaryotic ancestors, resulting from the appearance of a plethora of new cell features, including a nucleus and centrioles, as well as endosymbiotic bacteria that evolved toward mitochondria and plastids; peroxisomes, Golgi complex and endoplasmic reticulum; the rearrangement of genetic information in a "fragmented" fashion (i.e., interrupted genes) and packed into multiple linear chromosomes inside the nucleus; cilia; cytoskeleton and motors for vesicle and molecules transportation; sex, mitosis- and meiosis-based cell division; expansion of genome size; expansion of cell size; and in many phyla the emergence of multicellularity resulting out of different developmental programs.

Interestingly, despite the well-established idea that eukaryotes evolved from archaeal ancestors [74, 77–83], phylogenomic analyses have shown different roots for the cellular components of eukaryotes. While they inherited from Archaea the informational machineries, namely replication, transcription and translation, the metabolic and energetic enzymes are mostly of bacterial origin [60, 74, 82]. Consistent with this notion, genome-based phylogenetic analyses as well as structural and biochemical studies have shown that archaeal translation factors [8, 9, 31, 59, 84, 85], ribosomal proteins [18, 86–88], aminoacyl-tRNA synthetases (despite extensive horizontal gene transfer undergone among the three domains of life) [89, 90] and ribosomal RNAs [55, 62, 74, 78–80, 91, 92] have their closest homologs in eukaryotes rather than in bacteria.

The evolutionary emergences of the nucleus and interrupted genes were paramount events of eukaryote genesis. Crucially, they caused the interruption of genetic information of host cells and led to the spatio-temporal separation of transcription and translation. Therefore, upon their emergence eukaryotes needed the prompt evolution of nuclear machineries for intron splicing, for nucleocytoplasmic export and for mRNA protection to ensure that transcripts synthesized in the nucleus reach both the ribosomes and the storage bodies in the cytoplasm. Surveillance systems such as nonsense-mediated mRNA decay (NMD) also evolved to discard aberrant mRNAs [3, 4, 93, 94]. The arousal of eukaryotic cells also led to the evolution of novel features in the translation apparatus, mechanisms and regulation so that gene expression could take place. The major changes summarized are the following.

(1) Eukaryotic ribosomes are much bigger and more complex than their prokaryotic counterparts. Ribosomes evolved toward the eukaryotic 40S and 60S ribosomal subunits from prokaryotic 30S and 50S, respectively. This was due to the addition of several rRNA expansion segments, peptide additions to most ribosomal proteins, as well as the addition of extra eukaryotic-specific ribosomal proteins and the 5.8S rRNA. Thus, while bacterial 70S ribosomes contain $\sim$4500 nucleotides of rRNA, eukaryotic 80S ribosomes contain >5500 nucleotides of rRNA [88, 95–98]. The number of ribosomal proteins increased from 57 (23 in the small ribosomal subunit and 34 in the large subunit) in Bacteria and 68 (28;40) in Archaea to 78 (32;46) in Eukarya [18, 55, 86, 97].

(2) The initiation step of translation underwent a substantial increase in terms of the complexity and number of initiation factors as compared to prokaryotes, i.e., while in Bacteria and Archaea it is assisted by 3 and 6 factors, respectively, eukaryotes need the interplay of at least 14 factors. Thus, novel, eukaryotic-specific initiation factors evolved, namely eIF3 (all subunits), eIF4B, eIF4E, eIF4G, eIF4H and eIF5 [8, 9, 14, 31, 59, 85]. Except for eIF5, all of them recognize the mRNA 5′-UTR for recruitment into the ribosome.

(3) mRNAs also underwent profound changes during the transition from prokaryotic to eukaryotic cells. (a) They acquired a novel structure, i.e., monocistronic, capped, polyadenylated and with long UTRs. Moreover, eukaryotic mRNA 5′-UTRs are devoid of an SD motif and, for some lineages, the AUG start codon is surrounded by a context sequence instead. For vertebrate mRNAs, the optimal context is termed the "Kozak motif," which consists of the consensus sequence G/AXXATGG [99]. Experimental and in silico studies of a few mRNAs from some species suggest that this sequence is not conserved across eukaryotes [100–106]. Recently, a genome-wide in silico analysis of 48 species found that the preferred sequence around the start codon significantly varies across species of all eukaryotic kingdoms [107]. However, no experimental validation of this observation has been performed. (b) They acquired a novel life cycle, being transcribed, capped, polyadenylated, spliced and exported from the nucleus, and further stored, transported, translated and degraded in the cytoplasm. And (c) They acquired a novel functional conformation when engaged in translation, i.e., a circular shape displaying a functional crosstalk between both the 5′- and 3′-ends [4, 5, 93, 108].

(4) New mechanisms for translation regulation evolved in different lineages, including a plethora of eIF4E-interacting proteins (4E-IPs), the TOR pathway, microRNAs, different cytoplasmic granules, eIF2alpha kinases and the control of mRNA circularization by poly(A) tail shortening, among others [4–8, 31, 59].

4.1 A Closer Look at the Untranslated Regions of Eukaryotic mRNAs

Among the key features that evolved in eukaryotic mRNAs are the UTRs, as mRNA stability, transport and translation rates are tightly controlled by *cis*-acting elements located on them. Indeed, both 5′- and 3′-UTRs are critical targets of different networks of *trans*-acting factors for finely tuning gene expression at different levels. Notwithstanding, there are remarkable functional differences between both UTRs, as most *cis*-acting regulatory elements regulating mRNA polyadenylation, degradation, storage, localization and transport of mRNAs are localized at the 3′-UTR. In contrast, 5′-UTR is key for ribosomal landing, scanning and binding of diverse RNA-binding proteins regulating scanning and ATG codon recognition during translation initiation [66, 93, 108–119]. As a consequence of this, while the mean length of 5′-UTR remains remarkably constant in most eukaryotic phyla

(70 − 200 nucleotides), 3′-UTR mean length increases as morphological complexity increases [93, 108, 109, 112, 120–125].

According with the crucial roles UTRs play in post-transcriptional regulation of gene expression, leaderless mRNAs are rather seldom in eukaryotes [93, 108, 109, 114, 120–123], and extremely short 5′-UTRs have been reported only among mRNAs from the unicellular protists *Giardia lamblia*, with 5′-UTRs in the range of 0–14 nucleotides [126], and *Entamoeba histolytica*, which possesses some 5′-UTRs as short as 5 nucleotides [127]. However, this feature could be due to their parasitic life. Recently, some human mRNAs have also been reported to contain short 5′-UTRs with a median length of 12 nucleotides within a translational element termed TISU (which stands for translation initiator of short 5′-UTR) [128–130]. Yet, the frequency of this element in other species remains to be determined.

5 The Transition from Prokaryotic to Eukaryotic Translation

Several evolutionary forces played crucial roles in the transition from the ancestral, prokaryotic mode of translation toward the establishment of the predominant cap-dependent translation of eukaryotes. It is well established that the last common ancestor of extant eukaryotes had a genome with a high intron density, most likely as a result of an invasion of group II introns from the new mitochondrial endosymbionts into the genes of the host organism [94, 131–134]. The emergences of the nuclear membrane and interrupted genes were probably some of the primordial selection forces to overcome in the first eukaryotes [60, 94], raising the immediate need for developing systems for the protection and nucleocytoplasmic export of mRNAs, for intron splicing and for the removal of aberrant transcripts.

Moreover, because eukaryotic mRNAs lack both SD sequences and RPS1 protein, they cannot efficiently recruit the ribosome directly to the initiation codon. Most probably this was the most important selection pressure that led early eukaryotes to develop a novel mechanism to ensure the correct landing of the ribosome at the 5′-end of mRNAs, i.e., the cap-dependent initiation. These events led to the stepwise increase in sophistication during eukaryogenesis. Hence, although eukaryotes inherited from their archaeal ancestors eIF1, eIF1A, eIF2 (all subunits), eIF2B (but only the alpha, beta and delta and not the gamma or epsilon subunits), eIF4A, eIF5B and eIF6 [9, 59, 84, 85, 135–137], eIF3, eIF4G, eIF4E, eIF4B and PABP evolved exclusively in eukaryotes because of the need to recruit capped and polyadenylated transcripts possessing long 5′-UTRs devoid of SD sequences [3, 4, 6, 9, 14, 59, 84, 85]. Thus, the crucial question arises of how ribosomes from early eukaryotes might have recruited mRNAs to initiate translation in the absence of both eIF4 factors and PABP.

5.1 What Was the Mechanism of mRNA Recruiting in the Early Eukaryotes?

Phylogenomic analyses have recently shown that eukaryotes emerged from the so-called TACK *superphylum* within the Archaea domain, which comprises the Thaum-, Aig-, Cren- and Korarcaeota groups of archaea [74, 77, 81–83]. This means that the closest relative of the eukaryotic lineage is among the species of the *superphylum* TACK. Therefore, a close look to the mRNA structure of these lineages might shed light on the type of mRNA (i.e., SD-containing, leaderless or possessing a 5′-UTR devoid of SD motif) the first eukaryotes might have possessed. However, the current knowledge does not allow elucidating what species of the *superphylum* TACK eukaryotes evolved from, as well as what type of mRNA these particular species use.

Although genome-wide studies of hundreds of species have shown that the major pathway to initiate translation in Archaea might involve mostly leaderless mRNAs [26, 27, 29, 30, 33–37, 42], early eukaryotes might have synthesized transcripts possessing long 5′-UTRs devoid of the SD motif as happens in virtually all present-day eukaryotes. Based on this notion, and given the fact that archaea and eukaryotes lack an *RPS1* gene involved in recruitment of bacterial leadered mRNAs devoid of an SD motif, here we propose three possible mechanisms for mRNA recruitment by ribosome in early eukaryotes.

(1) *mRNAs used a variety of non-SD sequences that interacted with different internal regions of the rRNAs on the 30S ribosomal subunit*. This idea is supported by evidence proving sequence complementarity and interaction between hundreds of mRNAs and different segments of the 18S and 28S rRNA from different eukaryote species with a potential role in translation regulation [138–141]. Indeed, RPS1 is also missing in some bacterial lineages, which led G.E. Fox to suggest that RPS1 was added to bacterial ribosomes only after the Archaea-Bacteria divergence happened [55].

(2) *Alternate ribosomal proteins might have been responsible for mRNA recruitment.* Eukaryotes evolved a whole set of novel, eukaryote-specific ribosomal proteins [18, 86–88], making conceivable that some of them might have evolved because of enhanced mRNA recruitment. For example, ribosomal proteins such as RPS5 and RPS15, which along with eIF2α contact mRNA positions −3 and +4 of the AUG context sequence [142], might have been involved in driving mRNA recruitment in early eukaryotes.

(3) *Existing initiation factors promoted mRNA recruitment.* RPS1 contains six copies of an RNA-binding fragment that is known as the S1 domain. Many proteins possess one or more S1 domains, including the translation initiation factor IF1 and its eukaryotic equivalent eIF1A, as well as the eukaryotic eIF2α [18, 55]. Since the S1 motif is found in all three domains of life and factors IF-1/eIF1A are universally distributed, Fox [55] has suggested that IF-1/eIF1A might be the original source of

the S1 motif, possibly derived from the initiation machinery. Thus, in the absence of the RSP1 and SD motif, eIF1A or eIF2α might have been involved in mRNA recruitment in early eukaryotes.

6 The Natural History of the Cap Structure, eIF4s and PABP Sheds Light on the Evolution of the Cap-Dependent Translation

Because present-day cap-dependent translation is a highly sophisticated process, it cannot have appeared fully formed, but arose by stepwise addition of components and regulatory steps. So, what were the possible mechanisms underlying evolution of translation initiation in early eukaryotes? As with all evolutionary studies, we can infer the ancient nature of any current biological process by studying its present-day components and looking at their "ancestral" features. Here we think that the contemporary features of the cap structure, eIF4G, eIF4E eIF3 and PABP, all of them of eukaryotic origin, as well as the more ancient eIF4A, shed light on the evolutionary history of eukaryotic translation initiation. Analysis of these molecules argues for a stepwise addition of factors into the initiation step of translation by a mechanism of molecular tinkering [143], i.e., by recruiting more ancient components from other, already present cellular processes to perform a novel function into translation initiation.

Francoise Jacob first proposed the concept of "molecular tinkering" 40 years ago to explain one of the most creative forces of evolution, i.e., transforming a feature that evolved to perform a specific function to give it new functions [143]. This concept was afterwards applied by Gould and Vrba in 1982 to the evolution of morphological features that now enhance fitness but were not built by natural selection for their current role. For them, a morphological feature or structure, previously shaped by natural selection for a specific function (an adaptation), but later utilized for a new use is called a "exaptation" [144]. In the following, we analyze current features of different molecules to infer their evolutionary history and, finally, reconstruct the whole evolutionary history of the translation initiation in eukaryotes. The evidence supports the notion that some of the eukaryotic initiation factors are indeed molecular exaptations.

6.1 Origin of the Cap Structure of mRNAs

The m^7GpppN cap structure of eukaryotic mRNAs plays a crucial role in mRNA biogenesis and stability. It is essential for efficient splicing, mRNA export and translation. Interestingly, all nuclear processes of mRNA biogenesis (namely transcription, capping, polyadenylation, splicing nuclear export and stability) are

tightly intertwined [145–152]. During transcription, which is performed by RNA polymerase II (Pol II), the cap addition is the first modification that occurs to all eukaryotic pre-mRNAs. It is co-transcriptionally added after 20–30 nucleotides have been polymerized in virtue of the interaction of the capping enzymes with the carboxyl-terminal domain (CTD) of the largest subunit of Pol II. Once transcription reaches the transcript's end, a signal triggers polyadenylation of the pre-mRNA by the poly(A) polymerase and, right after the transcript is released, a process that is dependent on the presence of Pol II CTD. The synthesized transcript is then recognized by the nuclear cap-binding protein CBP20 in complex with CBP80 (forming the so-called nuclear CBC) for both intron splicing and nucleo-cytoplasm export to happen. Upon phosphorylation, Pol II CTD enhances the overall rate of splicing.

The extensive coupling of all process for mRNA biogenesis [145–149, 151, 152], the finding that the cap structure is recognized by many proteins belonging to different processes of RNA metabolism [149, 153] and the discovery of a strong dependence of most of mRNA degradation pathways on the cap structure (namely AU-rich element decay, bulk 5′–3′ decay, NMD, miRNA-mediated decay, and deadenylation-mediated mRNA decay) [147–149] support the hypothesis that the cap structure has been involved in different aspects of RNA metabolism ever since eukaryotes originated. It also supports the idea that among the very first components and processes that appeared in eukaryotes were the Pol II CTD, the cap structure and the CBC to provide a "platform" to assemble the splicing, nuclear export, mRNA protection and NMD machineries [3, 4]. Thus, both the cap and poly (A) tail of mRNAs might have played no role in translation during eukaryogenesis, being incorporated into the translation process later in evolution only after eIF4E, eIF4G and PABP had evolved.

6.2 Origin of Eukaryotic Initiation Factors 3, 4G and 4E

The scaffold eIF3 is the largest and functionally most complex of initiation factors, with a composition across eukaryotes from 6 to 13 different subunits. Among its activities, eIF3 binds to and coordinates the interaction between eIF4G and the 40S ribosomal subunit, thereby enhancing most of the reactions of the translation initiation pathway [154–157]. Structural and sequence studies have shown that eIF3, the 'lid' subcomplex of the 26S proteasome (involved in protein degradation) and the COP9 signalosome or CSN complex (involved in the ubiquitin-proteasome pathway, DNA-damage response and cell cycle control) share a similar architecture composed of multiple subunits possessing the PCI domain (Proteasome, CSN, eIF3) [156–161]. Since PCI proteins are crucial scaffolds for the assembly of multiprotein complexes, these observations support the hypothesis that an ancestral core of eIF3 evolved from a versatile PCI-containing multimeric complex involved in different cellular processes other than translation. The finding that some eIF3 subunits also play roles not related to translation, such as the cell cycle, apoptosis,

protein turnover, mRNA deadenylation or decay, 20S pre-rRNA processing or NMD [158–161], supports this hypothesis. Thus, an ancestral, multisubunit eIF3 was perhaps a scaffold that gradually incorporated additional subunits from other cellular machineries and was incorporated into translation initiation later in evolution because it improved the efficiency and regulation of the mRNA recruitment [4].

eIF4G is a scaffold, modular protein that possess binding sites for different proteins involved in translation initiation, such as PABP, eIF4E, eIF3 and eIF4A. The C-terminal third of all eIF4Gs contains of one, two or three consecutive α-helical domains called HEAT (Huntington, Elongation factor 3, A subunit of protein phosphatase 2A, and Target of rapamycin) [162, 163]. Homologs of the HEAT domain named HEAT-1 [162] also exist in Upf2/NMD2, a component of the NMD system, and in CBP80, indicating that they may have evolved from a common ancestral protein [85, 162, 164, 165]. For instance, the consecutive HEAT-1, HEAT-2 and HEAT-3 domains of eIF4G are present in CBP80 as well, meaning that both proteins descended from an ancestor protein that already contained the three consecutive HEAT domains [162].

HEAT-containing proteins participate in a wide variety of cellular processes that are dependent on assembling large multiprotein complexes [166, 167]. HEAT domains are part of central adapters driving processes such as mRNA processing, translation and degradation [85]. Since the complexes eIF4F, NMD and nuclear CBC each include a HEAT-1-containing protein (eIF4G, Upf2/NMD and CBP80, respectively) [164], it has been suggested that early in eukaryotic evolution a versatile ancestral protein containing the HEAT-1 domain served as an adapter in different RNA processes that subsequently diverged and evolved toward distinct binding specificities [85, 162, 164, 165]. Therefore, this protein may have first appeared in the nucleus as a proto-CBP80 to provide, together with the cap, a "platform" for splicing factors and for mRNA protection during nuclear export. Later in evolution, it might have diverged in the cytoplasm into the Upf2/NMD2 when NMD evolved, and also into a proto-eIF4G, a scaffold that facilitated a more efficient initiation of translation by bringing the mRNAs into the close proximity of the ribosomes. Therefore, and similar to eIF3, these features suggest that eIF4G might have appeared in early eukaryotes for functions different from in translation and that it was incorporated into this process later in evolution because it also conferred a better efficiency of mRNA recruitment [3, 4, 85, 162, 164, 165]. Cap-dependent initiation of translation could only then have evolved after sites to bind eIF4E, PABP, eIF4A and eIF3 appeared in the proto-eIF4G [3, 4].

eIF4E has long been known to play its main role in translation initiation through cap recognition [168] and is also of eukaryotic origin [85]. Interestingly, eIF4E is found being part of different cytoplasmic granules where it is involved in mRNA decay or storage [169, 170]. In addition, a fraction of this protein localizes inside the nucleus in several eukaryotes where it mediates the export of certain mRNAs to the cytoplasm [171–173]. These findings suggest that eIF4E is versatile enough to utilize the features required for cap-binding activity in different cellular processes [4, 171].

Although most probably eIF4E emerged as a translation factor, it has been discussed that other evolutionary scenarios are also possible [4, 171]. For instance, it may first have appeared in early eukaryotes either as a mediator of nuclear export, thus enhancing mRNA stability, or as a mediator of cytoplasmic storage of mRNAs, but playing no role in translation [4, 171]. An example of this possible scenario is provided by one of the eIF4E isoforms from *Giardia lamblia*, as it binds only to nuclear noncoding small RNAs and plays no role in translation [174]. The findings that the cap and eIF4E confer stability to mRNAs by protecting them from 5′ exonucleases and decapping enzymes [175] suggest that the appearance of both the cap and eIF4E could have been a big evolutionary leap by protecting mRNAs from degradation. Since 5′ exoribonucleases emerged after eukaryogenesis [176], and the enzymes for the for capping of mRNAs, namely 5′ triphosphatase, guanylyltransferase and guanine-N7-methyl-transferase are of eukaryotic origin [177, 178], it was suggested that the 5′ exoribonucleases evolved in early eukaryotes following the emergence of mRNA capping for cell protection from RNA viruses or viroids [177]. The appearance of eIF4E could have followed this evolution by further increasing mRNA stability, since in the absence of any means of interacting directly with the ribosome itself, it could not be involved in translation. eIF4E should have been incorporated into the translation process only after a scaffold protein emerged (namely eIF4G), able to coordinate eIF4E activity. Because the absence of eIF4E precludes the existence of the cap-dependent translation, the emergence of the ancestral eIF4E implies that its own mRNA was most likely translated in a cap-independent, IRES-dependent manner [3, 4, 171].

6.3 *Origin of PABP and the Evolution of mRNA Circularization*

PABPs are scaffold proteins of eukaryotic origin that evolved into two main families, nuclear and cytoplasmic. They interact with many proteins and participate in different events of mRNA biogenesis both inside the nucleus and in the cytoplasm. In the nucleus, PABPs play essential roles in mRNA polyadenylation and stability, and they may be involved in the mRNA shuttle to the cytoplasm. In the cytoplasm, PABPs either protect mRNAs from decay or trigger transcript decay by promoting mRNA interactions with deadenylase complex proteins. By interacting with eIF4G, PABP also promotes circularization of the mRNA, a conformation that is critical for translation initiation since it provides an effective means for the protein synthesis apparatus to selectively translate only intact mRNAs, i.e., those that harbor both a cap and a poly(A) tail. In addition, translation termination happens at a ‘correct’ stop codon, as opposed to a premature termination codon, only if the ribosome is close enough to the poly(A) tail. The signal indicating this proximity is the interaction of the terminating ribosome with PABP. In the absence of this signal, upframeshift protein (UPF) 1 binds eukaryotic releasing factors (eRF) 1 and 3 in the

terminating ribosome, triggering NMD. Finally, PABPs also play a role in mRNA transport and localization [179–184].

PABPs interact with poly(A) tails via their RNA-recognition motifs (RRMs). These are present in one to four repeats plus a carboxy-terminal domain (CTD) that interacts with factors regulating translation initiation and termination, polyadenylation and deadenylation [182]. The RRM is the most prevalent eukaryotic RNA-binding domain and is involved in all aspects of RNA metabolism. This is an ancient and versatile RNA-binding domain present in all eukaryotes and many bacteria [185]. RRM-containing proteins, including PABPs, evolved from successive duplications of a single RRM-carrying gene with the addition of auxiliary motifs during their diversification in eukaryotes [185].

Hernández has proposed that the poly(A) tail and an PABP first arose in early eukaryotes as part of the primary adaptive responses to the emergence of nuclear membrane and split genes, but initially they might have had no role in translation [4, 6]. Afterwards, mutations in PABP that allowed binding to eIF4G, thereby promoting mRNA circularization, underwent a strong positive selection because they (1) increased mRNA stability, (2) ensured a more efficient recruitment of the 40S ribosomal subunit by the mRNA and (3) mRNA circularity represents a checkpoint that determines to initiate translation only in intact mRNAs [4, 6].

6.4 *Origin of eIF4A and the Evolution of the Scanning Process*

Sequence comparison and biochemical analyses show that eIF4A is the most ancient of eIF4 factors. Orthologs are found in Archaea [9, 14, 59, 85, 186], Bacteria [187–189] and Eukaryotes [171, 186, 190–193], indicating that eIF4A evolved before eukaryotes appeared. eIF4A belongs to the extensive DEAD-box family of RNA proteins, which is a wide and versatile family of ATP-dependent RNA helicases that exists across all phyla of Bacteria, Archaea and Eukaryotes and that is involved in many aspects of RNA metabolism, including translation, nonsense-mediated mRNA decay (NMD), splicing, RNA transport and ribosomal biogenesis [190–192]. This indicates that RNA unwinding by RNA helicases already existed before the eukaryotes appeared and that eIF4A evolved from RNA helicases already present in the archaeal ancestor of eukaryotes [3, 4, 6].

Eukaryotes possess mRNAs with long 5′-UTRs with energetically stable secondary structures that would prevent scanning and hence translation. Therefore, the translation machinery requires RNA helicases to unwind these structures. In contrast to bacterial ribosomes, which possess intrinsic mRNA helicase activity [194], in eukaryotes RNA unwinding is mainly performed by eIF4A. Remarkably, other RNA helicases belonging to the asp-glu-ala-asp (DEAD)-box or DEAD/asp-glu-x-his (DExH)-box families also stimulate or repress translation by performing RNA unwinding during different steps of translation initiation, including the scanning

step. This is the case of helicases DDX3/Ded1, Dhh1/RCK, VASA/DDX4, RHA/DHX9 and DHX29. Interestingly, these helicases also play various roles in different processes of RNA metabolism other than translation, such as RNA export and pre-mRNA splicing and transport [190, 192, 193]. Since the RNA helicases are a family of proteins that participate in many processes of RNA metabolism in both the nucleus and the cytoplasm [191], it has been proposed that the early eukaryotic RNA helicases were versatile proteins with broad substrate specificities involved in different RNA processes, and this probably included translation initiation. Later in evolution, they diversified into more specific enzymes, some of them specializing in translation [3, 4, 171]. The finding that eIF4A-III, a highly related eIF4A-cognate, participates in NMD, RNA splicing and mRNA localization, but not in translation [192], supports this hypothesis.

Thus, an evolutionary scenario is possible where a proto-eIF4A with broad substrate specificity might have existed, performing its function in diverse aspects of the RNA metabolism, from which it was afterwards incorporated into translation. Crucially, the evolution of eIF4G and the incorporation of diverse RNA helicases, including a proto-eIF4A, into translation initiation allowed both the incorporation of eIF4E and the establishment of the scanning process in the translation mechanism. These events enabled the translation machinery to efficiently translate mRNAs with more complex 5′-UTRs, resulting in the current widespread cap-dependent translation initiation mechanism.

7 A Timeline for the Emergence of the Cap-Dependent Translation Initiation

We can summarize the evidence discussed above and outline a brief timeline hypothesis on the origin and early evolution of the cap-dependent translation initiation in early eukaryotes. *Overall, the evidence discussed in this chapter supports the notion that molecular tinkering* [143] *has played a crucial role underlying the establishment of the cap-dependent initiation of translation,* i.e., *by gradually recruiting into translation more ancient, already existing molecules involved in different cellular processes.* This notion is supported by the current existence of a diversity of viruses performing translation with a wide variety of requirements of the translation factors that, indeed, might represent intermediary steps of this evolutionary process [3, 4, 6].

Hernández (4) has proposed that upon eukaryote emergence, perhaps there was a transition period before the arousal of the cap-dependent translation when monocistronic mRNAs with long 5′-UTRs and devoid of SD sequences recruited the 40S ribosomal subunit in a cap-independent manner and in the absence of eIF3, PABP and eIF4 factors, becoming thus the first examples of an IRES. In other words, early eukaryotes inherited a functional translational apparatus from archaeal ancestors that recruited mRNAs in a cap-independent, IRES-dependent manner. The cap

structure and the poly(A) tail of mRNA, as well as a PABP and perhaps eIF4E, already existed, but they played no role in translation. They might have appeared for functions in RNA metabolism that emerged among the primary adaptive responses to the emergence of the nuclear membrane (i.e., the need for nucleocytoplasmic mRNA export and protection) and the appearance of interrupted genes, but initially had no role in translation [3, 4]. In this scenario, *present-day IRES are rather relicts of the past* [3]. *Discistroviridae* IRESs represent an example of the minimal level of complexity in terms of dependence on proteins to initiate translation (185). They show that some mRNAs could drive recognition of the AUG start site by the ribosome in the total absence of other factors, including tRNA. For other mRNAs, at least eIF2 and eIF5B of archaeal origin, were involved in binding of the Met-tRNA$_i^{Met}$ to the initiator codon and the assembly of 80S complexes, respectively, as the mechanism used by the some picornaviruses to initiate translation, such as the porcine teschovirus type 1 (186). In this virus, the 40S ribosomal subunit can actually be recruited directly to its mRNA by an IRES with only the further requirement of the eIF2- GTP-Met-tRNA$_i^{Met}$ ternary complex for 48S pre-initiation complex formation.

The incorporation into translation of novel scaffold molecules with coordinator abilities, such as an ancestral HEAT-containing domain protein (a proto-eIF4G), perhaps picked up from other cellular processes such as NMD or mRNA nuclear export, further improved the efficiency and regulation of the ribosomal subunit recruitment by the mRNA. Evidence for this possible evolutionary stage is provided by the translation driven by the encephalomyocarditis virus and other picornavirus IRESs, which requires nearly all the canonical initiation factors and the middle part of eIF4G, but neither eIF4E nor the cap structure is required [195, 196].

Later on in evolution, a minimal core of eIF3 (i.e., a proto-eIF3) could have been derived from other, more ancient cellular processes such as the ubiquitin-proteasome and protein degradation pathways and incorporated into translation. Translation initiation thus became more dependent on new factors like eIF3, which by bridging eIF4G and the 40S ribosomal subunit enhanced the efficiency and accuracy of mRNA recruitment. This hypothetical evolutionary stage is similar to what happens in the translation of messages from different viruses, including hepatitis C virus, pestiviruses and *Rhopalosiphum padi virus*, where direct binding of the 40S ribosomal subunit to the mRNA is driven by the IRES [197, 198]. HCV and pestivirus mRNAs have the additional requirement of eIF3 and eIF2-GTP-Met-tRNA$_i^{Met}$ ternary complex to form the 48S-initiation complex. In *Rhopalosiphum padi virus* mRNA, the binding of the 40S ribosomal subunit absolutely requires eIF3, but it occurs in the absence of the eIF4 group of factors [197, 198].

In all evolutionary stages, existing proto-eIF4A helicases, perhaps performing activity in different RNA metabolism activities, could help RNA unwinding. The incorporation of a proto-eIF4A along with eIF4E improved both the efficiency and the regulatory possibilities of mRNA recruitment even more, leading ultimately to the cap-dependent mechanism to initiate translation.

8 Concluding Remarks

One of the enigmas of modern biology is how eukaryotic translation emerged. We have discussed evidence supporting the notion that tinkering [143] might have played a crucial role in the origin and evolution of the cap-binding mechanism in eukaryotes [4, 6]. According to Jacob, "*…natural selection does not work as an engineer works. It works like a tinkerer—a tinkerer who does not know exactly what he is going to produce but uses whatever he finds around him whether it be pieces of string, fragments of wood, or old cardboards*" [143]. "*…Evolution would slowly modify his work, unceasingly retouching it, cutting here, lengthening there, seizing the opportunities to adapt it progressively to its new use…It works on what already exists, either transforming a system to give it new functions of combining several systems to produce a more elaborate one*" [143].

We have discussed that early eukaryotes inherited a core of translation machinery and that, in the absence of SD sequences in mRNAs and RPS1 in ribosomes, the first eukaryotic mRNAs were translated in a cap-independent, IRES-driven manner that was then superseded in evolution by the cap-dependent mechanism, rather than vice versa. Thus, the contemporary cellular IRESs might be relics of the past. This hypothesis is supported by the observations that (1) IRES-dependent, but not cap-dependent translation can take place in the absence of not only a cap, but also many initiation factors and (2) eIF4E and eIF4G, molecules absolutely required for cap-dependent translation, are among the most recently evolved translation factors.

Afterwards, the evolution of the translation machinery followed a gradual addition of scaffold proteins, namely eIF3, eIF4G, PABP as well of eIF4A and eIF4E, which highly improved the efficiency and regulation of mRNA binding to the 40S ribosomal subunit [3, 4]. Indeed, eIF3, eIF4G, eIF4A, PABP, the cap structure and the polyadenylation of mRNAs and perhaps eIF4E might be molecular exaptations. The rudiments of these molecules might have first arisen during eukaryogenesis with no role in translation before the cap-dependent initiation of translation appeared, performing activities other than translation, perhaps involved in mRNA nuclear export, splicing and stability, and were gradually added into the initiation of translation by a process of molecular tinkering later in evolution [143]. The diversity of viruses infecting present-day cells with a variety of needs of translation factors and cap that might represent the different evolutionary steps discussed here supports this hypothesis.

Finally, there are still many open questions on the evolution of translation in early eukaryotes. For example, we still lack satisfying explanations for the evolutionary origin of monocistronic transcripts, for the mechanism of mRNA recruiting in the early eukaryotes, for the origin of most ribosomal proteins and RNA extensions of rRNAs, and for the archaeal lineage that originated the early eukaryotes.

Acknowledgments This work was supported by the National Institute of Cancer (INCan), Mexico, and by the National Council of Science and Technology (CONACyT, grant no. 168154 to G.H.), Mexico.

References

1. Mathews MB, Sonenberg N, Hershey JWB. Origins and principles of translational control. In: Sonenberg N, Hershey JWB, Mathews MB, editors. Translational control of gene expression. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press; 2000. p. 1–31.
2. Mathews MB, Sonenberg N, Hershey JWB. Origins and principles of translational control. In: Mathews MB, Sonenberg N, Hershey JWB, editors. Translational control in biology and medicine. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press; 2007. p. 1–40.
3. Hernández G. Was the initiation of translation in early eukaryotes IRES-driven? Trends Biochem Sci. 2008;33:58–64.
4. Hernández G. On the origin of the cap-dependent initiation of translation in eukaryotes. Trends Biochem Sci. 2009;34:166–75.
5. Hernández G. On the emergence and evolution of the eukaryotic translation apparatus. In: Biyani M, editor. Cell-free protein synthesis. Rijekra, Croatia: InTech; 2012. p. 31–50.
6. Hernández G, Altmann M, Lasko P. Origins and evolution of the mechanisms regulating translation initiation in eukaryotes. Trends Biochem Sci. 2010;35:63–73.
7. Hernández G, Proud CG, Preiss T, Parsyan A. On the diversification of the translation apparatus across eukaryotes. Comp Funct Genom. 2012;2012:256848.
8. Benelli D, Londei P. Translation initiation in Archaea: conserved and domain-specific features. Biochem Soc Trans. 2011;19:89–93.
9. Londei P. Evolution of translational initiation: news insights from the archaea. FEMS Microbiol Rev. 2005;29:185–200.
10. Laursen BS, Sørensen HP, Mortensen KK, Sperling-Petersen HU. Initiation of protein synthesis in bacteria. Microbiol Mol Biol Rev. 2005;69:101–23.
11. Shine J, Dalgarno L. The 3′-terminal sequence of Escherichia coli 16S ribosomal RNA: complementary to nonsense triplets and ribosome binding sites. Proc Natl Acad Sci USA. 1974;71:1342–6.
12. Shine J, Dalgarno L. Determinant of cistron specificity in bacterial ribosomes. Nature. 1975;254:34–8.
13. Jacob WF, Santer M, Dahlberg AE. A single base change in the Shine-Dalgarno region of 16S rRNA of Escherichia coli affects translation of many proteins. Proc Natl Acad Sci USA. 1987;84:4757–61.
14. Dennis PP. Ancient ciphers: translation in archaea. Cell. 1997;89:1007–10.
15. Band L, Henner DJ. Bacillus subtilis requires a “stringent” Shine-Dalgarno region for gene expression. DNA. 1984;3:17–21.
16. Steitz JA, Jakes K. How ribosomes select initiator regions in mRNA: base-pair formation between the 3’ terminus of 16 s rRNA and the mRNA during initiation of protein synthesis in Escherichia coli. Proc Natl Acad Sci USA. 1975;72:4734–8.
17. Jackson RJ. A comparative view of initiation site selection mechanisms. In: Sonenberg N, Hershey JWB, Mathews MB, editors. Translational control of gene expression. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory press; 2000. p. 127–83.
18. Wilson DN, Nierhaus KH. Ribosomal proteins in the spotlight. Crit Rev Biochem Mol Biol. 2005;40:243–67.

19. Moll I, Grill S, Gründling A, Bläsi U. Effects of ribosomal proteins S1, S2 and the DeaD/CsdA DEAD-box helicase on translation of leaderless and canonical mRNAs in Escherichia coli. Mol Microbiol. 2002;44:1387–96.
20. Delvillani F, Papiani G, Dehò G, Briani F. S1 ribosomal protein and the interplay between translation and mRNA decay. Nucleic Acid Res. 2011;39:7702–15.
21. Sengupta J, Agrawal RK, Frank J. Visualization of protein S1 within the 30S ribosomal subunit and its interaction with messenger RNA. Proc Natl Acad Sci USA. 2001;98:11991–6.
22. Kaberdin VR, Bläsi U. Translation initiation and the fate of bacterial mRNAs. FEMS Microbiol Rev. 2006;30:967–79.
23. Komarova AV, Tchufistova LS, Supina EV, Boni IV. Protein S1 counteracts the inhibitory effect of the extended Shine-Dalgarno sequence on translation. RNA. 2002;8:1137–47.
24. Shell SS, Wang J, Lapierre P, Mir M, Chase MR, Pyle MM, Gawande R, Ahmad R, Sarracino DA, Ioerger TR, Fortune SM, Derbyshire KM, Wade JT, Gray T. Leaderless transcripts and small proteins are common features of the mycobacterial translational landscape. PLoS Genet. 2015;11:e1005641.
25. Ma J, Campbell A, Karlin S. Correlation between Shine-Dalgarno sequence and gene features such as predicted expression levels and operon structure. J Bacteriol. 2002;184:5733–45.
26. Torarinsson E, Klenk HP, Garret RA. Divergent transcriptional and translational signals in Archaea. Environ Microbiol. 2005;7:45–54.
27. Sensen CW. Organizational characteristics and information content of an archaeal genome: 156 kb of sequence from Sulfolobus solfataricus P2. Mol Microbiol. 1996;22:175–91.
28. Weiner J, Herrmann R, Browning GF. Transcription in Mycoplasma pneumoniae. Nucleic Acid Res. 2000;28:4488–96.
29. Chang B, Halgamuge S, Tang SL. Analysis of SD sequences in completed microbial genomes: non-SD-led genes are as common as SD-led genes. Gene. 2006;373:90–9.
30. Nakagawa S, Niimura Y, Miura KI, Gojobori T. Dynamic evolution of translation initiation mechanisms in prokaryotes. Proc Natl Acad Sci USA. 2010;107:6382–7.
31. Benelli D, Londei P. Begin at the beginning: evolution of translational initiation. Res Microbiol. 2009;160:493–501.
32. Brenneis M, Hering O, Lange C, Soppa J. Experimental characterization of Cis-acting elements important for translation and transcription in halophilic archaea. PLoS Genet. 2007;3:e229.
33. Zheng X, Hu GQ, She ZS, Zhu H. Leaderless genes in bacteria: clue to the evolution of translation initiation mechanisms in prokaryotes. BMC Genom. 2011;12:361–70.
34. Slupska MM, King AG, Fitz-Gibbon S, Besemer J, Borodovsky M, Miller JH. Leaderless transcripts of the crenarchaeal hyperthermophile Pyrobaculum aerophilum. J Mol Biol. 2001;309:347–60.
35. Tolstrup N, Sensen CW, Garrett RA, Clausen IG. Two different and highly organized mechanisms of translation initiation in the archaeon Sulfolobus solfataricus. Extremophiles. 2000;4:175–9.
36. Jäger D, Förstner KU, Sharma CM, Santangelo TJ, Reeve J. Primary transcriptome map of the hyperthermophilic archaeon Thermococcus kodakarensis. BMC Genom. 2014;15:684.
37. Cortes T, Schubert OT, Rose G, Arnvig KB, Comas I, Aebersold R, Young DB. Genome-wide mapping of transcriptional start sites defines an extensive leaderless transcriptome in Mycobacterium tuberculosis. Cell Rep. 2013;5:1121–31.
38. Kramer P, Gäbel K, Pfeiffer F, Soppa J. Haloferax volcanii, a prokaryotic species that does not use the Shine Dalgarno mechanism for translation initiation at 5'-UTRs. PLoS ONE. 2014;9:e94979.
39. Sullivan MJ, Curson AR, Shearer N, Todd JD, Green RT, Johnston AW. Unusual regulation of a leaderless operon involved in the catabolism of dimethylsulfoniopropionate in Rhodobacter sphaeroides. PLoS ONE. 2011;6:e15972.

40. Krishnan KM, Van Etten WJ, Janssen GR. Proximity of the start codon to a leaderless mRNA's 5' terminus is a strong positive determinant of ribosome binding and expression in Escherichia coli. J Bacteriol. 2010;192:6482–5.
41. Vesper O, Amitai S, Belitsky M, Byrgazov K, Kaberdina AC, Engelberg-Kulka H, Moll I. Selective translation of leaderless mRNAs by specialized ribosomes generated by MazF in Escherichia coli. Cell. 2011;147:147–57.
42. La Teana A, Benelli D, Londei P, Bläsi U. Translation initiation in the crenarchaeon Sulfolobus solfataricus: eukaryotic features but bacterial route. Biochem Soc Trans. 2013;41:350–5.
43. Hering O, Brenneis M, Beer J, Suess B, Soppa J. A novel mechanism for translation initiation operates in haloarchaea. Mol Microbiol. 2009;71:1451–63.
44. Omotajo D, Tate T, Cho H, Choudhary M. Distribution and diversity of ribosome binding sites in prokaryotic genomes. BMC Genom. 2015;16:e604.
45. Balakin AG, Skripkin EA, Shatsky IN, Bogdanov AA. Unusual ribosome binding properties of mRNA encoding bacteriophage lambda repressor. Nucleic Acid Res. 1992;20:563–71.
46. Brock JE, Pourshahian S, Giliberti J, Limbach PA, Janssen G. Ribosomes bind leaderless mRNA in Escherichia coli through recognition of their 5′-terminal AUG. RNA. 2008;14:2159–69.
47. Van Etten WJ, Janssen GR. An AUG initiation codon, not codon-anticodon complementarity, is required for the translation of unleadered mRNA in Escherichia coli. Mol Microbiol. 1998;27:987–1001.
48. Grill S, Gualerzi CO, Londei P, Bläsi U. Selective stimulation of translation of leaderless mRNA by initiation factor 2: evolutionary implications for translation. EMBO J. 2000;19:4101–10.
49. O'Donnell SM, Janssen GR. The initiation codon affects ribosome binding and translational efficiency in Escherichia coli of cI mRNA with or without the 5' untranslated leader. J Bacteriol. 2001;183:1277–83.
50. O'Donnell SM, Janssen GR. Leaderless mRNAs bind 70S ribosomes more strongly than 30S ribosomal subunits in Escherichia coli. J Bacteriol. 2002;184:6730–3.
51. Moll I, Grill S, Gualerzi CO, Bläsi U. Leaderless mRNAs in bacteria: surprises in ribosomal recruitment and translational control. Mol Microbiol. 2002;43:239–46.
52. Moll I, Hirokawa G, Kiel MC, Kaji A, Bläsi U. Translation initiation with 70S ribosomes: an alternative pathway for leaderless mRNAs. Nucleic Acid Res. 2004;32:3354–63.
53. Kudla G, Murray AW, Tollervey D, Plotkin JB. Coding-sequence determinants of gene expression in Escherichia coli. Science. 2009;324:255–8.
54. Scharff LB, Childs L, Walther D, Bock R. Local absence of secondary structure permits translation of mRNAs that lack ribosome-binding site. PLoS Genet. 2011;7:e1002155.
55. Fox GE. Origin and evolution of the ribosome. Cold Spring Harb Perspect Biol. 2012;2: e003483.
56. Golshani A, Krogan NJ, Xu J, Pacal M, Yang XC, Ivanov I, Providenti MA, Ganoza MC, Ivanov IG. Abou Haidar MG. Escherichia coli mRNAs with strong Shine/Dalgarno sequences also contain 5' end sequences complementary to domain #17 on the 16S ribosomal RNA. Biochem Biophys Res Commun. 2004;316:978–83.
57. Loechel S, Inamine JM, Hu PC. A novel translation initiation region from Mycoplasma genitalium that functions in Escherichia coli. Nucleic Acids Res. 1991;19:6905–11.
58. Jenner L, Romby P, Rees B, Schulze-Briese C, Springer M, Ehresmann C, Ehresmann B, Moras D, Yusupova G, Yusupov M. Translational operator of mRNA on the ribosome: how repressor proteins exclude ribosome binding. Science. 2005;308:120–3.
59. Kyrpides NC, Woese CR. Universally conserved translation initiation factors. Proc Natl Acad Sci USA. 1998;95:224–8.
60. Koonin EV. The origin and early evolution of eukaryotes in the light of phylogenomics. Genome Biol. 2010;11:209.

61. Koumandou VL, Wickstead B, Ginger ML, van der Giezen M, Dacks JB, Field MC. Molecular paleontology and complexity in the last eukaryotic common ancestor. Crit Rev Biochem Mol Biol. 2013;48:373–96.
62. Woese CR. On the evolution of cells. Proc Natl Acad Sci USA. 2002;99:8742–7.
63. Wu CJ, Janssen GR. Translation of vph mRNA in Streptomyces lividans and Escherichia coli after removal of the 5' untranslated leader. Mol Microbiol. 1996;22:339–55.
64. Sonenberg N, Hinnebusch A. Regulation of translation initiation in eukaryotes: mechanisms and biological targets. Cell. 2009;136:731–45.
65. Jackson RJ, Hellen CU, Pestova TV. The mechanism of eukaryotic translation initiation and principles of its regulation. Nat Rev Mol Cell Biol. 2010;11:113–27.
66. Hinnebusch AG. Molecular mechanism of scanning and start codon selection in eukaryotes. Microbiol Mol Biol Rev. 2011;75:434–67.
67. Pelletier J. Sonenberg N Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. Nature. 1988;334:320–5.
68. Jang SK, Kräusslich HG, Nicklin MJ, Duke GM, Palmenberg AC, Wimmer E. A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. J Virol. 1988;62:2636–43.
69. Macejak DG, Sarnow P. Internal initiation of translation mediated the 5'UTR leader of a cellular mRNA. Nature. 1991;353:90–4.
70. Martínez-Salas E, Piñeiro D, Fernández N. Alternative mechanisms to initiate translation in eukaryotic mRNAs. Comp Funct Genom. 2012;2012:391546.
71. Elroy-Stein O, Merrick WC. Translation initiation via cellular internal ribosome entry sites. In: Mathews MB, Sonenberg N, Hershey JWB, editors. Translational control in biology and medicine. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press; 2007. p. 155–72.
72. Stoneley M, Willis AE. Cellular internal ribosome entry segments: structures, trans-acting factors and regulation of gene expression. Oncogene. 2004;23:3200–7.
73. Dounda JA, Sarnow P. Translation initiation by viral internal ribosome entry sites. In: Mathews MB, Sonenberg N, Hershey JWB, editors. Translational control in biology and medicine. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press; 2007. p. 129–53.
74. Guy L, Saw JH, Ettema TJG. The archaeal legacy of eukaryotes: a phylogenomic perspective. In: Keeling PJ, Koonin EV, editors. The origin and evolution of eukaryotes. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press; 2014. p. 97–128.
75. Poole AM, Gribaldo S. Eukaryotic origins: how and when was the mitochondrion acquired? In: Keeling PJ, Koonin EV, editors. The origin and evolution of eukaryotes. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press; 2014. p. 129–40.
76. McFadden GI. Origin and evolution of plastids and photosynthesis in eukaryotes. In: Keeling PJ, Koonin EV, editors. The origin and evolution of eukaryotes. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press; 2014. p. 263–70.
77. Williams TA, Foster PG, Cox CJ, Embley M. An archaeal origin of eukaryotes supports only two primary domains of life. Nature. 2013;504:231–6.
78. Woese CR, Fox GE. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proc Natl Acad Sci USA. 1977;74:5088–90.
79. Woese CR, Fox GE. The concept of cellular evolution. J Mol Evol. 1977;10:1–6.
80. Woese CR. Interpreting the universal phylogenetic tree. Proc Natl Acad Sci USA. 2000;97:8392–6.
81. Cox CJ, Foster PG, Hirt RP, Harris SR, Embley TM. The archaebacterial origin of eukaryotes. Proc Natl Acad Sci USA. 2008;105:20356–61.
82. Williams TA, Foster PG, Nye TM, Cox CJ, Embley TM. A congruent phylogenomic signal places eukaryotes within the Archaea. Proc Biol Sci. 2012;279:4870–9.
83. Guy L, Ettema TJG. The archaeal 'TACK' superphylum and the origin of eukaryotes. Trends Microbiol. 2011;19:580–7.

84. Pedulla N, Palermo R, Hasenöhr D, Bläsi U, Cammarano P, Londei P. The archaeal eIF2 homologue: functional properties of an ancient translation initiation factor. Nucleic Acid Res. 2005;33:1804–12.
85. Aravind L, Koonin EV. Eukaryotic-specific domains in translation initiation factors: implications for translation regulation and evolution of the translation system. Genome Res. 2000;10:1172–84.
86. Lecompte O, Ripp R, Thierry JC, Moras D, Poch O. Comparative analysis of ribosomal proteins in complete genomes: an example of reductive evolution at the domain scale. Nucleic Acid Res. 2002;30:5382–90.
87. Wilson DN, Cate JHD. The structure and function of the eukaryotic ribosome. In: Hershey JWB, Sonenberg N, Mathews MB, editors. Protein synthesis and translational control. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press; 2012. p. 11–27.
88. Dresios J, Panopoulos P, Synetos D. Eukaryotic ribosomal proteins lacking a eubacterial counterpart: important players in ribosomal function. Mol Microbiol. 2006;59:1651–63.
89. Wolf YI, Aravind L, Grishin N, Koonin EV. Evolution of aminoacyl-tRNA synthetases-analysis of unique domain architectures and phylogenetic trees reveals a complex history of horizontal gene transfer events. Genome Res. 1999;9:689–710.
90. O'Donoghue P, Luthey-Schulten Z. On the Evolution of Structure in Aminoacyl-tRNA Synthetases. Microbiol Mol Biol Rev. 2003;67:550–3.
91. Petrov AS, Bernier CR, Hsiao C, Norris AM, Kovacs NA, Waterbury CC, Stepanov VG, Harvey SC, Fox GE, Wartell RM, Hud NV, Williams LD. Evolution of the ribosome at atomic resolution. Proc Natl Acad Sci USA. 2014;111:10251–6.
92. Roberts E, Sethi A, Montoya J, Woese CR, Luthey-Schulten Z. Molecular signatures of ribosomal evolution. Proc Natl Acad Sci USA. 2008;105:13953–8.
93. Lynch M. The origins of genome architecture. Sunderland, Massachusetts: Sinauer Associates, Inc., Publishers; 2007.
94. Martin W, Koonin EV. Introns and the origin of nucleus-cytosol compartmentalization. Nature. 2006;440:41–5.
95. Taylor DJ, Frank J, Kinzy TG. Structure and function of the eukaryotic ribosome and elongation factors. In: Mathews MB, Sonenberg N, Hershey JWB, editors. Translational control in biology and medicine. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press; 2007. p. 59–85.
96. Yokoyama T, Suzuki T. Ribosomal RNAs are tolerant toward genetic insetions: evolutionary origin of the expansion segments. Nucleic Acid Res. 2008;36:3539–51.
97. Hartman H, Favaretto P, Smith TF. The archaeal origins of the eukaryotic translational system. Archaea. 2006;2:1–9.
98. Klinge S, Voigst-Hoffmann F, Leibundgut M, Ban N. Atomic structures of the eukaryotic ribosome. Trends Biochem. 2012;37:189–98.
99. Kozak M. Regulation of translation via mRNA structure in prokaryotes and eukaryotes. Gene. 2005;361:13–37.
100. Lukaszewicz M. M. F, Jerouville B, Stas A, Boutry M. In vivo evaluation of the context sequence of the translation initiation codon in plants. Plant Sci. 2000;154:89–98.
101. Lutcke HA, Chow KC, Mickel FS, Moss KA, Kern HF, Scheele GA. Selection of AUG initiation codons differs in plants and animals. EMBO J. 1987;6:43–8.
102. Cigan M, Donahue TF. Sequence and structural features associated with translational initiator regions in yeast - a review. Gene. 1987;59:1–18.
103. Donahue TF, Cigan M. Sequence and structural requirements for efficient translation in yeast. Meth Enzymol. 1990;185:366–71.
104. Cavener DR. Eukaryotic start and stop translation sites. Nucleic Acid Res. 1991;19:3185–92.
105. Kim Y, Lee G, Jeon E, Sohn EJ, Lee Y, Kang H, Lee Dw, Kim DH, Hwang I. The immediate upstream region of the 5′-UTR from the AUG start codon has a pronounced effect on the translational efficiency in Arabidopsis thaliana. Nucleic Acid Res. 2014;14:485–98.

106. Kawaguchi R, Bailey-Serres J. mRNA sequence features that contribute to translational regulation in Arabidopsis. Nucleic Acid Res. 2005;33:955–65.
107. Nakagawa S, Niimura Y, Gojobori T, Tanaka H, Miura K. Diversity of preferred nucleotide sequences around the translation initiation codon in eukaryote genomes. Nucleic Acids Res. 2008;36:861–71.
108. Lynch M, Scofield DG, Hong X. The evolution of transcription-initiation sites. Mol Biol Evol. 2005;22:1137–46.
109. Mignone F, Gissi C, Liuni S, Pesole L. Untranslated regions of mRNAs. Genome Biol. 2002;3:1–10.
110. Mazumder B, Seshadri V, Fox PL. Translational control by the 3'-UTR: the ends specify the means. Trends in Biochem Sci. 2003;28:91–8.
111. Hentze MW, Gebauer F, Preiss T. cis-Regulatory sequences and trans-actin factors in translational control. In: Mathews MB, Sonenberg N, Hershey JWB, editors. Translational control in biology and medicine. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press; 2007. p. 269–96.
112. Pesole G, Mignone F, Gissi C, Grillo G, Licciulli F, Liuni S. Structural and functional features of eukaryotic mRNA untranslated regions. Gene. 2001;276:73–81.
113. Svitkin YV, Ovchinnikov LP, Dreyfuss G, Sonenberg N. General RNA binding proteins render translation cap dependent. EMBO J. 1996;15:7147–55.
114. Araujo PR, Yoon K, Ko D, Smith AD, Qiao M, Suresh U, Burns SC, Penalva LOF. Before it gets started: regulating translation at the 5′-UTR. Comp Funct Genom. 2012;12:1–8.
115. Alekhina AM, Vassilenko KS. Translation initiation in eukaryotes: versatility of the scanning model. Biochemistry (Moscow). 2012;77:1465–77.
116. van der Velden A, Thomas AAM. The role of the 5′untranslated region of an mRNA is translation regulation during development. Int J Biochem Cell Biol. 1999;31:87–106.
117. Faye MD, Holcik M. The role of IRES trans-acting factors in carcinogenesis. Biochim Biophys Acta. 2015;1849:887–97.
118. Komar AA, Hatzoglou M. Cellular IRES-mediated translation: the war of ITAFs in pathophysiological states. Cell Cycle. 2011;10:229–40.
119. King HA, Cobbold LC, Willis AE. The role of IRES trans-acting factors in regulating translation initiation. Biochem Soc Trans. 2010;38:1581–6.
120. Davuluri RV, Suzuki Y, Sugano S, Zhang MQ. CART Classification of Human 5′UTR Sequences. Genome Res. 2000;10:1807–16.
121. Resch AM, Ogurstov AY, Rogozin IB, Shabalina SA, Koonin EV. Evolution of alternative and constitutive regions of mammalian 5'UTRs. BMC Genom. 2009;10:162.
122. Moshonov S, Elfakes R, Golan-Mashiach M, Sinvani H, Dikstein R. Links between core promoter and basic gene features influence gene expression. BMC Genom. 2008;9:92.
123. Grillo G, Turi A, Licciulli F, Mignone F, Liuni S, Banfi S, Gennarino VA, Horner DS, Pavesi G, Picardi E, Pesole G. UTRdb and UTRsite (RELEASE 2010): a collection of sequences and regulatory motifs of the untranslated regions of eukaryotic mRNAs. Nucleic Acid Res. 2010;38:D75–80.
124. Vogel C, Abreu RS, Ko D, Le SY, Shapiro BA, Burns SC, Sandhu D, Boutz DR, Marcotte EM, Penalva LO. Sequence signatures and mRNA concentration can explain two-thirds of protein abundance variation in a human cell line. Mol Syst Biol. 2010;6:400.
125. de Sousa R, Penalva LO, Marcotte EM, Vogel C. Global signatures of protein and mRNA expression levels. Mol BioSyst. 2009;5:1512–26.
126. Adam RD. The Giardia lamblia genome. Int J Parasitol. 2000;30:475–84.
127. Bruchhaus I, Leippe M, Lioutas C, Tannich E. Unusual gene organization in the protozoan parasite Entamoeba histolytica. DNA Cell Biol. 1993;12:925–33.
128. Elfakes R, Dikstein R. A translation initiation element specific to mRNAs with very short 5UTR that Also regulates transcription. PLoS ONE. 2008;3:e3094.
129. Elfakes R, Sinvani H, Haimov O, Svitkin Y, Sonenberg N, Dikstein R. Unique translation initiation of mRNAs-containing TISU element. Nucleic Acid Res. 2011;39:7598–609.

130. Sinvani H, Haimov O, Svitkin Y, Sonenberg N, Tamarkin-Ben-Harush A, Viollet B, Dikstein R. Translational tolerance of mitochondrial genes to metabolic energy stress involves TISU and eIF1-eIF4G1 cooperation in start codon selection. Cell Metab. 2015;21:479–92.
131. Zhaxybayeva O, Gogarten P. Spliceosomal introns: new insights into their evolution. Curr Biol. 2003;13:R764–6.
132. Lambowitz AM, Zimmerly S. Mobile group II introns. Annu Rev Genet. 2004;38:1–35.
133. Rogozin IB, Carmel L, Csuros M, Koonin EV. Origin and evolution of spliceosomal introns. Biol Direct. 2012;16:11.
134. Irimia M. Roy SW (2014) Origin of spliceosomal introns and alternative splicing. In: Keeling PJ, Koonin EV, editors. The origin and evolution of eukaryotes. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press; 2014. p. 295–316.
135. Maone E, Di Stefano M, Berardi A, Benelli D, Marzi S, La Teana A, Londei P. Functional analysis of the translation factor aIF2/5B in the thermophilic archaeon Sulfolobus solfataricus. Mol Microbiol. 2007;65:700–13.
136. Lee JH, Choi SK, Roll-Mecak A, Burley SK, Dever TE. Universal conservation in translation initiation revealed by human and archaeal homologs of bacterial translation initiation factor IF2. Proc Natl Acad Sci USA. 1999;96:4342–7.
137. Kyrpides NC, Woese CR. Archaeal translation initiation revisited: the initiation factor 2 and eukaryotic initiation factor 2B a- b-d subunit families. Proc Natl Acad Sci USA. 1998;95:3726–30.
138. Pánek J, Klolár M, Vohradsky J, Valásek LS. An evolutionary conserved pattern of 18S rRNA sequence complementary to mRNA 5′UTRs and its implications for eukaryotic gene translation regulation. Nucleic Acid Res. 2013;41:7625–34.
139. Matveeva OV, Shabalina S. Intermolecular mRNA-rRNA hybridization and the distribution of potential interaction regions in murine 18S rRNA. Nucleic Acid Res. 1993;21:1007–11.
140. Mauro VP, Edelman GM. rRNA-like sequences occur in diverse primary transcripts: implications for the control of gene expression. Proc Natl Acad Sci USA. 1997;94:422–7.
141. Tranque P, Hu MC, Edelman GM, Mauro VP. rRNA complementarity within mRNAs: a possible basis for mRNA-ribosome interactions and translational control. Proc Natl Acad Sci USA. 1998;95:12238–43.
142. Pisarev AV, Kolupaeva VG, Pisareva VP, Merrick WC, Hellen CUT, Pestova T. Specific functional interactions of nucleotides at key −3 and +4 positions flanking the initiation codon with components of the mammalian 48S translation initiation complex. Genes Dev. 2006;20:624–36.
143. Jacob F. Evolution and Tinkering. Science. 1977;196:1161–6.
144. Gould SJ, Vrba ES. Exaptation - a missing term in the science of form. Paleobiology. 1982;8:4–15.
145. Nojima T, Hirose T, Kimura H, Hagiwara M. The interaction between cap-binding complex and RNA export factor is required for intronless mRNA export. J Biol Chem. 2007;282:15645–51.
146. Vinciguerra P, Stutz F. mRNA export: an assembly line from genes to nuclear pores. Curr Opin Cell Biol. 2004;16:285–92.
147. Cougot N, van Dijk E, Babajko S, Séraphin B. Cap-tabolism. Trends Biochem Sci. 2004;29:436–44.
148. Ling SH, Qamra R, Song H. Structural and functional insights into eukaryotic mRNA decapping. Wiley Interdiscip Rev RNA. 2011;2:193–208.
149. Topisirovic I, Svitkin YV, Sonenberg N, Shatkin AJ. Cap and cap-binding proteins in the control of gene expression. Wiley Interdiscip Rev RNA. 2010;2:277–98.
150. Ghosh A, Lima CD. Enzymology of RNA cap synthesis. Wiley Interdiscip Rev RNA. 2010;1:152–72.
151. Martinez-Rucobo FW, Kohler R, van de Waterbeemd M, Heck AJ, Hemann M, Herzog F, Stark H, Cramer P. Molecular basis of transcription-coupled pre-mRNA capping. Mol Cell. 2015;58:1079–89.

152. Reed R. Coupling transcription, splicing and mRNA export. Curr Opin Cell Biol. 2003;15:326–31.
153. Rhoads RE. eIF4E: new family members, new binding partners, new roles. J Biol Chem. 2009;284:16711–5.
154. Rezende AM, Assis LA, Nunes EC, Lima TD, Marchini FK, Freire ER, Reis CR, de Melo Neto OP. The translation initiation complex eIF3 in trypanosomatids and other pathogenic excavates-identification of conserved and divergent features based on orthologue analysis. BMC Genom. 2014;15:1175.
155. Hinnebusch AG. eIF3: a versatile scaffold for translation initiation complexes. Trends Biochem Sci. 2006;31:553–60.
156. des Georges A, Dhote V, Kuhn L, Hellen CU, Pestova TV, Frank J, Hashem Y. Structure of mammalian eIF3 in the context of the 43S preinitiation complex. Nature. 2015;525:491–5.
157. Querol-Audi J, Sun C, Vogan JM, Smith MD, Gu Y, Cate JH, Nogales E. Architecture of human translation initiation factor 3. Structure. 2013;21:920–8.
158. Pereira RV, de Gomes MS, Jannotti-Passos LK, Borges WC, Guerra-Sá R. Characterisation of the COP9 signalosome in Schistosoma mansoni parasites. Parasitol Res. 2013;112:2245–53.
159. Schwechheimer C. The COP9 signalosome (CSN): an evolutionary conserved proteolysis regulator in eukaryotic development. Biochim Biophys Acta. 2004;1695:45–54.
160. Scheel H, Hofmann K. Prediction of a common structural scaffold for proteasome lid, COP9-signalosome and eIF3 complexes. BMC Bioinformatics. 2005;6:71.
161. Wei N, Serino G, Deng XW. The COP9 signalosome: more than a protease. Trends Biochem Sci. 2008;33:592–600.
162. Marintchev A, Wagner G. eIF4G and CBP80 share a common origin and similar domain organization: implications for the structure and function of eIF4G. Biochemistry. 2005;44:12265–72.
163. Marcotrigiano J, Lomakin IB, Sonenberg N, Pestova TV, Hellen CUT, Burley SK. A conserved HEAT domain within eIF4G directs assembly of the translation initiation machinery. Mol Cell. 2001;7:193–203.
164. Pontig CP. Novel eIF4G domain homologues linking mRNA translation with nonsense-mediated mRNA decay. Trends Biochem Sci. 2000;25:423–6.
165. Mendell JT, Medghalchi SM, Lake RG, Noensie EN, Dietz HC. Novel Upf2p orthologues suggest a functional link between translation initiation and nonsense surveillance complexes. Mol Cell Biol. 2000;20:8944–57.
166. Kobe B, Gleichmann T, Horne J, Jennings IG, Scotney PD, Teh T. Turn up the HEAT. Structure. 1999;7:R91–7.
167. Andrade MA, Petosa C, O'Donoghue SI, Muller CW, Bork P. Comparison of ARM and HEAT protein repeats. J Mol Biol. 2001;309:1–18.
168. Sonenberg N. eIF4E, the mRNA cap-binding protein: from basic discovery to translational research. Biochem Cell Biol. 2008;86:178–83.
169. Layana C, Ferrero P, Rivera-Pomar R. Cytoplasmic ribonucleoprotein foci in eukaryotes: hotspots of bio(chemical) diversity. Comp Funct Genom. 2012;2012:504292.
170. Thomas MG, Loschi M, Desbats MA, Boccaccio GL. RNA granules: the good, the bad and the ugly. Cell Sign. 2011;23:324–34.
171. Hernández G, Vazquez-Pianzola P. Functional diversity of the eukaryotic translation initiation factors belonging to eIF4 families. Mech Dev. 2005;122:865–76.
172. Strudwick S, Borden KLB. The emerging roles of translation factor eIF4E in the nucleus. Differentiation. 2002;70:10–22.
173. Osborne MJ, Borden KLB. The eukaryotic translation initiation factor eIF4E in the nucleus: taking the road less traveled. Immunol Rev. 2015;263:210–23.

174. Li L, Wang CC. Identification in the ancient protist Giardia lamblia of two eukaryotic translation initiation factor 4E homologues with distinctive functions. Eukaryot Cell. 2005;4:948–59.
175. Ramirez CV, Vilela C, Berthelot K, McCarthy JE. Modulation of eukaryotic mRNA stability via the cap-binding translation complex eIF4F. J Mol Biol. 2002;318:951–62.
176. Deutscher MP, Li Z. Exoribonucleases and their multiple roles in RNA metabolism. Prog Nucleic Acid Res Mol Biol. 2001;66:67–105.
177. Shuman S. The mRNA capping apparatus as drug target and guide to eukaryotic phylogeny. Cold Spring Harb Symp Quant Biol. 2001;66:301–12.
178. Anantharaman V, Koonin EV, Aravind L. Comparative genomics and evolution of proteins involved in RNA metabolism. Nucleic Acid Res. 2002;30:1427–64.
179. Muhlemann O. Recognition and elimination of nonsense mRNA. Biochim Biophys Acta. 2008;1779:538–49.
180. Eliseeva IA, Lyabin DN, Ovchinnikov LP. Poly(A)-binding proteins: structure, domain organization, and activity regulation. Biochemistry (Moscow). 2013;78:1377–91.
181. Derry MC, Yanagiya A, Martineau Y, Sonenberg N. Regulation of poly(A)-binding protein through PABP-interacting proteins. Cold Spring Harb Quant Biol. 2006;71:537–43.
182. Kuhn U, Wahle E. Structure and function of poly(A) binding proteins. Biochim Biophys Acta. 2004;1678:67–84.
183. Vazquez-Pianzola P, Suter B. Conservation of the RNA transport machineries and their coupling to translation control across eukaryotes. Comp Funct Genom. 2012;2012:287852.
184. Vazquez-Pianzola P, Urlaub H, Suter B. Pabp binds to the osk 3'UTR and specifically contributes to osk mRNA stability and oocyte accumulation. Dev Biol. 2011;357:404–18.
185. Maris C, Dominguez C, Allain FH. The RNA recognition motif, a plastic RNA- binding platform to regulate post-transcriptional gene expression. FEBS J. 2005;272:2118–31.
186. Lu WT, Wilczynska A, Smith E, Bushell M. The diverse roles of the eIF4A family: you are the company you keep. Biochem Soc Trans. 2014;42:166–72.
187. Toone WM, Rudd KE, Friesen JD. deaD, a new Escherichia coli gene encoding a presumed ATP-dependent RNA helicase, can suppress a mutation in rpsB, the gene encoding ribosomal protein S2. J Bacteriol. 1991;173:3291–302.
188. Nishi K, Morel-Deville F, Hershey JW, Leighton T, Schnier J. An eIF-4A-like protein is a suppressor of an Escherichia coli mutant defective in 50S ribosomal subunit assembly. Nature. 1988;336:496–8.
189. Lu J, Aoki H, Ganoza MC. Molecular characterization of a prokaryotic translation factor homologous to the eukaryotic initiation factor eIF4A. Int J Biochem Cell Biol. 1999;31:215–29.
190. Linder P, Jankowsky E. From unwinding to clamping - the DEAD box RNA helicase family. Nature Rev Mol Cell Biol. 2011;12:505–16.
191. Lüking A, Stahl U, Schmidt U. The protein family of RNA helicases. Crit Rev Biochem Mol Biol. 1998;33:259–96.
192. Parsyan A, Svitkin Y, Shahbazian D, Gkogkas C, Lasko P, Merrick WC, Sonenberg N. mRNA helicases: the tacticians of translational control. Nat Rev Mol Cell Biol. 2011;12:235–45.
193. Marintchev A. Roles of helicases in translation initiation: a mechanistic view. Biochim Biophys Acta. 2015;1829:799–809.
194. Takyar S, Hickerson RP, Noller HF. mRNA helicase activity of the ribosome. Cell. 2005;120:49–58.
195. Pestova TV, Shatsky IN, Hellen CU. Functional dissection of eukaryotic initiation factor 4F: the 4A subunit and the central domain of the 4G subunit are sufficient to mediate internal entry of 43S preinitiation complexes. Mol Cell Biol. 1996;16:6870–8.

196. Pestova TV, Hellen CU, Shatsky IN. Canonical eukaryotic initiation factors determine translation by internal ribosomal entry. Mol Cell Biol. 1996;16:6859–69.
197. Pestova TV, Shatsky IN, Fletcher SP, Jackson RJ, Hellen CU. A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initiation of hepatitis C and classical swine fever virus RNAs. Genes Dev. 1998;12:67–83.
198. Terenin IM, Dmitriev SE, Andreev DE, Royall E, Belsham GJ, Roberts LO, Shatsky IN. A cross-kingdom internal ribosome entry site reveals a simplified mode of internal ribosome entry. Mol Cell Biol. 2005;25:7879–88.

Evolution of Translation in Mitochondria

Aldo E. García-Guerrero, Angélica Zamudio-Ochoa, Yolanda Camacho-Villasana, Rodolfo García-Villegas, Adrián Reyes-Prieto and Xochitl Pérez-Martínez

1 Introduction

Around 1.5 billion years ago a bacterial cell related to modern α-proteobacteria established a symbiosis with a eukaryote that originated mitochondria [1]. It is well-established that mitochondrial origin is monophyletic (i.e., it happened only once in evolution) and that the organelle arose from an α-proteobacterium with identity yet to be established [2, 3]. The symbiotic event was followed by extensive reduction of the organelle's genetic material, either by gene loss or gene transfer to the nuclear genome. In addition, mitochondrial DNA (mtDNA) from different lineages diverged extensively in shape, size, content, mutation rate, and gene expression mechanisms. What mitochondria from different lineages have in common is that more than 1000 proteins are present in the organelle [4–7]. However, only a very limited number of proteins are encoded in mtDNA. For example, mtDNAs from the *Phylum* Apicomplexa have only three protein-coding genes [8], animal mitochondria code (in general) for 13 proteins [9], land plants code for more than 30–40 proteins [10], and members of the jakobid protists, which are considered to be relics of the endosymbiont bacterial ancestor, code around 65 proteins [11]. Thus, the majority of proteins necessary for function are

Aldo E. García-Guerrero and Angélica Zamudio-Ochoa have contributed equally to this work.

A.E. García-Guerrero · A. Zamudio-Ochoa · Y. Camacho-Villasana · R. García-Villegas · X. Pérez-Martínez (✉)
Departamento de Genética Molecular, Instituto de Fisiología Celular, Universidad Nacional Autónoma de México, 04510 Mexico City, Mexico
e-mail: xperez@ifc.unam.mx

A.E. García-Guerrero
e-mail: aldog.2602@gmail.com

A. Reyes-Prieto
Department of Biology, University of New Brunswick, Fredericton, New Brunswick, Canada

G. Hernández and R. Jagus (eds.), *Evolution of the Protein Synthesis Machinery and Its Regulation*, DOI 10.1007/978-3-319-39468-8_6

imported into mitochondria from the cytosol (for a review, see [12]). The organization of mtDNA among lineages has also diverged. Some organisms have extended non-coding regions, including type I and type II introns, as is the case for land plants, while others, like metazoan, have very compact mitochondrial genomes with only a few hundreds non-coding regions [9]. While fungal and apicomplexan mtDNAs are lineal molecules, animals and some protists have circular mtDNA, trypanosome mtDNA is composed of minicircles and maxicircles, which are topologically intertwined [13], and some *Amoebidium* have several hundred linear DNA molecules with different gene contents in each molecule [14]. In land plants, mtDNA is arranged in circular molecules of DNA whose composition varies constantly as a high frequency of recombination events occurs in this clade [10].

Independent of the shape, coding capacity and size of mtDNA, these organelles contain a complete gene expression system that comprises DNA replication and maintenance, transcription, post-transcriptional processing, translation and post-translation functions, such as protein assembly and prosthetic group additions. Much of the mitochondrial expression machinery is nucleus-encoded, while only a limited set of mtDNA genes is coded in the organelle. The *Phylum* Apicomplexa has only two ribosomal RNAs coded by the mtDNA. Metazoans have around 22 tRNAs and the small and large subunits rRNAs, while land plant mtDNA in addition to tRNAs and rRNAs codes for a varied number of ribosomal proteins. Protists from the jakobid lineage code in addition a translation factor (TufA), 5S rRNA, RNA polymerase and a sigma factor, and three chaperones for protein processing [11]. In general, mitochondrial genomes code for subunits of respiratory complexes and ATP synthase. Apicomplexan mtDNAs code for only subunits 1 and 3 of cytochrome *c* oxidase and cytochrome *b* from complex *bc1* [8]. In contrast, the jakobid *Phylum* codes for 12 subunits from complex I (NADH dehydrogenase), 3 for complex II (succinate dehydrogenase), 1 for *bc*1 complex, 3 for cytochrome *c* oxidase and 6 for ATP synthase [11].

Since mitochondria evolved from an α-proteobacterial ancestor, one might expect that the mtDNA expression mechanisms have conserved bacterial features. Even when this is the case, many novel mechanisms to control mtDNA expression have emerged and diverged among the eukaryotic groups. Some are conserved among certain lineages, but others appeared later during eukaryote divergence. In the present chapter, we describe the most prominent features of the mitochondrial translation machinery across different eukaryotic lineages. This knowledge allows us to better understand the evolution of the translation process in mitochondria.

2 The Mitochondrial Genetic Code

Translation in jakobid and land plant mitochondria uses the universal genetic code in mitochondria [11, 15]. However, at least 27 genetic code alterations (i.e., codon reassignments) are detected in mitochondrial systems of diverse eukaryote lineages (reviewed in [16, 17]). One of the most common changes in mitochondrial genetic code is the reassignment of termination codons to sense codons, such as the use of

Table 1 Mitochondrial genetic code in different organisms

	Group/organism	Genetic code	Comments
Protists	Jakobida	Standard	
	Euglenozoa Alveolata Rhizaria Amoebozoa Malawimonads	*UGA* Stop → Trp	Alternative initiation codons *Trypanosoma spp*: *UUA*, *UUG*, *CUG* *Leishmania spp*: *AUU*, *AUA* *Tetrahymena spp*: *AUU*, *AUA*, *AUG* *Paramecium sp: AUA, AUU,AUC, GUG, GUA*
Fungi	All fungi	*UGA* Stop → Trp	
	Saccharomyces cerevisiae, Candida glabrata, *Hansenula saturnus*, *Kluyveromyces thermotolerans*	*AUA* Ile → Met *CUU* Leu → Thr *CUC* Leu → Thr *CUA* Leu → Thr *CUG* Leu → Thr *UGA* Stop → Trp *CGA* Arg → Absent *CGC* Arg → Absent	*AUA* is frequently used in *VAR1* gene
Metazoa	Invertebrates	*AGA* Arg → Ser *AGG* Arg → Ser *AUA* Ile → Met *UGA* Stop → Trp *UAA* Stop → Tyr	In some flat and round worms: *AAA* Lys → Asn In ascidians: *AGA*, *AGG* Arg → Gly
	Vertebrates	*AUA* Ile → Met *UGA* Stop → Trp	Alternative initiation codons: *Bos taurus*: *AUA* *Homo sapiens*: *AUA*, *AUU* *Mus musculus*: *AUA*, *AUU*, *AUC* *Gallus gallus*: *GUG*
Non-chloropycean algae	Rodophyta Haptophyta	*UGA* Stop → Trp	
Viridiplantae	Chlorophyta	*UAG* Stop → Leu	Also found in *Scenedesmus obliquus*: *UCA* Ser → Stp
	Embryophyta	Standard	Alternative initiation codons: Gymnosperms and angiosperms: *ACG* *Cycas taitungensis*: *GCG* Alternative termination codons: Gymnosperms and angiosperms: *CGA*

Based in http://www.ncbi.nlm.nih.gov/Taxonomy/Utils/wprintgc.cgi

the canonical *UGA* stop codon to decode tryptophan in numerous biological groups [18] (Table 1). Other cases include the use of the typical *UAG* triplet as a leucine codon in chlorophycean algae [19] and the codon *UAA,* which decodes tyrosine in the nematode *Radopholus similis* [20]. Non-standard stop codons are used in mitochondria from some lineages. For example, the chlorophycean algae *Scenedesmus obliquus* uses the *TCA* codon as a translation stop signal [21]. In bryophytes and vascular plants, the codons *CAA*, *CGA* and *GGU* are reassigned stop codons, while *AAA* and *AAU* are recognized as stop codons in *Oryza sativa* [22]. In vertebrate mitochondria, the *AGA* and *AGG* codons, which are universally assigned to arginine, were thought to become stop codons [23]. However, recent studies indicate that these codons are unassigned [24]. Other prevalent reassignment is the use of *AUA* in *Saccharomyces cerevisiae*, vertebrates and some invertebrates to decode methionine instead of the canonical isoleucine [25, 26]. The standard arginine codons *AGA/AGG* were reassigned to serine in certain invertebrate groups (Nematoda, Arthropoda) and decode glycine in Ascidians [27]. In *S. cerevisiae,* the typical arginine codons *CGA/CGC* are unassigned, and the triplets *CUU/CUC/CUA/CUG* are used for threonine instead of leucine [28]. In some invertebrates (flat and round worms), *AAA* was reassigned, from lysine to asparagine (for an example, see [29]). Atypical start codons are also present in mitochondrial systems. For example, humans use *AUA* and *AUU* as start alternatives [24]; other cases of alternative start codons occur in trypanosomatids, which use *UUA, UUG, CUG* and the ciliate *Tetrahymena* with *AUU, AUA* or *AUG* [30] and certain nematodes that use the *UUG* triplet to initiate protein translation [31].

Why did mitochondria acquire modified codon assignments during evolution? One explanation is that codon reassignments might be a consequence of the organelle genome reduction, which encodes for a small set of proteins, and in most cases for a small number of tRNAs [8–10]. The diversity of mitochondrial genetic codes across eukaryotic groups might also reflect differential mutational rates in mtDNAs, a general increase in AT content and a diversification of genome expression mechanisms [17]. Interestingly, in silico studies suggest that genome size is not correlated to incident mutations that could lead to codon reassignments (i.e., the size of mitochondrial genomes does not correlate with mutation rates) [16]. The tRNAs' structure, the mitochondrial-targeted aminoacyl tRNA synthetases and in general the translation machinery are adjusted to the mitochondrial genetic code of each eukaryotic group. For example, reassignment of *UGA* for tryptophan (instead of the stop codon) is mediated by a tRNA where the wobble position carries a modified uridine. Modifications include 5-taurinomethyluridine (τm^5U), 5-carboxymethylaminomethyl-2-thio-uridine ($cmnm^5s^2U$) or 5-carboxymethylaminomethyluridine ($cmnm^5U$). These modifications expand the decoding capacity to R-ending codons, enabling the decoding of UGG and UGA as tryptophan [32]. Decoding of mammalian *AUG* and *AUA* as methionine is possible because the met-$tRNA^{Met(CAU)}$ has a 5-formylcytidine (f^5C) in the wobble position [33, 34]. Some theories try to explain how reassignments in the mitochondrial genetic code might have occurred during evolution. Two of the most established theories are the *Codon Capture* and *Ambiguous Intermediate* models.

The *Codon Capture,* also termed the *Codon Disappearance* theory, proposed by Osawa and Jukes in 1989, postulates that genetic code alterations are the result of neutral changes associated with the GC/AT content balance [35–38]. The theory posits that the disappearances of both the codon and the decoding tRNA are fundamental steps for further codon reassignment. Later, the "lost codon" can be reintroduced into the system by new mutations, but now is decoded at a relatively low efficiency by a different, noncognate tRNA, but with a similar anticodon sequence that allows the "capture" of the recently reestablished codon. Some reassignments are consistent with this *Codon Capture* model, such as the case of the frequent reuse of the *UGA* triplet to decode tryptophan [16]. A prediction derived from this model is that in mitochondrial genomes, which are high in AT content, GC-rich codons disappear at higher frequencies than AT-rich codons [36, 38, 39]. However, some codon reassignments in mtDNA do not follow the predictions of the GC/AT content balance. Thus, the *Codon Capture* theory does not explain satisfactorily the use of GC-rich codons in genomes with high AT content or the fact that some codons seem to be unassigned in some mtDNA genetic systems.

The *Ambiguous Intermediate* theory, proposed by Schultz and Yarus [40, 41], suggests that codon reassignment is the result of selective mechanisms that favor ambiguity in codon recognition during protein translation. The model postulates that codon recognition ambiguity, associated with structural changes in the tRNA molecules, is fundamental for the codon reassignment. The idea is that the codon in the spotlight is suddenly decoded by two different tRNAs, namely the "original" and the new "mutant," which is now able to form a cognate pair with the codon. Later the "mutant" tRNA takes over the codon in a selection-driven process. Thus, the triplet is reassigned to a new amino acid. During mitochondrial evolution, many repeated tRNAs for each amino acid were lost, and in general mitochondria contain only one tRNA for each amino acid [42]. In contrast to the *Codon Capture theory*, in this model the initial loss of the codon before the reassignment is not necessary [37]. Some examples consistent with the *Ambiguous Intermediate* theory are the reassignments of leucine to threonine in yeast mitochondria [17] and from serine to lysine in Arthropoda [43].

Overall, both models are not mutually exclusive, as reassignments might have arisen from combinatory events during evolution [17, 44]. Some changes in the mitochondrial genetic code are explained by the *Codon Capture* theory, while others by the *Ambiguous Intermediate* theory.

3 Mitochondrial tRNAs

Translation of mitochondrial mRNAs requires around 20 tRNAs, but the exact number varies depending on the wobble rules and the genetic code in each species. Mitochondrial tRNAs have nuclear and mitochondrial origins. Depending on the organism, the proportion of nuclear and mitochondrial tRNAs varies. While human and the jakobid *Andalucia godoyi* [9, 11] encode a complete set of mitochondrial

encoded tRNAs for reading all codons in mtDNA, protist-like *Trypanosoma brucei* and *Plasmodium falciparum* have no mtDNA-coded tRNAs [8, 45] and therefore have to import all tRNAs necessary for translation. Interestingly, the number of mtDNA-derived tRNAs among closely related organisms is variable. For example, in chlorophycean algae, *Chlamydomonas reinhardtii* codes for only 3 tRNAs in mitochondria, while *Nephroselmis olivacea* codes for a full set of 26 tRNAs [46]. It is expected that mitochondria would import only the necessary number of tRNAs to complete the ~22 tRNAs necessary for translation. However, in some cases, import of redundant tRNAs can take place. For example, in the yeast *S. cerevisiae* a $tRNA^{Lys}_{(CUU)}$ is imported from cytosol even when mtDNA codes for the full set of tRNAs necessary to decode all codons [28, 47]. This tRNA is particularly important to decode codons under stress conditions [48]. Mammalian mitochondria can also import redundant cytosolic tRNAs [49]. The unicellular algae *C. reinhardtii* imports 31 tRNAs instead of the expected 22 tRNAs necessary to decode all codons [50]. The mechanisms to import cytosolic tRNAs are particular to each eukaryotic group, indicating that import of tRNAs into mitochondria is a process that emerged independently several times during evolution (for a review, see [51]). Delivery of tRNAs to mitochondria is mediated by proteins, usually with a previously described function. *S. cerevisiae* Eno2 (involved in glycolysis) delivers the charged $tRNA^{Lys(CUU)}$ to the mitochondrial surface, where the mitochondrial lysyl-tRNA synthetase binds it and co-transports it via the general import machinery. In land plants, aminoacyl-tRNA synthetases might be involved in the delivery of tRNAs to mitochondria, and the Voltage Dependent Anion Channel (VDAC), together with the outer membrane receptors Tom20 and Tom40, functions in tRNA import. In *Trypanosoma*, the cytosolic EF1a, together with the import component Tim17 and Hsp70, Hsp60 and Hsp20 might participate in the delivery and import of cytosolic tRNAs.

The structure, sequence and post-transcriptional modifications of mitochondrial tRNAs have conserved features with cytosolic RNAs. However, many of these features have amazingly diverged in different eukaryotic groups and among specific tRNAs from the same organism. According to the structural characteristics, mitochondrial tRNAs are classified into five groups, named 1–5 [51]. Group 1 shares the most conserved features with cytosolic tRNAs. They carry canonical T and D arms, anticodon and acceptor arms, and L1/l2 connectors (involved in joining the acceptor and anticodon helices) [52]. This class of tRNAs is present in mitochondria from amoebozoans, alveolates, plants and fungi. Group 2 carries conserved anticodon and acceptor arms. However, T/D arms may be smaller in size and may have less conservation on bases involved in D/T-loop interactions (mainly bases G18, G19, U55 and C56). These tRNAs are present in amoebozoans, alveolates, plants, fungi and some metazoans (including mammals). Group 3 consists of tRNAs where the acceptor arm may be 1–3 nucleotides shorter; they are T-armless and carry a shorter D-loop. The L2 connector is also shorter (6–7 nt instead of 21–30 nt). This class of tRNAs is present in some nematodes, bryozoan and arachnid species. Group 4 is represented by some insect and bryozoan species and by mammals. They are D-armless and carry shorter T arms. The L1 connector is also shorter (5–12 nt instead of 19–20 nt). Group 5 carries both shorter anticodon and

acceptor arms; they are T- and D-armless and have shorter L1 and L2 connectors. These minimalist tRNAs are found in acaria and some nematodes. In this group, the only conserved features with cytosolic tRNAs are the presence of an acceptor arm with the 3′-single-stranded CCA terminus and an anticodon arm with the canonical anticodon loop of seven nucleotides. The shortest mitochondrial tRNA so far is 54-nt long ($tRNA^{Ser(UCU)}$) from the nematode *Ascaris suum* [53, 54].

As a universal feature, mitochondrial tRNAs are also post-transcriptionally modified to define the structure and decoding capabilities. The best understood model is *Bos taurus*, where all mitochondrial tRNAs were isolated and analyzed. There are 15 types of modifications at 118 positions (representing 7.5 % abundance in mitochondrial tRNA bases) [55]. However, the occurrence of modified nucleotides can be as low as one residue in mitochondrial $tRNA^{Ser}$ of the rodent *Mesocricetus auratus* (representing 1.7 % abundance) [56]. To date, 15 out of 18 conserved modifications (present throughout kingdoms of life) are observed in mitochondrial tRNAs, with the exception of ac^4C, m^3U and $m^{66}A$, which are not yet detected (reviewed in [51]). Comparative analyses of tRNA sequences indicate that mitochondria have the highest number of modified positions that are not universally conserved. The acceptor stem is particularly rich in Ψ residues, and the number of modifications located in positions 46–50, 5′ to the T arm, is also relatively low in mitochondrial tRNAs [55, 56]. There are mitochondria-specific base modifications, like τm^5U and τm^5s^2U, discovered in ascidian mitochondria [32], f^5C, f^5Cm, present at the wobble position 34 in bovine and the nematode *A. suum* [33, 54], and k^2C in potato [57].

4 Mitochondrial mRNAs

Mitochondrial mRNAs have conserved some prokaryotic features, but some other characteristics have diverged. Mitochondrial mRNAs from some lineages, such as jakobid protists, have a putative Shine-Dalgarno-like sequence to locate the ribosome at the correct AUG start codon [11]. Other lineages lack a Shine-Dalgarno-like sequence and therefore must have different, unknown mechanisms to initiate translation. This is the case for flowering plants [58] and mammal mitochondria [59]. Similar to what is observed in prokaryotes, mitochondrial mRNAs do not have a 7-methylguanylate cap (5′-cap), as is found in cytosol mRNAs. Moreover, mitochondrial mRNAs undergo post-transcriptional modifications before they are ready for translation. The major post-transcriptional RNA processing events in mitochondria include 3′-end polyadenylation, intron/exon splicing and editing. Polyadenylation of RNA is present in all kingdoms of life and is a near-universal feature of RNA metabolism, although it can trigger different signals among cells and organelles. Today, the function of polyadenylation in mitochondrial gene expression is not fully understood. An additional interesting feature of mitochondrial mRNAs from some lineages is the requirement of RNA edition before translation. RNA editing might be important to correct transcript sequences that otherwise would affect the translation product's function [59, 60].

An example is editing in land plants, where the amino acid encoded by an edited mRNA is frequently more conserved than the one predicted from the gene sequence.

4.1 Polyadenylation of Mitochondrial mRNAs

Polyadenylation is the non-template addition of adenosine residues to the 3′ end of RNAs. In the eukaryotic cytoplasm, the majority of nuclear-encoded mRNAs require a poly(A) tail for stability, nuclear export and translatability (for reviews, see [61, 62]). In contrast, in prokaryotes, RNA polyadenylation functions to tag the mRNA for exonucleolytic degradation [63, 64]. Although mitochondria have a monophyletic origin, many features of polyadenylation have extensively diverged within eukaryotes.

In mammalian mitochondria, 12 out of 13 mRNAs have stable poly(A) tails of 45 nt on average. However, there are slight variations between cell types and between transcripts within the same cell type [65]. For example, only the *ND6* transcript lacks a poly(A) tail [66]. The precise function of polyadenylation is not entirely understood. However, one function of polyadenylation is to complete the UAA codon, since several mammalian RNAs contain incomplete translational stop codons. The same feature is observed in general in metazoans, where some coding regions lack a complete UAA stop codon, suggesting that polyadenylation also plays an important role in translation [67, 68]. Although polyadenylation produces stable transcripts [66, 69], truncated, adenylated transcripts may coexist, suggesting that human mitochondria use transient poly(A) tails to degrade RNA [70]. The mechanism of a possible differential polyadenylation on stabilizing and destabilizing RNAs remains to be elucidated. In plants, similarly to the bacterial system, addition of a poly(A) tail targets exonucleolytic degradation of RNA [71]. In trypanosomatid mitochondria, most protein-coding transcripts suffer a massive edition (insertion or deletion of uridines) necessary to render translatable mRNAs [59]. The addition of a poly(A) tail in these organisms seems to render both stable and unstable transcripts. Polyadenylation in these organisms has an intricate relation to edition and translation. Poly(A) tails are 20–200 nt long, and the length of the tail seems to correlate with the state of edition. Short tails ($\sim$20 nt) stabilize edited or non-edited mRNAs. Long (100–200 nt) poly(A/U) tails are added to fully edited RNAs, and this extension might render the transcript translationally competent [72, 73].

Yeast mitochondria are so far the only organelles that do not polyadenylate their mRNAs. This was found to be the case in *S. cerevisiae, Schizosaccharomyces pombe* and *Candida albicans* [74–77], suggesting that it might be a general phenomenon of fungal mitochondria. Instead, the 3′ ends of some, but not all fungal mitochondrial mRNAs possess a conserved dodecamer sequence that is encoded in the mitochondrial genome and seems to be vital for mRNA stability and translatability [75, 76, 78, 79].

4.2 Edition of Mitochondrial mRNAs

Some organisms require mitochondrial (and plastid) transcript edition before they can be translated. RNA editing consists of nucleotide substitutions, post-transcriptional or co-transcriptional insertion/deletions. These three processes occur in very different taxonomic groups, suggesting that they arose as several independent acquisitions [80]. This process is present in dinoflagellates (variable one-nucleotide substitutions), excavates (U insertions/deletions), unikonts (co-transcriptional insertion of 1 or 2 nucleotides), metazoa (U to C substitution) and archaeplastida (U to C and C to U substitutions) [80]. Editions throughout a transcript can be limited in number, as is the case for land plants [60, 81]. In other cases, extensive edition of a transcript is required to transform an unrecognizable sequence into a conserved protein sequence, as is the case of trypanosomatids [59, 82] and calcaronean sponges [83].

In land plants, C to U (and less frequently U to C) editing often results in changes of the amino acid sequence from what the genomic sequence predicts. This process evolved in land plants [84] and was most likely subsequently lost in some marchantiid liverworts [85]. The number of edited nucleotides among plant lineages is: *Physcomitrella patens* edits 11 sites [86], *Arabidopsis thaliana* edits 600 cytidines [87], while the lycophytes *Isoetes engelmanii* and *Selaginella moellendorfii* edit more than 1,700 and 2,100 nucleotides, respectively [88, 89]. The composition of the RNA editosome is not yet fully understood, although *cis*- and *trans*-factors are essential for the editing process. The *cis* elements that specify the editing of the C target are present in close proximity to the edition site. *Trans*-factors include members of the pentatricopeptide repeat (PPR) motif-containing family, which are site-specific recognition factors. While the cytidine deaminase catalyzing C-to-U conversion has not been identified, considerable evidence points to the C-terminal DYW domain found on some PPR proteins, which exhibits sequence similarity to known cytidine deaminase motifs (for reviews, see [60, 80, 81]).

In trypanosomatids, edition is a post-transcriptional process, where uridines are inserted or deleted from mRNA precursors [90]. Edition introduces start and stop codons, restores frame shifts and often completes the coding sequence of mRNAs. Mitochondrial editing can occur at different extensions: transcripts that are never edited, transcripts where edition is restricted to a small region, with minimal edition, and transcripts that are extensively edited or pan-edited, where a single mRNA is altered by 553 insertions and 89 deletions [80]. The process in trypanosomatids includes mRNA cleavage, U deletion or insertion, and mRNA ligation [91]. The maxicircle molecules of mtDNA code for guide RNAs (gRNAs), which are derived from scattered intergenic regions. A partial hybrid is formed between the 5′ portion of the gRNA and the complementary sequence on the pre-edited mRNA. Cleavage of the mRNA at the 3′ end of the first base that is not paired with the gRNA leaves a free 3′OH. The uridine addition or deletion is followed by immediate relegation of the two molecules. Many proteins have been implicated in the edition process (reviewed in [59, 82]). However, these proteins are not related to the proteins involved on plant edition.

5 The Mitoribosome

Mitochondrial ribosomes (mitoribosomes) are located in the matrix, and are closely associated with the inner membrane [92, 93]. This location facilitates the insertion of newly synthesized products, which are mainly hydrophobic proteins. All mitochondrial genomes currently sequenced encode ribosomal RNAs (rRNAs). In contrast, almost all mitochondrial ribosomal proteins (MRPs) are nuclear encoded. Thus, assembly of functional ribosomes requires a coordinated expression of both genomes and a proper import of the necessary components into the organelle [94, 95]. The mechanism of this process is almost unknown, but evidence supports that several MRPs assemble with rRNAs in a co-transcriptional fashion [96, 97].

In contrast to the cytosolic ribosomes, mitoribosome composition is highly variable between different eukaryotic lineages. Their sedimentation coefficient ranges from 80S in ciliates, to 70–74S in fungi, to 77–78S in vascular plants and 55S in animals. These variable sedimentation values are the result of the difference in the protein:RNA ratio, while bacterial ribosomes contain a protein:RNA proportion of 1:2, in mitoribosomes this proportion varies from 1:1 in yeast to 2:1 in bovine [98].

The α-proteobacterial ribosome is composed of 54 proteins [99], which were also likely to be present in the ancestor of mitochondria. It is proposed that, in the earliest stage of eukaryotic evolution, several novel proteins were recruited for ribosomal function, and only one, Rps20, was lost, resulting in an ancestral mitoribosome of 72 proteins (Fig. 1) [100]. An interesting feature of several mitoribosomal proteins of bacterial origin is that they increased in length sequence. Accordingly, this stage in mitoribosome evolution is known as the "constructive phase", as the total size of the ribosome was increased considerably [101].

The cause of the constructive phase of the mitoribosome is proposed to be the accumulation of slightly deleterious mutations on the mitochondrial genome, as this genome, with the exception of land plants, presents a higher mutation rate than the nuclear one [103, 104]. Slightly deleterious mutations could trigger the recruitment of new proteins because a mutation in an original component of the complex is compensated by the interaction with a new component [105]. This process is called Constructive Neutral Evolution (CNE), a universal evolutionary ratchet that leads to complexity [106]. Accordingly, genes coding for MRPs show higher levels of amino acid replacements than cytoplasmic ribosomal proteins, which suggests a compensatory modification [107, 108]. Gain of complexity throughout the evolution of mitochondria is not exclusive to the mitoribosome. The respiratory chain complexes have also acquired new proteins that are usually important for regulation, assembly and stability [101]. These eukaryotic subunits are in general localized in the peripheral regions of the enzymes. This feature is also observed for the mitoribosomes [109, 110]. The extensive gain of protein mass observed for mitoribosomes does not reflect the fate of all endosymbiotic organelles, as the plastid ribosomes only gained approximately 170 kDa [111, 112]. Several evolutionary mechanisms have led to the increase of protein mass in the mitoribosome.

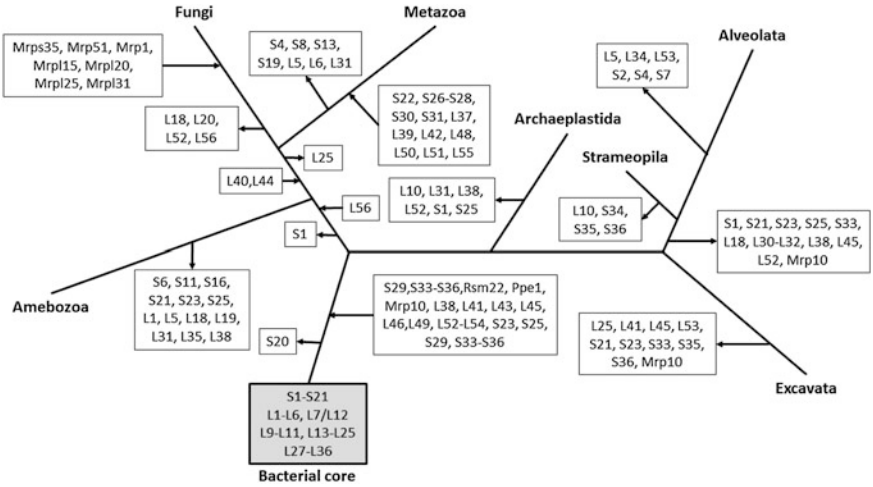


Fig. 1 Reconstruction of the evolutionary history of the mitochondrial ribosome proteome. Incoming and outgoing arrows indicate the gains and losses of the ribosomal proteins that are showed in the box. This figure is based on the data given by [102] and [100]. The models considered for the construction of this figure were: for fungi *Neurospora crassa, Aspergillus fumigatus, Saccharomyces cerevisiae, Schizosaccharomyces pombe, Cryptococcus neoformans, Ustilago maydis* and *Encephalitozoon cuniculi*; for metazoa *Mus musculus, Homo sapiens, Danio rerio, Drosophila melanogaster, Caenorhabditis elegans* and *Monosiga brevicolis*; for amoebozoa *Dictyostelium discoideum* and *Entamoeba histolytica*; for Archaeplastida *Arabidopsis thaliana, Oryza sativa, Chlamydomonas reinhardtii, Ostreococcus tauri* and *Cyanidioschyzon merolae*; for Strameopila *Thalassiosira pseudonana and Phytophthora ramorum*; for Alveolata *Tetrahymena thermophila, Paramecium tetraurelia, Theileria annulata, Plasmodium falciparum, Plasmodium yoelii* and *Cryptosporidium parvum* and for Excavata *Leishmania brasilensis, Leishmania infantum, Leishmania major, Trypanosoma cruzi, Trypanosoma brucei, Naegleria gruberi, Trichomonas vaginalis, Giardia lamblia* and *Reclinomonas americana*

One of them was the recruitment of existent proteins, such as the case of Mrpl45, a homolog of Tim44 (a subunit of the mitochondrial protein translocase machinery), which is present in several bacteria but is not part of the prokaryote ribosome [102]. Mrpl39, a metazoan protein, was recruited later in evolution and is homologous to threonyl-tRNA synthetases [113]. It is proposed that addition of Mrpl39 to the mitoribosome compensated for the loss of bacterial proteins involved in tRNA binding [102]. Numerous new ribosomal proteins emerged through gene duplication. For instance, Mrps10 gave rise to Mrpl48 through this process in metazoans. Interestingly, the duplicated gene product became part of the other ribosomal subunit [102]. Another case is Mrps18, which in *Caenorhabditis elegans* has three variants originated by gene duplication. It is believed that each ribosome contains only one copy of the protein, suggesting that mitoribosomes exist in heterogeneous populations [102].

The increment in protein mass in mitoribosomes is not only due to the addition of new subunits, but also to the gain of new domains in the prokaryotic proteins.

MRPs are sometimes almost twice the size of their bacterial counterparts [102]. In *P. falciparum*, Mrpl4 has an AAA domain, which is not present in the bacterial counterpart. This domain is known to participate in chaperone-like functions [114]. Another case is the presence of an RRM (RNA recognition motif) domain in Mrps19 of *A. thaliana*, which could be involved in the association of the protein with rRNAs [115]. In the yeast *S. cerevisiae*, the carboxyl-terminal end of Mrp20, which is mitochondria-specific, plays a role in ribosome assembly [116].

The evolution of mitochondria involved numerous independent losses of ribosomal proteins in different lineages (Fig. 1). Bacterial-exclusive S20 protein seems to have been lost early during mitoribosome evolution. This protein is not essential for bacterial growth. However, its absence causes a decrease in the association of the ribosomal subunits [117, 118]. In contrast, S1 protein, which was lost early in the evolution of unikonts, is an essential protein in bacteria [119]. Moreover, there is no apparent pattern favoring protein loss from either bacterial or eukaryotic origin, suggesting that there is no tendency in protein dispensability [100].

Whereas protein gain in mitoribosomes is a general phenomenon in all lineages, the rRNA content varies greatly. While bacteria have an rRNA content of 1.4 MDa, in mitochondria this number varies from 0.5 MDa in *C. elegans* to 1.6 MDa in *Neurospora crassa*. Since animals show an important reduction of rRNA, it was previously thought that the proteins acquired during mitoribosome evolution replaced the lost helices of rRNA [120]. However, now it is clear that the high content of proteins in mitoribosomes is not a consequence of the lower concentration of rRNA, as the increase in MRPs occurred previously to the reductive phase of rRNA [101]. This is consistent with the structural data in which the extra proteins of the ribosome do not substitute the lost portions of rRNA [121–123]. Furthermore, it is proposed that rRNA reduction might be driven by the reduction of the mitochondrial genome size and not necessarily by adaptive changes of the translational machinery [124].

Reduction of rRNA had triggered an important mitoribosome remodeling. For example, the bacterial ribosomal protein L24 contacts the helices H7 and H19 of the 23S rRNA, stabilizing its binding to the 39S subunit. The mammalian mitochondrial counterpart, Mrpl24, lacks both helices. However, mitochondria-specific protein elements maintain Mrpl24 in the same place and orientation as the prokaryotic counterpart [125].

An almost general phenomenon in mitochondria is the loss of 5S rRNA, which is present only in plants and some algae. An extension of the 23S rRNA replaces the resulting gap in *N. crassa*. On the contrary, in the mammalian mitoribosome this space is occupied by protein [121–123].

An interesting aspect of the evolution of mitoribosomes is their assembly mechanisms. This process has probably evolved differently in each lineage, as the components of the ribosome are partially different among eukaryote groups. In some lineages mitochondrial-encoded rRNAs are fragmented, need edition or lack 5S rRNA [8, 9, 126]. As stated above, the composition of proteins also diverged among eukaryote lineages. The understanding of ribosome assembly, the order of

rRNA processing and protein addition, the chaperones involved in such events and the role of mitochondrial RNA granules in ribosome biogenesis are just starting to emerge, especially in mammals and yeast models [97, 127].

6 Mitochondrial Translation Initiation

Translation initiation in bacteria is carried out by three conserved factors: IF1, IF2 and IF3 [128]. There are important differences between prokaryote and mitochondrial initiation factors. While mitochondrial $IF2_{mt}$ is universally present, $IF3_{mt}$ is semi-universal and $IF1_{mt}$ was completely lost from the mitochondrial machinery [129]. In addition, there are important structural variations in the mitochondrial initiation factors. In agreement with the prokaryotic origin, mitochondria seem to initiate translation with formylated methionine, at least for the studied cases. Initially, by in vitro experiments, it was demonstrated that the initiation machinery in mammals does not need a formylated Met-tRNA. However, recent experiments demonstrate that a failure in formylation is a cause of disease in humans [130, 131]. In the yeast *S. cerevisiae* it was previously shown that a mutant *Δfmt1* (coding for a methionyl-tRNA formyltransferase) does not affect translation initiation [132]. However, an accessory factor, Aep3, was compensating the lack of *Δfmt1*. Double mutant *Δfmt1* and *Δaep3* affect respiratory growth [133]. The mechanisms of translation initiation regulation have extensively diverged from the bacterial counterpart. Despite the differences between bacterial and mitochondrial translation initiation factors, the general steps for initiation are conserved.

6.1 Structural and Functional Conservation of $IF2_{mt}$

In bacteria, IF2 interacts with initiator fMet-tRNA and promotes binding with the small ribosomal subunit and with mRNA. It also contains a GTPase activity to release all initiation factors from the completely assembled ribosome into the mRNA. IF2 triggers the binding of tRNA to the incomplete P site on the 30S subunit. After binding of the 50S subunit to the initiation complex, IF2 GTP hydrolysis assists the release of all initiation factors from the completely assembled ribosome [128]. Bacterial IF2 contains six domains (I–VI). To date, the function of domain I is not completely understood. Domain II stabilizes the interaction of IF2 with the ribosomal 30S subunit; this region is not conserved among bacterial species. Domain III is a linker between domains II and IV. Domain IV contains the GTPase activity. Domain V interacts with the ribosomal 30S subunit, and domain VI recognizes the fMet-tRNA [134]. In mitochondria, $IF2_{mt}$ consists only of domains III–VI (Fig. 2). This short version of $IF2_{mt}$ is not particular for mitochondrial systems, as shorter versions of IF2 factors are also present in some bacterial groups, like extremophiles [135]. Instead, in mammals $IF2_{mt}$ interaction

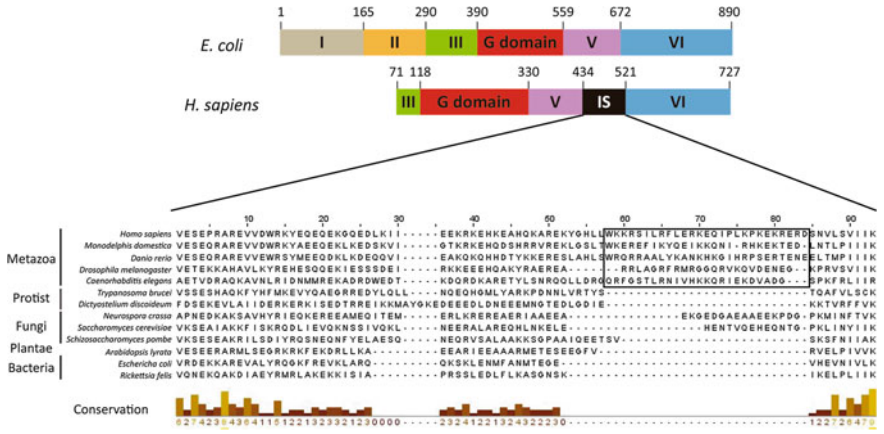


Fig. 2 Alignment of the insertion sequence on mIF2 among species from different phyla. Variations in domain composition between *E. coli* IF2 and *Homo sapiens* mIF2 are presented in the upper panel. Numbers indicate amino acid positions. The insertion sequence (IS) is amplified, and an alignment among different species is shown. Alignments of the insertion sequences were made with the MAFFT software, with the Blosum70 matrix as in [129]

with the 28S is performed by domain III, and this interaction is even stronger when GTP is bound to domain V [136]. Domain III of $IF2_{mt}$ is not conserved in all eukaryote lineages, but its function might be compensated by the differences in the small ribosomal subunit protein and RNA content among eukaryotes [137]. Domain IV is the most conserved region of $IF2_{mt}$ in structure and sequence similarity, sharing 99 % (metazoans) to 50 % (fungi) identity with bacterial counterparts [138–140]. Domain V in $IF2_{mt}$ is modestly conserved with the bacterial IF2, sharing identity of 35–50 %. However, the function of this region is not completely understood [140]. It might be important for interaction with the small ribosomal subunit because structural modeling of $IF2_{mt}$ suggests that this region is similar to domain II of EF-Tu and EF-G. This region is important for contact with small ribosomal subunits [141]. Interestingly, domain V in metazoa $IF2_{mt}$ contains an insertion of variable length and sequence (Fig. 2). This region might perform the same function as IF1 [142] (discussed below). Domain VI in bacteria and mitochondria is divided into subdomains C1 and C2. Subdomain C2 is important for binding of IF2 to the fMet-tRNA [143]. Mutagenic analysis in *Bacillus stearothermophilus* shows that there are two critical cysteines at the 668 and 714 positions necessary for this interaction [144, 145]. These amino acids are usually present in $IF2_{mt}$, suggesting that $IF2_{mt}$ subdomain C2 conserved the same function as in bacteria. $IF2_{mt}$ contains the C1 subdomain. However, as in bacteria, the function of this domain is still unknown. By NMR studies it was suggested that this subdomain from *Bacillus stearothermophilus* has a similar structure as domain III from eukaryotic eIF5B. This region is implicated in transmitting and amplifying structural changes to the G-domain after GTP binding [144, 146].

6.2 The Mystery of the Lost IF1 in Mitochondria

In bacteria, IF1 plays an important role in the recognition of the correct *AUG* initiation codon. IF1 interacts with the A site of the 30S ribosomal subunit and prevents binding of the initiator aminoacyl-tRNA to the A site. In addition, bacterial IF1 increases the affinity of IF2 for 30S, has a role in small subunit dissociation and assists the release of IF2 form the 70S complex [128, 147]. So far, biochemical and bioinformatic approaches have failed to identify mitochondrial IF1. This suggests that IF1 was lost at the earliest stage of eukaryotic evolution [129, 148]. Mitochondria may be able to bypass the need for IF1: Experimental evidence indicates that, in the presence of mammalian $IF2_{mt}$ and $IF3_{mt}$, the bacterial ribosome does not need IF1 for the formation of the 70S particle or translation in general [149]. As discussed above, metazoan $IF2_{mt}$ contains an insertion between domains V and VI [129] (Fig. 2). Even though there is no conservation of the insertion sequence among eukaryotic $IF2_{mt}$s, it is possible that this insertion substitutes the function of IF1, at least in metazoa. Through cryo-electron microscopy and nuclease digestion experiments, it was observed that bacterial IF2 associates with the interphase of the 30S subunit [134, 148]. Modeling of mammalian $IF2_{mt}$ suggests that the extension is close to the small subunit A site, similar to bacterial IF1 [140]. How the need for IF1 is bypassed in other eukaryotic lineages remains an open question.

6.3 Mitochondrial $IF3_{mt}$

Bacterial IF3 plays a critical role in translation initiation. It binds the 30S ribosomal subunit in order to prevent association with the 50S subunit. This interaction is necessary for the initiation complex to recognize the Shine-Dalgarno sequence in the mRNA and enhances the interaction and activity of IF2 [140, 150]. Simple BLAST-P analysis failed to detect orthologs of IF3 in mitochondria. However, the existence of $IF3_{mt}$ was hypothesized because orthologs of the bacterial ribosomal proteins S7, S11 and S18, which are in proximity to IF3, are present in mitochondria [129]. More sensitive searching algorithms, like PSI-BLAST, identified $IF3_{mt}$ in fungi, animal, plant and excavates mitochondria [129, 140]. Structural data show that the IF3 C-terminal end is necessary for interaction with the 30S subunit through residues in two helical segments, designated H3 and H4. In most $IF3_{mt}$ orthologs the C-terminal domain is the least conserved region of the protein. However, some residues from the H3 segments are conserved in $IF3_{mt}$ [140, 151, 152]. Biochemical and structural approaches will clarify the mechanism of action of $IF3_{mt}$.

6.4 How Is the AUG Start Codon Recognized in Mitochondria?

As discussed above, most mitochondrial mRNAs seem to lack Shine-Dalgarno-like sequences to direct the ribosome to the AUG start codon. This is the case for metazoans, flowering plants and fungi [57, 58, 153]. In the case of metazoans it is even more puzzling because the start codon locates at or very near the 5′ end of the mRNA. This implicates that cells developed different mechanisms to localize the ribosome to the correct start codon.

Mammalian mitochondria have developed an initiation codon selection that relies on leaderless mRNAs. Addition of three nucleotides prior to the *COX2* 5′ AUG decreased translation by 40 %, and addition of 12 nucleotides reduced translation by 80 % [154]. It is proposed that the movement of the ribosome is

Table 2 Orthologs of known translational activators in mitochondria. Taken from [134, 157, 165, 168, 169, 225]

Mitochondrial target gene	Translational activator(s)	Reported species	Orthologs	Conserved function?
COB	Cbs1	*S. cerevisiae*	No	
	Cbs2	*S. cerevisiae*	No	
	Cbp3	*S. cerevisiae*	No	
	Cbp6	*S. cerevisiae, S. pombe*	Yes (only in fungi)	No, in *S. pombe* is only a chaperone
	Cbp1	*S. cerevisiae*	Yes (only in fungi)	Not known
COX1	Pet309	*S. cerevisiae, S. pombe*	Yes (only in fungi)	No, in *S. pombe* Ppr5 is activator and Ppr4 is repressor
	TACO1	*S. cerevisiae, H. sapiens*	Yes	Not known in *S. cerevisiae*
	C12orf62	*H. sapiens, S. cerevisiae*	Yes	No, in *S. cerevisiae* is a chaperone (Cox14)
	Mss51	*S. cerevisiae, S. pombe, M. musculus*	Yes	No, in *S. pombe* is only a chaperone and in mammals is a metabolic regulator
COX2	Pet111	*S. cerevisiae, H. sapiens*	Yes	No, in *H. sapiens* is a nuclear protein
COX3	Pet54	*S. cerevisiae*	No	
	Pet122	*S. cerevisiae*	No	
	Pet494	*S. cerevisiae*	No	
ATP6/8	Atp22	*S. cerevisiae*	No	
ATP9	Aep1	*S. cerevisiae*	No	
	Aep2	*S. cerevisiae*	No	
Accessory factors in translation initiation				
Aep3		*S. cerevisiae*	No	
Rsm28		*S. cerevisiae*	No	
Rmd9		*S. cerevisiae*	No	

paused after the first 17 nucleotides of the mRNA enter the ribosome. The small subunit then inspects the mRNA 5′ end. If there is a start codon at the P site, then a stable initiation complex is formed [58, 154].

Study of the mechanisms for initiation codon selection in the yeast *S. cerevisiae* has made important progress. A group of proteins, named translational activators, plays a role in the localization of the mitoribosome in the correct AUG start codon. Each one of these proteins interacts with specific mitochondrial mRNAs and with the ribosome to pose it on the start codon [155–157]. In addition, translational activators interact with each other and with the mitochondrial inner membrane, probably to tether translation initiation to the site where nascent peptides will be inserted [158–160] (Table 2). Many of these proteins are members of the pentatricopeptide repeat (PPR) family or RNA recognition motif (RRM). Other translational activators have no detectable RNA-binding motifs whatsoever. Many efforts have been made to find orthologs of these proteins in other organisms. Some translational activators may be present in other fungi [161–163] and probably also in humans [164]: However, in mammalian mitochondria, the mechanisms of action of the putative activators remains to be elucidated, as human mRNAs have either very short or no 5′-UTRs. Translational activation is also observed in plastids [165–167], suggesting that this mechanism arose several times during eukaryotic evolution.

7 Translation Elongation

Translation elongation in mitochondria is highly conserved with bacteria. During this process three elongation factors (EF) assist the mitoribosome for addition of new residues to the nascent polypeptide chain. EF-Tu_{mt} forms a ternary complex with the aminoacylated tRNA and GTP and enters the mitoribosome A site. Cognate codon-anticodon pairing triggers GTP hydrolysis by EF-Tu_{mt} and release of EF-Tu_{mt}-GDP. The mitoribosome catalyzes the peptide bond formation at the PTC. Thus, deacetylated tRNA is left in the P site and the elongated peptidyl-tRNA in the A site of the ribosome. This process is assisted by EF-$G1_{mt}$, which catalyzes the translocation of peptidyl-tRNA from the A to the P site, and removing the deacetylated tRNA from the ribosome. EF-Ts_{mt} exchanges GDP to GTP from EF-Tu_{mt} to allow a new round of elongation [58, 140]. Mitochondrial elongation, at least for mammalian and yeast models, seems to be a more conserved process than initiation and ribosome recycling. However, many components of the translation machinery have extensively diverged in different phyla, leading to adaptations of the elongation machinery. For example, as previously discussed, the structure of tRNAs and of mitoribosomes has diverged from the bacterial ancestor. In the next section, we discuss the main changes observed in the elongation factors.

7.1 Mitochondrial EF-Tu$_{mt}$

EF-Tu$_{mt}$ must be able to bind tRNA with canonical conformations (e.g., fungi, plants and some protist lineages) and shorter tRNAs versions (metazoans) [51]. In some cases the tRNAs reduction is so extensive that EF-Tu$_{mt}$ should use alternative binding modes. The divergence of EF-Tu is evident inside the nematode group: nematodes have 2 EF-Tu$_{mt}$ homologs [53]. While EF-Tu1$_{mt}$ is unable to bind cloverleaf type tRNAs, it is the only factor that binds T-armless tRNAs [170, 171]. *C. elegans* EF-Tu1$_{mt}$ has a C-terminal extension of around 60 amino acids that likely interacts with the D arm of T-armless tRNAs [172]. In the *Trichinella* lineage EF-Tu1$_{mt}$ binds T-armless tRNAs, D-armless tRNA and cloverleaf type tRNAs [170, 171]. Nematode EF-Tu2$_{mt}$ has a short C-terminal extension of 7–15 amino acids that is necessary for interaction with the D-armless tRNA^{Ser} [170]. *C. elegans* mt EF-Tu2$_{mt}$ is unique because it interacts with phosphates on the T arm on the opposite side from where canonical EF-Tu binds [173].

In trypanosomatids, EF-Tu$_{mt}$ has a highly charged insertion of approximately 30 amino acids near the C terminus. This trypanosomatid-specific motif is dispensable for the union of EF-Ts$_{mt}$, but critical for EF-Tu$_{mt}$ function [174]. This extension might be necessary for interaction with tRNAs or with the mitoribosome, which has less RNA content than mammalian ribosomes [175]. Since trypanosomatid mitochondrial tRNAs are imported, EF-Tu$_{mt}$ has evolved to interact with eukaryotic-type tRNAs, suggesting that the appearance of this motif is an adaptation of the mitochondrial machinery to recognize imported tRNAs [174]. Interestingly, complete loss of tRNA genes from mtDNA is also observed in apicomplexans [176, 177], and therefore their EF-Tu$_{mt}$s have to bind imported tRNAs. However, in this case EF-Tu$_{mt}$ is closer to the bacterial factor, indicating that each group has their own mechanisms for imported tRNA-EF-Tu$_{mt}$ binding [174]. Another distinctive feature of EF-Tu$_{mt}$ is found in hemi-ascomycete yeasts, where EF-Ts$_{mt}$ seems to be lost [178]. *S. cerevisiae* EF-Tu$_{mt}$ displays greater affinity for GTP, like the self-recycling GTPases EF-G or IF2 [179]. It is functionally equivalent to the *S. pombe* EF-Tu$_{mt}$/EF-Ts$_{mt}$ [180].

7.2 Mitochondrial EF-G1$_{mt}$ and EF-G2$_{mt}$

Bacterial EF-G participates in translation elongation and ribosome recycling. However, in some bacterial groups these functions are separated in two specialized paralogs. This is the case of the majority of Spirochaetes, Planctomycetes, Lentisphaera and some species of δ-proteobacteria. Mitochondria of most organisms have specialized EF-G$_{mt}$s paralogs as well, and these are phylogenetically related to the specialized bacterial EF-Gs. EF-G$_{mt}$s probably were acquired before the eukaryotic last common ancestor, since at least one EF-G$_{mt}$ paralog is present in all mitochondriate eukaryotes [181].

So far, aerobic eukaryotes that possess a unique EF-G_{mt} (which is an EF-$G1_{mt}$ paralog) are the plastid/apicoplast-carrying eukaryotes: Archaeplastida, stramenopile algae and Apicomplexa [181]. Interestingly, instead of a second EF-G_{mt}, these species have a plastid/apicoplast-targeted EF-G (termed EF-G_{cp} or EF-G_{api}). Outside of these groups, some *Cryptococcus* species have only one EF-$G1_{mt}$. Curiously, they do not have plastids. Until now, it is estimated that all mitochondriate eukaryotes have either two specialized EF-Gs or one EF-$G1_{mt}$ and an EF-G_{cp}/EF-G_{api} [181, 182]. Until now, it is estimated that all mitochondriate eukaryotes have either two specialized EF-G_{mt}s or one EF-G_{mt} and an EF-G_{cp}/EF-G_{api} [181, 182].

There are limited studies about the function of each mitochondrial paralog. Mammalian EF-$G1_{mt}$ is specialized in translation elongation, while EF-$G2_{mt}$ participates in ribosome recycling [183]. In contrast, *A. thaliana* EF-$G1_{mt}$ carries both functions, translocation and ribosome recycling [182]. *P. falciparum* EF-$G1_{mt}$ participates in ribosome recycling, although its translocation activity was not investigated [184, 185].

8 Termination and Ribosomal Recycling

Translation ends when the ribosome reaches one of three stop codons, *UGA*, *UAA*, or *UAG*. These codons are recognized by releases factors (RF) that enter the ribosomal A site and induce release of the nascent peptide (class-I RFs). Bacterial RF1 recognizes *UAA* and *UAG*, while RF2 recognizes *UAA* and *UGA* [186]. These factors assist the hydrolysis of ester bonds on the peptidyl-tRNA, which is located in the ribosomal P site, releasing the newly synthesized protein. Bacterial class-II RFs are GTPases that trigger dissociation of the class-I RF from the ribosome after peptide release. RFs have a conserved GGQ motif that is involved in ester bond hydrolysis (peptidyl-hydrolase domain, PTH), whereas RF1 has a PAT or PVT motif and RF2 SPF motif important for stop codon recognition (codon-recognition domain, CR) (reviewed in [187]). As mentioned in a previous section, mitochondria recognize non-conventional codons as stop codons (Table 1). Thus, understanding the mechanisms of termination and stop codon recognition is a fertile ground for research. Despite recent advances (mostly in mammalian mitochondria), it is still unclear how translation terminates in mitochondria. The most challenging subject is to understand how non-standard stop codons are decodified in mitochondria.

Mitochondrial release factors divide in five distinct subfamilies: mtRF1a, mtRF2a and ICT1, derived from bacterial ancestors, C12orf65 and mtRF1, so far found only in vertebrates. While mtRF1a, mtRF1 and mtRF2a conserved both the PTH and CR domains, ICT1 and C12orf65 have lost the CR domain [188]. Because the release factor family seems particularly prone to genetic expansion and functional divergence [188], there are high probabilities that the mechanism of translation termination varies among different phyla.

- mtRF1a is present in every eukaryotic organism and evolved from an α-protobacterial ancestor [188]. This protein recognizes *UAA/UAG* stop codons, both in vitro and in vivo [189, 190].
- mtRF1 is a vertebrate-specific mitochondrial protein [190], and it may originate from duplication of the mtRF1a gene at the root of this clade [191]. The function of mtRF1 is controversial. It may recognize the non-standard stop codons AGA and AGG [191]. However, posterior structural predictions and experimental data could not find evidence that mtRF1 recognizes any of the stop codons used in mitochondria [190, 192, 193]. In human mitochondria, the AGG and AGA codons may not function as stop codons. Instead, they promote a −1 frameshift that creates a standard TGA stop codon that can be decodified by mtRF1a [24]. Nonetheless, an analysis from all the vertebrate genomes showed that a −1 frameshift (or even a −2 frameshift) could not originate a canonical TGA stop in every ORF ending in AGG or AGA [188].
- ICT1 (immature colon carcinoma transcript-1) is widely distributed in mitochondria from all eukaryotic phyla [188]. ICT1 is a codon-independent release factor that lost the CR domain. In addition, it is an integral component of the mitoribosome and a crucial component for its assembly [194]. ICT1 is the eukaryotic ortholog of bacterial ArfB. This protein is a rescue factor of stalled ribosomes in prematurely truncated mRNAs and is also part of the bacterial ribosome [195, 196]. ICT1's role in mitochondrial translation is still not completely understood. The position of ICT1 in the mitoribosome is incompatible with the mechanism used by ArfB [193, 195]. In fact, the ICT1 integrated to the mitoribosome has no release factor activity [192]. This protein can rescue ribosomal complexes not only at the ends of mRNAs, but also in the middle of mRNAs, and even can rescue ribosomes depleted of mRNAs [192, 194]. ICT1 might terminate translation of ORFs ending in AGG and AGA since these codons are unassigned in mammalian mitochondria, and mitoribosomes stalled at AGG/AGA codons might be recognized by ICT1 [192, 193].
- C12orf65 is a release factor that probably derived from ICT1. It has a wide phylogenetic distribution and is only absent in viridiplantae [188]. Contrary to ICT1, C12orf65 is a mitochondrial soluble matrix protein that does not exhibit ribosomal-dependent peptidyl hydrolase activity. However, ICT1 overexpression partially complements C12orf65's absence, indicating that both proteins must have some overlapping functions [197].
- mtRF2s lack experimental data about their function or mitochondrial localization. mtRF2 has a narrow phylogenetic distribution, consistently found in land plants, red algae, dictyosteliida and some stramenopiles (brown algae, oomycetes and Blastocystis). It has been lost at least five times during eukaryotic evolution, in concordance with the reassignment of the UGA codon to Trp [188].

The final step of mitochondrial translation consists of recycling of the mitoribosomes. Once the nascent chain is released, the ribosome recycling factor 1 ($RRF1_{mt}$) and the specialized $EF\text{-}G2_{mt}$ (see above) separate ribosome subunits to

allow new cycles of translation [181, 198]. In addition, $IF3_{mt}$ may attach to the ribosomal SSU to prevent futile association of the mitoribosome until an initiation complex is formed [199].

9 Mitochondrial Translation Is Coupled to Protein Assembly

The majority of proteins encoded by mitochondrial DNA are subunits of the respiratory chain complexes and the ATP synthase. These proteins are usually hydrophobic, with two or more transmembrane stretches. Thus, it is expected that mitochondrial translation machinery is physically coupled to the mitochondrial inner membrane. Mitochondrial and cytoplasmic subunits have to assemble and acquire the necessary prosthetic groups in coordination to make active enzymes. Indeed, major progress in the field has come from beaker's yeast *S. cerevisiae*. Defects in the coordination of mitochondrial-encoded subunit synthesis and assembly are proposed to affect the cell physiology. When cytochrome *c* oxidase is not assembled, then Cox1 synthesis in mitochondria is downregulated [200, 201]. Cox1 is part of the central core of the enzyme and has 12 transmembrane stretches. Downregulation of Cox1 synthesis may prevent generation of pro-oxidant species, because the poorly assembled heme *a* present in Cox1 has peroxidase activity [202]. When the ATPase F1 sector is not assembled, then translation of the *ATP8/ATP6* transcript is downregulated [203, 204]. This prevents accumulation on the membrane of Atp6–Atp9 rings that could interfere with the membrane potential [205, 206]. Translation of the mitochondrial *COB* mRNA, coding for the cytochrome *b* subunit from the *bc*1 complex is also tightly linked to enzyme assembly [207]. Coordination of mitochondrial translation and assembly is an intricate process that requires many factors, many of them specific for each mitochondrial-coded protein [159, 203, 208, 209] (Fig. 3). In general, initiation of mitochondrial mRNA translation is achieved by translational activators (discussed above), which assist in the positioning of the ribosome on the initiator *AUG*. Translational activators are themselves associated with the inner membrane and with other translational activators to tether translation initiation to the localization of assembly of nascent peptides [155, 156]. Some of these mRNA-specific activators act exclusively on the 5′-UTR of the target mRNA and are no longer required in downstream events after translation initiation. Some other translational activators have dual functions. Together with specific chaperones, they physically interact with the newly synthesized peptide and are key players in the coordination of translation and assembly. In general, it is proposed that once the respiratory complex proceeds to assembly, these chaperones/translational activators are released from the assembly intermediary and recycled for new rounds of translation [159, 208]. Equivalent processes were described for *C. reinhardtii* photosynthetic complexes in chloroplasts [165].

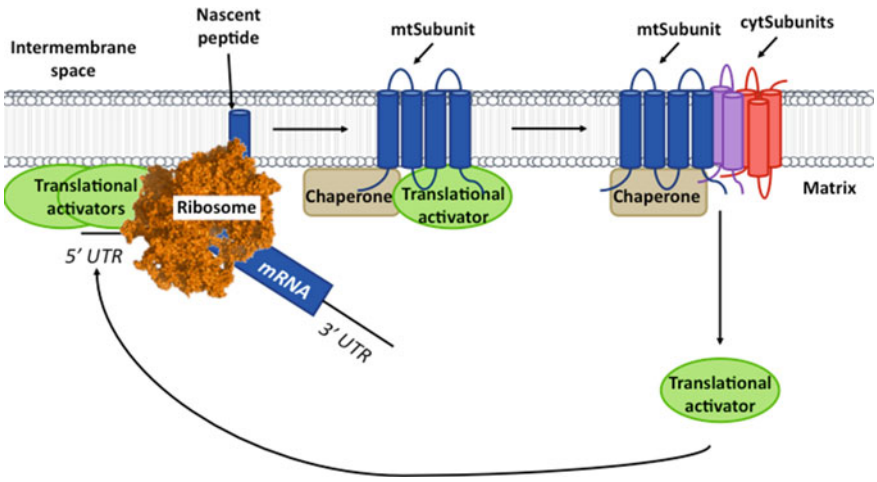


Fig. 3 General model for the coupling of protein synthesis and membrane assembly in yeast mitochondria. Translational activators assist the ribosome in localization of the AUG start codon through recognition of specific sequences within the 5′-UTR of each mRNA Cbs1/Cbs2/Cbp3/Cbp6 (cytochrome *b* synthesis) [207, 210]; Pet309/Mss51 (Cox1 synthesis) [157, 200, 201, 211]; Pet111 (Cox2 synthesis) [212]; Pet494/Pet122/Pet54 (Cox3 synthesis) [213–215], Atp22 (Atp6 and Atp8 synthesis) [204]; Aep1, Aep2 (Atp9 synthesis) [216, 217]. Some of these activators play a second role in coordination of translation/assembly. They physically interact with newly made peptides (mtSubunits) and with additional chaperones to form assembly intermediaries. This is the case for Cbp3/Cbp6 [207] and Mss51 [201]. Once the respiratory complexes assemble with cytosolic, imported subunits (CytSubunits), then the chaperones/translational activators are released and recycled. Translational activators are now ready for new rounds of translation

In human mitochondria, the scenario is not as clear as with yeast. However, some orthologs of the yeast translational activators and chaperones are present in humans, and some of them may have similar roles [164, 209, 218, 219].

10 Concluding Remarks

Mitochondrial translation evolution is an exciting field in biology. Many efforts have been made to understand mitochondrial translation by studying yeast, mammals, plant and trypanosomatid species, and to a lesser degree apicomplexans, nematodes and *Drosophila*. Since the mitochondrial translation machinery has extensively diverged among eukaryotes, it is necessary to study it in many disparate groups of eukaryotes in order to understand its evolution. To date, many questions remain open in this field. For instance: (1) *W*hat are the role and mechanism of action in the translation of mitochondrial-encoded small and large non-coding RNAs discovered a few years ago [220–222]? (2) How do mRNA polyadenylation and edition regulate stability, editing and translation in different species? (3) Are

there specialized ribosomes devoted to translating a specific mRNA in mitochondria? If this is the case, how are populations of each type of ribosome regulated in different conditions or cell types? (4) What are the mechanism and regulation of the recently discovered "programmed translational bypass"? This process was first discovered in the yeast *Magnusiomyces capitatus* [223, 224]. In this organism, almost all protein-coding genes in mtDNA have insertions of 27–55 nucleotides, called byps. Translation of these insertions would lead to frameshifts and premature termination of translation, so a precise mechanism to bypass these elements is necessary. Is translation bypass present in different eukaryotic groups? (5) What is the role of the recently discovered mitochondrial RNA granules in translation [127]? Are they present exclusively on mammalian mitochondria, or do they have a broader prevalence in eukaryotes? (6) Since Shine-Dalgarno sequences are absent in many eukaryotic lineages, how is the start codon *AUG* recognized by the initiation complex among different phyla? (7) Are the translation and assembly of nascent peptides prevalent processes in eukaryotes? The field of mitochondrial evolution awaits answers to these exiting questions.

Acknowledgments This work was supported by a research grant from Consejo Nacional de Ciencia y Tecnología (47514 to XP-M) and fellowships 255917 (to AE. G-G), 298954 (to AZ-O) and 250726 (to RG-V); Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT), UNAM (IN204414 to XP-M). AE.G-G and RG-V are students of the Programade Doctoradoen Ciencias Biomédicas, UNAM.

References

1. Hori H, Osawa S. Origin and evolution of organisms as deduced from 5S ribosomal RNA sequences. Mol Biol Evol. 1987;4:445–72.
2. Gray MW, Burger G, Lang BF. Mitochondrial evolution. Science. 1999;283:1476–81.
3. Degli Esposti M. Bioenergetic evolution in proteobacteria and mitochondria. Genome Biol Evol. 2014;6:3238–3251. doi:10.1093/gbe/evu257.
4. Millar AH, Heazlewood JL, Kristensen BK, Braun HP, Moller IM. The plant mitochondrial proteome. Trends Plant Sci. 2005; 10:36-43. doi:10.1016/j.tplants.2004.12.002.
5. Premsler T, Zahedi RP, Lewandrowski U, Sickmann A. Recent advances in yeast organelle and membrane proteomics. Proteomics. 2009;9:4731–43. doi:10.1002/pmic.200900201.
6. Szklarczyk R, Huynen MA. Mosaic origin of the mitochondrial proteome. Proteomics. 2010;10:4012–24. doi:10.1002/pmic.201000329.
7. Gawryluk RM, Chisholm KA, Pinto DM, Gray MW. Compositional complexity of the mitochondrial proteome of a unicellular eukaryote (Acanthamoeba castellanii, supergroup Amoebozoa) rivals that of animals, fungi, and plants. J Proteomics. 2014;109:400–16. doi:10.1016/j.jprot.2014.07.005.
8. Hikosaka K, Kita K, Tanabe K. Diversity of mitochondrial genome structure in the Phylum Apicomplexa. Mol Biochem Parasitol. 2013;188:26–33. doi:10.1016/j.molbiopara.2013.02.006.
9. Boore JL. Animal mitochondrial genomes. Nucleic Acids Res. 1999;27:1767–80. doi:10.1093/nar/27.8.1767.
10. Gualberto JM, Mileshina D, Wallet C, Niazi AK, Weber-Lotfi F, Dietrich A. The plant mitochondrial genome: dynamics and maintenance. Biochimie. 2014;100:107–20. doi:10.1016/j.biochi.2013.09.016.

11. Burger G, Gray MW, Forget L, Lang BF. Strikingly bacteria-like and gene-rich mitochondrial genomes throughout jakobid protists. Genome Biol Evol. 2013;5:418–38. doi:10.1093/gbe/evt008.
12. Bohnert M, Pfanner N, van der Laan M. Mitochondrial machineries for insertion of membrane proteins. Curr Opin Struct Biol. 2015;33:92–102.
13. Stuart K, Feagin JE. Mitochondrial DNA of kinetoplastids. Int Rev Cytol. 1992;141:65–88.
14. Burger G, Forget L, Zhu Y, Gray MW, Lang BF. Unique mitochondrial genome architecture in unicellular relatives of animals. Proc Natl Acad Sci USA. 2003;100:892–7. doi:10.1073/pnas.0336115100.
15. Campbell WH, Gowri G. Codon usage in higher plants, green algae, and cyanobacteria. Plant Physiol. 1990;92:1–11.
16. Knight RD, Landweber LF, Yarus M. How mitochondria rede-fine the code. J Mol Evol. 2001;53:299–313.
17. Bezerra AR, Guimaraes AR, Santos MA. Non-standard genetic codes define new concepts for protein engineering. Life (Basel). 2015;5:1610–28.
18. Inagaki Y, Ehara M, Watanabe KI, Hayashi-Ishimaru Y, Ohama T. Directionally evolving genetic code: the UGA codon from stop to tryptophan in mitochondria. J Mol Evol. 1998;47:378–84.
19. Hayashi-Ishimaru Y, Ohama T, Kawatsu Y, Nakamura K, Osawa S. UAG is a sense codon in several chlorophycean mitochondria. Curr Genet. 1996;30:29–33.
20. Jacob JE, Vanholme B, Van Leeuwen T, Gheysen G. A unique genetic code change in the mitochondrial genome of the parasitic nematode Radopholus similis. BMC Res Notes. 2009;2:192.
21. Nedelcu AM, Lee RW, Lemieux C, Gray MW, Burger G. The complete mitochondrial DNA sequence of Scenedesmus obliquus reflects an intermediate stage in the evolution of the green algal mitochondrial genome. Genome Res. 2000;10:819–31.
22. Xu W, Xing T, Zhao M, Yin X, Xia G, Wang M. Synonymous codon usage bias in plant mitochondrial genes is associated with intron number and mirrors species evolution. PLoS One. 2015;10:e0131508.
23. Osawa S, Ohama T, Jukes TH, Watanabe K. Evolution of the mitochondrial genetic code. I. Origin of AGR serine and stop codons in metazoan mitochondria. J Mol Evol. 1989;29:202–7.
24. Temperley R, Richter R, Dennerlein S, Lightowlers RN, Chrzanowska-Lightowlers ZM. Hungry codons promote frameshifting in human mitochondrial ribosomes. Science. 2010;327:301. doi:10.1126/science.1180674.
25. Clark-Walker GD, Weiller GF. The structure of the small mito-chondrial DNA of Kluyveromyces thermotolerans is likely to reflect the ancestral gene order in fungi. J Mol Evol. 1994;38:593–601.
26. Desjardins P, Morais R. Nucleotide sequence and evolution of coding and noncoding regions of a quail mitochondrial genome. J Mol Evol. 1991;32:153–61.
27. Yokobori S, Ueda T, Feldmaier-Fuchs G, Paabo S, Ueshima R, Kondow A, Nishikawa K, Watanabe K. Complete DNA sequence of the mitochondrial genome of the ascidian Halocynthia roretzi (Chordata, Urochordata). Genetics. 1999;153:1851–62.
28. Foury F, Roganti T, Lecrenier N, Purnelle B. The complete sequence of the mitochondrial genome of Saccharomyces cerevisiae. FEBS Lett. 1998;440:325–31.
29. Jacobs HT, Elliott DJ, Math VB, Farquharson A. Nucleotide sequence and gene organization of sea urchin mitochondrial DNA. J Mol Biol. 1988;202:185–217.
30. Ziaie Z, Suyama Y. The cytochrome oxidase subunit I gene of Tetrahymena: a 57 amino acid NH2-terminal extension and a 108 amino acid insert. Curr Genet. 1987;12:357–68.
31. Pritchard AE, Seilhamer JJ, Mahalingam R, Sable CL, Venuti SE, Cummings DJ. Nucleotide sequence of the mitochondrial genome of Paramecium. Nucleic Acids Res. 1990;18:173–80.
32. Suzuki T, Miyauchi K, Suzuki T, Yokobori S, Shigi N, Kondow A, Takeuchi N, Yamagishi A, Watanabe K. Taurine-containing uridine modifications in tRNA anticodons are

required to decipher non-universal genetic codes in ascidian mitochondria. J Biol Chem. 2011;286:35494–8.
33. Moriya J, Yokogawa T, Wakita K, Ueda T, Nishikawa K, Crain PF, Hashizume T, Pomerantz SC, McCloskey JA, Kawai G, et al. A novel modified nucleoside found at the first position of the anticodon of methionine tRNA from bovine liver mitochondria. Biochemistry. 1994;33:2234–9.
34. Takemoto C, Spremulli LL, Benkowski LA, Ueda T, Yokogawa T, Watanabe K. Unconventional decoding of the AUA codon as me-thionine by mitochondrial tRNAMet with the anticodon f5CAU as revealed with a mitochondrial in vitro translation system. Nucleic Acids Res. 2009;37:1616–27.
35. Osawa S, Jukes TH. Evolution of the genetic code as affected by anticodon content. Trends Genet. 1988;4:191–8.
36. Osawa S, Ohama T, Jukes TH, Watanabe K, Yokoyama S. Evo-lution of the mitochondrial genetic code. II. Reassignment of codon AUA from isoleucine to methionine. J Mol Evol. 1989;29:373–80.
37. Sengupta S, Yang X, Higgs PG. The mechanisms of codon reassignments in mitochondrial genetic codes. J Mol Evol. 2007;64:662–88.
38. Silva RM, Miranda I, Moura G, Santos MA. Yeast as a model organism for studying the evolution of non-standard genetic codes. Brief Funct Genomic Proteomic. 2004;3:35–46.
39. Miranda I, Silva R, Santos MA. Evolution of the genetic code in yeasts. Yeast. 2006;23: 203–13.
40. Schultz DW, Yarus M. Transfer RNA mutation and the malleability of the genetic code. J Mol Biol. 1994;235:1377–80.
41. Schultz DW, Yarus M. On malleability in the genetic code. J Mol Evol. 1996;42:597–601.
42. Jia W, Higgs PG. Codon usage in mitochondrial genomes: distinguishing context-dependent mutation from translational selection. Mol Biol Evol. 2008;25:339–51.
43. Abascal F, Posada D, Knight RD, Zardoya R. Parallel evolution of the genetic code in arthropod mitochondrial genomes. PLoS Biol. 2006;4:e127.
44. Chin JW. Expanding and reprogramming the genetic code of cells and animals. Annu Rev Biochem. 2014;83:379–408.
45. Hancock K, Hajduk SL. Sequence of Trypanosoma brucei tRNA genes encoding cytosolic tRNAs. Nucleic Acids Res. 1992;20:2602.
46. Turmel M, Lemieux C, Burger G, Lang BF, Otis C, Plante I, Gray MW. The complete mitochondrial DNA sequences of Nephroselmis olivacea and Pedinomonas minor. Two radically different evolutionary patterns within green algae. Plant Cell. 1999;11:1717–30.
47. Martin RP, Schneller JM, Stahl AJ, Dirheimer G. Import of nuclear deoxyribonucleic acid coded lysine-accepting transfer ribonucleic acid (anticodon C-U-U) into yeast mitochondria. Biochemistry. 1979;18:4600–5.
48. Kamenski P, Kolesnikova O, Jubenot V, Entelis N, Krasheninnikov IA, Martin RP, Tarassov I. Evidence for an adaptation mechanism of mitochondrial translation via tRNA import from the cytosol. Mol Cell. 2007;26:625–37.
49. Rubio MA, Rinehart JJ, Krett B, Duvezin-Caubet S, Reichert AS, Soll D, Alfonzo JD. Mammalian mitochondria have the innate ability to import tRNAs by a mechanism distinct from protein import. Proc Natl Acad Sci USA. 2008;105:9186–91.
50. Vinogradova E, Salinas T, Cognat V, Remacle C, Marechal-Drouard L. Steady-state levels of imported tRNAs in Chlamydomonas mitochondria are correlated with both cytosolic and mitochondrial codon usages. Nucleic Acids Res. 2009;37:1521–8.
51. Salinas-Giege T, Giege R, Giege P. tRNA biology in mitochondria. Int J Mol Sci. 2015;16:4518–59. doi:10.3390/ijms16034518.
52. Watanabe Y, Suematsu T, Ohtsuki T. Losing the stem-loop structure from metazoan mitochondrial tRNAs and co-evolution of interacting factors. Front Genet. 2014;5:109.
53. Giege R, Juhling F, Putz J, Stadler P, Sauter C, Florentz C. Structure of transfer RNAs: similarity and variability. Wiley Interdiscip Rev RNA. 2012;3:37–61. doi:10.1002/wrna.103.

54. Suzuki T, Suzuki T. A complete landscape of post-transcriptional modifications in mammalian mitochondrial tRNAs. Nucleic Acids Res. 2014;42:7346–57. doi:10.1093/nar/gku390.
55. Machnicka MA, Olchowik A, Grosjean H, Bujnicki JM. Distribution and frequencies of post-transcriptional modifications in tRNAs. RNA Biol. 2014;11:1619–29. doi:10.4161/15476286.2014.992273.
56. Weber F, Dietrich A, Weil JH, Marechal-Drouard L. A potato mitochondrial isoleucine tRNA is coded for by a mitochondrial gene possessing a methionine anticodon. Nucleic Acids Res. 1990;18:5027–30.
57. Hazle T, Bonen L. Comparative analysis of sequences preceding protein-coding mitochondrial genes in flowering plants. Mol Biol Evol. 2007;24:1101–12.
58. Christian BE, Spremulli LL. Mechanism of protein biosynthesis in mammalian mitochondria. Biochim Biophys Acta. 2012;1819:1035–54. doi:10.1016/j.bbagrm.2011.11.009.
59. Aphasizhev R, Aphasizheva I. Mitochondrial RNA editing in trypanosomes: small RNAs in control. Biochimie. 2014;100:125–31. doi:10.1016/j.biochi.2014.01.003.
60. Takenaka M, Zehrmann A, Verbitskiy D, Hartel B, Brennicke A. RNA editing in plants and its evolution. Annu Rev Genet. 2013;47:335–52.
61. Norbury CJ. Cytoplasmic RNA: a case of the tail wagging the dog. Nat Rev Mol Cell Biol. 2013;14:643–53.
62. Tian B, Manley JL. Alternative cleavage and polyadenylation: the long and short of it. Trends Biochem Sci. 2013;38:312–20.
63. Silva IJ, Saramago M, Dressaire C, Domingues S, Viegas SC, Arraiano CM. Importance and key events of prokaryotic RNA decay: the ultimate fate of an RNA molecule. Wiley Interdiscip Rev RNA. 2011;2:818–36.
64. Bandyra KJ, Luisi BF. Licensing and due process in the turnover of bacterial RNA. RNA Biol. 2013;10:627–35.
65. Temperley RJ, Wydro M, Lightowlers RN, Chrzanowska-Lightowlers ZM. Human mitochondrial mRNAs–like members of all families, similar but different. Biochim Biophys Acta. 2010;1797:1081–5.
66. Mercer TR, Neph S, Dinger ME, Crawford J, Smith MA, Shearwood AM, Haugen E, Bracken CP, Rackham O, Stamatoyannopoulos JA, Filipovska A, Mattick JS. The human mitochondrial transcriptome. Cell. 2011;146:645–58.
67. Bernt M, Braband A, Schierwater B, Stadler PF. Genetic aspects of mitochondrial genome evolution. Mol Phylogenet Evol. 2012;69:328–38.
68. Bernt M, Donath A, Juhling F, Externbrink F, Florentz C, Fritzsch G, Putz J, Middendorf M, Stadler PF. MITOS: improved de novo metazoan mitochondrial genome annotation. Mol Phylogenet Evol. 2013;69:313–9.
69. Ojala D, Montoya J, Attardi G. tRNA punctuation model of RNA processing in human mitochondria. Nature. 1981;290:470–4.
70. Slomovic S, Laufer D, Geiger D, Schuster G. Polyadenylation and degradation of human mitochondrial RNA: the prokaryotic past leaves its mark. Mol Cell Biol. 2005;25:6427–35. doi:10.1128/MCB.25.15.6427-6435.2005.
71. Gagliardi D, Stepien PP, Temperley RJ, Lightowlers RN, Chrzanowska-Lightowlers ZM. Messenger RNA stability in mitochondria: different means to an end. Trends Genet. 2004;20:260–7. doi:10.1016/j.tig.2004.04.006.
72. Etheridge RD, Aphasizheva I, Gershon PD, Aphasizhev R. 3' adenylation determines mRNA abundance and monitors completion of RNA editing in T. brucei mitochondria. EMBO J. 2008;27:1596–608. doi:10.1038/emboj.2008.87.
73. Schuster G, Stern D. RNA polyadenylation and decay in mitochondria and chloroplasts. Prog Mol Biol Transl Sci. 2009;85:393–422. doi:10.1016/S0079-6603(08)00810-6.
74. Dziembowski A, Piwowarski J, Hoser R, Minczuk M, Dmochowska A, Siep M, van der Spek H, Grivell L, Stepien PP. The yeast mitochondrial degradosome. Its composition, interplay between RNA helicase and RNase activities and the role in mitochondrial RNA metabolism. J Biol Chem. 2003;278:1603–1611.

75. Turk EM, Das V, Seibert RD, Andrulis ED. The mitochondrial RNA landscape of Saccharomyces cerevisiae. PLoS One. 2013;8:e78105.
76. Schafer B, Hansen M, Lang BF. Transcription and RNA-processing in fission yeast mitochondria. RNA. 2005;11:785–95. doi:10.1261/rna.7252205.
77. Kolondra A, Labedzka-Dmoch K, Wenda JM, Drzewicka K, Golik P. The transcriptome of Candida albicans mitochondria and the evolution of organellar transcription units in yeasts. BMC Genomics. 2015;16:827.
78. Butow RA, Zhu H, Perlman P, Conrad-Webb H. The role of a conserved dodecamer sequence in yeast mitochondrial gene expression. Genome. 1989;31:757–60.
79. Zhu H, Conrad-Webb H, Liao XS, Perlman PS, Butow RA. Functional expression of a yeast mitochondrial intron-encoded protein requires RNA processing at a conserved dodecamer sequence at the 3' end of the gene. Mol Cell Biol. 1989;9:1507–12.
80. Chateigner-Boutin AL, Small I. Organellar RNA editing. Wiley Interdiscip Rev RNA. 2011;2:493–506.
81. Takenaka M, Verbitskiy D, Zehrmann A, Hartel B, Bayer-Csaszar E, Glass F, Brennicke A. RNA editing in plant mitochondria-connecting RNA target sequences and acting proteins. Mitochondrion. 2014;19 Pt B:191–197.
82. Read LK, Lukes J, Hashimi H. Trypanosome RNA editing: the complexity of getting U in and taking U out. Wiley Interdiscip Rev RNA. 2016;7:33–51.
83. Lavrov DV, Adamski M, Chevaldonne P, Adamska M. Extensive mitochondrial mRNA editing and unusual mitochondrial genome organization in calcaronean sponges. Curr Biol. 2016;26:86–92.
84. Gray MW. Evolutionary origin of RNA editing. Biochemistry. 2012;51:5235–42.
85. Knoop V. Plant mitochondrial genome peculiarities evolving in the earliest vascular plant lineages. J Syst Evol. 2013;51:1–12.
86. Rudinger M, Funk HT, Rensing SA, Maier UG, Knoop V. RNA editing: only eleven sites are present in the Physcomitrella patens mitochondrial transcriptome and a universal nomenclature proposal. Mol Genet Genomics. 2009;281:473–81.
87. Bentolila S, Oh J, Hanson MR, Bukowski R. Comprehensive high-resolution analysis of the role of an Arabidopsis gene family in RNA editing. PLoS Genet. 2013;9:e1003584.
88. Grewe F, Herres S, Viehover P, Polsakiewicz M, Weisshaar B, Knoop V. A unique transcriptome: 1782 positions of RNA editing alter 1406 codon identities in mitochondrial mRNAs of the lycophyte Isoetes engelmannii. Nucleic Acids Res. 2011;39:2890–902.
89. Hecht J, Grewe F, Knoop V. Extreme RNA editing in coding islands and abundant microsatellites in repeat sequences of Selaginella moellendorffii mitochondria: the root of frequent plant mtDNA re-combination in early tracheophytes. Genome Biol Evol. 2011;3:344–58.
90. Benne R, Van den Burg J, Brakenhoff JP, Sloof P, Van Boom JH, Tromp MC. Major transcript of the frameshifted coxII gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. Cell. 1986;46:819–26.
91. Blum B, Bakalara N, Simpson L. A model for RNA editing in kinetoplastid mitochondria: "guide" RNA molecules transcribed from maxicircle DNA provide the edited information. Cell. 1990;60:189–98.
92. Liu M, Spremulli L. Interaction of mammalian mitochondrial ribosomes with the inner membrane. J Biol Chem. 2000;275:29400–6. doi:10.1074/jbc.M002173200.
93. Pfeffer S, Woellhaf MW, Herrmann JM, Forster F. Organization of the mitochondrial translation machinery studied in situ by cryoelectron tomography. Nat Commun. 2015;6:6019. doi:10.1038/ncomms7019.
94. Chacinska A, Koehler CM, Milenkovic D, Lithgow T, Pfanner N. Importing mitochondrial proteins: machineries and mechanisms. Cell. 2009;138:628–44. doi:10.1016/j.cell.2009.08.005.
95. Tzagoloff A, Myers AM. Genetics of mitochondrial biogenesis. Annu Rev Biochem. 1986;55:249–85. doi:10.1146/annurev.bi.55.070186.001341.

96. Bogenhagen DF, Martin DW, Koller A. Initial steps in RNA processing and ribosome assembly occur at mitochondrial DNA nucleoids. Cell Metab. 2014;19:618–29. doi:10.1016/j.cmet.2014.03.013.
97. De Silva D, Tu YT, Amunts A, Fontanesi F, Barrientos A. Mitochondrial ribosome assembly in health and disease. Cell Cycle. 2015;14:2226–50. doi:10.1080/15384101.2015.1053672.
98. Kitakawa M, Isono K. The mitochondrial ribosomes. Biochimie. 1991;73:813–25. doi:10.1016/0300-9084(91)90061-5.
99. Ramulu HG, Groussin M, Talla E, Planel R, Daubin V, Brochier-Armanet C. Ribosomal proteins: toward a next generation standard for prokaryotic systematics? Mol Phylogenet Evol. 2014;75:103–17. doi:10.1016/j.ympev.2014.02.013.
100. Desmond E, Brochier-Armanet C, Forterre P, Gribaldo S. On the last common ancestor and early evolution of eukaryotes: reconstructing the history of mitochondrial ribosomes. Res Microbiol. 2011;162:53–70. doi:10.1016/j.resmic.2010.10.004.
101. van der Sluis EO, Bauerschmitt H, Becker T, Mielke T, Frauenfeld J, Berninghausen O, Neupert W, Herrmann JM, Beckmann R. Parallel structural evolution of mitochondrial ribosomes and OXPHOS complexes. Genome Biol Evol. 2015;7:1235–51. doi:10.1093/gbe/evv061evv061.
102. Smits P, Smeitink JA, van den Heuvel LP, Huynen MA, Ettema TJ. Reconstructing the evolution of the mitochondrial ribosomal proteome. Nucleic Acids Res. 2007;35:4686–703. doi:10.1093/nar/gkm441.
103. Brown WM, George M Jr, Wilson AC. Rapid evolution of animal mitochondrial DNA. Proc Natl Acad Sci USA. 1979;76:1967–71.
104. Haag-Liautard C, Coffey N, Houle D, Lynch M, Charlesworth B, Keightley PD. Direct estimation of the mitochondrial DNA mutation rate in Drosophila melanogaster. PLoS Biol. 2008;6:e204.
105. Neiman M, Taylor DR. The causes of mutation accumulation in mitochondrial genomes. Proc Biol Sci. 2009;276:1201–9. doi:10.1098/rspb.2008.1758.
106. Lukes J, Archibald JM, Keeling PJ, Doolittle WF, Gray MW. How a neutral evolutionary ratchet can build cellular complexity. IUBMB Life. 2011;63:528–37.
107. Barreto FS, Burton RS. Evidence for compensatory evolution of ribosomal proteins in response to rapid divergence of mitochondrial rRNA. Mol Biol Evol. 2013;30:310–4. doi:10.1093/molbev/mss228.
108. Rand DM, Haney RA, Fry AJ. Cytonuclear coevolution: the genomics of cooperation. Trends Ecol Evol. 2004;19:645–53. doi:10.1016/j.tree.2004.10.003.
109. Berry S. Endosymbiosis and the design of eukaryotic electron transport. Biochim Biophys Acta. 2003;1606:57–72. doi:S0005272803000847[pii].
110. Gabaldon T, Rainey D, Huynen MA. Tracing the evolution of a large protein complex in the eukaryotes, NADH:ubiquinone oxidoreductase (Complex I). J Mol Biol. 2005;348:857–70. doi:10.1016/j.jmb.2005.02.067.
111. Yamaguchi K, Subramanian AR. The plastid ribosomal pro-teins. Identification of all the proteins in the 50 S subunit of an orga-nelle ribosome (chloroplast). J Biol Chem. 2000;275:28466–82. doi:10.1074/jbc.M005012200.
112. Yamaguchi K, von Knoblauch K, Subramanian AR. The plastid ribosomal proteins. Identification of all the proteins in the 30 S sub-unit of an organelle ribosome (chloroplast). J Biol Chem. 2000;275:28455–65. doi:10.1074/jbc.M004350200.
113. Spirina O, Bykhovskaya Y, Kajava AV, O'Brien TW, Nierlich DP, Mougey EB, Sylvester JE, Graack HR, Wittmann-Liebold B, Fischel-Ghodsian N. Heart-specific splice-variant of a human mitochondrial ribosomal protein (mRNA processing; tissue specific splicing). Gene. 2000;261:229–34. doi:10.1016/S0378-1119(00)00504-7.
114. Mogk A, Haslberger T, Tessarz P, Bukau B. Common and specific mechanisms of AAA+ proteins involved in protein quality control. Biochem Soc Trans. 2008;36:120–5. doi:10.1042/BST0360120.
115. Daubner GM, Clery A, Allain FH. RRM-RNA recognition: NMR or crystallography and new findings. Curr Opin Struct Biol. 2013;23:100–8. doi:10.1016/j.sbi.2012.11.006.

116. Kaur J, Stuart RA. Truncation of the Mrp20 protein reveals new ribosome-assembly subcomplex in mitochondria. EMBO Rep. 2011;12:950–5. doi:10.1038/embor.2011.133.
117. Ryden-Aulin M, Shaoping Z, Kylsten P, Isaksson LA. Ribo-some activity and modification of 16S RNA are influenced by deletion of ribosomal protein S20. Mol Microbiol. 1993;7:983–92.
118. Tobin C, Mandava CS, Ehrenberg M, Andersson DI, Sanyal S. Ribosomes lacking protein S20 are defective in mRNA binding and subunit association. J Mol Biol. 2010;397:767–76. doi:10.1016/j.jmb.2010.02.004.
119. Sorensen MA, Fricke J, Pedersen S. Ribosomal protein S1 is required for translation of most, if not all, natural mRNAs in Escherichia coli in vivo. J Mol Biol. 1998;280:561–9. doi:10.1006/jmbi.1998.1909.
120. O'Brien TW. Evolution of a protein-rich mitochondrial ribo-some: implications for human genetic disease. Gene. 2002;286:73–9. doi:S0378111901008083 [pii].
121. Amunts A, Brown A, Toots J, Scheres SH, Ramakrishnan V. Ribosome. The structure of the human mitochondrial ribosome. Science. 2015;348:95–8. doi:10.1126/science.aaa1193.
122. Greber BJ, Bieri P, Leibundgut M, Leitner A, Aebersold R, Boehringer D, Ban N. Ribosome. The complete structure of the 55S mammalian mitochondrial ribosome. Science. 2015;348:303–8. doi:10.1126/science.aaa3872
123. Sharma MR, Koc EC, Datta PP, Booth TM, Spremulli LL, Agrawal RK. Structure of the mammalian mitochondrial ribosome reveals an expanded functional role for its component proteins. Cell. 2003;115:97–108. doi:S0092867403007621 [pii].
124. Schneider A, Ebert D. Covariation of mitochondrial genome size with gene lengths: evidence for gene length reduction during mitochondrial evolution. J Mol Evol. 2004;59:90–6.
125. Greber BJ, Boehringer D, Leitner A, Bieri P, Voigts-Hoffmann F, Erzberger JP, Leibundgut M, Aebersold R, Ban N. Architecture of the large subunit of the mammalian mitochondrial ribosome. Nature. 2014;505:515–9. doi:10.1038/nature12890.
126. Simpson L, Thiemann OH, Savill NJ, Alfonzo JD, Maslov DA. Evolution of RNA editing in trypanosome mitochondria. Proc Natl Acad Sci USA. 2000;97:6986–93. doi:97/13/6986 [pii].
127. Antonicka H, Shoubridge EA. Mitochondrial RNA granules are centers for posttranscriptional RNA processing and ribosome biogenesis. Cell Rep. 2015;S2211–1247:00055–8. doi:10.1016/j.celrep.2015.01.030.
128. Gualerzi CO, Pon CL. Initiation of mRNA translation in bacteria: structural and dynamic aspects. Cell Mol Life Sci. 2015;72:4341–67.
129. Atkinson GC, Kuzmenko A, Kamenski P, Vysokikh MY, Lakunina V, Tankov S, Smirnova E, Soosaar A, Tenson T, Hauryliuk V. Evolutionary and genetic analyses of mitochondrial translation initiation factors identify the missing mitochondrial IF3 in S. cerevisiae. Nucleic Acids Res. 2012;40:6122–34.
130. Hinttala R, Sasarman F, Nishimura T, Antonicka H, Brunel-Guitton C, Schwartzentruber J, Fahiminiya S, Majewski J, Faubert D, Ostergaard E, Smeitink JA, Shoubridge EA. An N-terminal formyl methionine on COX 1 is required for the assembly of cytochrome c oxidase. Hum Mol Genet. 2015;24:4103–13.
131. Tucker EJ, Hershman SG, Kohrer C, Belcher-Timme CA, Patel J, Goldberger OA, Christodoulou J, Silberstein JM, McKenzie M, Ryan MT, Compton AG, Jaffe JD, Carr SA, Calvo SE, RajBhandary UL, Thorburn DR, Mootha VK. Mutations in MTFMT underlie a human disorder of formylation causing impaired mitochondrial translation. Cell Metab. 2011;14:428–34.
132. Li Y, Holmes WB, Appling DR, RajBhandary UL. Initiation of protein synthesis in Saccharomyces cerevisiae mitochondria without formylation of the initiator tRNA. J Bacteriol. 2000;182:2886–92.
133. Lee C, Tibbetts AS, Kramer G, Appling DR. Yeast AEP3p is an accessory factor in initiation of mitochondrial translation. J Biol Chem. 2009;284:34116–25.
134. Marzi S, Knight W, Brandi L, Caserta E, Soboleva N, Hill WE, Gualerzi CO, Lodmell JS. Ribosomal localization of translation initiation factor IF2. RNA. 2003;9:958–69.

135. Moreno JM, Sorensen HP, Mortensen KK, Sperling-Petersen HU. Macromolecular mimicry in translation initiation: a model for the initiation factor IF2 on the ribosome. IUBMB Life. 2000;50:347–54.
136. Spencer AC, Spremulli LL. Interaction of mitochondrial initiation factor 2 with mitochondrial fMet-tRNA. Nucleic Acids Res. 2004;32:5464–70.
137. Tibbetts AS, Oesterlin L, Chan SY, Kramer G, Hardesty B, Appling DR. Mammalian mitochondrial initiation factor 2 supports yeast mitochondrial translation without formylated initiator tRNA. J Biol Chem. 2003;278:31774–80.
138. Koc EC, Burkhart W, Blackburn K, Koc H, Moseley A, Spremulli LL. Identification of four proteins from the small subunit of the mammalian mitochondrial ribosome using a proteomics approach. Protein Sci. 2001;10:471–81.
139. Koc EC, Burkhart W, Blackburn K, Moyer MB, Schlatzer DM, Moseley A, Spremulli LL. The large subunit of the mammalian mitochondrial ribosome. Analysis of the complement of ribosomal proteins present. J Biol Chem. 2001;276:43958–69.
140. Spremulli LL, Coursey A, Navratil T, Hunter SE. Initiation and elongation factors in mammalian mitochondrial protein biosynthesis. Prog Nucleic Acid Res Mol Biol. 2004;77:211–61.
141. Laursen BS, Siwanowicz I, Larigauderie G, Hedegaard J, Ito K, Nakamura Y, Kenney JM, Mortensen KK, Sperling-Petersen HU. Characterization of mutations in the GTP-binding domain of IF2 resulting in cold-sensitive growth of Escherichia coli. J Mol Biol. 2003;326:543–51.
142. Gaur R, Grasso D, Datta PP, Krishna PD, Das G, Spencer A, Agrawal RK, Spremulli L, Varshney U. A single mammalian mito-chondrial translation initiation factor functionally replaces two bacterial factors. Mol Cell. 2008;29:180–90.
143. Roll-Mecak A, Cao C, Dever TE, Burley SK. X-Ray structures of the universal translation initiation factor IF2/eIF5B: conformational changes on GDP and GTP binding. Cell. 2000;103:781–92.
144. Guenneugues M, Caserta E, Brandi L, Spurio R, Meunier S, Pon CL, Boelens R, Gualerzi CO. Mapping the fMet-tRNA(f)(Met) binding site of initiation factor IF2. Embo J. 2000;19:5233–40.
145. Misselwitz R, Welfle K, Krafft C, Welfle H, Brandi L, Caserta E, Gualerzi CO. The fMet-tRNA binding domain of translational initiation factor IF2: role and environment of its two Cys residues. FEBS Lett. 1999;459:332–6.
146. Wienk H, Tomaselli S, Bernard C, Spurio R, Picone D, Gualerzi CO, Boelens R. Solution structure of the C1-subdomain of Bacillus stearothermophilus translation initiation factor IF2. Protein Sci. 2005;14:2461–8.
147. Grunberg-Manago M, Dessen P, Pantaloni D, Godefroy-Colburn T, Wolfe AD, Dondon J. Light-scattering studies showing the effect of initiation factors on the reversible dissociation of Escherichia coli ribosomes. J Mol Biol. 1975;94:461–78.
148. Yassin AS, Haque ME, Datta PP, Elmore K, Banavali NK, Spremulli LL, Agrawal RK. Insertion domain within mammalian mitochondrial translation initiation factor 2 serves the role of eubacterial initiation factor 1. Proc Natl Acad Sci USA. 2011;108:3918–23.
149. Koc EC, Spremulli LL. Identification of mammalian mitochondrial translational initiation factor 3 and examination of its role in initiation complex formation with natural mRNAs. J Biol Chem. 2002;277:35541–9.
150. Dallas A, Noller HF. Interaction of translation initiation factor 3 with the 30S ribosomal subunit. Mol Cell. 2001;8:855–64.
151. Pioletti M, Schlunzen F, Harms J, Zarivach R, Gluhmann M, Avila H, Bashan A, Bartels H, Auerbach T, Jacobi C, Hartsch T, Yonath A, Franceschi F. Crystal structures of complexes of the small ribosomal subunit with tetracycline, edeine and IF3. Embo J. 2001;20:1829–39.
152. Sette M, Spurio R, van Tilborg P, Gualerzi CO, Boelens R. Identification of the ribosome binding sites of translation initiation factor IF3 by multidimensional heteronuclear NMR spectroscopy. RNA. 1999;5:82–92.

153. Dieckmann CL, Staples RR. Regulation of mitochondrial gene expression in Saccharomyces cerevisiae. Int Rev Cytol. 1994;152:145–81.
154. Christian BE, Spremulli LL. Preferential selection of the 5'-terminal start codon on leaderless mRNAs by mammalian mitochondrial ribosomes. J Biol Chem. 2010;285:28379–86.
155. Fox TD. Genetics of mitochondrial translation. In: Hershey JWB, Matthews MB, Sonenberg N, editors. Translational control. Cold Spring Harbor, NY: Cold Spring Harbor Press; 1996. p. 733–58.
156. Herrmann JM, Woellhaf MW, Bonnefoy N. Control of protein synthesis in yeast mitochondria: the concept of translational activators. Biochim Biophys Acta. 2013;1833:286–94.
157. Zamudio-Ochoa A, Camacho-Villasana Y, Garcia-Guerrero AE, Perez-Martinez X. The Pet309 pentatricopeptide repeat motifs mediate efficient binding to the mitochondrial COX1 transcript in yeast. RNA Biol. 2014;11:953–67.
158. Naithani S, Saracco SA, Butler CA, Fox TD. Interactions among COX1, COX2, and COX3 mRNA-specific translational activator proteins on the inner surface of the mitochondrial inner membrane of Saccharomyces cerevisiae. Mol Biol Cell. 2003;14:324–33.
159. Ott M, Amunts A, Brown A. Organization and regulation of mitochondrial protein synthesis. Annu Rev Biochem. 2016; In press. doi:10.1146/annurev-biochem-060815-014334.
160. Sanchirico ME, Fox TD, Mason TL. Accumulation of mitochondrially synthesized Saccharomyces cerevisiae Cox2p and Cox3p depends on targeting information in untranslated portions of their mRNAs. Embo J. 1998;17:5796–804.
161. Coffin JW, Dhillon R, Ritzel RG, Nargang FE. The Neurospora crassa cya-5 nuclear gene encodes a protein with a region of homology to the Saccharomyces cerevisiae PET309 protein and is required in a post-transcriptional step for the expression of the mitochondrially encoded COXI protein. Curr Genet. 1997;32:273–80.
162. Costanzo MC, Bonnefoy N, Williams EH, Clark-Walker GD, Fox TD. Highly diverged homologs of Saccharomyces cerevisiae mitochondrial mRNA-specific translational activators have orthologous functions in other budding yeasts. Genetics. 2000;154:999–1012.
163. Kuhl I, Dujeancourt L, Gaisne M, Herbert CJ, Bonnefoy N. A genome wide study in fission yeast reveals nine PPR proteins that regulate mitochondrial gene expression. Nucleic Acids Res. 2011;39:8029–41.
164. Szklarczyk R, Wanschers BF, Cuypers TD, Esseling JJ, Riemersma M, van den Brand MA, Gloerich J, Lasonder E, van den Heuvel LP, Nijtmans LG, Huynen MA. Iterative orthology predic-tion uncovers new mitochondrial proteins and identifies C12orf62 as the human ortholog of COX14, a protein involved in the assembly of cytochrome c oxidase. Genome Biol. 2012;13:R12.
165. Choquet Y, Wollman FA. Translational regulations as specific traits of chloroplast gene expression. FEBS Lett. 2002;529:39–42.
166. Lefebvre-Legendre L, Choquet Y, Kuras R, Loubery S, Douchi D, Goldschmidt-Clermont M. A nucleus-encoded chloroplast protein regulated by iron availability governs expression of the photosystem I subunit PsaA in Chlamydomonas reinhardtii. Plant Physiol. 2015;167:1527–40.
167. Prikryl J, Rojas M, Schuster G, Barkan A. Mechanism of RNA stabilization and translational activation by a pentatricopeptide repeat protein. Proc Natl Acad Sci USA. 2011;108:415–20.
168. Weraarpachai W, Sasarman F, Nishimura T, Antonicka H, Aure K, Rotig A, Lombes A, Shoubridge EA. Mutations in C12orf62, a factor that couples COX I synthesis with cytochrome c oxidase assembly, cause fatal neonatal lactic acidosis. Am J Hum Genet. 2011;90:142–51.
169. Williams EH, Butler CA, Bonnefoy N, Fox TD. Translation initiation in Saccharomyces cerevisiae mitochondria: functional interactions among mitochondrial ribosomal protein Rsm28p, initiation factor 2, methionyl-tRNA-formyltransferase and novel protein Rmd9p. Genetics. 2007;175:1117–26.
170. Arita M, Suematsu T, Osanai A, Inaba T, Kamiya H, Kita K, Sisido M, Watanabe Y, Ohtsuki T. An evolutionary 'intermediate state' of mitochondrial translation systems found in

Trichinella species of parasitic nematodes: co-evolution of tRNA and EF-Tu. Nucleic Acids Res. 2006;34:5291–9.
171. Ohtsuki T, Watanabe Y, Takemoto C, Kawai G, Ueda T, Kita K, Kojima S, Kaziro Y, Nyborg J, Watanabe K. An "elongated" translation elongation factor Tu for truncated tRNAs in nematode mitochondria. J Biol Chem. 2001;276:21571–7.
172. Sakurai M, Watanabe Y, Watanabe K, Ohtsuki T. A protein ex-tension to shorten RNA: elongated elongation factor-Tu recognizes the D-arm of T-armless tRNAs in nematode mitochondria. Biochem J. 2006;399:249–56. doi:10.1042/BJ20060781.
173. Suematsu T, Sato A, Sakurai M, Watanabe K, Ohtsuki T. A unique tRNA recognition mechanism of Caenorhabditis elegans mitochondrial EF-Tu2. Nucleic Acids Res. 2005;33:4683–91.
174. Cristodero M, Mani J, Oeljeklaus S, Aeberhard L, Hashimi H, Ramrath DJ, Lukes J, Warscheid B, Schneider A. Mitochondrial translation factors of Trypanosoma brucei: elongation factor-Tu has a unique subdomain that is essential for its function. Mol Microbiol. 2013;90:744–55.
175. Sharma MR, Booth TM, Simpson L, Maslov DA, Agrawal RK. Structure of a mitochondrial ribosome with minimal RNA. Proc Natl Acad Sci USA. 2009;106:9637–42.
176. Esseiva AC, Naguleswaran A, Hemphill A, Schneider A. Mito-chondrial tRNA import in Toxoplasma gondii. J Biol Chem. 2004;279:42363–8.
177. Pino P, Aeby E, Foth BJ, Sheiner L, Soldati T, Schneider A, Soldati-Favre D. Mitochondrial translation in absence of local tRNA aminoacylation and methionyl tRNA Met formylation in Apicomplexa. Mol Microbiol. 2010;76:706–18.
178. Gaillardin C, Duchateau-Nguyen G, Tekaia F, Llorente B, Casaregola S, Toffano-Nioche C, Aigle M, Artiguenave F, Blandin G, Bolotin-Fukuhara M, Bon E, Brottier P, de Montigny J, Dujon B, Durrens P, Lepingle A, Malpertuy A, Neuveglise C, Ozier-Kalogeropoulos O, Potier S, Saurin W, Termier M, Wesolowski-Louvel M, Wincker P, Souciet J, Weissenbach J. Genomic exploration of the hemiascomycetous yeasts: 21. Comparative functional classification of genes. FEBS Lett. 2000;487:134–49. doi:S0014579300022924 [pii].
179. Rosenthal LP, Bodley JW. Purification and characterization of Saccharomyces cerevisiae mitochondrial elongation factor Tu. J Biol Chem. 1987;262:10955–9.
180. Chiron S, Suleau A, Bonnefoy N. Mitochondrial translation: elongation factor tu is essential in fission yeast and depends on an exchange factor conserved in humans but not in budding yeast. Genetics. 2005;169:1891–901. doi:10.1534/genetics.104.037473.
181. Atkinson GC, Baldauf SL. Evolution of elongation factor G and the origins of mitochondrial and chloroplast forms. Mol Biol Evol. 2011;28:1281–92. doi:10.1093/molbev/msq316.
182. Suematsu T, Watanabe O, Kita K, Yokobori S, Watanabe Y. Arabidopsis thaliana mitochondrial EF-G1 functions in two different translation steps. J Biochem. 2013;155:107–14. doi:10.1093/jb/mvt105.
183. Tsuboi M, Morita H, Nozaki Y, Akama K, Ueda T, Ito K, Nierhaus KH, Takeuchi N. EF-G2mt is an exclusive recycling factor in mammalian mitochondrial protein synthesis. Mol Cell. 2009;35:502–10.
184. Gupta A, Mir SS, Jackson KE, Lim EE, Shah P, Sinha A, Siddiqi MI, Ralph SA, Habib S. Recycling factors for ribosome disassembly in the apicoplast and mitochondrion of Plasmodium falciparum. Mol Microbiol. 2013;88:891–905. doi:10.1111/mmi.12230.
185. Johnson RA, McFadden GI, Goodman CD. Characterization of two malaria parasite organelle translation elongation factor G proteins: the likely targets of the anti-malarial fusidic acid. PLoS One. 2011;6:e20633. doi:10.1371/journal.pone.0020633.
186. Scolnick E, Tompkins R, Caskey T, Nirenberg M. Release factors differing in specificity for terminator codons. Proc Natl Acad Sci USA. 1968;61:768–74.
187. Klaholz BP. Molecular recognition and catalysis in translation termination complexes. Trends Biochem Sci. 2011;36:282–92.
188. Duarte I, Nabuurs SB, Magno R, Huynen M. Evolution and diversification of the organellar release factor family. Mol Biol Evol. 2012;29:3497–512. doi:10.1093/molbev/mss157.

189. Nozaki Y, Matsunaga N, Ishizawa T, Ueda T, Takeuchi N. HMRF1L is a human mitochondrial translation release factor involved in the decoding of the termination codons UAA and UAG. Genes Cells. 2008;13:429–38. doi:10.1111/j.1365-2443.2008.01181.x.
190. Soleimanpour-Lichaei HR, Kuhl I, Gaisne M, Passos JF, Wydro M, Rorbach J, Temperley R, Bonnefoy N, Tate W, Lightowlers R, Chrzanowska-Lightowlers Z. mtRF1a is a human mitochondrial translation release factor decoding the major termination codons UAA and UAG. Mol Cell. 2007;27:745–57. doi:10.1016/j.molcel.2007.06.031.
191. Young DJ, Edgar CD, Murphy J, Fredebohm J, Poole ES, Tate WP. Bioinformatic, structural, and functional analyses support release factor-like MTRF1 as a protein able to decode nonstandard stop codons beginning with adenine in vertebrate mitochondria. RNA. 2010;16:1146–55.
192. Akabane S, Ueda T, Nierhaus KH, Takeuchi N. Ribosome rescue and translation termination at non-standard stop codons by ICT1 in mammalian mitochondria. PLoS Genet. 2014;10: e1004616. doi:10.1371/journal.pgen.1004616.
193. Lind C, Sund J, Aqvist J. Codonreading specificities of mitochondrial release factors and translation termination at non-standard stop codons. Nat Commun. 2013;4:2940. doi:10.1038/ncomms3940.
194. Richter R, Rorbach J, Pajak A, Smith PM, Wessels HJ, Huynen MA, Smeitink JA, Lightowlers RN, Chrzanowska-Lightowlers ZM. A functional peptidyl-tRNA hydrolase, ICT1, has been recruited into the human mitochondrial ribosome. EMBO J. 2010;29:1116–25. doi:10.1038/emboj.2010.14.
195. Gagnon MG, Seetharaman SV, Bulkley D, Steitz TA. Structur-al basis for the rescue of stalled ribosomes: structure of YaeJ bound to the ribosome. Science. 2012;335:1370–2. doi:10.1126/science.1217443.
196. Handa Y, Inaho N, Nameki N. YaeJ is a novel ribosome-associated protein in Escherichia coli that can hydrolyze peptidyl-tRNA on stalled ribosomes. Nucleic Acids Res. 2010;39:1739–48. doi:10.1093/nar/gkq1097.
197. Antonicka H, Ostergaard E, Sasarman F, Weraarpachai W, Wibrand F, Pedersen AM, Rodenburg RJ, van der Knaap MS, Smeitink JA, Chrzanowska-Lightowlers ZM, Shoubridge EA. Mutations in C12orf65 in patients with encephalomyopathy and a mitochondrial translation defect. Am J Hum Genet. 2011;87:115–22.
198. Rorbach J, Richter R, Wessels HJ, Wydro M, Pekalski M, Farhoud M, Kuhl I, Gaisne M, Bonnefoy N, Smeitink JA, Lightowlers RN, Chrzanowska-Lightowlers ZM. The human mitochondrial ribosome recycling factor is essential for cell viability. Nucleic Acids Res. 2008;36:5787–99. doi:10.1093/nar/gkn576.
199. Christian BE, Spremulli LL. Evidence for an active role of IF3mt in the initiation of translation in mammalian mitochondria. Biochemistry. 2009;48:3269–78. doi:10.1021/bi8023493.
200. Barrientos A, Zambrano A, Tzagoloff A. Mss51p and Cox14p jointly regulate mitochondrial Cox1p expression in Saccharomyces cerevisiae. Embo J. 2004;23:3472–82.
201. Perez-Martinez X, Broadley SA, Fox TD. Mss51p promotes mitochondrial Cox1p synthesis and interacts with newly synthesized Cox1p. Embo J. 2003;22:5951–61.
202. Khalimonchuk O, Bird A, Winge DR. Evidence for a pro-oxidant intermediate in the assembly of cytochrome oxidase. J Biol Chem. 2007;282:17442–9.
203. Rak M, Su CH, Xu JT, Azpiroz R, Singh AM, Tzagoloff A. Regulation of mitochondrial translation of the ATP8/ATP6 mRNA by Smt1p. Mol Biol Cell. 2016.
204. Rak M, Tzagoloff A. F1-dependent translation of mitochondrially encoded Atp6p and Atp8p subunits of yeast ATP synthase. Proc Natl Acad Sci USA. 2009;106:18509–14.
205. Godard F, Tetaud E, Duvezin-Caubet S, di Rago JP. A genetic screen targeted on the FO component of mitochondrial ATP synthase in Saccharomyces cerevisiae. J Biol Chem. 2011;286:18181–9.
206. Rak M, Gokova S, Tzagoloff A. Modular assembly of yeast mitochondrial ATP synthase. Embo J. 2011;30:920–30.

207. Gruschke S, Kehrein K, Rompler K, Grone K, Israel L, Imhof A, Herrmann JM, Ott M. Cbp3-Cbp6 interacts with the yeast mitochondrial ribosomal tunnel exit and promotes cytochrome b synthesis and assembly. J Cell Biol. 2011;193:1101–14.
208. Mick DU, Fox TD, Rehling P. Inventory control: cytochrome c oxidase assembly regulates mitochondrial translation. Nat Rev Mol Cell Biol. 2011;12:14–20.
209. Dennerlein S, Rehling P. Human mitochondrial COX1 assembly into cytochrome c oxidase at a glance. J Cell Sci. 2015;128:833–7.
210. Rodel G. Translational activator proteins required for cytochrome b synthesis in Saccharomyces cerevisiae. Curr Genet. 1997;31:375–9.
211. Manthey GM, McEwen JE. The product of the nuclear gene PET309 is required for translation of mature mRNA and stability or production of intron-containing RNAs derived from the mitochondrial COX1 locus of Saccharomyces cerevisiae. Embo J. 1995;14: 4031–43.
212. Mulero JJ, Fox TD. PET111 acts in the 5'-leader of the Saccha-romyces cerevisiae mitochondrial COX2 mRNA to promote its translation. Genetics. 1993;133:509–16.
213. Costanzo MC, Fox TD. Product of Saccharomyces cerevisiae nuclear gene PET494 activates translation of a specific mitochondrial mRNA. Mol Cell Biol. 1986;6:3694–703.
214. Costanzo MC, Seaver EC, Fox TD. At least two nuclear gene products are specifically required for translation of a single yeast mitochondrial mRNA. Embo J. 1986;5:3637–41.
215. Costanzo MC, Fox TD. Specific translational activation by nuclear gene products occurs in the 5' untranslated leader of a yeast mitochondrial mRNA. Proc Nat Acad Sci USA. 1988;85:2677–81.
216. Finnegan PM, Payne MJ, Keramidaris E, Lukins HB. Characterization of a yeast nuclear gene, AEP2, required for accumulation of mitochondrial mRNA encoding subunit 9 of the ATP synthase. Curr Genet. 1991;20:53–61.
217. Payne MJ, Finnegan PM, Smooker PM, Lukins HB. Characterization of a second nuclear gene, AEP1, required for expression of the mitochondrial OLI1 gene in Saccharomyces cerevisiae. Curr Genet. 1993;24:126–35.
218. Mimaki M, Wang X, McKenzie M, Thorburn DR, Ryan MT. Understanding mitochondrial complex I assembly in health and disease. Biochim Biophys Acta. 2012;1817:851–62.
219. Wanschers BF, Szklarczyk R, van den Brand MA, Jonckheere A, Suijskens J, Smeets R, Rodenburg RJ, Stephan K, Helland IB, Elkamil A, Rootwelt T, Ott M, van den Heuvel L, Nijtmans LG, Huynen MA. A mutation in the human CBP4 ortholog UQCC3 im-pairs complex III assembly, activity and cytochrome b stability. Hum Mol Genet. 2014;23: 6356–65.
220. Duarte FV, Palmeira CM, Rolo AP. The role of microRNAs in mitochondria: small players acting wide. Genes (Basel). 2014;5:865–86.
221. Rackham O, Shearwood AM, Mercer TR, Davies SM, Mattick JS, Filipovska A. Long noncoding RNAs are generated from the mitochondrial genome and regulated by nuclear-encoded proteins. RNA. 2011;17:2085–93.
222. Richter-Dennerlein R, Dennerlein S, Rehling P. Integrating mitochondrial translation into the cellular context. Nat Rev Mol Cell Biol. 2015;16:586–92.
223. Lang BF, Jakubkova M, Hegedusova E, Daoud R, Forget L, Brejova B, Vinar T, Kosa P, Fricova D, Nebohacova M, Griac P, Tomaska L, Burger G, Nosek J. Massive programmed translational jumping in mitochondria. Proc Natl Acad Sci USA. 2014;111:5926–31.
224. Nosek J, Tomaska L, Burger G, Lang BF. Programmed translational bypassing elements in mitochondria: structure, mobility, and evolutionary origin. Trends Genet. 2015;31:187–94.
225. Watanabe Y, Tsurui H, Ueda T, Furushima R, Takamiya S, Kita K, Nishikawa K, Watanabe K. Primary and higher order structures of nematode (Ascaris suum) mitochondrial tRNAs lacking either the T or D stem. J Biol Chem. 1994;269:22902–6.

eIF4Es and Their Interactors from Yeast Species

Daniela Ross and Michael Altmann

1 Introduction

The kingdom Fungi is divided into six phyla, namely Microspora, Zygomycota, Glomeromycota, Ascomycota, Basidiomycota and Lichens (see Fig. 1a). Phylum Ascomycota is divided into the subphyla Taphrinomycotina, Saccharomycotina and Pezizomycotina [1]. In this chapter we describe the variety of eIF4Es and associated proteins in unicellular fungi belonging to the subphyla Taphrinomycotina and Saccharomycotina, which are generally quoted as yeasts. The third subphylum, Pezizomycotina (consisting mostly of mould species such as Penicillium and Aspergillus), is not clearly defined as uni- or multicellular and will not be further analysed in this chapter.

The best known yeast species are *Saccharomyces cerevisiae* (budding yeast) and *Schizosaccharomyces pombe* (fission yeast). Both species belong to the subphyla Saccharomycotina/"true yeasts" and Taphrinomycotina, respectively (for more details on phylogenetic classification data; see Fig. 1b). Though unicellular, yeast populations possess—especially when under stress—multicellular-like properties such as pseudohyphenation [2, 3] and quorum sensing, thereby increasing their survival potential [4].

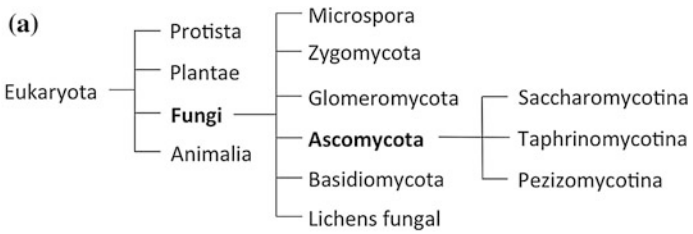
The genomes of baker's budding yeast *Saccharomyces cerevisiae* and fission yeast *Schizosaccharomyces pombe* were sequenced about 15–20 years ago and, as versatile model organisms, have rendered plenty of biological information to our current knowledge on eukaryotic processes. The genomic sequencing of less known unicellular yeast species has been pursued in recent years, thereby allowing comparative studies of all gene encoding translation factors. In this chapter, we will

D. Ross · M. Altmann (✉)
Institut für Biochemie und Molekulare Medizin (IBMM), University of Bern, Bühlstr. 28, 3012 Bern, Switzerland
e-mail: michael.altmann@ibmm.unibe.ch

D. Ross
e-mail: daniela.ross@ibmm.unibe.ch

G. Hernández and R. Jagus (eds.), *Evolution of the Protein Synthesis Machinery and Its Regulation*, DOI 10.1007/978-3-319-39468-8_7

(a)



(b)

Subphylum	Class	Order	Family	Genus	*Species*	
Taphrinomycotina	Schizosaccharomycetes	Schizosaccharomycetales	Schizosaccharomycetaceae	Schizosaccharomyces	*S. cryophilus*	SCHCR
					S. japonicus	SCHJY
					S. octosporus	SCHOY
					S. pombe	SCHPO
	Pneumocystidomycetes	Pneumocystidales	Pneumocystidaceae	Pneumocystis	*P. murina*	PNEMU
	Taphrinomycetes	Taphrinales	Taphrinaceae	Taphrina	*T. deformans*	TAPDE
Saccharomycotina	Saccharomycetes	Saccharomycetales	Trichomonascaceae	Blastobotrys	*B. adeninivorans*	BLAAD
			Dipodascaceae	Yarrowia	*Y. lipolytica*	YARLI
			Pichiaceae	Pichia	*Issatchenkia orientalis*	ISSOR
				Brettanomyces	*Dekkera bruxellensis*	DEKBR
				Ogataea	*O. parapolymorpha*	OGAPD
			Saccharomycetaceae	Eremothecium	*Ashbya gossypii*	ASHGO
					Ashbya aceri	ASHAC
					E. cymbalariae	ERECY
				Lachancea	*L. thermotolerans*	LACTC
				Kluyveromyces	*K. lactis*	KLULA
					K. marxianus	KLUMA
				Kazachstania	*K. africana*	KAZAF
					K. naganishii	KAZNA
				Tetrapisispora	*T. phaffii*	TETPH
					T. blattae	TETBL
				Nakaseomyces/Candida	*C. glabrata*	CANGA
				Naumovozyma	*N. dairenensis*	NAUDC
					N. castellii	NAUCC
				Torulaspora	*T. delbrueckii*	TORDC
				Saccharomyces	*S. cerevisiae*	YEAST
					S. arboricola	SACAR
					S. kudriavzevii	SACK1
				Zygosaccharomyces	*Z. bailii*	ZYGBA
					Z. rouxii	ZYGRC
				Vanderwaltozyma	*V. polyspora*	VANPO
				Hanseniaspora	*H. uvarum*	HANUV
			Phaffomycetaceae	Komagataella/Pichia	*P. pastoris*	PICPG
				Cyberlindnera	*C. fabianii*	CYBFA
				Wickerhamomyces	*W. ciferrii*	WICCF
			Debaryomycetaceae	Candida/Lodderomyces	*C. albicans*	CANAL
					C. tropicales	CANTT
					C. maltosa	CANMX
					C. dubliniensis	CANDC
					C. orthopsilosis	CANO9
					C. parapsilosis	CANPC
					L. elongisporus	LODEL
				Yamadazyma	*Candida tenuis*	CANTC
				Millerozyma	*Pichia sorbitophila*	PICSO
				Spathaspora	*S. passalidarum*	SPAPN
				Scheffersomyces	*S. stipitis*	PICST
				Meyerozyma	*M. guilliermondii*	PICGU
				Debaryomyces	*D. hansenii*	DEBHA
			Metschnikowiaceae	Clavispora	*C. lusitaniae*	CLAL4

Fig. 1 **a** Phylogenetic classification of the different kingdoms, the six fungi phyla and the three Ascomycota subphyla. **b** Table showing the phylogenetic classification of subphyla Saccharomycotina/"true yeasts" and Taphrinomycotina. All species analysed in this study are named with corresponding abbreviations (identifiers): 6 species from the subphylum Taphrinomycotina (*dark grey*) and 43 species from the subphylum Saccharomycotina (*light grey*) classified according to class, order, family and genus [1]

review the current knowledge on the following initiation factors from yeasts: the cap-binding protein eIF4E and its interactor eIF4G, which acts as a scaffold protein for further initiation factors (eIF4E and eIF4G form together with the helicase eIF4A the eIF4F complex), as well as two further eIF4E interactors termed p20 and Eap1. We also will comment on the evolutionary implications of the new findings mentioned here.

2 eIF4E

In 1997, the 3D structure of *S. cerevisiae* and murine eIF4E (bound to a cap analogue) was solved [5, 6]. It resembles a baseball glove where the convex part (the palm of the glove) consisting of eight antiparallel beta sheets (b1–b8) forms the cap-binding groove and is covered by three long anti-parallel helices (a2, a4 and a5). Three short alpha-helical structures (a1, a3 and a6) are inserted in loops connecting beta sheets and contribute to the concave part of the protein to which the eIF4E interactors bind. The unstructured amino terminal part of yeast eIF4E (aa 1–36) was originally missing in structural studies and was later solved when yeast eIF4E was obtained in a complex with the central part of eIF4G (aa 393–490), which when bound to eIF4E changes its structure [7]. As seen in Fig. 2, the amino terminus of yeast eIF4E—though unstructured—carries an additional short helix (aa T25–S30).

While the amino terminus of yeast eIF4E (aa 1–20) can be deleted without noticeable consequences, further deletions (up to aa 1–35) render a strong, slow growth phenotype and further amino terminal deletions result in a lethal phenotype. This can be due to eIF4E's reduced interaction with the central core of eIF4G [7]. Thus, the amino terminus of eIF4E—though mostly unstructured—plays an important role in its proper interaction with eIF4G.

In eukaryotic species, a conserved core of 160–170 amino acids of eIF4E contains eight aromatic residues (mostly tryptophanes, which are named Trp1 to Trp8 according to their position; see Fig. 3b). The aromatic residues of Trp3 and

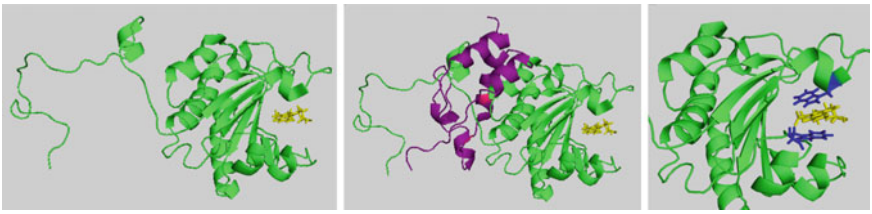
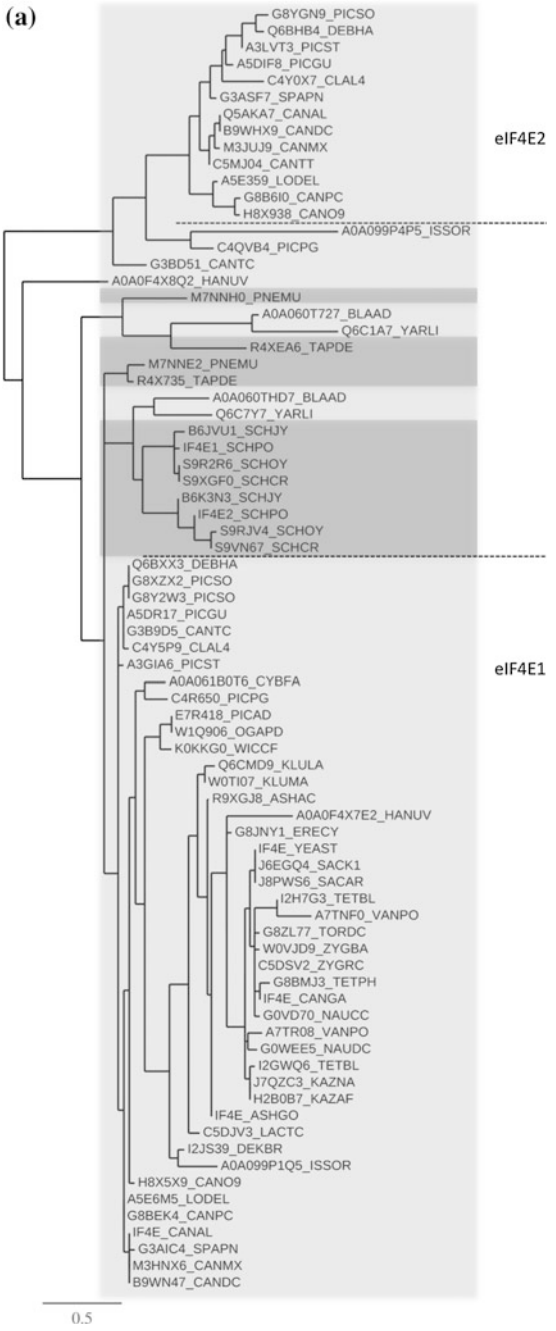


Fig. 2 3D-structure of yeast eIF4E. *Left panel* full-length yeast eIF4E (213 amino acids, *green/light grey*) interacting with m^7GDP (*yellow sticks*); *middle panel* full-length yeast eIF4E in a complex with eIF4G aa 405–488 (*purple/dark grey*); *right panel* close-up of the concave mRNA cap-binding site highlighting Trp3 and Trp5 (*blue/dark grey sticks*); 3D structure as determined by Gross et al. [7] (PDB structure 1RF8)


(a)
G8YGN9_PICSO
Q6BHB4_DEBHA
A3LVT3_PICST
A5DIF8_PICGU
C4Y0X7_CLAL4
G3ASF7_SPAPN
Q5AKA7_CANAL
B9WHX9_CANDC
M3JUJ9_CANMX
C5MJ04_CANTT
A5E359_LODEL
G8B6I0_CANPC
H8X938_CANO9
A0A099P4P5_ISSOR
C4QVB4_PICPG
G3BD51_CANTC
A0A0F4X8Q2_HANUV
M7NNH0_PNEMU
A0A060T727_BLAAD
Q6C1A7_YARLI
R4XEA6_TAPDE
M7NNE2_PNEMU
R4X735_TAPDE
A0A060THD7_BLAAD
Q6C7Y7_YARLI
B6JVU1_SCHJY
IF4E1_SCHPO
S9R2R6_SCHOY
S9XGF0_SCHCR
B6K3N3_SCHJY
IF4E2_SCHPO
S9RJV4_SCHOY
S9VN67_SCHCR
eIF4E2
Q6BXX3_DEBHA
G8XZX2_PICSO
G8Y2W3_PICSO
A5DR17_PICGU
G3B9D5_CANTC
C4Y5P9_CLAL4
A3GIA6_PICST
A0A061B0T6_CYBFA
C4R650_PICPG
E7R418_PICAD
W1Q906_OGAPD
K0KKG0_WICCF
Q6CMD9_KLULA
W0TI07_KLUMA
R9XGJ8_ASHAC
A0A0F4X7E2_HANUV
G8JNY1_ERECY
IF4E_YEAST
J6EGQ4_SACK1
J8PWS6_SACAR
I2H7G3_TETBL
A7TNF0_VANPO
G8ZL77_TORDC
W0VJD9_ZYGBA
C5DSV2_ZYGRC
G8BMJ3_TETPH
IF4E_CANGA
G0VD70_NAUCC
A7TR08_VANPO
G0WEE5_NAUDC
I2GWQ6_TETBL
J7QZC3_KAZNA
H2B0B7_KAZAF
IF4E_ASHGO
C5DJV3_LACTC
I2JS39_DEKBR
A0A099P1Q5_ISSOR
H8X5X9_CANO9
A5E6M5_LODEL
G8BEK4_CANPC
IF4E_CANAL
G3AIC4_SPAPN
M3HNX6_CANMX
B9WN47_CANDC
eIF4E1
0.5

◀ **Fig. 3** **a** Dendrogram of eIF4Es from subphyla Saccharomycotina/"true yeasts" and Taphrinomycotina. Subdivisions of Saccharomycotina (*light grey*) with eIF4E2, Taphrinomycotina (*dark grey*) and Saccharomycotina (*light grey*) with one or two eIF4E1s are marked. **b** eIF4E1 from different yeast species. Tryptophanes 1–8 (*velvet/dark grey*) are marked; Trp1 corresponds to Trp-43, Trp2 to Trp-46, Trp3 to Trp-58, Trp4 to Trp-75, Trp5 to Trp-104, Trp6 to Trp-115, Trp7 to Trp-130 and Trp8 to Trp-166 of the yeast amino acid sequence. SCHPO = *S. pombe*; CANAL = *C. albicans*; PICPG = *Pichia pastoris*; TETBL = *T. blattae*; VANPO = *V. polyspora*; CANGA = *Candida glabrata*; YEAST = *S. cerevisiae*). Besides *S. pombe* eIF4E1, all aligned eIF4Es belong to species from the subphylum Saccharomycotina. The sequence alignment was produced with Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/). **c** Sequence alignment of yeast eIF4E and *S. pombe* eIF4E1 and eIF4E2. Conserved Trp1-8 (velvet/dark grey) and D29 of *S. pombe* eIF4E2 (light grey), which was used to introduce an amino terminal deletion of *S. pombe* eIF4E2 (see text), are marked

(b)

```
sp|P78954|IF4E1_SCHPO1      MQTEQPPKESQTENTVSEPQEKALRTVFDDKINFNLKHPLARPWTLWFLMPPTP-GLEWN
sp|Q9P975|IF4E_CANAL        -MSEELAQKT-EEL------SLDSKTVFDSKEEFNAKHPLNSRWTLWYTKPQTNKSENWH
tr|C4R650|C4R650_PICPG      --------------MSETENAPHEKSLLDDQVDFTVKHPLNSEWTLWYTKPAVHEQESWS
tr|I2H7G3|I2H7G3_TETBL      MSADEVAQKL-EKTTL-----------------NSVSHPLNTKWTLWYTKPAIDKSESWS
tr|A7TR08|A7TR08_VANPO      MSVEEVTQKL-ESVAVDD---NSPKTVLSDSSNFEVKHPLNSKWTLWYTKPAVDKSETWS
tr|A7TNF0|A7TNF0_VANPO      MSVEEVSKKF-EETVSVS-------------GKVQVQHPLNSKWTLWYTKPAVDKSESWS
tr|I2GWQ6|I2GWQ6_TETBL      MSVDEVTQKL-DETKLNE---DTPKTILSDSTQFEYKHPLNTKWTLWYTKPAVDKSESWS
sp|P07260|IF4E_YEAST        MSVEEVSKKF-EENVSVDDTTATPKTVLSDSAHFDVKHPLNTKWTLWYTKPAVDKSESWS
sp|Q9P974|IF4E_CANGA        MSVDEVTKKF-EETVSVD----GPKTVLSDAKDFEVKHPLNTKWTLWYTKPAVDKSESWS
                                                              .***   ****:  *       *

sp|P78954|IF4E1_SCHPO1      ELQKNIITFNSVEEFWGIHNNINPASSLPIKSDYSFFREGVRPEWEDVHNKTGGKWAFQN
sp|Q9P975|IF4E_CANAL        DLLKPVITFSSVEEFWGIYNSIPPANQLPLKSDYHLFKEGIRPEWEDEANSKGGKWQFSF
tr|C4R650|C4R650_PICPG      DLLRPIISFNSVEEFWGIFNSIPKASDLPVKSDYHLFRDGIKPEWEDERNANGGKLTYQF
tr|I2H7G3|I2H7G3_TETBL      DLLRPVTSFKTVEEFWAVQQNVPEPRDLPMKSDYHVFRNDIRPEWEDAANAKGGKWSAQI
tr|A7TR08|A7TR08_VANPO      DLLRPVTSFETVEEFWAIIQNIPEPHELPMKSDYHVFRDNIRPEWEDPANSKGGKWSFHI
tr|A7TNF0|A7TNF0_VANPO      DLLRPVTSFGSVEEFWAVHENVPVPHELPLKSDYHIFRDQIRPEWEDTANAKGGKWSFQL
tr|I2GWQ6|I2GWQ6_TETBL      DLLRPVTSFETVEEFWAIFKNIPTPIDLPLKSDYHVFRNDIRPEWEDEANSKGGKWSFQV
sp|P07260|IF4E_YEAST        DLLRPVTSFQTVEEFWAIIQNIPEPHELPLKSDYHVFRNDVRPEWEDEANAKGGKWSFQL
sp|Q9P974|IF4E_CANGA        DLLRPVTSFQSVEEFWAIVQNIPEPHELPLKSDYHVFRNDIRPEWEDSANAKGGKWSFQV
                            :* : : :* :*****.: :.:    .**:**** .*:: ::***** * .***

sp|P78954|IF4E1_SCHPO1      KGRG--GNALDEMWLTTVLAAIGETLDPTG--QEVMGVVINMRKGFYRLAVWTKSCNNRE
sp|Q9P975|IF4E_CANAL        NKKSEVNPIINDLWLRGLLAVIGETIEDEE--NEVNGIVLNIRKQAYRVGIWTKDCD-ES
tr|C4R650|C4R650_PICPG      AYR---KVDINELWSRGLLSVIGETIQDDDGETEVNGIVLSIRRAALKIALWTKSKD-EA
tr|I2H7G3|I2H7G3_TETBL      RGNK--HSEINQLWLKAVLSVIGETVDEDE--SEVNGVVLSVRKGNFKIALWTKSES-KE
tr|A7TR08|A7TR08_VANPO      RSK---SSDIDELWLRALLAVIGETIDEDD--SQINGVVLSIRKGTSKFALWTKADDNTE
tr|A7TNF0|A7TNF0_VANPO      RGG---HGDVDELWLRALLSVIGETIDEED--SQINGIVLSVRRGTYKFALWTKSQE-KE
tr|I2GWQ6|I2GWQ6_TETBL      RGK---NNNVEELWMRSLMAVIGETIDEDE--YQINGVVLSVRKGTYKFALWTKSED-RE
sp|P07260|IF4E_YEAST        RGK---GADIDELWLRTLLAVIGETIDEDD--SQINGVVLSIRKGGNKFALWTKSED-KE
sp|Q9P974|IF4E_CANGA        RGR---GAEIDELWLRTLLAVIGETIDEED--SQINGVVLNVRKGGNRFALWTKSCD-KE
                                      ::::*   :::.****::      :: *:*:.:*:   :..:***  .

sp|P78954|IF4E1_SCHPO1      VLMEIGTRFKQVLNLPRSETIEFSAHEDSSKSGSTRAKTRMSV
sp|Q9P975|IF4E_CANAL        KLKTVGERLKKVLQLNDEQKVEFMSHDASNTRGAEPQIVL---
tr|C4R650|C4R650_PICPG      VLRPIGERFKKVLKLTARDHIDFQPHKSGAGSKNVPSFSI---
tr|I2H7G3|I2H7G3_TETBL      PLQKIGARFKKILELEDDEHLDFFPHSIANVKHPQPAISL---
tr|A7TR08|A7TR08_VANPO      ALTNVGTKFKQVLKLGDEDKIEFVPHSSTG-KYVQATITL---
tr|A7TNF0|A7TNF0_VANPO      ALCNIGARFKEVLKLGEEEKIEFFPHSASTAKHAQPAIIL---
tr|I2GWQ6|I2GWQ6_TETBL      PLFEIGKKFKQVLKLNDENQLEFFPHSSANNKHPQPSITL---
sp|P07260|IF4E_YEAST        PLLRIGGKFKQVLKLTDDGHLEFFPHSSANGRHPQPSITL---
sp|Q9P974|IF4E_CANGA        PLSNIGARFKQVLKLADEDTLEFFPHSTANDRHSQPTITL---
                             *  :* ::*::*:*     ::*  *.
```

Fig. 3 (continued)

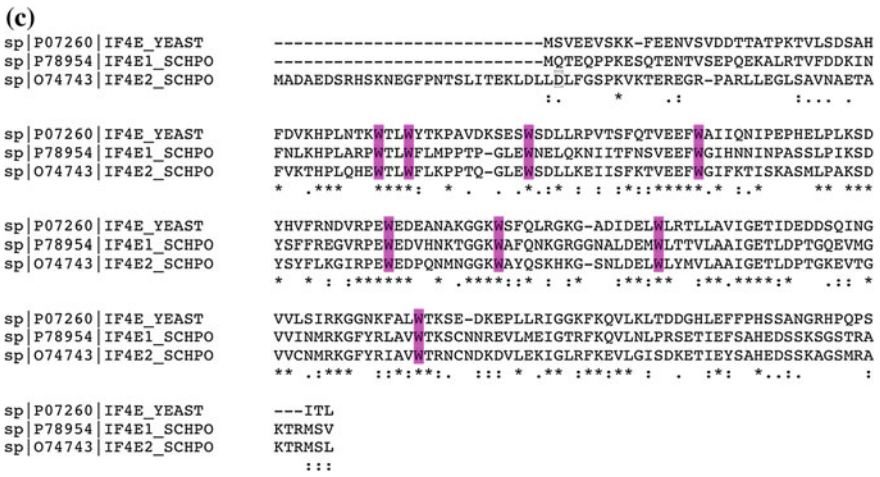


Fig. 3 (continued)

tryptophane Trp5 are positioned in a way that can form a stack with the ^{7}m-G purine residue of the cap structure (see Fig. 2; right panel).

Variations in the arrangement of the conserved tryptophane molecules have allowed for the classification of metazoan eIF4Es as class I, II and III [8, 9]. Class I members all carry Trp1 (corresponding to Trp-43 of murine eIF4E) and Trp3 (corresponding to Trp-56 of murine eIF4E), while class II members possess Trp1 -> Tyr/Phe/Leu and Trp3 -> Tyr/Phe exchanges, and class III (rather seldom) carries Trp1 but Trp3 -> Cys/Tyr. A clear functional assignment of these three classes is still lacking, though at least one class II eIF4E found in *Drosophila* and other insect species and termed eIF4E-8 or 4EHP has been shown to act as an mRNA-specific inhibitor of eIF4E activity [10–12].

3 eIF4E in Lower Eukaryotes

Almost 30 years ago, the cloning of the gene encoding for *S. cerevisiae* eIF4E was shown to be essential for the survival of haploid cells [13], allowing to obtain conditionally lethal yeast eIF4E mutants where the endogenous version of eIF4E was replaced by orthologue versions of the gene from different organisms [11, 14] or temperature-sensitive alleles of yeast eIF4E [15, 16]. This offered interesting insights into the properties of eIF4E [13, 14]. It came as a surprise that eIF4E from such differing organisms as humans, *Drosophila* and plants could cross-complement for the lack of essential eIF4E activity in *S. cerevisiae* [11, 14, 17]. It should be mentioned that not all eIF4Es are capable of substituting for yeast eIF4E, e.g. *Trypanosoma* eIF4Es do not complement in yeast [18]. This might be due to the unusual trimethylated cap structure of *Trypanosoma* mRNAs.

Later discoveries of different variants of eIF4E in different eukaryotic species (such as human/murine, plant or *Drosophila* eIF4E [9, 11, 19]) allowed for the following conclusion: three different eIF4E versions may coexist in one species: namely, type A is required for general translation of capped mRNAs. This form of eIF4E exists in all eukaryotic organisms. All type A eIF4Es belong to class I of the classification mentioned above; type B is required for translation of particular mRNAs and exists beside type A. As an example we will discuss properties of *S. pombe* eIF4E2. Type C is an inhibitor of translation of specific mRNAs, identified in insects and mammalian species as 4EHP [10–12]. Not surprisingly, while type A cross-complements for *S. cerevisiae* eIF4E, type B and type C (which are summarised as eIF4E2s) do not. In principal, all here-studied eIF4Es from around 50 unicellular yeast species belong to type A, but a second eIF4E2 is found in some species (see dendrogram; Fig. 3a), which carries an extended amino and/or carboxy terminus. Whether this difference (maybe in conjunction with other differences along the eIF4E amino acid sequence) plays a functional role is still unclear. eIF4E classification results in unicellular fungi species carrying three possible eIF4E "combinations":

(1) Only one essential eIF4E gene (e.g. *S. cerevisiae*)
(2) Two eIF4E1 copies, termed eIF4E1A and eIF4E1B, which might both be essential for survival (*Vanderwaltozyma polyspora* and *Tetrapisispora blattae*; see Fig. 3b).
(3) One essential eIF4E1 and one eIF4E2 gene copy thought to be required for translating specific mRNAs or to act as an eIF4E inhibitor (e.g. in *C. albicans*).

In the case of *S. pombe*, two eIF4Es have been identified and have been termed eIF4E1 and eIF4E2 [20]. For both, the amino acid sequence is very conserved (see Fig. 3c). While eIF4E1 is essential for the survival of *S. pombe*, eIF4E2 is not and may fulfil specific functions such as translation of specific mRNAs under stress conditions [21].

4 eIF4E1 Is Highly Conserved Among Species

All eIF4E1 variants from different yeasts studied so far are very conserved and cross-complement *S. cerevisiae* eIF4E (our unpublished results). As an example, we present an alignment of selected eIF4E1s from some yeast species. They show a high conservation of amino acid sequences, especially of Trp1 to Trp8 (the only exception is Trp6 -> Leu in *Pichia pastoris* eIF4E1) and surrounding amino acid sequences (see Fig. 3b). *Vanderwaltozyma polyspora* and *Tetrapisispora blattae* carry two copies of eIF4E1, which both complement for yeast eIF4E (our unpublished results).

5 eIF4E2 Differs Considerably from eIF4E1

Not all yeast species from the subphylum Saccharomycotina/"true yeasts" carry two eIF4Es, but for those that do, a second version, eIF4E2, which can be clearly separated from eIF4E1, is observed (see dendrogram; Fig. 3a). eIF4E2s mostly carry extended amino and/or carboxy termini and show several gaps and/or insertions when aligned with eIF4E1. Importantly, Trp3, which is responsible for forming a stack with the capped structure, is lacking or may be replaced by Tyr or Phe. This allows the question of whether eIF4E2 can efficiently bind to capped mRNAs.

Though only tested for *C. albicans* (our unpublished results), the prediction is that eIF4E2 will not cross-complement yeast eIF4E. It remains unclear what the function of eIF4E2 is, either an inhibitor or an activator of translation of specific mRNAs. A further significant feature of this eIF4E2 family is its absolutely conserved amino terminal amino acid sequence (motif MSENLKRAESLFNRIMN; not shown), which is not observed for eIF4E1. This motif might be important in conjunction with other mutations on eIF4E2 (such as the lacking/substituted Trp3) to fulfil specialised functions.

6 eIF4Es from Taphrinomycotina

Yeast species from this subphylum are quite distant relatives from those of the subphylum Saccharomycotina/"true yeasts". Taphrinomycotina yeast species carry two copies of yeast eIF4E, which for *S. pombe* have been termed eIF4E1 and eIF4E2 [20]. While *S. pombe* eIF4E1 is essential and cross-complements for yeast eIF4E, eIF4E2 has been shown to be non-essential for the survival of *S. pombe* cells. *S. pombe* eIF4E2 shows reduced interaction with eIF4G in vivo and is probably required for the translation of specific stress mRNAs. As opposed to *S. pombe* eIF4E1, eIF4E2 does not cross-complement for yeast eIF4E [20–22]. Besides an extended amino terminus, *S. pombe* eIF4E2 is otherwise quite conserved when compared to yeast or *S. pombe* eIF4E1; especially all eight tryptophanes and surrounding sequences are conserved (see Fig. 3c).

We have experimentally addressed the properties of *S. pombe* eIF4E2 by shortening its amino terminus and deleting the first 28 amino acids. The shortened version of *S. pombe* eIF4E2 is capable of complementing yeast eIF4E (our unpublished results), indicating that indeed the amino terminus inhibits cross-complementation. Probably, the amino terminal part of *S. pombe* eIF4E2 plays a mitigating role in its interaction with eIF4G, but otherwise the protein is fully functional. This finding is not completely unexpected as certain variants of *Arabidopsis* eIF4E with extended amino termini have also been shown to impede/attenuate its interaction with plant eIF4G [23].

Quite different than for Schizosaccharomyces species (such as *S. pombe*) is the situation for eIF4E2 from the species *Taphrina* and *Pneumocystis*. A clear divergence between eIF4E1 and eIF4E2 from *Taphrina* and *Pneumocystis* is observed in the dendrogram (see Fig. 3a). Accordingly, all eIF4Es from Schizosaccharomyces species carry the eight conserved tryptophanes, while a distinct feature of *Taphrina* and *Pneumocystis* eIF4E2 is that Trp1/2, Trp3 and Trp5 are modified to Phe or Tyr. We anticipate that while amino terminal-deleted versions of eIF4E2 from different Schizosaccharomyces species cross-complement yeast eIF4E, this will not be the case for *Taphrina* and *Pneumocystis* eIF4E2, which have been previously assigned as class II eIF4Es [9].

7 eIF4E Interactors

Three proteins interacting with eIF4E in yeast have been described: eIF4G, p20 and Eap1. eIF4G is a scaffold protein that attracts several other initiation factors and allows for the circularisation of translated mRNAs. p20 and Eap1 rather than general inhibitors of capped mRNA translation are modulators that affect the translation of particular mRNAs. As they all share the short amino acid motif YxxxxLL/YFMI, which is necessary but not sufficient for interaction with eIF4E, binding to eIF4E is mutually exclusive. As this short amino acid motif is present in many proteins (in yeast probably around 15 % of all proteins), it is not possible to predict which proteins will bind to eIF4E by in silico alignments. Also, it is not known whether further eIF4E-binding proteins that do not carry the canonical element exist in yeast (such examples exist in higher cells). If so, binding of more than one protein to eIF4E simultaneously would be possible.

The canonical eIF4E-binding domain is located at very different positions relative to each other for eIF4G, p20 and Eap1 (see Fig. 4). A particular arrangement of negatively and positively charged motifs around the eIF4E-binding domain has been observed, suggesting that eIF4E-binding proteins, when bound to eIF4E, could electrostatically embrace capped mRNAs, thereby stabilising mutual interactions [24]. Structural data of yeast eIF4E bound to capped mRNAs in complex with eIF4E interactors are still lacking to confirm or discard this interesting hypothesis.

8 eIF4G

As shown in Fig. 4, eIF4G is a scaffold protein that binds several initiation factors such as eIF4E and eIF4A but also carries an interaction domain for poly(A)-binding protein and thereby allows the circularisation of translating mRNAs. Especially important is the HEAT domain of eIF4G (aa 597–871 of yeast eIF4G1 and aa 557–832 of yeast eIF4G2), a crescent-shaped domain consisting of ten alpha helices

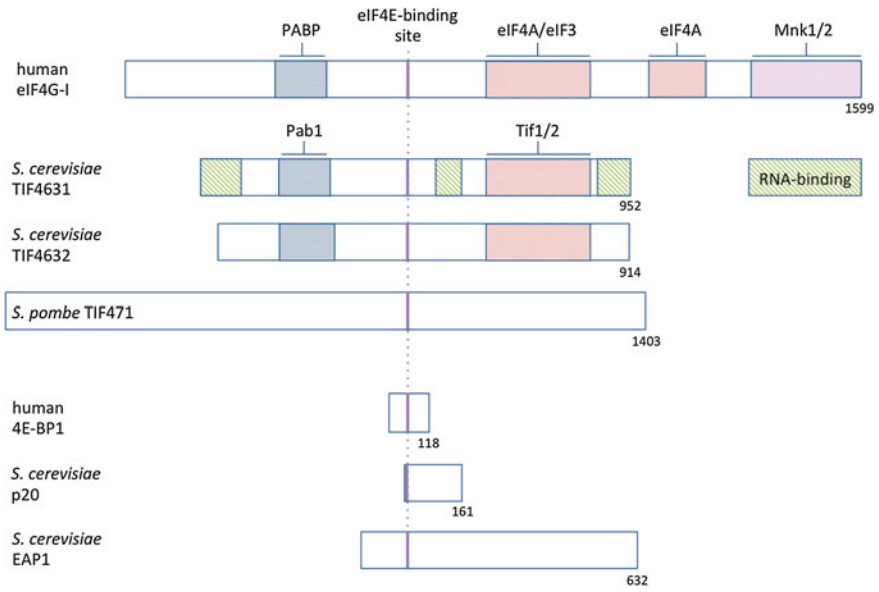


Fig. 4 Cartoon comparing different eIF4Gs (eIF4G-I, TIF4631, TIF4632, TIF471) and eIF4E interactors such as 4E-BP1, p20 and Eap1. The canonical YxxxxLY/F/L/I motif found in all these proteins and its position are indicated. Binding sites for poly(A)-binding protein PABP/Pab1p, for eIF4A/Tif1/2, for eIF3, for Mnk1/2 and for RNA are marked

arranged in five antiparallel HEAT repeats [25], which is found in several proteins. It serves as an interaction platform for other proteins such as eIF4A (see Fig. 4). In higher eukaryotes, eIF4G carries at its carboxy terminus a second binding site for eIF4A and a binding site for the multi-subunit complex eIF3 and the MAP kinase Mnk1/2. In all yeasts, the extended carboxy terminus on eIF4G is lacking and eIF3 probably does not physically interact with eIF4G. Besides eIF4A, the HEAT domain of eIF4G also interacts with eIF1 and eIF5, thereby reinforcing stringent AUG selection [26]. As opposed to mammals, eIF4G does not bind to eIF3, which independently and together with eIF2 can bind to 40S subunits [27]. Besides several binding sites for proteins, yeast eIF4G carries at least three distinct RNA-binding sites [28]. eIF4G also serves as a modulator of translation (and not only as initiation factor) as it binds to Scd6 (a translation repressor whose interaction with eIF4G depends on its RGG motif) and forms a translationally repressed mRNP complex [29].

S. cerevisiae carries two copies encoding for eIF4G, namely TIF4631 for eIF4G1 and TIF4632 encoding for eIF4G2. Deletion of one or both genes is non-essential for yeast, but deletion of both is lethal, indicating that both eIF4G1 and eIF4G2 are functionally redundant [30]. This assumption is reinforced by experiments where the ORFs encoding for both eIF4G1 and eIF4G2 have been interchanged in a yeast strain without causing any phenotype [31]. Studies silencing

eIF4G activity in yeast indicate that it is required for effective translation of all capped mRNAs and that it is involved in recognition by 43S-preinitiation complexes of the capped 5′-UTR but not for the scanning of the 5′UTR region of mRNAs [32].

Saccharomycetales species exist with only one or with two copies for the eIF4G gene. While the amino terminal part of eIF4G can vary considerably among species, the carboxy terminus is quite conserved in sequence and length. Phylogenetically, eIF4Gs from yeast species can be clearly grouped as follows: those with one eIF4E and those with two eIF4E genes form two separate groups (not shown).

9 p20

p20 is a small, acidic, non-essential protein that is the target for multiple phosphorylations at several serine and threonine residues and that is expressed in *S. cerevisiae* at similar levels as eIF4E (around 10,000–20,000 copies per cell; see also, *Saccharomyces* Genome Database, http://www.yeastgenome.org/). p20 only exists in yeast species from the subphylum Saccharomycotina (see Figs. 1b and 6).

Initially, p20 was assumed to be a general inhibitor of translation by competing with eIF4G for binding to eIF4E. Surprisingly, a knock-out haploid Δp20 yeast strain did not show any severe translational defect or phenotype under laboratory conditions [33]. But, in accordance with an inhibitory effect on cap-dependent translation, overexpression of p20 in yeast strains carrying mutations in components of eIF4F leads to a more pronounced phenotype of those mutations [34].

More recent data on the translational profile of a p20 knock-out yeast strain have led to the conclusion that it acts as a modulator of translation affecting preferentially cap-dependent translation of certain mRNAs [35]. A very recent study indicates that p20 might exert its inhibitory function by binding, besides to eIF4E, to specific RNA motifs localised in the 3′UTR region of certain mRNAs, thereby leading to circularisation and stabilisation of translationally repressed mRNAs [36].

It should be emphasised that besides the conserved YxxxxLL/F/M/I/Y 4E-binding motif, p20 has no homology in its primary amino acid sequence to 4E-BPs from higher organisms. However, it fulfils similar functions as 4E-BPs from higher organisms by modulating the activity of specific mRNAs. Thus, it represents an example of evolutionary convergence (no homology of the primary amino acid sequence, but a similar function) [37]. Though not related in its primary amino acid sequence, p20 could adopt a similar three-dimensional structure to 4E-BPs when in complex with eIF4E. In accordance, NMR data were obtained for yeast eIF4E in a complex with mammalian 4E-BP2 and contact points on its surface could be determined [5]. Mammalian 4E-BPs, when expressed in yeast, only exerted an inhibitory effect on translation when yeast eIF4E was replaced by human eIF4E as the only source of cap-binding protein [14, 38]. This led to the conclusion that the affinity of mammalian 4E-BPs for yeast eIF4E is much smaller than for mammalian eIF4E [38]. To our knowledge, structural features determining the

different strengths of interaction between yeast and higher eIF4Es with 4E-BPs have not yet been determined. Unfortunately, no 3D structure of any p20, alone or in a complex with eIF4E, has been obtained so far.

It is difficult to associate a clear phenotype to the loss of p20 in *S. cerevisiae*, as we only observe a mild temperature-sensitive phenotype in our laboratory strains (reduced growth at 37 °C), which makes its functional study more difficult. As mentioned, only yeast species from the subphylum Saccharomycotina/"true yeasts" but not from other subphyla carry p20. So, *S. pombe* (and other Taphrinomycotina species), which are genetically very distant from *S. cerevisiae*, do not have p20. This allows the conclusion that p20's function modulating the translation of particular mRNAs is very specific for certain yeast species or that its function can be accomplished by other proteins—maybe by eIF4E2—in other yeast species.

The dendrogram of p20-carrying yeast species shows a division into two branches (not shown). Especially conserved among all p20s is the amino terminus, which carries the eIF4E-binding motif and the carboxy terminus consisting of the hydrophobic motif FNAFEAL followed by several acidic residues. Some differences between both branches become evident when comparing p20 sequences (Fig. 5a). All group 1 yeast species, which share a very conserved p20 form, carry one eIF4E1 (or in some cases two eIF4E1 genes), while group 2 yeast species, which all have besides eIF4E1 also eIF4E2, carry a somewhat different p20 with an amino acid motif insertion in the middle part of the protein. Does this correlation just reflect the phylogenetic distance and/or does it have a functional significance? Preliminary experiments from our laboratory indicate that both variants of p20 are functional in yeast, as *C. albicans* p20 can complement for the lack of endogenous yeast p20 in our knock-out yeast strains with a mild temperature-sensitive phenotype (poor growth at 37 °C; unpublished results).

10 Eap1

Eap1 is a further interactor of yeast eIF4E [39] that competes with eIF4G in translation initiation. Eap1 is a 70-kD protein with several functions ascribed. It is assumed to interact with or form part of the spindle pole body (SPB; the equivalent in yeast to the mammalian centrosome). When insertion of the newly formed SPB into the nuclear envelope fails, Eap1 specifically inhibits initiation of the translation of POM34 mRNA, which encodes for an integral membrane protein of the nuclear pore [40]. A knock-out of EAP1 is non-lethal but causes an increased rate of aneuploidy and a temperature-sensitive phenotype. Its function in chromosome separation during mitosis seems to be separate from its second property in competing with eIF4G for eIF4E binding [41]. Eap1 also plays a role in a regulatory circuit where translation initiation is down-regulated when lipid synthesis is repressed [42, 43]. Additionally, Eap1 has been described to accelerate degradation by increasing decapping of certain mRNAs through its interaction with eIF4E [44, 45].

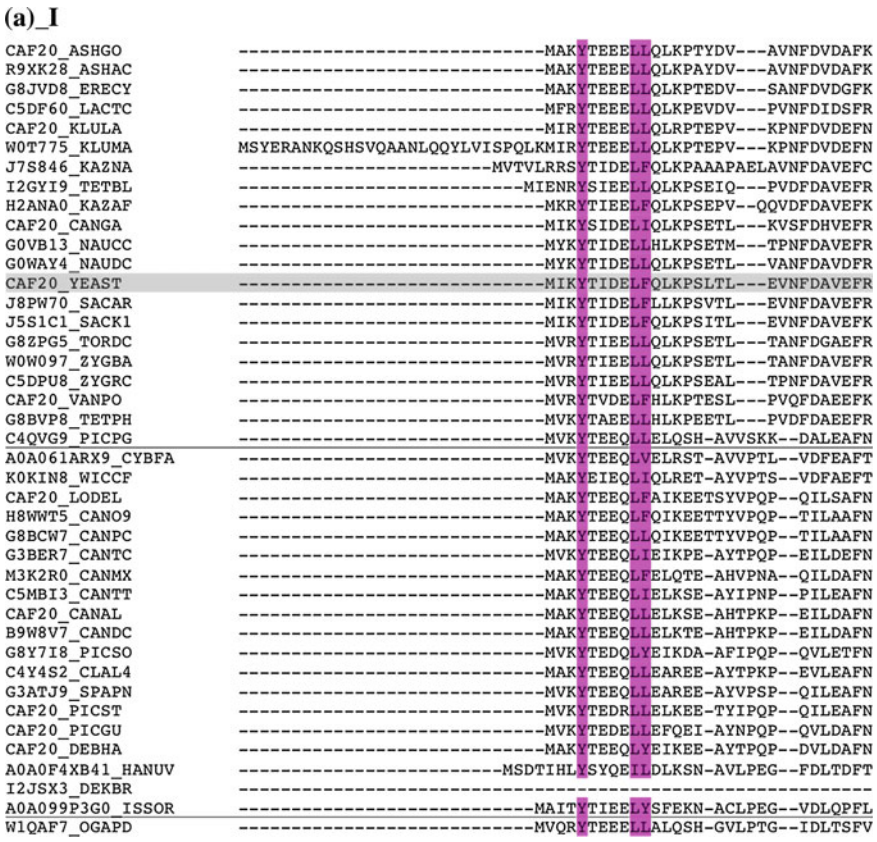


Fig. 5 **A** Sequence alignment of Saccharomycotina species carrying p20. p20s from 41 species are divided in two groups (*separating lines*): species with only eIF4E1 (e.g. *S. cerevisiae*, marked in *light grey*) and species with an additional eIF4E2 (identifiers are found in Fig. 1b or 6). *Ogataea parapolymorpha* (OGAPD) is the only species that does not cluster accordingly in this alignment as it only has one eIF4E1 and no eIF4E2. The canonical amino terminal eIF4E-binding motif YxxxxLL/I/Y/F is labelled (*velvet/dark grey*). **b** Sequence alignment of yeast Eap1 and p20. The conserved canonical eIF4E-binding motif YxxxxLY/F is labelled (*velvet/grey*)

Eap1 (as p20) only exists in Saccharomycotina. Recent data with a knock-out or a TAP-tagged version of Eap1 suggest that several hundred mRNAs are interacting and/or affected by Eap1 action. The interactome of Eap1 is mostly different from that of p20 [35, 36]. Besides the eIF4E-binding domain, Eap1 only shares limited homology with p20 (see Fig. 5b).

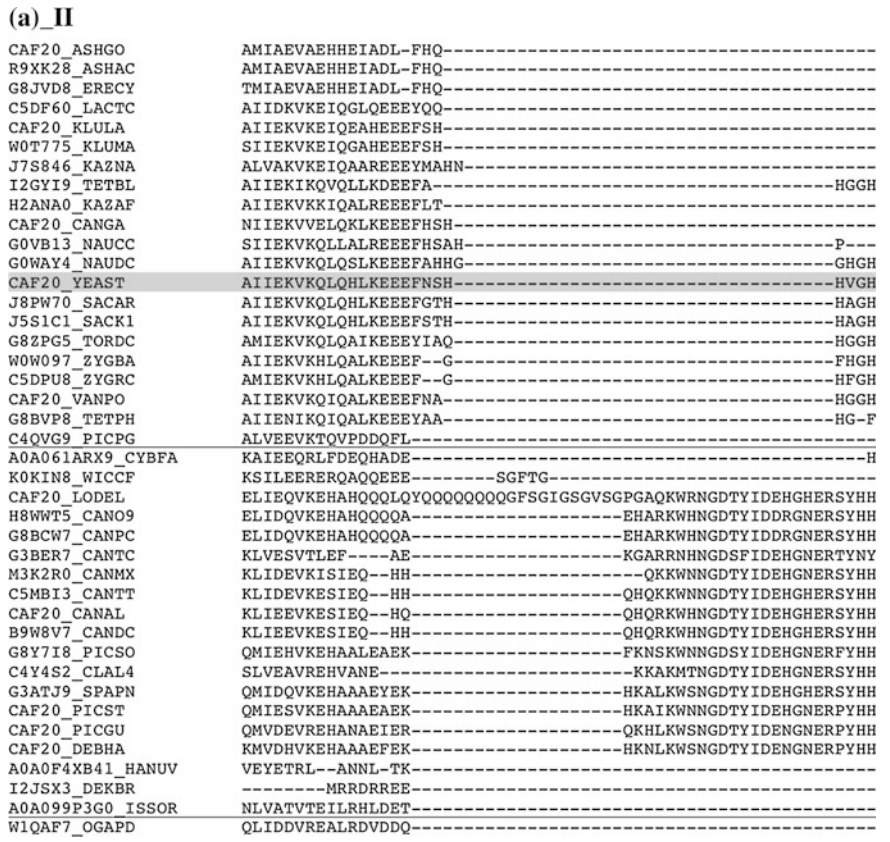


Fig. 5 (continued)

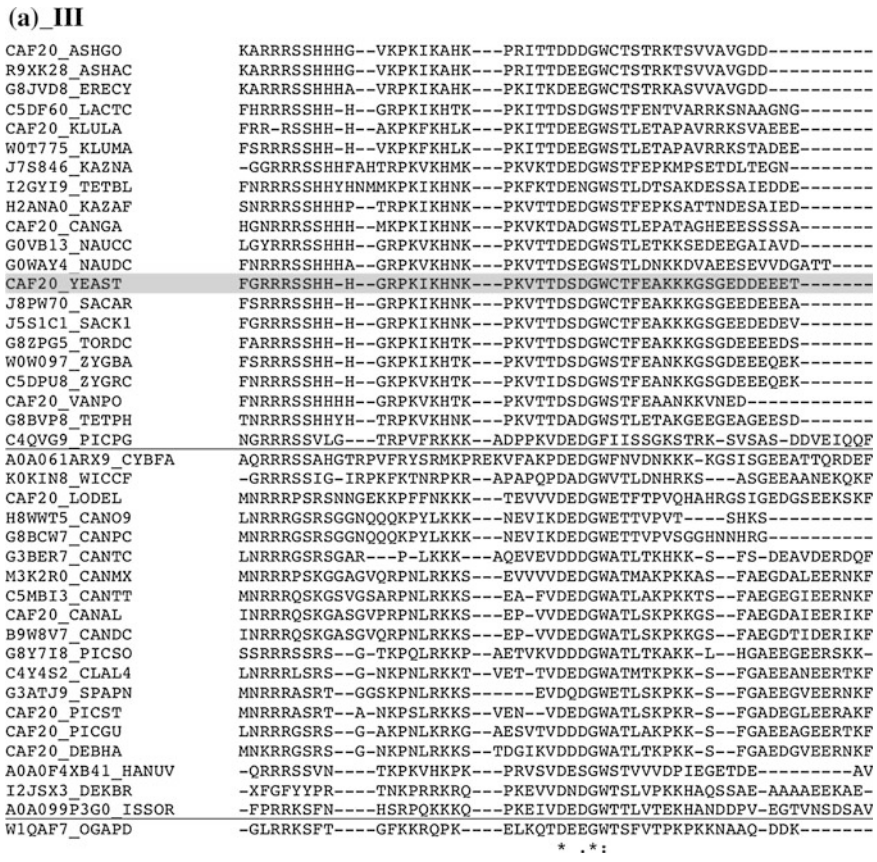


Fig. 5 (continued)

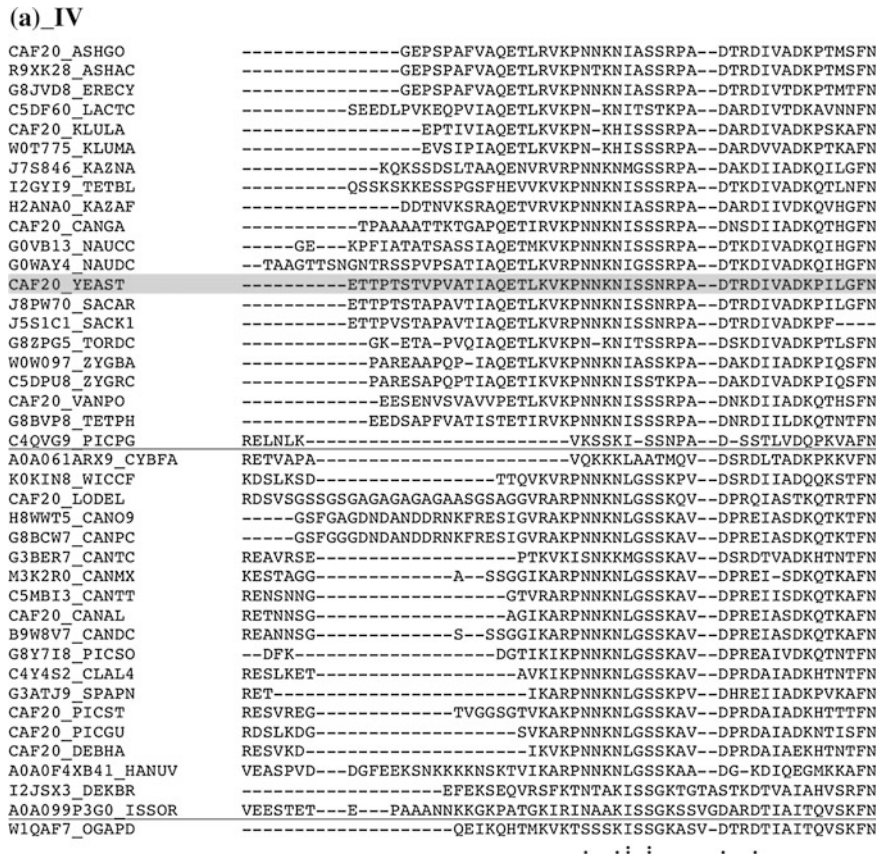


Fig. 5 (continued)

(a)_V

```
CAF20_ASHGO             AFAALESDDEDQF-----
R9XK28_ASHAC            AFAALESDEEDEF-----
G8JVD8_ERECY            AFAALESDEEDQV-----
C5DF60_LACTC            AFAALESDEEDE------
CAF20_KLULA             AFAALESDEEEEQE----
W0T775_KLUMA            AFAALGSDEEEEEDDDEE
J7S846_KAZNA            AFAALESESESDNE----
I2GYI9_TETBL            AFAALESESDDE------
H2ANA0_KAZAF            AFAALESDEEEEE-----
CAF20_CANGA             AFAALEDEEDE-------
G0VB13_NAUCC            TFAALESDDDE-------
G0WAY4_NAUDC            AFAALESDEDDE------
CAF20_YEAST             AFAALESEDEDDEA----
J8PW70_SACAR            AFAALESEDEDEEA----
J5S1C1_SACK1            ------------------
G8ZPG5_TORDC            AFAALESEDEDEDEDEE-
W0W097_ZYGBA            AFAALESEEEDDE-----
C5DPU8_ZYGRC            AFAALESEEEDDE-----
CAF20_VANPO             AFAALESDEEEET-----
G8BVP8_TETPH            AFAALESDEDE-------
C4QVG9_PICPG            AFDALMDSDGDSE-----
A0A061ARX9_CYBFA        AFAALDSEEDESDEDED-
K0KIN8_WICCF            AFAALGDESDEESEEDDE
CAF20_LODEL             AFEALEGNDDEDDDE---
H8WWT5_CANO9            AFEALEGDDDDDDE----
G8BCW7_CANPC            AFEALEGDDDDDNE----
G3BER7_CANTC            AFEALDD-----------
M3K2R0_CANMX            AFAALDDDDEDEDEDDE-
C5MBI3_CANTT            AFAALDDDEEDDDE----
CAF20_CANAL             AFAALGDEDDDDEDDE--
B9W8V7_CANDC            AFAALGDEDDDE------
G8Y7I8_PICSO            AFEALGDEDDDEDEDEE-
C4Y4S2_CLAL4            AFEALGGDDDDDE-----
G3ATJ9_SPAPN            AFAALEDDDDDEE-----
CAF20_PICST             AFEALGDGDDDDDDE---
CAF20_PICGU             AFEALGDDSDDDDE----
CAF20_DEBHA             AFEALGDGDDSDDE----
A0A0F4XB41_HANUV        AFDALALEDDE-------
I2JSX3_DEKBR            AFDALNEEDDSDQE----
A0A099P3G0_ISSOR        AFDALALDDDE-------
W1QAF7_OGAPD            AFDALNAEDEAIESDEE-
```

Fig. 5 (continued)

Fig. 5 (continued)

11 Conclusions

Despite their simplicity as unicellular organisms, yeast species show different eIF4E combinations, which include species such as *S. pombe* that contain more than one "general" eIF4E required for capped mRNA translation. eIF4E2s in some yeast species were probably "invented" to allow for translation of specific mRNAs under conditions when general translation is down-regulated because of a variety of possible stress conditions. "Simple" eIF4E2s such as that of *S. pombe* show an attenuated eIF4G interaction, which was probably introduced during evolution by the addition of an extended amino terminus. "Complex" eIF4E2s such as that of *C. albicans* probably diverged from an original eIF4E1 to fulfil other functions such as selection or repression of specific mRNAs.

The repertoire for potential regulation of translation was augmented with the acquisition of eIF4E interactors such as p20 and Eap1, which only exist in Saccharomycotina/true yeasts (for a summary, see Fig. 6) and select/repress for the translation of a subset of mRNAs by mechanisms that are not well understood yet and probably include the interaction of these proteins with other proteins bound to specific sequences at the 3′UTR of certain mRNAs [36]. It is surprising that knock-outs of p20 in *S. cerevisiae* do not lead to the observation of strong phenotypes under laboratory conditions. Possible reasons for this are: (1) more complex multivariate conditions are encountered by yeast cells in free nature (simultaneous changes of temperature, humidity, food supply, competitors) and are not simple to mimic in the laboratory; (2) *S. cerevisiae* (and other yeast species) have developed more than one parallel pathway to encounter stress conditions that allow them to swap from one to the other depending on the availability of factors involved in each pathway.

Species		4E-1	4E-2	p20	EAP1
Schizosaccharomyces cryophilus	SCHCR	x	x		
Schizosaccharomyces japonicus	SCHJY	x	x		
Schizosaccharomyces octosporus	SCHOY	x	x		
Schizosaccharomyces pombe	SCHPO	x	x		
Pneumocystis murina	PNEMU	x	x		
Taphrina deformans	TAPDE	x	x		
Blastobotrys adeninivorans	BLAAD	x	x		x
Yarrowia lipolytica	YARLI	x	x		x
Issatchenkia orientalis	ISSOR	x	x	x	
Dekkera bruxellensis	DEKBR	x		x	
Ogataea parapolymorpha	OGAPD	x		x	
Ashbya gossypii	ASHGO	x		x	x
Ashbya aceri	ASHAC	x		x	x
Eremothecium cymbalariae	ERECY	x		x	x
Lachancea thermotolerans	LACTC	x		x	x
Kluyveromyces lactis	KLULA	x		x	x
Kluyveromyces marxianus	KLUMA	x		x	x
Kazachstania africana	KAZAF	x		x	x
Kazachstania naganishii	KAZNA	x		x	x
Tetrapisispora phaffii	TETPH	x		x	x
Tetrapisispora blattae	TETBL	2x		x	x
Candida glabrata	CANGA	x		x	x
Naumovozyma dairenensis	NAUDC	x		x	x
Naumovozyma castellii	NAUCC	x		x	x
Torulaspora delbrueckii	TORDC	x		x	x
Saccharomyces cerevisiae	YEAST	x		x	x
Saccharomyces arboricola	SACAR	x		x	x
Saccharomyces kudriavzevii	SACK1	x		x	x
Zygosaccharomyces bailii	ZYGBA	x		x	x
Zygosaccharomyces rouxii	ZYGRC	x		x	x
Vanderwaltozyma polyspora	VANPO	2x		x	x
Hanseniaspora uvarum	HANUV	x	x	x	
Pichia pastoris	PICPG	x	x	x	x
Cyberlindnera fabianii	CYBFA	x	x	x	
Wickerhamomyces ciferrii	WICCF	x		x	x
Candida albicans	CANAL	x	x	x	x
Candida tropicales	CANTT		x	x	x
Candida maltosa	CANMX	x	x	x	x
Candida dubliniensis	CANDC	x	x	x	x
Candida orthopsilosis	CANO9	x	x	x	x
Candida parapsilosis	CANPC	x	x	x	x
Lodderomyces elongisporus	LODEL	x	x	x	x
Candida tenuis	CANTC	x	x	x	
Pichia sorbitophila	PICSO	x	x	x	x
Spathaspora passalidarum	SPAPN	x	x	x	x
Scheffersomyces stipitis	PICST	x	x	x	
Meyerozyma guilliermondii	PICGU	x	x	x	x
Debaryomyces hansenii	DEBHA	x	x	x	x
Clavispora lusitaniae	CLAL4	x	x	x	x

Fig. 6 Summary of yeast species carrying eIF4Es, p20 and Eap1. The table summarises the presence of one or two copies (2x) of eIF4E1, of eIF4E2, of p20 and of Eap1 for 6 species from the subphylum Taphrinomycotina (*dark grey*) and 43 species from the subphylum Saccharomycotina (*light grey*). Note that for one Saccharomycotina species no eIF4E1 (CANTT) and for two further Saccharomycotina species no p20 (BLAAD and YARLI) have been reported

A further potential for regulation arises from eIF4G itself. Though most yeast species carry two copies of eIF4G, which show considerable variations (especially in their amino termini), there is so far no evidence for specialised functions of eIF4Gs in yeast. So, the conclusion is that the amino acid divergence of eIF4Gs does not affect essential parts of those proteins (especially localised in their HEAT domain) and that this divergence just reflects the accumulation of mutations ever since the original genome of a yeast ancestor species was duplicated.

Acknowledgments This work was supported by Swiss National Foundation grant 31003A_146722/1.

References

1. Margulis L, Chapman MJ. Kingdoms and domains. An illustrated guide to the phyla of life on Earth. Amsterdam: Academic Press; 2010.
2. Ross D, Saxena M, Altmann M. eIF4E is an important determinant of adhesion and pseudohyphal growth of the yeast S. cerevisiae. PLoS One. 2012;7:e50773. doi:10.1371/journal.pone.0050773.
3. Cullen PJ, Sprague GF Jr. The regulation of filamentous growth in yeast. Genetics. 2012;190:23–49. doi:10.1534/genetics.111.127456.
4. Youk H, Lim WA. Secreting and sensing the same molecule allows cells to achieve versatile social behaviors. Science. 2014;343:1242782. doi:10.1126/science.1242782.
5. Matsuo H, Li H, McGuire AM, Fletcher CM, Gingras AC, Sonenberg N, Wagner G. Structure of translation factor eIF4E bound to m7GDP and interaction with 4E-binding protein. Nat Struct Biol. 1997;4:717–24.
6. Marcotrigiano J, Gingras AC, Sonenberg N, Burley SK. Cocrystal structure of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP. Cell. 1997;89:951–61.
7. Gross JD, Moerke NJ, von der Haar T, Lugovskoy AA, Sachs AB, McCarthy JE, Wagner G. Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E. Cell. 2003;115:739–50.
8. Jones GD, Williams EP, Place AR, Jagus R, Bachvaroff TR. The alveolate translation initiation factor 4E family reveals a custom toolkit for translational control in core dinoflagellates. BMC Evol Biol. 2015;15:14. doi:10.1186/s12862-015-0301-9.
9. Joshi B, Lee K, Maeder DL, Jagus R. Phylogenetic analysis of eIF4E-family members. BMC Evol Biol. 2005;5:48. doi:10.1186/1471-2148-5-48.
10. Tettweiler G, Kowanda M, Lasko P, Sonenberg N, Hernandez G. The distribution of eIF4E-family members across Insecta. Comp Funct Genomics. 2012;2012:960420. doi:10.1155/2012/960420.
11. Hernandez G, Altmann M, Sierra JM, Urlaub H, Diez del Corral R, Schwartz P, Rivera-Pomar R. Functional analysis of seven genes encoding eight translation initiation factor 4E (eIF4E) isoforms in Drosophila. Mech Dev. 2005;122:529–43. doi:10.1016/j.mod.2004.11.011.
12. Cho PF, Poulin F, Cho-Park YA, Cho-Park IB, Chicoine JD, Lasko P, Sonenberg N. A new paradigm for translational control: inhibition via 5'-3' mRNA tethering by Bicoid and the eIF4E cognate 4EHP. Cell. 2005;121:411–23. doi:10.1016/j.cell.2005.02.024.
13. Altmann M, Handschin C, Trachsel H. mRNA cap-binding protein: cloning of the gene encoding protein synthesis initiation factor eIF-4E from Saccharomyces cerevisiae. Mol Cell Biol. 1987;7:998–1003.

14. Altmann M, Muller PP, Pelletier J, Sonenberg N, Trachsel H. A mammalian translation initiation factor can substitute for its yeast homologue in vivo. J Biol Chem. 1989;264:12145–7.
15. Altmann M, Sonenberg N, Trachsel H. Translation in Saccharomyces cerevisiae: initiation factor 4E-dependent cell-free system. Mol Cell Biol. 1989;9:4467–72.
16. Altmann M, Trachsel H. Altered mRNA cap recognition activity of initiation factor 4E in the yeast cell cycle division mutant cdc33. Nucleic Acids Res. 1989;17:5923–31.
17. Altmann M, Linder P. Power of yeast for analysis of eukaryotic translation initiation. J Biol Chem. 2010;285:31907–12. doi:10.1074/jbc.R110.144196.
18. Yoffe Y, Zuberek J, Lerer A, Lewdorowicz M, Stepinski J, Altmann M, Darzynkiewicz E, Shapira M. Binding specificities and potential roles of isoforms of eukaryotic initiation factor 4E in Leishmania. Eukaryot Cell. 2006;5:1969–79. doi:10.1128/EC.00230-06.
19. Hernandez G, Vazquez-Pianzola P. Functional diversity of the eukaryotic translation initiation factors belonging to eIF4 families. Mech Dev. 2005;122:865–76. doi:10.1016/j.mod.2005.04.002.
20. Ptushkina M, Berthelot K, von der Haar T, Geffers L, Warwicker J, McCarthy JE. A second eIF4E protein in Schizosaccharomyces pombe has distinct eIF4G-binding properties. Nucleic Acids Res. 2001;29:4561–9.
21. Ptushkina M, Malys N, McCarthy JE. eIF4E isoform 2 in Schizosaccharomyces pombe is a novel stress-response factor. EMBO Rep. 2004;5:311–6. doi:10.1038/sj.embor.7400088.
22. Ptushkina M, Fierro-Monti I, van den Heuvel J, Vasilescu S, Birkenhager R, Mita K, McCarthy JE. Schizosaccharomyces pombe has a novel eukaryotic initiation factor 4F complex containing a cap-binding protein with the human eIF4E C-terminal motif KSGST. J Biol Chem. 1996;271:32818–24.
23. Patrick RM, Mayberry LK, Choy G, Woodard LE, Liu JS, White A, Mullen RA, Tanavin TM, Latz CA, Browning KS. Two Arabidopsis loci encode novel eukaryotic initiation factor 4E isoforms that are functionally distinct from the conserved plant eukaryotic initiation factor 4E. Plant Physiol. 2014;164:1820–30. doi:10.1104/pp.113.227785.
24. Cawley A, Warwicker J. eIF4E-binding protein regulation of mRNAs with differential 5'-UTR secondary structure: a polyelectrostatic model for a component of protein-mRNA interactions. Nucleic Acids Res. 2012;40:7666–75. doi:10.1093/nar/gks511.
25. Marcotrigiano J, Lomakin IB, Sonenberg N, Pestova TV, Hellen CU, Burley SK. A conserved HEAT domain within eIF4G directs assembly of the translation initiation machinery. Mol Cell. 2001;7:193–203.
26. He H, von der Haar T, Singh CR, Ii M, Li B, Hinnebusch AG, McCarthy JE, Asano K. The yeast eukaryotic initiation factor 4G (eIF4G) HEAT domain interacts with eIF1 and eIF5 and is involved in stringent AUG selection. Mol Cell Biol. 2003;23:5431–45.
27. Jivotovskaya AV, Valasek L, Hinnebusch AG, Nielsen KH. Eukaryotic translation initiation factor 3 (eIF3) and eIF2 can promote mRNA binding to 40S subunits independently of eIF4G in yeast. Mol Cell Biol. 2006;26:1355–72. doi:10.1128/mcb.26.4.1355-1372.2006.
28. Berset C, Zurbriggen A, Djafarzadeh S, Altmann M, Trachsel H. RNA-binding activity of translation initiation factor eIF4G1 from Saccharomyces cerevisiae. RNA. 2003;9:871–80.
29. Rajyaguru P, She M, Parker R. Scd6 targets eIF4G to repress translation: RGG motif proteins as a class of eIF4G-binding proteins. Mol Cell. 2012;45:244–54. doi:10.1016/j.molcel.2011.11.026.
30. Goyer C, Altmann M, Lee HS, Blanc A, Deshmukh M, Woolford JL Jr, Trachsel H, Sonenberg N. TIF4631 and TIF4632: two yeast genes encoding the high-molecular-weight subunits of the cap-binding protein complex (eukaryotic initiation factor 4F) contain an RNA recognition motif-like sequence and carry out an essential function. Mol Cell Biol. 1993;13:4860–74.
31. Clarkson BK, Gilbert WV, Doudna JA. Functional overlap between eIF4G isoforms in Saccharomyces cerevisiae. PLoS ONE. 2010;5:e9114. doi:10.1371/journal.pone.0009114.

32. Park EH, Zhang F, Warringer J, Sunnerhagen P, Hinnebusch AG. Depletion of eIF4G from yeast cells narrows the range of translational efficiencies genome-wide. BMC Genom. 2011;12:68. doi:10.1186/1471-2164-12-68.
33. Altmann M, Schmitz N, Berset C, Trachsel H. A novel inhibitor of cap-dependent translation initiation in yeast: p20 competes with eIF4G for binding to eIF4E. EMBO J. 1997;16:1114–21. doi:10.1093/emboj/16.5.1114.
34. de la Cruz J, Iost I, Kressler D, Linder P. The p20 and Ded1 proteins have antagonistic roles in eIF4E-dependent translation in Saccharomyces cerevisiae. Proc Natl Acad Sci USA. 1997;94:5201–6.
35. Cridge AG, Castelli LM, Smirnova JB, Selley JN, Rowe W, Hubbard SJ, McCarthy JE, Ashe MP, Grant CM, Pavitt GD. Identifying eIF4E-binding protein translationally-controlled transcripts reveals links to mRNAs bound by specific PUF proteins. Nucleic Acids Res. 2010;38:8039–50. doi:10.1093/nar/gkq686.
36. Castelli LM, Talavera D, Kershaw CJ, Mohammad-Qureshi SS, Costello JL, Rowe W, Sims PF, Grant CM, Hubbard SJ, Ashe MP, Pavitt GD. The 4E-BP Caf20p mediates both eIF4E-dependent and independent repression of translation. PLoS Genet. 2015;11:e1005233. doi:10.1371/journal.pgen.1005233.
37. Hernandez G, Altmann M, Lasko P. Origins and evolution of the mechanisms regulating translation initiation in eukaryotes. Trends Biochem Sci. 2010;35:63–73. doi:10.1016/j.tibs.2009.10.009.
38. Hughes JM, Ptushkina M, Karim MM, Koloteva N, von der Haar T, McCarthy JE. Translational repression by human 4E-BP1 in yeast specifically requires human eIF4E as target. J Biol Chem. 1999;274:3261–4.
39. Cosentino GP, Schmelzle T, Haghighat A, Helliwell SB, Hall MN, Sonenberg N. Eap1p, a novel eukaryotic translation initiation factor 4E-associated protein in Saccharomyces cerevisiae. Mol Cell Biol. 2000;20:4604–13.
40. Sezen B, Seedorf M, Schiebel E. The SESA network links duplication of the yeast centrosome with the protein translation machinery. Genes Dev. 2009;23:1559–70. doi:10.1101/gad.524209.
41. Chial HJ, Stemm-Wolf AJ, McBratney S, Winey M. Yeast Eap1p, an eIF4E-associated protein, has a separate function involving genetic stability. Curr Biol. 2000;10:1519–22.
42. Deloche O, de la Cruz J, Kressler D, Doere M, Linder P. A membrane transport defect leads to a rapid attenuation of translation initiation in Saccharomyces cerevisiae. Mol Cell. 2004;13:357–66.
43. Meier KD, Deloche O, Kajiwara K, Funato K, Riezman H. Sphingoid base is required for translation initiation during heat stress in Saccharomyces cerevisiae. Mol Biol Cell. 2006;17:1164–75. doi:10.1091/mbc.E05-11-1039.
44. Blewett NH, Goldstrohm AC. A eukaryotic translation initiation factor 4E-binding protein promotes mRNA decapping and is required for PUF repression. Mol Cell Biol. 2012;32:4181–94. doi:10.1128/MCB.00483-12.
45. Rendl LM, Bieman MA, Vari HK, Smibert CA. The eIF4E-binding protein Eap1p functions in Vts1p-mediated transcript decay. PLoS ONE. 2012;7:e47121. doi:10.1371/journal.pone.0047121.

Expansion of eIF4E and 4E-BP Family Members in Deuterostomes

Kathleen M. Gillespie, Tsvetan R. Bachvaroff and Rosemary Jagus

1 Introduction

1.1 Deuterostome Phylogeny

Recent progress in the genomic analysis of fish has provided a broad range of species for analysis [1–11]. Analysis of eIF4E sequences from these genomes, as well as those of select tetrapods, echinoderm (*Strongylocentrotus purpuratus*), tunicate (*Ciona intestinalis*), and cephalocordate (*Branchiostoma lanceolatum*), has allowed a glimpse of the origins and evolution of the eIF4E and 4E-BP families in vertebrates.

The emergence of vertebrates has been accompanied by duplication within the eIF4E family. The Cambrian explosion, also called the Cambrian radiation, was the relatively short evolutionary event, beginning around 550 million years ago (Ma) in the Cambrian Period, during which most major animal phyla appeared, as indicated by the fossil record [12]. Lasting for about the next 20–25 million years, it resulted in the divergence of most metazoan phyla [13, 14]. The earliest generally accepted deuterostome fossils, those of echinoderms, appeared in the Late Atdabanian (Cambrian, third stage) about 520 Ma [15]. 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 or lancelets; and the vertebrates,

K.M. Gillespie · T.R. Bachvaroff · R. Jagus (✉)
Institute of Marine and Environmental Technology, University of Maryland Center for Environmental Science, 701 E. Pratt Street, Baltimore, MD 21202, USA
e-mail: jagus@umces.edu

K.M. Gillespie
e-mail: kgillesp@umces.edu

G. Hernández and R. Jagus (eds.), *Evolution of the Protein Synthesis Machinery and Its Regulation*, DOI 10.1007/978-3-319-39468-8_8

including fish and tetrapods. Figure 1 illustrates deuterostome phylogeny [16]. Phylogenetic analysis based on assembled sequences of more than 200 nuclear-encoded proteins supports the pairing of echinoderms with hemichordates, corroborating morphological interpretations of larval similarities between these two groups [17].

Gene duplication is considered to be a major force of evolution because new copies may acquire new functions by mutation (known as neofunctionalization) [18]. It is generally accepted that two rounds of whole-genome duplication occurred during the evolution of vertebrates from their chordate ancestors probably before the divergence of gnathostomes between 550 (VGD1) and 500 (VGD2) Ma [19, 20]. The last common ancestor of all gnathostomes (jawed vertebrates) was the common ancestor of the Actinopterygii (ray-finned fish) and Sarcopterygii (lobe-finned fishes and tetrapods), which probably lived during the Silurian period, approximately 420 Ma [21–23]. Within vertebrates, lampreys and hagfish are closely related (Cyclostomata) and are the nearest relatives of jawed vertebrates. Molecular analyses suggest that cartilaginous fish are the most basal gnathostomes and that the coelacanth and lungfish (lobe-finned fish) are the closest living relatives of tetrapods. It has been said that tetrapods are in fact specialized lobefin fish that have adapted to life on land [8]. Within the ray-finned fish, the dominant extant group is the Teleostii comprised of over 20,000 species. The teleost lineage split from basal ray-finned fishes and started to diverge after another whole genome duplication event, referred to as the teleost-specific genome duplication (TGD) that took place 320–350 Ma [9–11]. The teleosts began a major evolutionary radiation in the Triassic, about 200 Ma, attributed to the TGD, and have since undergone massive diversification in morphology, physiology, and habitat. Their genomes did not remain static, and they are still evolving. Because of the genome duplications, determining the orthology of genes between teleosts and tetrapods can be challenging.

Genome duplication is always followed by gene loss, genomic rearrangement, and rediploidization [24]. The expansion of teleosts was accompanied by a phase of major genomic rearrangement [25, 26], an increase in the molecular evolutionary rate [27], and rediploidization of the tetraploid genome after the TGD. Rediploidization involves massive loss of gene copies and other genomic elements from a duplicated state back to a singleton state. Sequence divergence between TGD paralogs occurs in an asymmetric manner, usually with one of the two TGD duplicates evolving faster than the other [28, 29]. Divergence in paralog sequence and function often occur in a lineage-specific way so that not all lineages lose the same copy [30–32]. This means that annotation of genomes using BLAST-based methods is problematic because of lineage-specific reshuffling of teleost genomes following the TGD.

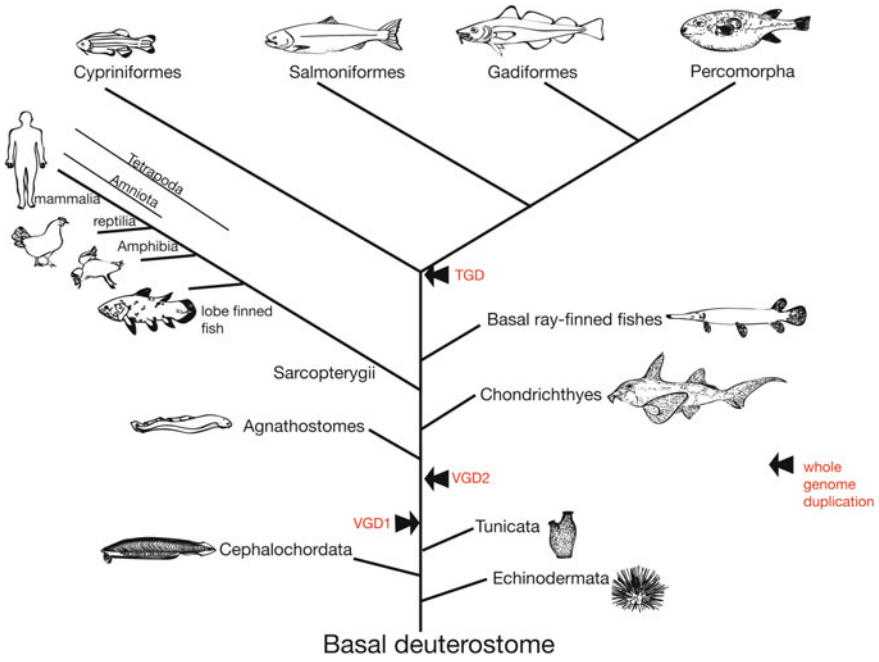
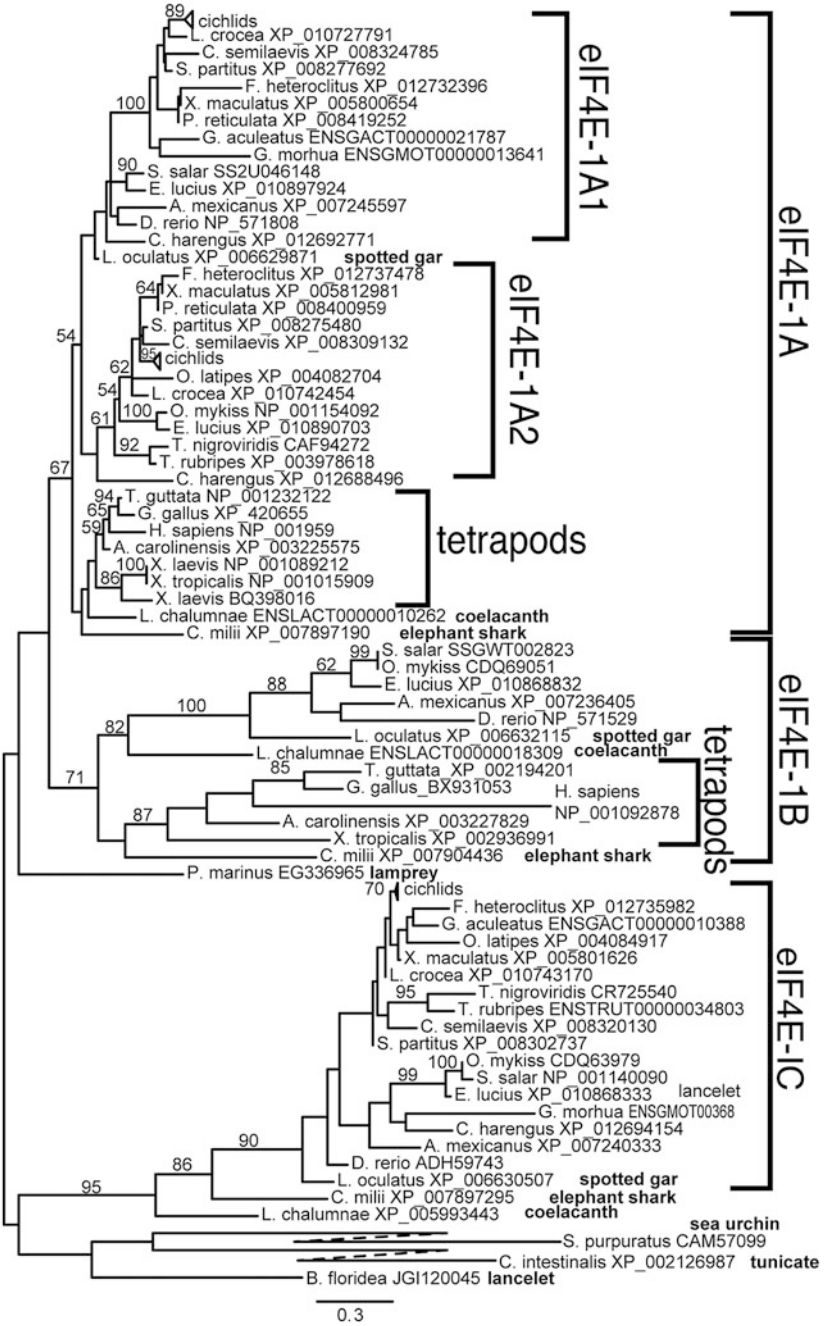


Fig. 1 Phylogeny of deuterostomes

2 The Radiation of the Vertebrates Is Reflected in the Expansion of Their eIF4E Family Members

Teleosts are characterized by many derived characteristics that are absent in basal ray-finned fish such as gar, sturgeon, and paddlefish. Within the teleosts, the Ostariophysi (such as zebrafish) retain many primitive characteristics and occupy a relatively basal position [33]. 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 emerging teleosts such as the percomorphs, medaka, stickleback, tilapia, fugu, and tongue sole [33–35]. However, with an evolutionary separation of less than 150 million years, the zebrafish is still closer to the more recently emerged fish species than any mammalian model organism such as the mouse, whose common ancestor with the teleosts lived over 400 million years ago [34]. However, in terms of connecting fish genes to those of tetrapods, the genome sequence of spotted gar, *Lepisosteus oculatus,* the lineage of which diverged from teleosts before the TGD, is providing a valuable resource [11].

All deuterostomes have one representative from each of the three classes; eIF4E-1, eIF4E-2, and eIF4E-3. In contrast, mammals have an additional eIF4E-1 cognate, eIF4E-1B, that functions to downregulate the translation of


cichlids
L. crocea XP_010727791
C. semilaevis XP_008324785
S. partitus XP_008277692
F. heteroclitus XP_012732396
X. maculatus XP_005800654
P. reticulata XP_008419252
G. aculeatus ENSGACT00000021787
G. morhua ENSGMOT00000013641
S. salar SS2U046148
E. lucius XP_010897924
A. mexicanus XP_007245597
D. rerio NP_571808
C. harengus XP_012692771
L. oculatus XP_006629871 spotted gar
eIF4E-1A1
F. heteroclitus XP_012737478
X. maculatus XP_005812981
P. reticulata XP_008400959
S. partitus XP_008275480
C. semilaevis XP_008309132
cichlids
O. latipes XP_004082704
L. crocea XP_010742454
O. mykiss NP_001154092
E. lucius XP_010890703
T. nigroviridis CAF94272
T. rubripes XP_003978618
C. harengus XP_012688496
eIF4E-1A2
T. guttata NP_001232122
G. gallus XP_420655
H. sapiens NP_001959
A. carolinensis XP_003225575
X. laevis NP_001089212
X. tropicalis NP_001015909
X. laevis BQ398016
tetrapods
L. chalumnae ENSLACT00000010262 coelacanth
C. milii XP_007897190 elephant shark
eIF4E-1A
S. salar SSGWT002823
O. mykiss CDQ69051
E. lucius XP_010868832
A. mexicanus XP_007236405
D. rerio NP_571529
L. oculatus XP_006632115 spotted gar
L. chalumnae ENSLACT00000018309 coelacanth
T. guttata XP_002194201
G. gallus_BX931053
H. sapiens NP_001092878
A. carolinensis XP_003227829
X. tropicalis XP_002936991
C. milii XP_007904436 elephant shark
tetrapods
eIF4E-1B
P. marinus EG336965 lamprey
cichlids
F. heteroclitus XP_012735982
G. aculeatus ENSGACT00000010388
O. latipes XP_004084917
X. maculatus XP_005801626
L. crocea XP_010743170
T. nigroviridis CR725540
T. rubripes ENSTRUT00000034803
C. semilaevis XP_008320130
S. partitus XP_008302737
O. mykiss CDQ63979
S. salar NP_001140090
E. lucius XP_010868333 lancelet
G. morhua ENSGMOT00368
C. harengus XP_012694154
A. mexicanus XP_007240333
D. rerio ADH59743
L. oculatus XP_006630507 spotted gar
C. milii XP_007897295 elephant shark
L. chalumnae XP_005993443 coelacanth
eIF4E-IC
sea urchin
S. purpuratus CAM57099
C. intestinalis XP_002126987 tunicate
B. floridea JGI120045 lancelet
0.3

◀ **Fig. 2** Phylogeny of deuterostome Class I eIF4E family members using the maximum likelihood under the Jones-Taylor-Thornton amino acid substitution matrix with a gamma rate correction using RAxML. The bootstrap values over 100 replicates are shown when greater than 50 %

mRNAs containing cytoplasmic polyadenylation elements (CPEs) in their 3′-UTRs [36, 37]. In addition to eIF4E-1B, zebrafish hve an additional Class I cognate, eIF4E-1C, as well as two Class II eIF4Es, eIF4E-2A and eIF4E-2B [38–40]. The availability of fully sequenced genomes from many deuterostome species has provided 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.

3 Database Searches

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 mouse (*Mus musculus*) genomes; in Ensembl (http://www.ensembl.org/index.html) for 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. The phylogenetic analysis used RAxML with 100 bootstrap iterations with the Jones-Taylor-Thornton amino acid substitution matrix with a gamma rate parameter [41]. Gene loci designations and orientations were determined using both the NCBI gene database and the Ensembl gene region-of-interest function. Spotted cat shark, *Scyliorhinus canicula*, transcriptomic sequences were provided by Dr. Helen Dooley and Anthony Redmond from the University of Aberdeen.

4 Phylogenetic Analysis of Deuterostome eIF4E Family Members

Previous phylogenetic analysis has divided Metazoa/Fungi/Viridiplantae eIF4E family members into three clades corresponding to Class I, Class II, and Class III eIF4E family members [38]. Figure 2 shows a phylogenetic analysis of deuterostome ClassI eIF4Es. The evidence of gene duplication is apparent from the number

of orthologs of each eIF4E class across the deuterostomes to give the subclasses, eIF4E-1A, eIF4E-1B, and eIF4C. Evidence of the duplication of Class I eIF4Es can be seen in the elephant shark (*Callorhinchus milii*), coelacanth (*Latimeria chalumnae*), and basal ray-finned fish, spotted gar (*Lepisosteus oculatus*), all of which have the three eIF4E-1 subclasses: eIF4E-1A, -1B, and -1C. The emergence of eIF4E-1 subclasses in these species is consistent with the duplication of Class I family members prior to the teleost-specific whole-genome duplication (TGD), probably at one of the vertebrate genome duplications that occurred at ∼550 Ma (VGD1) and 500 Ma (VGD2).

Deuterostome Class I eIF4Es: 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 of vertebrates. The eIF4E of lamprey, *Petromyzon marinus*, is intermediate between the eIF4E-1A and eIF4E-1B clades. The elephant shark (*Callorhinchus milii*), coelacanth (*Latimeria chalumnae*), and the basal ray-finned fish, spotted gar (*Lepisosteus oculatus*), all have three eIF4E-1 subclasses: eIF4E-1A, -1B, and -1C. When reviewing sequences, C-terminal motifs facilitated the recognition of Class I subtypes as well as separation from Class III eIF4Es. eIF4E-1As have the motif "SHAD," eIF4E-1Bs have "AHAD," eIF4E-1Cs have "SHDD," and eIF4E-3s have "PHEEHH."

Tetrapods and teleosts form two poorly supported clades within the eIF4E-1A clade (Figs. 2 and 3). Further examination shows a clear separation between the eIF4E-1As of early teleosts such as zebrafish, carp, and cavefish when compared to the more recently emerging teleosts such as cod, tilapia, medaka, tongue sole, and puffer fish. The eIF4E-1As from elephant shark, coelacanth, and spotted gar stand outside each cluster; the coelacanth and elephant shark eIF4E-1As fall closer to the tetrapod cluster, whereas the spotted gar eIF4E-1A falls closer to the teleost early cluster. Consistent with the clustering of sequences, in the *eif4e1a* gene loci, the elephant shark and coelacanth loci resemble that of human *eif4e1a* having nearby *Metap1* genes (Fig. 6). The loci for zebrafish and pike also have *Metap1*, but lack *tspan5* and *adh5*. They resemble each other in having *lingo2* near by, suggesting a re-shuffled gene order and consistent with the two poorly supported clades within the eIF4E-1A clade for tetrapods and teleosts (Figs. 2 and 3). The *eif4e1a2* gene, such as that in tongue sole, has one distinct signature gene, *Gar1* (data not shown), consistent with duplication of the eIF4E1A gene in some teleost lineages.

The neofunctionalized eIF4E1B does not function as a translational initiation factor, but as a regulator of mRNA recruitment. The expression of eIF4E-1B is confined to ovaries, oocytes, and early embryos in mice, zebrafish, and *Xenopus* and in zebrafish testis [36, 37, 39, 42, 43]. In *Xenopus* oocytes, eIF4E-1B has been identified as a component of the CPEB mRNP repressor complex along with the eIF4E-binding protein 4E-T, the Xp54/DDX6 RNA helicase, and the RNA-binding proteins Pat and Lsm14, as well as mRNAs containing 3′-untranslated sequences recognized by CPEB [36]. Neither recombinant nor oocyte eIF4E-1B from zebrafish or Xenopus is able to bind immobilized m^7GTP, in contrast to eIF4E1A [36, 39]. Within the eIF4E-1B sequences, tetrapods and teleosts form distinct clades that

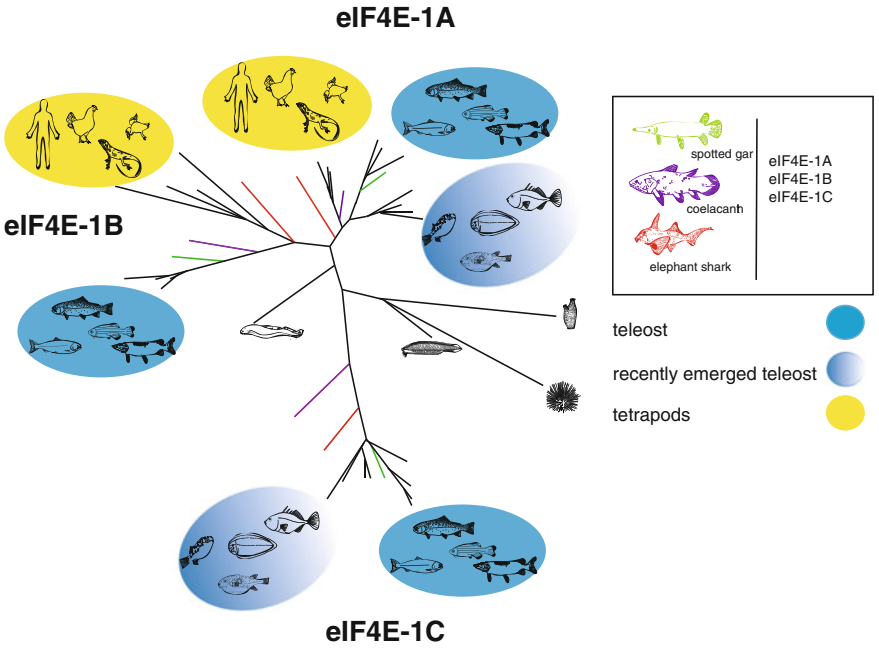


Fig. 3 Illustration of the distribution of deuterostome Class I eIF4Es derived from phylogenenetic analysis

are moderately supported (>80 % bootstrap). Again, the eIF4Es from elephant shark, speckled gar, and coelacanth stand outside each cluster with the sequences from spotted gar falling closer to the teleost cluster and the elephant shark falling closer to the tetrapods. It was previously reported that eIF4E-1B in zebrafish is not orthologous to the tetrapod form because the locus is not conserved between zebrafish and human [37]. However, this analysis was done before so many genomes were available and before they were so well annotated. Re-examination in the light of more fish genomes shows that in coelacanth and shark, the *eif4e1b* locus is the same as that found in tetrapods and is similar to the *eif4e1b* locus in gar (Fig. 6). In contrast, the teleost locus is clearly unique. It seems likely that after the duplication of the *eif4e1b* locus arising from the TGD, *eif4e1b* has been asymmetrically retained in teleosts such as the *eif4e1b* locus in zebrafish is different from that of tetrapods. In the *eif4e1b* loci, *tspan17* (tetraspanin) and *sncb* (synuclein) genes are found in the tetrapod, basal ray-finned fish, and shark. The representatives of *eif4e1b* genes in teleosts, zebrafish, and pike are instead located near the *casr* (calcium sensing receptor) gene (Fig. 6), consistent with the observations of Evsikov [37]. Interestingly, eIF4E-1B has been lost from the more recently emerging teleosts such as cod, tilapia, medaka, tongue sole, and puffer fish.

The eIF4E-1Cs are unique to the coelacanth, elephant shark, and gar and to the teleost fishes. The zebrafish eIF4E-1C has retained its function as a prototypical initiation factor by its ability to bind to cap analog, interact with eIF4G, and

complement a *S. cerevisiae* strain conditionally deficient in functional eIF4E [40]. Like eIF4E-1A, eIF4E-1C also interacts with the eIF4E-binding proteins (4E-BPs). There is a clear separation between the eIF4E-1Cs of early teleosts such as zebrafish, carp, and cavefish when compared to the more recently emerging teleosts such as cod, tilapia, medaka, tongue sole, and puffer fish. The eIF4E-1Cs of elephant shark and coelacanth fall outside the teleost clusters; the spotted gar eIF4E-1C falls close to the early teleost cluster. The loci for the eIF4E-1Cs are clearly distinct from that of eIF4E-1A with closeness to *tet1* in common (Fig. 6).

Deuterostome Class II eIF4Es: eIF4E family members of Class II fall within two discrete clusters within the phylogenetic tree (Fig. 4). The eIF4E designated as eIF4E-2 (eIF4E-2A) comprises the majority of the Class II eIF4Es within the deuterostomes. The eIF4E-2B cluster is represented primarily by the ray-finned fish, gar, and some teleosts. There is a second Class II eIF4E in *Xenopus*, which clusters with the fish eIF4E-2Bs in the phylogenetic analysis, but which is found on an entirely different locus (data not shown). Its phylogenetic relationship to the fish eIF4E-2Bs is currently unclear. As with the Class I eIF4Es, the Class II eIF4Es tend to segregate consistently with their evolutionary relationships, although it should be pointed out that the eIF4E-2 from elephant shark and coelacanth cluster more closely with the tetrapod eIF2A. Like eIF4E-2A, zebrafish eIF4E-2B does not bind m^7GTP well or eIF4G [40]. Interestingly, eIF4E-2B has been lost in some higher teleosts. The gene loci of eIF4E-2A show a characteristic signature of the genes *chrnd*, *chrng*, and *prss56* (Fig. 6). The eIF4E-2B is clearly a unique gene, the locus of which is near *mink1* (misshapen-like kinase 1), *rangrf* (RAN guanine nucleotide recycling factor), and *gp1ba* (glycoprotein 1ba).

Deuterostome Class III eIF4Es: These are the most conserved eIF4E family members across the deuterostomes. However, in the percomorph teleosts and some cyprinodontiforms there is an additional eIF4E-3 member. Phylogenetic analysis supports two clades we have termed eIF4E-3A and -3B (Fig. 5). The tetrapod eIF4E-3s fall within the eIF4E-3A clade but cluster more closely with each other than with the fish eIF4E-3As. eIF4E-3A from gar and coelacanth clusters more closely with the teleost eIF4E-3As. In contrast, elephant shark eIF4E-3A clusters more closely with the tetrapod eIF4E-3s. The loci for eIF4E-3A (Fig. 6) are close to *gpr27* (G protein coupled receptor), *rybp* (RING1 and YY1 binding protein), and *prok2* (prokineticin) (Fig. 6), while *eif4e3b* is a distinct gene locus from eIF4E-3A, close to the *foxp1*-like (forkhead box P1) and the *mitf* (microphthalmia-associated transcription factor) (data not shown). The distribution of eIF4E-3A and -3B in the teleosts varies. Some cyprinodontiforms and some percomorphs (tetradontiformes and pleuronectiformes) such as fugu, pufferfish, and tongue sole have both eIF4E-3A and -3B, as do medaka, southern platyfish, mummychug, and guppy. However, zebrafish, stickleback, and cod have only eIF4E-3A, while blind cave fish and Atlantic salmon have only eIF4E-3B.

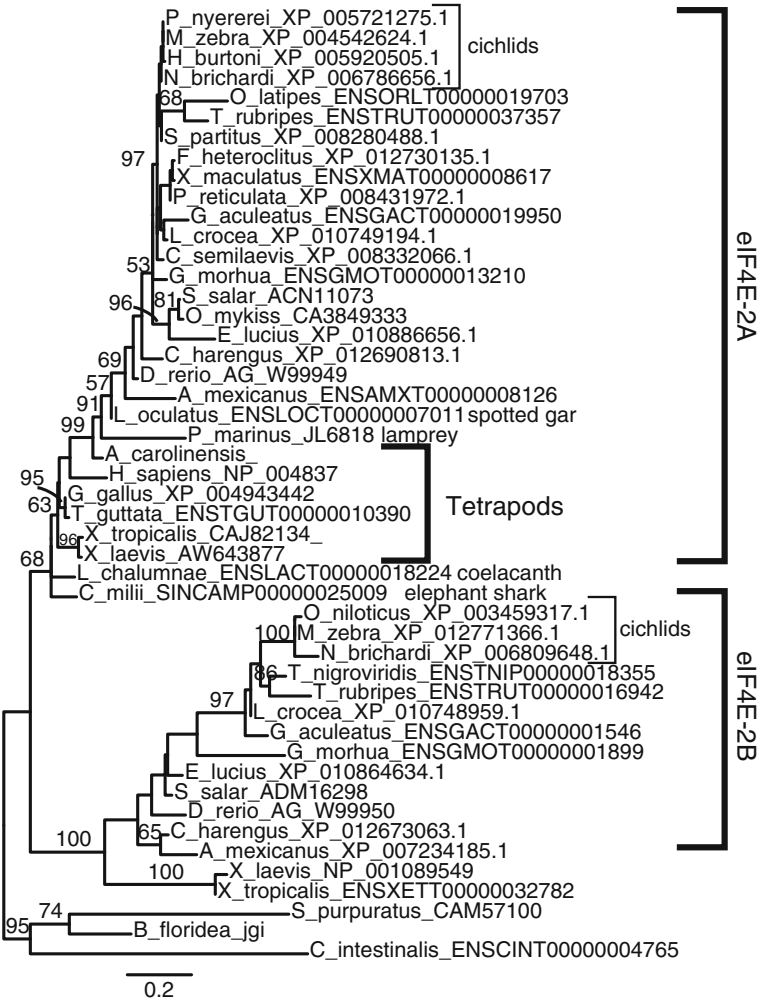


Fig. 4 Phylogeny of deuterostome Class II eIF4Es using the maximum likelihood under the Jones-Taylor-Thornton amino acid substitution matrix with a gamma rate correction using RAxML. The bootstrap values over 100 replicates are shown when greater than 50 %

5 Summary of Deuterostome eIF4E Family Members

A representation of the distribution of different eIF4E family members is provided in Fig. 7. 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, although teleosts may have retained different gene copies from tetrapods. eIF4E-1C is lost in tetrapods. eIF4E1B represents a neofunctionalized duplication of the

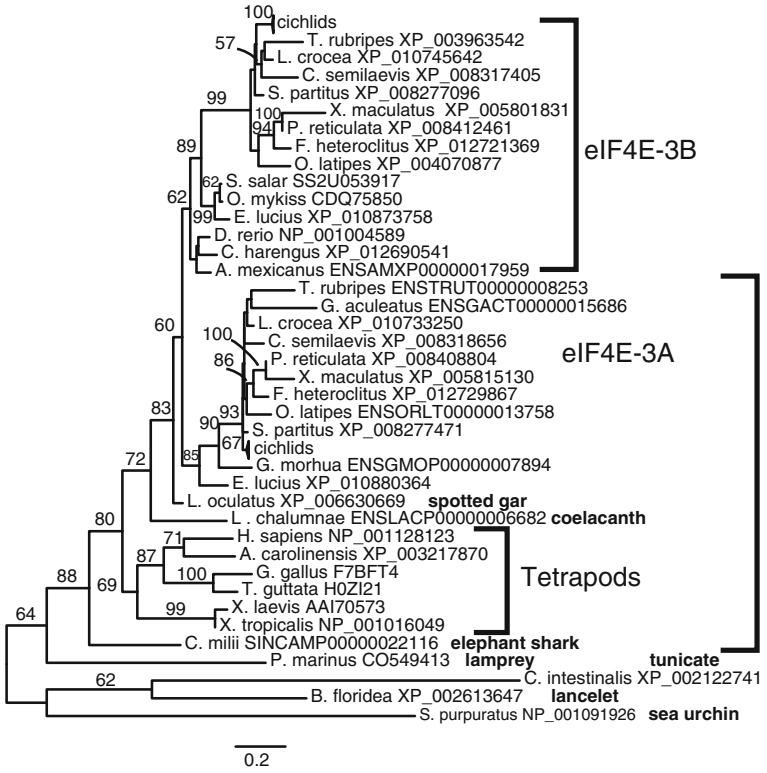


Fig. 5 Phylogeny of deuterostome Class III eIF4Es using the maximum likelihood under the Jones-Taylor-Thornton amino acid substitution matrix with a gamma rate correction using RAxML. The bootstrap values over 100 replicates are shown when greater than 50 %

ancestral eIF4E-1. It has been lost in the percomorph teleosts, but retained in sharks, basal ray-finned fish, lower teleosts, and tetrapods. The eIF4E-1B of teleosts is not orthologous to that of tetrapods. Gar and many teleosts have a second eIF4E-2, which we have termed eIF4E-2B. The eIF4E-2B has been lost in the amniotes and coelacanths, but retained in basal ray-finned fish. *Xenopus* also has a second eIF4E-2, the sequence of which clusters with the teleost eIF4E-2Bs, although it is not orthologous. In general, it is the *eif4e* genes in elephant shark and coelacanth that appear to share the same loci as human. Northern pike is a genetic wild card of sorts, in that it has all eight known deuterostome eIF4E family members, eIF4E-1A1, -1A2, eIF4E-1B, eIF4E-1C, eIF4E-2A, -2B, eIF4E-3A, and -3B.

Overall, the main expansion of the deuterostome suite of eIF4Es preceded the diversification of the teleosts to give the subclasses eIF4E-1A, -1B, and -1C arising from the VGDs. The two VGDs are thought to have occurred prior to the lamprey-gnathostome split, based on analysis of selected families of gene duplicates [44, 45]. Such a scenario would predict that lamprey should also have

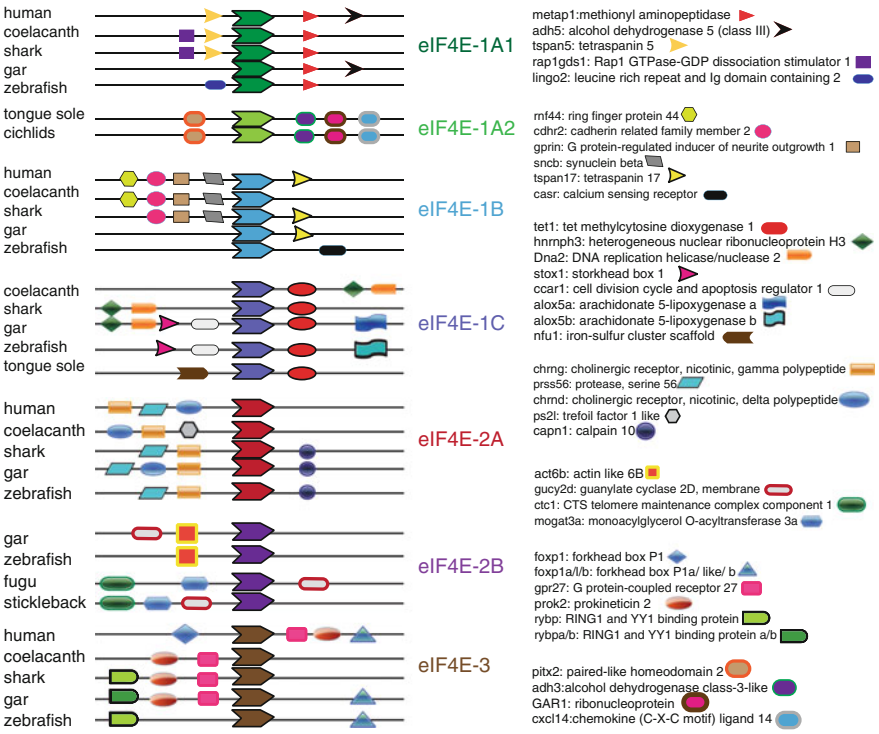


Fig. 6 Gene loci of the vertebrate eIF4E family members: Genes proximal to the *eif4e* gene within 0.1–0.4 Mb on contig or chromosome. Color scheme is coordinated with identical genes. Identity of genes proximal to *eif4e* genes

eIF4E-1A, -1B, and -1C, which it does not. However, it seems that lamprey has thrown out more and different duplications than the gnathostomes [20, 44, 45]. During teleost evolution, expansion of the subclasses took place in certain lineages arising from the TGD. Reduction of eIF4E family members accompanied the evolution of the amniotes as well as the more recently emerging teleosts. Overall, duplication within the different classes of eIF4E occurred early in vertebrate evolution and after the TGD in different subclasses with some neofunctionalization, as well as asymmetric losses in different vertebrate classes.

6 Expansion of 4E-BPs in vertebrates

The key role of eIF4E regulation in protein synthesis is underscored by the presence of eIF4E-binding protein, 4E-BP, a protein that binds specifically to eIF4E to inhibit cap-dependent translation initiation by competing with eIF4G for binding to eIF4E (reviewed [46, 47]). The competition between 4E-BPs and eIF4G is

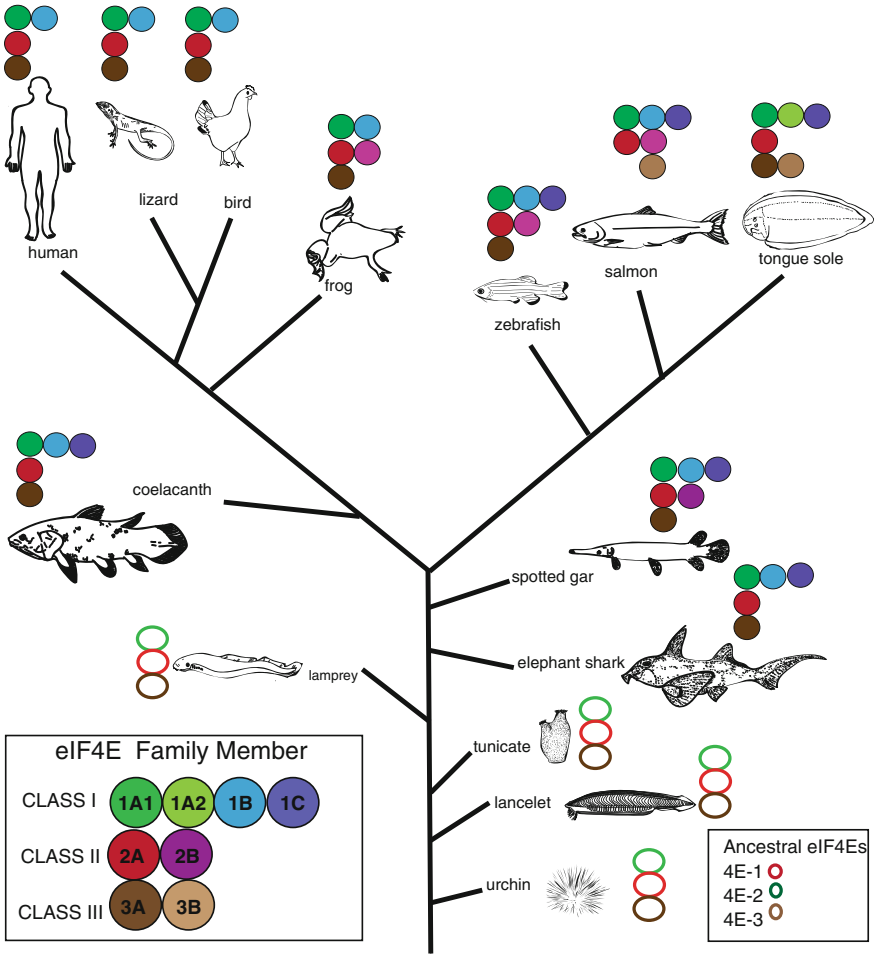


Fig. 7 Distribution of eIF4E family members among the deuterostomes

explained by the presence of a common eIF4E-binding motif in 4E-BP and eIF4G, YXXXLΦ, where Φ is any hydrophobic amino acid. In addition to the eIF4E-binding site, each 4E-BP harbors two identical canonical TOR phosphorylation sites, TPGGT, and several *trans*regulatory phosphorylation sites. Once activated, in mouse, mTORC1 phosphorylates Thr37 and Thr46 in human 4E-BP1, which are priming sites for subsequent phosphorylation at Ser65 and Thr70 [48, 49]. The binding of 4E-BP to eIF4E is controlled by the phosphorylation state. The underphosphorylated forms of 4E-BPs interact with eIF4E, whereas the hyperphosphorylated forms do not [50–52]. Upon cell stimulation with serum, growth factors, or hormones, 4E-BP1 becomes hyperphosphorylated and dissociates from eIF4E to relieve the translational inhibition [52, 53]. Phylogenetic analysis suggests

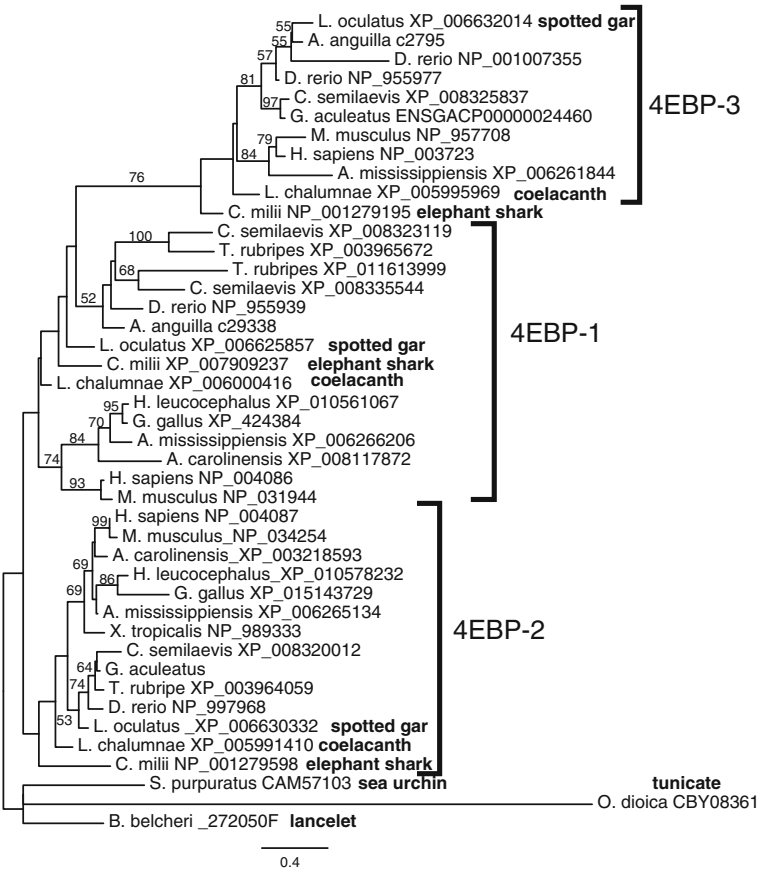


Fig. 8 Phylogeny of deuterostome 4E-BPs using the maximum likelihood under the Jones-Taylor-Thornton amino acid substitution matrix with a gamma rate correction using RAxML. The bootstrap values over 100 replicates are shown when greater than 50 %

that 4E-BP appeared as a single-copy gene in the last common ancestor of Amoebozoa, Glaucocystophyta, Fungi, and Metazoa (Hernandez et al., this book). It is found in all Metazoa except Nematoda, but is only sparsely represented in the protist lineages.

As with the eIF4E family, early deuterostome 4E-BP has been duplicated in vertebrates, with up to six cognates found in teleosts. Figure 8 shows a phylogenetic analysis of the 4E-BPs across the deuterostomes. Unlike the single 4E-BP found in sea urchin, urochordates, and cephalochordates, vertebrates have three 4E-BPs, 4E-BP1, 4E-BP2, and 4E-BP3. These are first seen in elephant shark, likely as a result of the one or more of the VGDs, and persist throughout the jawed vertebrates. In teleosts, further expansion of each the three 4E-BP classes takes place, probably as a result of the TGD, to give six eIF4E cognates termed here

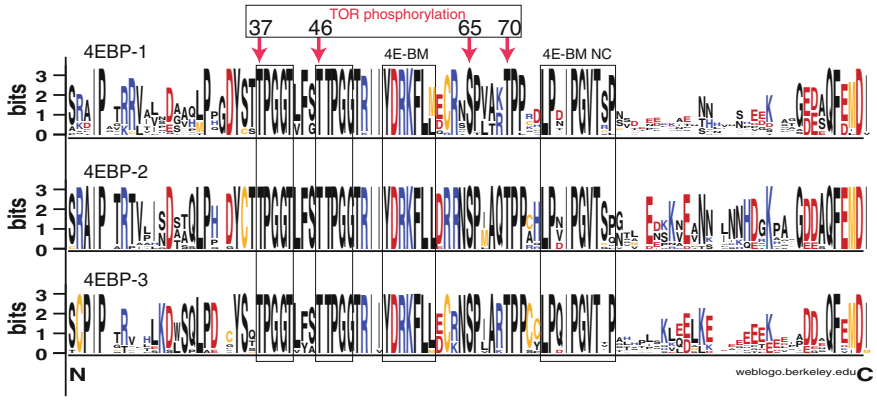


Fig. 9 The sequence logo for vertebrate 4E-BPs calculated using weblogo. The positively charged amino acids (R, K, H) are *colored blue*, the negatively charged amino acids (D, E) are *colored red*, and sulfur-containing amino acids (C, M) are *colored orange. Boxed regions contain the conservd motifs described in the text*

4E-BP1A, 4E-BP1B, 4E-BP2A, 4E-BP2B, 4E-BP3A, and 4E-BP3B. So far no one fish species has been found to have more than four 4E-BP cognates. Duplicated 4E-BP genes seem to have been lost, but in a lineage-specific way so that not all fish lineages lose the same copy.

Figure 9 represents sequence logos that were created from the alignments of the core sequences of 4E-BP1, 4E-BP2, and 4E-BP3 from a range of deuterostome species. A sequence logo is a graphical technique for displaying a summary of a set of aligned sequences [54]. 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. The 4E-BPs form three distinctive clades: 4E-BP1, 4E-BP2, and 4E-BP3. Across these three clades the central 35-amino acid region of the protein is nearly absolutely conserved between clades. In addition to the eIF4E-binding site, each 4E-BP harbors two canonical TOR phosphorylation sites, TPGGT, and several *trans*regulatory phosphorylation sites. Once activated, mTORC1 phosphorylates Thr37 and Thr46 in human 4E-BP1, which are priming sites for subsequent phosphorylation at Ser65 and Thr70 [48, 49]. A comparison of 4E-BP1s, 4E-BP2s, and 4E-BP3s across the vertebrates shows that both TOR phosphorylation sites, TPGGT, are absolutely conserved. The residues equivalent to Ser65 and Thr70 in human 4E-BP1 are universally conserved. The eIF4E-binding motif (4EBM) is YDRKFLL/M in all 4E-BP classes. In addition, the non-canonical 4E-BP is recognizable in all 4E-BP classes [55–59]. Within the C-terminus region of 4E-BP classes, patterns of residues emerge in the 4EBD-NC to distinguish between these forms. 4E-BP1s show conservation of an asparagine doublet followed by a glutamic acid or aspartic acid; 4E-BP2s have a higher occurrence of asparagines, while 4E-BP3 has a predominance of negatively

charged aspartic acid and glutamic acid (D) residues. Overall, the different classes are highly conserved, and each class has distinctive charge patterns near the carboxy terminus. All 4E-BP classes have a conserved N-terminus that is a site of high conservation, QFEMDI.

7 Summary of Deuterostome 4E-BP Family Members

Despite the high conservation of the 4E-BPs and the possible redundancy of having three classes of 4E-BPs, duplication of teleost 4E-BP genes has resulted in some fish species having four 4E-BPs. Zebrafish have four 4E-BPs, 4E-BP1, 4E-BP2, 4E-BP3A, and 4E-BP3B (referred to as 4E-BP3L in GenBank, as per [39]). Two 4E-BP3s are found in many teleosts, presumably arising from the TGD. Two 4E-BP1s are found in tongue sole. Figure 10 shows the loci of *eif4ebp* genes in vertebrates. It is strikingly apparent that although the duplicated *eif4ebps* show little sequence variation, they are found at different loci. The tetrapod 4E-BP1s, along with those from coelacanth, elephant shark, and gar, show the highest

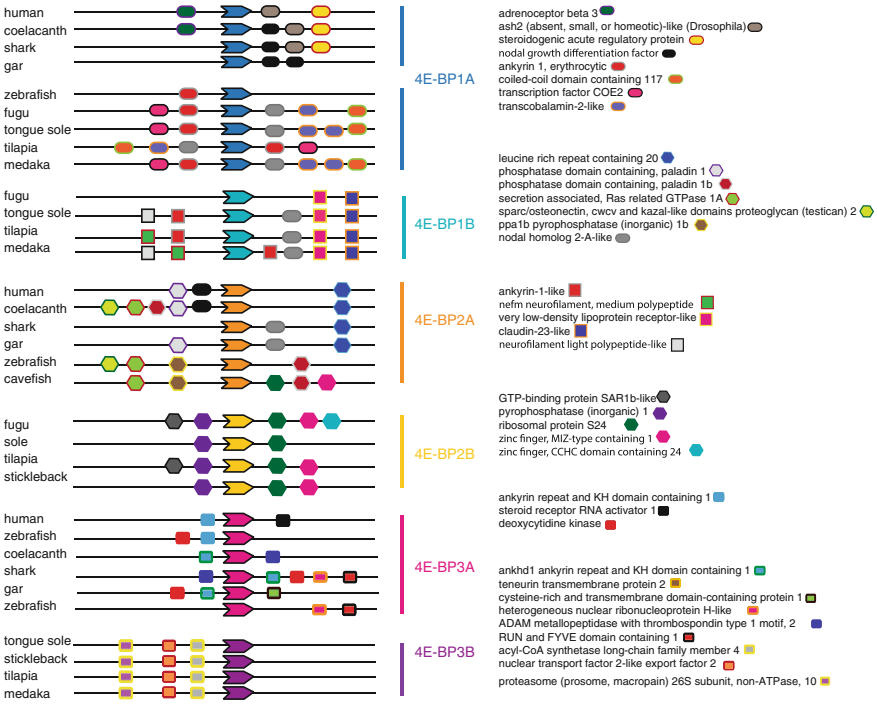


Fig. 10 Gene loci of the vertebrate 4E-BPs: genes proximal to the *eif4e* gene within 0.1–0.4 Mb on contig or chromosome. Color scheme is coordinated with identical genes. Identity of genes proximal to *eif4e* genes

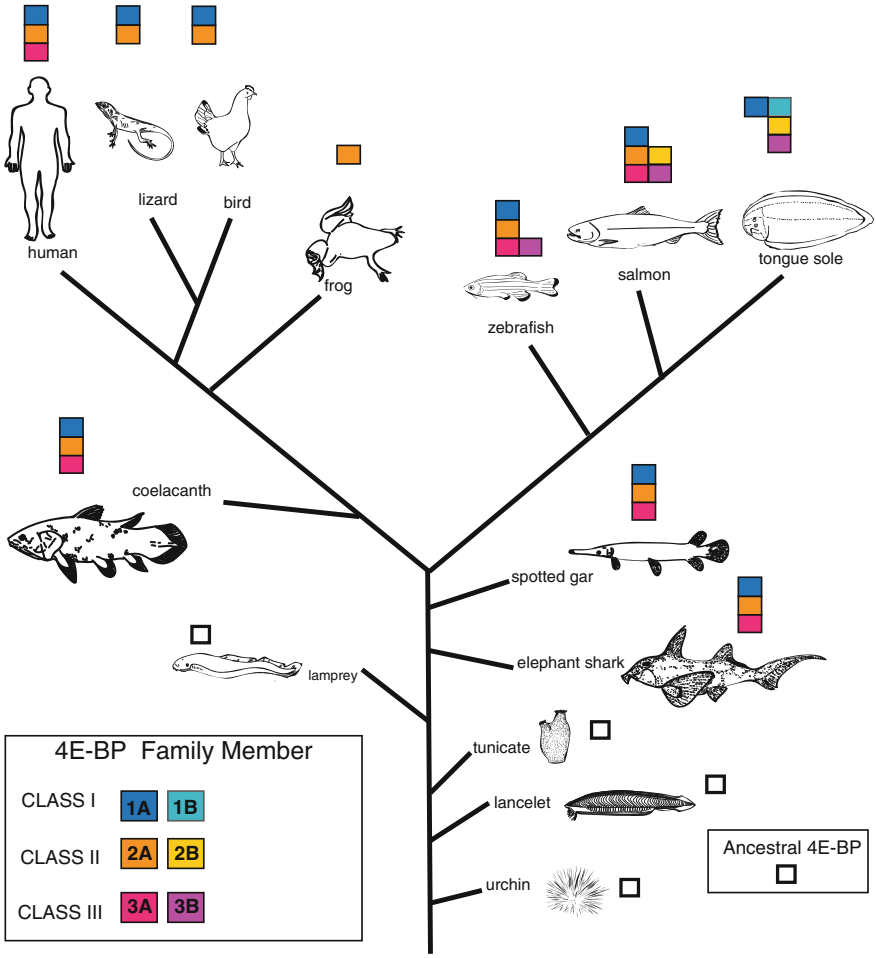


Fig. 11 Distribution of 4E-BPs among the deuterostomes

identity/similarity to the teleost 4E-BP1A. However, they are not orthologous to the teleost 4E-BP1As since they are at very different gene loci.

Figure 11 illustrates the distribution of 4E-BPs in deuterostomes. From a single 4E-BP gene in sea urchins, tunicates and lancelets, the VGDs gave rise to three classes of 4E-BPs, 4E-BP1, 4E-BP2, and 4E-BP3. These three classes gave rise to duplicated genes for each as a result of the TGD, although not all duplicates have been retained. For ease of discussion, these have been termed 4E-BP1B, 4E-BP2B, and 4E-BP3B. Tongue sole and some other percomorphs have retained both 4E-BP1A and 4E-BP1B. From the gene loci data, it is apparent that tongue sole has retained 4E-BP2B and 4E-BP3B, unlike zebrafish and other cyprinodontiforms. It is

anticipated that the duplicated genes will have different expression patterns and different roles in the regulation of gene expression.

Zebrafish have retained both 4E-BP3A and 4E-BP3B (referred to as 4E-BP3L in GenBank, as per [39]). Zebrafish 4E-BP3B, but not 4E-BP3A, has been shown to have a key effect in a zebrafish muscle inactivity model [60]. Inactivity was found to trigger upregulation of 4E-BP3B, which led to diminished myosin and actin content, myofibrilogenesis, and fiber growth. The changes were accompanied by preferential reduction of the muscle transcription factor Mef2c. Analysis of recruitment of *mef2c* mRNA to polysomes showed that the reduction in Mef2c levels was due to reduced translation of *mef2ca* mRNA. Loss of Mef2ca function reduced normal muscle growth and diminished the reduction in growth caused by inactivity. Blocking 4E-BP3B function increased Mef2ca translation and also prevented the decline in mef2ca translation caused by inactivity. Conversely, overexpression of active 4E-BP3B mimicked inactivity by decreasing the recruitment of *mef2ca* mRNA into polysomes. These findings have identified zebrafish 4E-BP3B, but not 4E-BP3A, as a key TOR-dependent regulator of muscle fiber size in response to activity. The results are consistent with the distinct patterns of *eif4ebp* tissue expression; *eif4bp1* and *eif4ebp2* are expressed widely and at high levels in head and neural tissue, *eif4ebp3a* is only abundant in pancreas, and *eif4ebp3b* is only expressed in eye, muscle and the branchial arch region [61].

Tetrapods show a different story. Although mammals retain the three 4E-BP classes, it seems that *Xenopus* has lost 4E-BP1 and 4E-BP3, and reptiles and birds have retained only 4E-BP1 and 4E-BP2.

8 Concluding Remarks

The important role of genome duplication in the evolution of the vertebrates was recognized by Ohno, who postulated that divergence of duplicated genes resulting from tetraploidization events drives the emergence of new gene functions [62]. In extant teleost genomes, around 20–25 % of protein coding genes are still retained as two TGD paralogs, increasingly referred to as ohnologs [30], with important differences among functional gene classes with regard to the retention and loss of their TGD paralogs [5, 11, 28]. Genes of high interest to the regulation of gene expression, such as transcription factors and other developmental genes, are over-represented among retained TGD paralogs [28, 63]. Sequence divergence between TGD paralogs often occurs in an asymmetric manner, with one of the two TGD duplicates evolving faster than the other [28, 29]. Furthermore, the processes of non-functionalization, rediploidization, and divergence in sequence and function often occur in a lineage-specific way [5, 30, 31].

In the vertebrate lineages, both eIF4Es and 4E-BPs show expansion as a result of the VGDs, as well as further expansion in teleosts as a result of the TGD. Multiple classes and subclasses of eIF4Es and 4E-BPs have arisen with some

neofunctionalization and asymmetric loss. Because of the genome duplications and asymmetric loss of loci, eIF4Es and 4E-BPs closely related in sequence may not show orthology between teleosts and tetrapods.

Acknowledgments RJ and TB are supported by NIH R01ES021949-01 and NSF OCE1313888 to RJ and Allen R. Place. KG was supported by a graduate fellowship from the NOAA-EPP-funded Living Marine Sciences Cooperative Science Center (LMRCSC), NA11SEC4810002. Thanks are extended to Dr. Helen Dooley and Anthony Redmond from the University of Aberdeen for providing transcriptomic sequences of eIF4Es and 4E-BPs from spotted cat shark, *Scyliorhinus canicula*. This work was begun and inspired by Dr. Bhavesh Joshi, currently chief scientific officer, BridgePath Scientific.

References

1. Davidson WS, Koop BF, Jones SJ, Iturra P, Vidal R, Maass A, Jonassen I, Lien S, Omholt SW. Sequencing the genome of the Atlantic salmon (*Salmo salar*). Genome Biol. 2010;11:403.
2. Star B, Nederbragt AJ, Jentoft S, Grimholt U, Malmstrøm M, Gregers TF, Rounge TB, Paulsen J, Solbakken MH, Sharma A, Wetten OF, Lanzén A, Winer R, Knight J, Vogel JH, Aken B, Andersen O, Lagesen K, Tooming-Klunderud A, Edvardsen RB, Tina KG, Espelund M, Nepal C, Previti C, Karlsen BO, Moum T, Skage M, Berg PR, Gjøen T, Kuhl H, Thorsen J, Malde K, Reinhardt R, Du L, Johansen SD, Searle S, Lien S, Nilsen F, Jonassen I, Omholt SW, Stenseth NC, Jakobsen KS. The genome sequence of Atlantic cod reveals a unique immune system. Nature. 2011;477:207–10.
3. Bernardi G, Wiley EO, Mansour H, Miller MR, Orti G, Haussler D, O'Brien SJ, Ryder OA, Venkatesh B. The fishes of Genome 10K. Mar Genomics. 2012;7:3–6.
4. Berthelot C, Brunet F, Chalopin D, Juanchich A, Bernard M, Noël B, Bento P, Da Silva C, Labadie K, Alberti A, Aury JM, Louis A, Dehais P, Bardou P, Montfort J, Klopp C, Cabau C, Gaspin C, Thorgaard GH, Boussaha M, Quillet E, Guyomard R, Galiana D, Bobe J, Volff JN, Genêt C, Wincker P, Jaillon O, Roest Crollius H, Guiguen Y. The rainbow trout genome provides novel insights into evolution after whole-genome duplication in vertebrates. Nat Commun. 2014;5:3657.
5. Schartl M, Walter RB, Shen Y, Garcia T, Catchen J, Amores A, Braasch I, Chalopin D, Volff JN, Lesch KP, Bisazza A, Minx P, Hillier L, Wilson RK, Fuerstenberg S, Boore J, Searle S, Postlethwait JH, Warren WC. The genome of the platyfish, *Xiphophorus maculatus*, provides insights into evolutionary adaptation and several complex traits. Nat Genet. 2013;45:567–72.
6. Chen S, Zhang G, Shao C, Huang Q, Liu G, Zhang P, Song W, An N, Chalopin D, Volff JN, Hong Y, Li Q, Sha Z, Zhou H, Xie M, Yu Q, Liu Y, Xiang H, Wang N, Wu K, Yang C, Zhou Q, Liao X, Yang L, Hu Q, Zhang J, Meng L, Jin L, Tian Y, Lian J, Yang J, Miao G, Liu S, Liang Z, Yan F, Li Y, Sun B, Zhang H, Zhang J, Zhu Y, Du M, Zhao Y, Schartl M, Tang Q, Wang J. Whole-genome sequence of a flatfish provides insights into ZW sex chromosome evolution and adaptation to a benthic lifestyle. Nat Genet. 2014;46:253–60.
7. Spaink HP, Jansen HJ, Dirks RP. Advances in genomics of bony fish. Brief Funct Genomics. 2014;13:144–56.
8. Braasch I, Peterson SM, Desvignes T, McCluskey BM, Batzel P, Postlethwait JH. A new model army: emerging fish models to study the genomics of vertebrate Evo-Devo. J Exp Zool B Mol Dev Evol. 2015;324:316–41.
9. Amemiya CT, Alföldi J, Lee AP, Fan S, Philippe H, Maccallum I, Braasch I, Manousaki T, Schneider I, Rohner N, Organ C, Chalopin D, Smith JJ, Robinson M, Dorrington RA,

Gerdol M, Aken B, Biscotti MA, Barucca M, Baurain D, Berlin AM, Blatch GL, Buonocore F, Burmester T, Campbell MS, Canapa A, Cannon JP, Christoffels A, De Moro G, Edkins AL, Fan L, Fausto AM, Feiner N, Forconi M, Gamieldien J, Gnerre S, Gnirke A, Goldstone JV, Haerty W, Hahn ME, Hesse U, Hoffmann S, Johnson J, Karchner SI, Kuraku S, Lara M, Levin JZ, Litman GW, Mauceli E, Miyake T, Mueller MG, Nelson DR, Nitsche A, Olmo E, Ota T, Pallavicini A, Panji S, Picone B, Ponting CP, Prohaska SJ, Przybylski D, Saha NR, Ravi V, Ribeiro FJ, Sauka-Spengler T, Scapigliati G, Searle SM, Sharpe T, Simakov O, Stadler PF, Stegeman JJ, Sumiyama K, Tabbaa D, Tafer H, Turner-Maier J, van Heusden P, White S, Williams L, Yandell M, Brinkmann H, Volff JN, Tabin CJ, Shubin N, Schartl M, Jaffe DB, Postlethwait JH, Venkatesh B, Di Palma F, Lander ES, Meyer A, Lindblad-Toh K. The African coelacanth genome provides insights into tetrapod evolution. Nature. 2013;496:311–6.

10. Venkatesh B, Lee AP, Ravi V, Maurya AK, Lian MM, Swann JB, Ohta Y, Flajnik MF, Sutoh Y, Kasahara M, Hoon S, Gangu V, Roy SW, Irimia M, Korzh V, Kondrychyn I, Lim ZW, Tay BH, Tohari S, Kong KW, Ho S, Lorente-Galdos B, Quilez J, Marques-Bonet T, Raney BJ, Ingham PW, Tay A, Hillier LW, Minx P, Boehm T, Wilson RK, Brenner S, Warren WC. Elephant shark genome provides unique insights into gnathostome evolution. Nature. 2014;505:174–9.
11. Braasch I, Gehrke AR, Smith JJ, Kawasaki K, Manousaki T, Pasquier J, Amores A, Desvignes T, Batzel P, Catchen J, Berlin AM, Campbell MS, Barrell D, Martin KJ, Mulley JF, Ravi V, Lee AP, Nakamura T, Chalopin D, Fan S, Wcisel D, Cañestro C, Sydes J, Beaudry FE, Sun Y, Hertel J, Beam MJ, Fasold M, Ishiyama M, Johnson J, Kehr S, Lara M, Letaw JH, Litman GW, Litman RT, Mikami M, Ota T, Saha NR, Williams L, Stadler PF, Wang H, Taylor JS, Fontenot Q, Ferrara A, Searle SM, Aken B, Yandell M, Schneider I, Yoder JA, Volff JN, Meyer A, Amemiya CT, Venkatesh B, Holland PW, Guiguen Y, Bobe J, Shubin NH, Di Palma F, Alföldi J, Lindblad-Toh K, Postlethwait JH. The spotted gar genome illuminates vertebrate evolution and facilitates human-teleost comparisons. Nat Genet. 2016.
12. Maloof AC, Porter SM, Moore JL, Dudas FO, Bowring SA, Higgins JA, Fike DA, Eddy MP. The earliest Cambrian record of animals and ocean geochemical change. Geol Soc Bull. 2010;122:1731–74.
13. Budd GE. At the origin of animals: the revolutionary cambrian fossil record. Curr Genomics. 2013;14:344–54.
14. Erwin DH, Laflamme M, Tweedt SM, Sperling EA, Pisani D, Peterson KJ. The Cambrian conundrum: early divergence and later ecological success in the early history of animals. Science. 2011;334:1091–7.
15. Dornbos SQ, Bottjer DJ. Evolutionary paleoecology of the earliest echinoderms: Helicoplacoids and the Cambrian substrate revolution. Geology. 2000;28:839–42.
16. Marlétaz F, Holland LZ, Laudet V, Schubert M. Retinoic acid signaling and the evolution of chordates. Int J Biol Sci. 2006;2:38–47.
17. Blair JE, Hedges SB. Molecular phylogeny and divergence times of deuterostome animals. Mol Biol Evol. 2005;22:2275–84.
18. Meyer A, Schartl M. 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. 1999;11:699–704.
19. Dehal P, Boore JL. Two rounds of whole genome duplication in the ancestral vertebrate. PLoS Biol. 2005;3:e314.
20. 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,

Li W. Sequencing of the sea lamprey (*Petromyzon marinus*) genome provides insights into vertebrate evolution. Nat Genet. 2013;45:415-21, 421e1-2.
21. Clark MS. Genomics and mapping of teleostei (bony fish). Comp Funct Genomics. 2003;4:182–93.
22. Ulloa PE, Iturra P, Neira C. Zebrafish as a model organism for nutrition and growth: towards comparative studies of nutritional genomics applied to aquacultured fish. Rev Fish Biol Fisheries. 2015.
23. Betancur RR, 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, Ortí G. The tree of life and a new classification of bony fishes. PLoS Curr. 2013;5.
24. Sémon M, Wolfe KH. Consequences of genome duplication. Curr Opin Genet Dev. 2007;17:505–12.
25. Sémon M, Wolfe KH. Rearrangement rate following the whole-genome duplication in teleosts. Mol Biol Evol. 2007;24:860–7.
26. Amores A, Catchen J, Ferrara A, Fontenot Q, Postlethwait JH. Genome evolution and meiotic maps by massively parallel DNA sequencing: spotted gar, an outgroup for the teleost genome duplication. Genetics. 2011;188:799–808.
27. Ravi V, Venkatesh B. Rapidly evolving fish genomes and teleost diversity. Curr Opin Genet Dev. 2008;18:544–50.
28. Brunet FG, Roest Crollius H, Paris M, Aury JM, Gibert P, Jaillon O, Laudet V, Robinson-Rechavi M. Gene loss and evolutionary rates following whole-genome duplication in teleost fishes. Mol Biol Evol. 2006;23:1808–16.
29. Steinke D, Salzburger W, Braasch I, Meyer A. Many genes in fish have species-specific asymmetric rates of molecular evolution. BMC Genom. 2006;7:20.
30. Braasch I, Postlethwait JH. Polyploidy and genome evolution. Springer; 2012. P. 341–83.
31. Postlethwait J, Amores A, Cresko W, Singer A, Yan YL. Subfunction partitioning, the teleost radiation and the annotation of the human genome. Trends Genet. 2004;20:481–90.
32. Sémon M, Wolfe KH. Reciprocal gene loss between *Tetraodon* and zebrafish after whole genome duplication in their ancestor. Trends Genet. 2007;23:108–12.
33. Lauder G, Liem K. The evolution and interrelationships of the Actinopterygian fishes. Bull Mus Comp Zool. 1989;150:95–187.
34. Dahm R, Geisler R. Learning from small fry: the zebrafish as a genetic model organism for aquaculture fish species. Mar Biotechnol (NY). 2006;8:329–45.
35. Glasauer SM, Neuhauss SC. Whole-genome duplication in teleost fishes and its evolutionary consequences. Mol Genet Genomics. 2014;289:1045–60.
36. Minshall N, Reiter MH, Weil D, Standart N. CPEB interacts with an ovary-specific eIF4E and 4E-T in early *Xenopus* oocytes. J Biol Chem. 2007;282:37389–401.
37. Evsikov AV, Marín de Evsikova C. Evolutionary origin and phylogenetic analysis of the novel oocyte-specific eukaryotic translation initiation factor 4E in Tetrapoda. Dev Genes Evol. 2009;219:111–8.
38. Joshi B, Lee K, Maeder DL, Jagus R. Phylogenetic analysis of eIF4E-family members. BMC Evol Biol. 2005;5:48.
39. Robalino J, Joshi B, Fahrenkrug SC, Jagus R. Two zebrafish eIF4E family members are differentially expressed and functionally divergent. J Biol Chem. 2004;279:10532–41.
40. Gillespie KM. Characterization of the eukaryotic translation initiation factor eIF4E (eIF4E) family members in zebrafish (*Danio rerio*). 2015.
41. Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 2006;22:2688–90.
42. Evsikov AV, Marín de Evsikova C. Gene expression during the oocyte-to-embryo transition in mammals. Mol Reprod Dev. 2009;76:805–18.
43. Kubacka D, Miguel RN, Minshall N, Darzynkiewicz E, Standart N, Zuberek J. Distinct Features of Cap Binding by eIF4E1b Proteins. J Mol Biol. 2015;427:387–405.

44. Kuraku S. Insights into cyclostome phylogenomics: pre-2R or post-2R. Zoolog Sci. 2008;25:960–8.
45. Kuraku S, Meyer A, Kuratani S. Timing of genome duplications relative to the origin of the vertebrates: did cyclostomes diverge before or after? Mol Biol Evol. 2009;26:47–59.
46. Topisirovic I, Svitkin YV, Sonenberg N, Shatkin AJ. Cap and cap-binding proteins in the control of gene expression. Wiley Interdiscip Rev RNA. 2011;2:277–98.
47. Gingras AC, Raught B, Sonenberg N. eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. Annu Rev Biochem. 1999;68:913–63.
48. Gingras AC, Raught B, Gygi SP, Niedzwiecka A, Miron M, Burley SK, Polakiewicz RD, Wyslouch-Cieszynska A, Aebersold R, Sonenberg N. Hierarchical phosphorylation of the translation inhibitor 4E-BP1. Genes Dev. 2001;15:2852–64.
49. Gingras AC, Kennedy SG, O'Leary MA, Sonenberg N, Hay N. 4E-BP1, a repressor of mRNA translation, is phosphorylated and inactivated by the Akt(PKB) signaling pathway. Genes Dev. 1998;12:502–13.
50. Haghighat A, Mader S, Pause A, Sonenberg N. Repression of cap-dependent translation by 4E-binding protein 1: competition with p220 for binding to eukaryotic initiation factor-4E. EMBO J. 1995;14:5701–9.
51. Richter JD, Sonenberg N. Regulation of cap-dependent translation by eIF4E inhibitory proteins. Nature. 2005;433:477–80.
52. Ma XM, Blenis J. Molecular mechanisms of mTOR-mediated translational control. Nat Rev Mol Cell Biol. 2009;10:307–18.
53. Fonseca BD, Smith EM, Yelle N, Alain T, Bushell M, Pause A. The ever-evolving role of mTOR in translation. Semin Cell Dev Biol. 2014;36:102–12.
54. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. Genome Res. 2004;14:1188–90.
55. Umenaga Y, Paku KS, In Y, Ishida T, Tomoo K. Identification and function of the second eIF4E-binding region in N-terminal domain of eIF4G: comparison with eIF4E-binding protein. Biochem Biophys Res Commun. 2011;414:462–7.
56. Paku KS, Umenaga Y, Usui T, Fukuyo A, Mizuno A, In Y, Ishida T, Tomoo K. 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. 2012;441:237–45.
57. Lukhele S, Bah A, Lin H, Sonenberg N, Forman-Kay JD. Interaction of the eukaryotic initiation factor 4E with 4E-BP2 at a dynamic bipartite interface. Structure. 2013;21:2186–96.
58. Igreja C, Peter D, Weiler C, Izaurralde E. 4E-BPs require non-canonical 4E-binding motifs and a lateral surface of eIF4E to repress translation. Nat Commun. 2014;5:4790.
59. Peter D, Igreja C, Weber R, Wohlbold L, Weiler C, Ebertsch L, Weichenrieder O, Izaurralde E. Molecular architecture of 4E-BP translational inhibitors bound to eIF4E. Mol Cell. 2015;57:1074–87.
60. Yogev O, Williams VC, Hinits Y, Hughes SM. eIF4EBP3L acts as a gatekeeper of TORC1 in activity-dependent muscle growth by specifically regulating Mef2ca translational initiation. PLoS Biol. 2013;11:e1001679.
61. Thisse B, Heyer V, Lux A, Alunni V, Degrave A, Seiliez I, Kirchner J, Parkhill JP, Thisse C. Spatial and temporal expression of the zebrafish genome by large-scale in situ hybridization screening. Methods Cell Biol. 2004;77:505–19.
62. Ohno S. Evolution by gene duplication. Heidelberg, Germany: Springer; 1970.
63. Kassahn KS, Dang VT, Wilkins SJ, Perkins AC, Ragan MA. Evolution of gene function and regulatory control after whole-genome duplication: comparative analyses in vertebrates. Genome Res. 2009;19:1404–18.

The Diversification of eIF4E Family Members in Plants and Their Role in the Plant-Virus Interaction

Tzvetanka D. Dinkova, León Martinez-Castilla and Miguel A. Cruz-Espíndola

1 Introduction

During translation initiation, the cap-binding protein eIF4E recognizes the cap structure present at the mRNA 5′ end. Its interaction with eIF4G enables it to recruit the mRNA for translation [1]. The eIF4G interacts with the multi-subunit eIF3, bringing together the mRNA and the 43S initiation complex formed by eIF3, the ternary complex (eIF2-Met- tRNAMet-GTP), the 40S ribosomal subunit and other initiation factors. The eIF4G protein also recruits a DEAD-box ATPase and ATP-dependent RNA helicase, eIF4A to unwind secondary structures in the 5′ untranslated region (5′UTR) of the mRNA and facilitate the ribosome scanning toward the AUG initiation codon. The interaction of eIF4G with eIF4E–5′ cap and the poly(A) binding protein (PABP) located at the 3′ end of the mRNA allows the transcript circularization for efficient translation re-initiation.

Particular eIF4E amino acids that interact with the 5′ cap structure are highly conserved across all eukaryotic organisms [2]. Its partner, eIF4G, interacts through a conserved YXXXXLΦ motif (where X is any amino acid and Φ is hydrophobic), and formation of the eIF4G/eIF4E complex (eIF4F) improves the ability to bind the 5′ cap [3, 4]. The eIF4E protein primarily functions in the initiation of translation as part of the eIF4F complex. However, interaction of eIF4E with proteins different from eIF4G may display different roles in the cell, such as nucleo-cytoplasmic transport, translational repression and turnover of mRNA [5].

Multiple eIF4E family members have been identified in a wide range of organisms including plants, flies, mammals, frogs, birds, nematodes and fish [2, 6]. Some of them have altered cap-binding affinities or interactions with eIF4G and other proteins, providing clues to their physiological roles. It has been suggested

T.D. Dinkova (✉) · L. Martinez-Castilla · M.A. Cruz-Espíndola
Departamento Bioquímica, Facultad de Química, Universidad Nacional Autónoma de México, 04510 Mexico City, Mexico
e-mail: cesy@unam.mx

G. Hernández and R. Jagus (eds.), *Evolution of the Protein Synthesis Machinery and Its Regulation*, DOI 10.1007/978-3-319-39468-8_9

that each organism has at least one eIF4E that is ubiquitous and constitutively expressed to carry out general translation and that the other family members are involved in specialized functions such as developmental cues, abiotic stress or anti-viral defense [7–13].

The success of viral infection greatly relies on the use of host protein translation machinery to produce specific proteins required for each stage of the virus cycle. Hence, important defense mechanisms involve the regulation of different components of the translation machinery to restrict the viral invasion. In plants, translation initiation factors of eIF4E and eIF4G families have been identified as recessive resistance markers for particular virus infections depending on the host. The presence of multiple variants of these factors in plants and the naturally occurring mutations evidenced selectivity in the interaction between viral proteins and the cellular translation factors.

Here we discuss recent advances in our knowledge about the evolution of different family members of eIF4E in plants, their impact on the plant-virus interplay and the possible impact of natural selection in these factors as a defense strategy to viruses.

2 Translation Initiation Factor 4E Family in Plants

Multiple eIF4E family members exist in plants [14]. eIF4E and eIF(iso)4E belong to class I, and the novel cap-binding protein (nCBP, currently termed 4EHP) belongs to class II, according to the metazoan nomenclature [2]. The classification in class I, class II and class III considered conservation of Trp43 and Trp56 (numbering according to the human eIF4E-1 sequence) in the protein. Members from class I conserve Trp43 and Trp56, from Class II have both residues substituted by Tyr or Phe, and from Class III only have substituted Trp56 with Tyr, Phe or Cys. Each class might display particular cap-binding affinities, a range of interactions with eIF4G or other proteins, selective tissue distribution or even cellular compartmentalization, providing diverse mechanisms in translation initiation regulation [15].

In plants, the eIF(iso)4E protein preferentially interacts with eIF(iso)4G forming the unique plant eIF(iso)4F complex, in addition to a canonical eIF4F complex formed by eIF4E and eIF4G [3, 16]. Each complex displays some selectivity in the recognition of mono- and di-methylated cap structures as well as in translating mRNAs structured in the 5′UTR [17, 18]. In most plant species, eIF(iso)4E shows about 50 % amino acid identity with eIF4E, and the relative abundance of each protein might vary depending on the developmental stage and the plant tissue [19, 20]. Plant 4EHP has been studied to a lesser extent than eIF4E and eIF(iso)4E. According to a biochemical characterization, it binds to mono-methylated cap and interacts with eIF(iso)4G, but poorly promotes in vitro translation [21]. The presence of class II eIF4Es is highly conserved among eukaryotes and is important during development, as demonstrated in animals [9, 10, 22].

In *Arabidopsis thaliana*, there are two additional eIF4E-related genes termed eIF4E1b and eIF4E1c [23]. It was demonstrated that eIF4Eb and eIF4Ec bind in vitro m^7GTP and eIF4G, albeit their affinity to eIF4G is much lower than that of eIF4E. *eIF4 EB* transcript is present at low levels in reproductive *Arabidopsis* tissues, and *eIF4EC* transcript has not been detected in global expression analyses. At present, it is not known whether the correspondent proteins are expressed in vivo and have some specialized function for the plant. The absence of striking phenotypes in mutants for each eIF4E or eIF(iso)4E in *Arabidopsis* suggested that their function could be partially redundant [24–26]. This is supported by the observation that in this species the absence of eIF(iso)4E induces higher eIF4E expression, probably to compensate for its function in general translation initiation. On the other hand, eIF4E and eIF(iso)4E natural or knockout mutants have been associated with recessive resistance to different RNA viruses in a wide host range [13, 27]. The mechanism of such resistance is thought to rely on interactions between the translation initiation factor and viral proteins. However, the specificity of the virus requirement for a particular isoform is not completely understood.

3 Translation Initiation Factor 4G Family in Plants

Translational activity of eIF4E depends on its interaction with eIF4G when forming eIF4F complex. Higher plants have at least two variants, eIF4G and eIF(iso)4G, each displaying high-affinity interactions with eIF4E and eIF(iso)4E, respectively. It was demonstrated that the subunits might be interchangeable in vitro, but with significantly lower affinity [26], supporting that in vivo, each complex is formed with high specificity [1]. In *Arabidopsis*, eIF(iso)4G is encoded by two different genes, *eif(iso)4g1* and *eif(iso)4g2*, which are distinct from eIF4G, but also functionally differ from each other. The single mutants *eifiso4g1* and *eifiso4g2* do not display obvious phenotypes, while double mutants on these genes are significantly affected in their growth and reproduction as well as in the stress response. Hence, the remaining eIF4G, forming a complex with eIF(iso)4E or not, is unable to substitute the eIF(iso)4F function during plant development [28].

The functional specialization of wheat eIF4G and eIF(iso)4G has been better supported experimentally than that of eIF4E and eIF(iso)4E. In vitro assays revealed that eIF4G preferentially stimulates cap-independent translation by binding a pseudoknot structure at the 5′ tobacco etch virus (TEV) UTR [29]. Also, mixed wheat eIF4G/eIF(iso)4E and eIF(iso)4G/eIF4E complexes display more similar activity to the native eIF4F and eIF(iso)4F in vitro translation experiments, suggesting that the large subunit of the complex might be driving mRNA preference [26]. A recent analysis of eIF4G isoform function, regarding reporter mRNA stability and translation, in the available *Arabidopsis* mutants revealed that eIF(iso)4G1 and eIF(iso)4G2 are functionally distinct, the activity of eIF(iso)4G2 being more similar to eIF4G [30]. This finding and the fact that the eIF(iso)4G2 sequence features are present only in *Brassicaceae* suggest that members of the eIF4G family

have developed specialized functions in mRNA recognition depending on 5′UTR characteristics and probably on the protein-protein interactions proper of each organism. *eif4 g* or *eif(iso)4 g* single mutants were also associated to virus resistance [31–33]. Noticeably, the reported virus specificity for eIF4E or eIF(iso)4E is not mirrored by eIF4G/eIF(iso)4G [13]. This suggests the existence of a complex interplay of each subunit in the eIF4F and eIF(iso)4F complexes with additional protein partners to exert specific regulatory mechanisms in plant development and in response to viral infection challenges.

4 Evolution of Plant eIF4E Family Members

Phylogenetic analyses on eIF4E and eIF4G sequences available from Viridiplantae sequenced genomes suggested that eIF(iso)4E appeared for the first time in flowering plants, while eIF(iso)4G is more ancient and all sequenced genomes contain at least one copy of this isoform [14]. In these analyses, two eIF4E family members were reported for *Gymnosperms*, one resembling the more conserved plant eIF4E and the other being different from the canonical eIF4E, but also distinct from eIF(iso)4E. Since both eIF4G and eIF(iso)4G are in *Gymnosperms*, it is possible that each eIF4E displays particular affinity for eIFiso4G or eIF4G. However, proving this hypothesis requires biochemical characterization of the *Gymnosperm* eIF4F subunit joining. A phylogenetic reconstruction with currently available eIF4E Viridiplantae sequences, based on the core region, is represented in Fig. 1. The log-likelihood of this tree under the JTT + I + G amino acid substitution model was −22230.103749. The tree was inferred from the amino acid alignment of the core region of plant and green algae eIF4E-like proteins. Confidence values for each internal split of the tree were computed from bootstrap replicates in which RAxML was allowed to halt bootstrapping automatically when convergence criteria were met (108 bootstrap replicates). The tree shows that the eIF4E family from vascular plants is divided into three main paralogous lineages: a clade of 4EHP-like sequences, an eIF4E-like clade and an eIF(iso)4E-like clade.

The eIF4E-like and eIF(iso)4E-like clades appear to be sister to each other, forming a monophyletic group (albeit with low bootstrap support), which in turn is sister to the 4EHP-like lineage. The monophyly of each one of these three groups is well supported by bootstrap values: 63 % for the eIF4E-like clade, 83 % for the monophyly of seed plants eIF(iso)4E-like sequences (although support drops to 51 % if sequences from bryophyta and lycopodiophyta are included in this lineage) and 100 % for the lineage of 4EHP-like sequences from vascular plants.

The longest branch of this phylogeny of plant eIF4E family is at the base of a monophyletic group of sequences from green algae and, intriguingly, *Amborella trichopoda*. If the tree is rooted at his longest branch, a grade that also contains sequences from green algae is formed adjacent to the 4EHP-like clade. However, in a previous analysis, 4EHP-like sequences were considered absent from green algae [14]. In our phylogenetic analysis the 4EHP-like clade contains sequences from

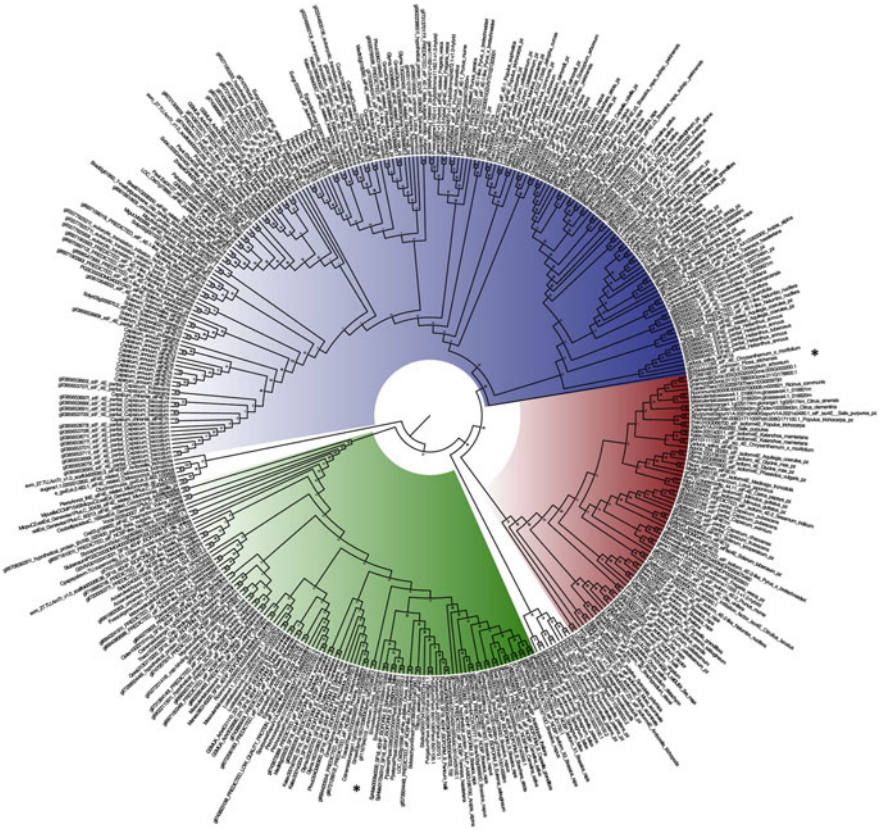


Fig. 1 Maximum likelihood phylogeny of eIF4E proteins from Viridiplantae. An initial reference alignment of members of the eIF4E protein family from plants was taken from the Pfam database (Pfam no. PF01652). From this alignment we discarded regions that did not corresponded to the well-conserved "core" of eIF4E-like proteins [2]. The conserved cores of *Arabidopsis* eIF4E paralogs were used as queries or pBLAST searches on NCBI's RefSeq database as well as on Phytozome (http://www.phylo.org/). The retrieved sequences were aligned to the core via the–addlong function of the MAFFT alignment program. The JTT model of amino acid substitution was picked as optimal. The maximum likelihood phylogenetic reconstruction was performed on the RaXML program, version 7.0 [101] running on the CIPRES Science Gateway [102]. The coloring of branches indicate three well-supported monophyletic clades: *green branches* correspond to 4EHP (formerly nCBP) sequences (bootstrap support for the monophyly of this clade 100 %), *red branches* correspond to eIF(iso)4E sequences (bootstrap support 83 %), and *blue branches* correspond to eIF4E-like sequences (support 63 %). *Asterisks* highlight the two sequences belonging to *Gymnosperm*

mosses and club mosses, a gymnosperm and angiosperms (including *Amborella trichopoda*). Interestingly, neither the eIF4E nor the eIF(iso)4E clades are as diverse as the 4EHP-like clade. For instance, eIF4E-like sequences include one from *Picea sitchensis* (a gymnosperm) but not from any moss or club moss, whereas the eIF

(iso)4E-like clade is closely allied to sequences from *Selaginella moellendroffii* and *Physcomitrella patens*, but none from gymnosperms.

The order of nesting of these monophyletic groups as well as the taxonomic provenance of their constituents is compatible with a scenario in which an ancestral cap-binding protein already existed in algae. Then, in the ancestor of vascular plants an original single copy gene suffered a duplication that gave rise to the 4EHP-like lineage and the eIF4E-like plus eIF(iso)4E-like lineage. The secondary duplication that gave rise to eIF4E and eIF(iso)4E appears to have followed shortly. Because of its closeness to algal sequences, it would appear that 4EHP retains characteristics akin to those of the ancestral cap-binding protein.

5 Evolutionary and Biochemical Interpretations for the Presence of Multiple eIF4E Family Members in Plants

Flowering plants encode eIF4E and eIFiso4E belonging to class I. At least one of them is needed for survival; 4EHP-like proteins, or even *Arabidopsis* eIF4E-related eIF4Eb and eIF4Ec proteins, if expressed, are unable to fulfill their function in plant translation. As part of eIF4F and eIFiso4F complexes, they exhibit distinct specificity for mRNA binding and translation in vitro [3] and in vivo [34]. Thus, a major relevance is attributed to eIF4G or eIF(iso)4G in each complex. This subunit prefers particular sequences and structures in the mRNA 5′UTR [18, 29], and eIF4(iso)G is required for *Arabidopsis* normal development [28]. eIF(iso)4G seems to have evolved before eIF4E diversification [14]. The phylogenetic tree from currently available eIF4E sequences (Fig. 1) supports the notion that eIF(iso)4G might have had as partner an original eIF4E sequence (i.e., the single eIF4E from modern green algae) sharing characteristics between the class I and class II eIF4E proteins. Upon divergence of eIF(iso)4E, a preference of interaction with this paralog was established for eIF(iso)4G.

How the eIF4F and eIF(iso)4F complexes support their selectivity in translation or other cellular functions in plants remains a mystery. In animals, there is a known regulatory mechanism in eIF4E activity regulation by the binding of proteins different from eIF4G through a similar motif (YXXXXLΦ). This mechanism selectively operates for different eIF4E family members, depending on protein-protein affinity and particular mRNA sequences [35]. The most studied of such proteins in animals are 4E-BP1, 2 and 3. They act as sensors of stress and nutrient availability through a phosphorylation cascade involving the mammalian target of rapamycin (mTOR) pathway [35, 36]. Stimulating the pathway by nutrients and growth factors phosphorylates 4E-BPs, releasing eIF4E for binding eIF4G and promoting translation. Under stress, unphosphorylated 4E-BP binds eIF4E, thus preventing its interaction with eIF4G and repressing translation. Other characterized eIF4E-interacting

proteins are development-related proteins that act on specific mRNA translation regulation [10, 37].

Although the TOR pathway is essentially conserved in plants and responds to nutrient, hormone and stress signals [38–40], 4E-BPs are missing. Instead, eIF4E and eIF(iso)4E protein expression patterns seem to be differentially influenced by TOR stimulation or inhibition in maize [41, 42]. Considering eIF4F and eIF(iso)4F mRNA selectivity, this could represent an alternative mechanism of translation regulation in response to metabolic states. Different expression patterns of eIF4F and eIF(iso)4F during plant development further support the possibility that each complex undertakes specific translational activity [19, 20, 43].

There are a couple of reports on plant proteins, bearing the eIF4E-binding motif, shown to weakly interact in vitro with eIF4E or eIFiso4E [44, 45]. However, the in vivo interaction of eIF4E family members with these proteins or their role in plant translation regulation has not been demonstrated. A proteomic characterization of cap-binding complexes in maize revealed the presence of new potential eIF4E/eIF(iso)4E or eIF4G/eIF(iso)4G interactors with a possible role during development and stress response [46]. Further explorations of these candidates are needed to establish whether they directly interact with eIF4E or eIFiso4E or not and whether any of them has some relevance in translation.

6 Plant eIF4E Impact on the Specificity in Host-Virus Interaction

Considering the early diversification of eIF4E family members in plants, differences in their function are expected. As described above, few reports have experimentally proven their function in vivo. However, it was evidenced that an important range of plant viruses requires one or another eIF4E family member in their infection cycle depending on the host [13, 27]. Naturally occurring mutants for either eIF4E or eIF(iso)4E were identified as markers of viral resistance, particularly (but not exclusively) for potyviruses (see Table 1). Potyvirus resistance was associated with amino acid changes in the loop near the cap recognition pocket of eIF4E (Fig. 2) in pepper, lettuce and pea [47–49]. On the other hand, natural or induced mutations that generate truncated or null eIF(iso)4E mutants were related to potyvirus resistance in *Arabidopsis*, pepper and plum [24, 50–52]. In some species potyviruses have adapted to use both eIF4E and eIFiso4E, and resistance is consequently associated with mutations in both genes. Such is the case of pepper veinal mottle virus (PVMV) [53].

The reported susceptibility/resistance associated with eIF4E or eIF(iso)4E is mostly based on the interaction between these proteins and the genome-linked viral protein (VPg) covalently attached at the 5′ end of the viral-positive RNA strand [78, 79]. However, the specific eIF4E-VPg interaction alone could not fully explain the observed relationships among the virus-host-eIF4E isoform [80–82]. Therefore, the

Table 1 Natural resistance against virus determined by eIF4E family members

Common name	Plant species	Virus	Resistance determinant	References
Chinese cabbage	*Brassica rapa*	PPV, TuMV	eIF(iso)4E (*retr01*)	[54–56]
Pepper	*Capsicum spp.*	PVY, TEV	eIF4E (*pvr1*; *pvr2*)	[57–61]
Pepper	*Capsicum spp.*	PVMV, ChiVMV	eIF(iso)4E (*pvr6*)	[53, 61]
Watermelon	*Citrullus lanatus*	ZYMV	eIF4E	[62]
Muskmelon	*Cucumis melo*	MNSV	eIF4E (*nsv1*)	[63, 64]
Barley	*Hordeum vulgare*	BaMMV, BaYMV	eIF4E (*rym4/rym5*)	[65–67]
Lettuce	*Lattuca sativa*	LMV	eIF4E (*mo1*)	[48, 68]
Green tomato	*Solanum habrochaites*	PVY, TEV	eIF4E (*pot1*)	[69]
Common bean	*Phaseolus vulgaris*	BCMV, ClYVV	eIF4E (*bc3*)	[70, 71]
Pea	*Pisum sativum*	PsBMV, BYMV, ClYVV	eIF4E (*sbm1/wlv/cyv2*)	[49, 72–75]
Apricot	*Prunus armeniaca*	PPV	eIF4E	[76]

This table only includes natural mutations of eIF4E family members as determinants of virus resistance. For a complete panorama on eIF4E/eIF(iso)4E-mediated resistance, see reference [13]. *BaMMV* barley mild mosaic virus; *BaYMV* barley yellow mosaic virus; *BCMV* bean common mosaic virus; *BYMV* bean yellow mosaic virus; *ChiVMV* chilli veinal mottle virus; *ClYVV* clover yellow vein virus; *LMV* lettuce mosaic virus; *MNSV* melon necrotic spot virus (*Tombusviridae*); *PPV* plum pox virus; *PsBMV* pea seed-borne mosaic virus; *PVMV* pepper veinal mottle virus; *PVY* potato virus Y; *TEV* tobacco etch virus; *TuMV* turnip mosaic virus; *ZYMV* zucchini yellow mosaic virus. Except noticed, all viruses belong to the *Potyviridae* family. BaMMV and BaYMV are bymoviruses, the rest are potyviruses

involvement of additional host and viral determinants is probably required to determine how the interaction between the initiation factors and the viral genome is specified and what role it performs during viral infection.

Several functions have been proposed for the eIF4E/eIF(iso)4E-VPg interaction during potyviral infection. First, it might play a role in viral RNA translation, since VPg mimics the 5′cap of cellular mRNAs, and through interaction with a particular eIF4E isoform it might recruit the translation machinery. Although this is feasible, it could not fully explain the requirement of a particular eIF4E isoform depending on the host-virus system. For example, tobacco etch virus (TEV) requires a functional eIF(iso)4E to infect *Arabidopsis thaliana*, but its RNA possesses a translational enhancer at the 5′UTR, which preferentially recruits eIF4G and is translated in a cap-independent fashion [29]. On the other hand, eIF4E, but not eIF(iso)4E, was found as relevant for TEV infection in tomato [83]. In another scenario, eIF(iso)4G represents a determinant in rice yellow mottle virus (RYMV) susceptibility/resistance, and it was demonstrated that VPg directly binds to eIFiso4G rather than eIF4E isoforms [84, 85]. Therefore, while specific eIF4E/eIF(iso)4E-VPg interactions may be

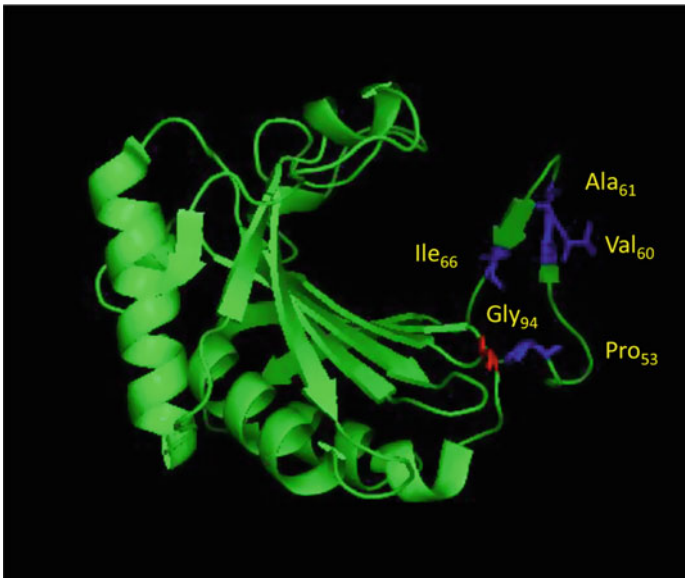


Fig. 2 A 3D ribbon diagram based on wheat eIF4E structure [77]. Amino acid positions relevant for potyvirus resistance are shown. Residue G_{94} shaded in *red* represents $G_{107}R$ mutation affecting both VPg and cap binding and is associated with virus resistance in pepper, tomato and pea. *Blue-shaded residues* represent other positions where substitutions affect only VPg binding

used in viral RNA translation for some virus-host combinations, it probably has additional roles in the infection cycle.

Second, a sequestration of eIF4E/eIF(iso)4E by VPg might inhibit global or specific cellular translation to favor viral RNA translation during the virus infection cycle [86, 87]. There are a number of viral strategies that impair the proper interaction between components of the translation machinery and decrease cellular translation upon virus infection [13]. However, global translation is not usually affected by the absence of one of eIF4E family members. Hence, sequestration of either eIF4E or eIF(iso)4E from translation would not greatly impact cellular protein synthesis. Nevertheless, it could promote changes in specific mRNA translation [34]. An indirect effect of sequestering a particular eIF4E isoform could be the translation inhibition of particular host factors used in plant defense against viral infection.

Third, specific interaction between eIF4E isoform-VPg and/or other viral and host factors might drive viral replication, location to particular cellular compartments and viral movement to systemic tissues to achieve a successful viral [80–82, 88]. A particular co-localization of the turnip mosaic virus (TuMV) precursor NIa (VPg + Pro) with eIF(iso)4E, eEF1A and PABP in membrane-associated viral replicating vesicles suggests a possible role for these factors in viral RNA replication [89–91]. In addition, mutation of eIF4E or eIF(iso)4E is related with restricted cell-to-cell or systemic movement for some potyviruses [49, 50].

Furthermore, the interaction between eIF(iso)4E and the TEV coat protein (CP) might facilitate the viral systemic movement in *Arabidopsis* [81]. A direct interaction between eIF4E or eIF(iso)4E and the viral helper component proteinase (HcPro) has also been demonstrated for potyviruses [92]. Hence, multiprotein interactions likely underlay host-virus specificity regarding the eIF4E isoform resistance determinants. A recent report found that complementation of TEV-resistant *Arabidopsis eifiso4e* mutant with a susceptible eIF4E allele from pepper was able to restore particular TEV accession susceptibility [93]. This was surprising since the pepper eIF4E sequence is much more similar to *Arabidopsis* eIF4E than to eIF(iso)4E (Fig. 1). A plausible explanation for the complementation outcome is that the pepper susceptible eIF4E was able to interact with particular viral determinants in a host-independent fashion. Supporting this, the overexpression of this allele, but not of *Arabidopsis* eIF(iso)4E, conferred susceptibility to a different TEV accession that is normally unable to infect *Arabidopsis*.

7 Positive Natural Selection for Virus Resistance in Plant eIF4E and eIF(iso)4E

Natural variation and functional analyses evidenced co-evolution between eIF4E and potyviral VPg [60, 67, 94]. The resistance-conferring mutations on eIF4E represent an important tool in plant defense mechanisms and adaptation [47–49, 57, 69]. Based on the resolved wheat eIF4E structure, mutated amino acids locate mainly in two regions on the protein surface, one near the cap-binding pocket and another rotated 90° from this pocket [77]. Most of the resistant mutants carry more than one amino acid substitution, making it difficult to understand the significance of each mutation. A nice functional analysis was performed for the pepper *pvr1* (eIF4E) resistant allele, where dissected alterations revealed the role of corresponding amino acids in affinity with potyviral VPg *in planta* [47]. The study found that $G_{107}R$ (G_{94} for wheat; Fig. 2) near the cap-binding pocket is able to disrupt both the VPg interaction and the cap binding, while $L_{79}R$ (I_{66} for wheat, Fig. 2), located within an external loop, strongly affected VPg, but not cap binding. Therefore, the optimal interaction of VPg and eIF4E requires at least two binding sites, and one is closely related to the cap-binding site. Mutation of glycine 107 is a critical amino acid substitution to gain resistance against viruses not only for pepper (*pvr1*; [59]), but also for lettuce (*mo1*; [48]), tomato (*pot-1*; [69]) and pea (*sbm1*; [49]). Additional mutations observed in all these mutant alleles could act to minimize the costs of cap-binding impairment or to reinforce the VPg-binding disruption associated with resistance.

In vitro interactions between eIF4E and cap analogs or potyviral VPg have shown that VPg interaction and cap binding are not mutually exclusive, although the binding of one ligand reduces the affinity for the other [78]. Taking together the mutational analysis, in vitro experiments and distribution of substituted amino acids

on eIF4E structure, it is proposed that the VPg-binding domain partially overlaps with the cap-binding domain, but does not necessarily affect the cap binding. The precise contact points between eIF4E and potyviral VPg are optimized for each potyvirus and host system. Distinct eIF4E isoforms are preferred depending on the host, and only specific combinations of mutations lead to virus resistance [74]. For example, the *pvr6* locus in pepper represents a null allele of eIF(iso)4E, which in combination with *pvr2* (eIF4E point mutations) allows PVMV resistance [53, 61]. Therefore, in some cases acquiring virus resistance might render defective functions for eIF4E family members at the cellular level. The presence of multiple eIF4E proteins in plants might compensate for the lack of function for one of the members [24, 25], allowing the evolution of resistance [68, 94, 95]. Some species, like pepper or *Arabidopsis*, might offset the cost of mutations on eIF4E by taking advantage of additional paralogs, different from eIF(iso)4E and 4EHP, although their function in cellular translation is uncertain [14, 23]. Nevertheless, even the presence of such paralogs might not be sufficient to fulfill expected functions of a particular eIF4E family member under non-optimal growth conditions.

Viruses have also evolved to overcome natural resistance mediated by eIF4E or eIF(iso)4E mutants [94, 96, 97]. Particularly, amino acids of the VPg central domain involved in eIF4E binding are subject to positive selection to restore an interaction with the mutated resistance protein [96]. A co-evolution of both the eIF4E resistance determinant and VPg avirulence factor might lead to the diversification of both genes [60, 95] in the context of particular plant-virus systems and geographical contexts [75]. Initial evolutionary studies were performed mostly at the intraspecific level or involved a small number of species with available data on resistance-related point mutations [67, 94, 96]. These studies revealed that although eIF4E is highly constrained [analyses that average the ratio of non-synonymous-synonymous substitutions (dN/ds) over the whole length of the sequence yield a strong signal of negative natural selection], particular amino acid positions appear under positive natural selection, including some of the amino acids relevant for cap binding. Testing the available eIF4E allele sequences for pepper, tomato and pea by two different methods, a phylogenetic analysis using maximum likelihood (PAML) and hypothesis testing using phylogenies (HyPhy), found ten positions as positively selected, from which seven corresponded to resistance mutations [94]. Between these, the G_{107} position in pepper and pea (G_{94} in wheat, Fig. 2) and $AA_{76–77}$ ($VA_{60–61}$ in wheat, Fig. 2) were found under positive selection by both methods.

In a more recent report, 22 eIF4E sequences belonging to different plant species including monocots and dicots were tested for natural selection, revealing only four positions under positive selection [95]. Two of these positions corresponded to reported resistance-related substitutions in pepper and pea, while the other two had not been previously associated with virus resistance. In this analysis, the position corresponding to G_{107} (G_{94} in wheat) did not appear under positive selection, suggesting that it might be relevant in host-virus interaction only for a narrow group of plant species. Since resistance-associated mutations were reported only for 7 out of 22 analyzed species, it is possible that positions not associated with known virus resistance are relevant for a similar purpose or for different adaptive processes in the

remaining species [95]. Alternatively, the possibility remains that the position corresponding to G_{107} indeed plays a role in virus resistance for any of the species, but that resistance has appeared independently as pressure was exerted only in particular contexts. If this is the case, only a series of analyses focusing on dN/dS rate patterns among alleles and paralogs within species would be able to detect the signature of positive natural selection.

Taking advantage of pea germplasm diversity and the systematic screening of eIF4E resistance alleles in a high number of accessions with known geographical origins, a recent analysis exposed the frequency of resistance-associated allelic diversity and their relationship with domestication [75]. At least in pea, the several virus-resistance eIF4E alleles did not appear in wild accessions, whereas they showed a particular clustering according to different regions of the cultivated crop. Similar results were reported for barley eIF4E haplotypes, where higher mutation frequency was also coincident with regions of early cultivation and high bymovirus incidence [67]. This supports that evolution of resistance proceeded independently in wild and crop species, suggesting that viruses had greater impact as agriculture developed.

Curiously, high evolutionary constraints were also found for potyviral VPg [60, 95, 96]. Most amino acid substitutions under positive selection belong to the central domain of the protein and are involved in overcoming eIF4E-mediated resistance with different infection spectrums [98, 99]. The co-evolutionary pattern of amino acid substitutions in eIF4E and VPg could have important consequences for each particular pathosystem. For eIF4E, mutations conferring resistance and maintaining generally unaltered cellular function are accepted within each specific genetic background. Several mutations impacting the resistance (VPg interaction) could probably contend with the reciprocal counter-evolution of viral VPg for a longer time. On the other hand, VPg mutations to overcome eIF4E resistance might impact its interaction with several targets at the same time, allowing pleiotropic pathogenicity effects [99].

The durability of some pepper eIF4E resistance alleles has been also evaluated in the context of co-evolutionary history for eIF4E and potyviral VPg [95]. This study evidenced a positive correlation between the resistance spectrum and its durability, suggesting that the spectrum of action of resistance alleles might be actually used as predictor of their potential durability in the pathosystem. Also, additional alleles, such as *pvr6* and the genetic background in pepper *pvr2* accessions, were shown to play an important role in the frequency of resistance breakdown [97, 100]. Therefore, evaluation of the spectrum broadness in reported resistance alleles, as well as the quantitative resistance in different genetic backgrounds and habitats, would provide better understanding of the evolution of eIF4E-mediated resistance against viruses.

8 Concluding Remarks

The greater availability of completely sequenced plant genomes increases our understanding on eIF4E evolution in combination with other components of the translation machinery. However, in vivo functional analyses are still scarce and unable to decipher whether paralogs within each species are required for specific functions in translation regulation or other cellular processes. Testing the available eIF4E family member natural or induced mutants under diverse stressful conditions, as well as uncovering their potential cellular targets and interactors, would shed light on the associated regulatory mechanisms and the relevance of their conservation throughout evolution. Regarding their pivotal role in plant-virus interaction, future work should aim at a deeper understanding of the mechanisms underlying family member-virus-host specificity and exploration of the role exerted by additional components associated with the eIF4E-VPg complexes. Evolutionary analyses for eIF4E variants within species, or groups of closely related species, and their relationship with viral VPg co-evolution are promising tools for virus-resistance field management. However, is there a cost for the spread of mutant eIF4E versions? In other words, how would plants with impaired eIF4E and/or eIF (iso)4E contend with other stressful conditions, such as temperature, salinity or drought? Combinatorial biotechnological, genetic and biochemical approaches would surely contribute to answering these questions, not only for the model plant Arabidopsis, but also for agriculturally relevant species.

Acknowledgments Research performed in Dr. Dinkova's laboratory is supported by grants from Consejo Nacional de Ciencia y Tecnología, CONACYT 238439, Facultad de Química PAIP 5000-9118, and Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica, PAPIIT IN211215. The authors acknowledge the help provided by Dr. Vasti T. Juárez-González in figure preparation.

References

1. Browning KS, Bailey-Serres J. Mechanism of Cytoplasmic mRNA Translation. In: The Arabidopsis book. doi:10.1199/tab.0176; 2015.
2. Joshi B, Lee K, Maeder DL, Jagus R. Phylogenetic analysis of eIF4E-family members. BMC Evol Biol. 2005;5:48.
3. Mayberry LK, Allen ML, Dennis MD, Browning KS. Evidence for variation in the optimal translation initiation complex: plant eIF4B, eIF4F, and eIF(iso)4F differentially promote translation of mRNAs. Plant Physiol. 2009;150:1844–54.
4. Kaye NM, Emmett KJ, Merrick WC, Jankowsky E. Intrinsic RNA binding by the eukaryotic initiation factor 4F depends on a minimal RNA length but not on the m^7G cap. J Biol Chem. 2009;284:17742–50.
5. Rhoads RE. eIF4E: new family members, new binding partners, new roles. J Biol Chem. 2009;284:16711–5.
6. Hernández G, Vazquez-Pianzola P. Functional diversity of the eukaryotic translation initiation factors belonging to eIF4 families. Mech Dev. 2005;122:865–76.

7. Amiri A, Keiper BD, Kawasaki I, Fan Y, Kohara Y, Rhoads RE, Strome S. An isoform of eIF4E is a component of germ granules and is required for spermatogenesis in *C. elegans*. Development. 2001;128:3899–912.
8. Ptushkina M, Malys N, McCarthy JE. eIF4E isoform 2 in *Schizosaccharomyces pombe* is a novel stress-response factor. EMBO Rep. 2004;5:311–6.
9. Dinkova TD, Keiper BD, Korneeva NL, Aamodt EJ, Rhoads RE. Translation of a small subset of *Caenorhabditis elegans* mRNAs is dependent on a specific eukaryotic translation initiation factor 4E isoform. Mol Cell Biol. 2005;25:100–13.
10. Cho PF, Poulin F, Cho-Park YA, Cho-Park IB, Chicoine JD, Lasko P, Sonenberg N. A new paradigm for translational control: inhibition via mRNA tethering by Bicoid and the eIF4E cognate 4EHP. Cell. 2005;121:411–23.
11. Syntichaki P, Troulinaki K, Tavernarakis N. eIF4E function in somatic cells modulates ageing in *Caenorhabditis elegans*. Nature. 2007;445:922–6.
12. Hernández G, Proud CG, Preiss T, Parsyan A. On the diversification of the translation apparatus across eukaryotes. Comp Funct Genomics. 2012;2012:256848.
13. Sanfaçon H. Plant translation factors and virus resistance. Viruses. 2015;7:3392–419.
14. Patrick RM, Browning KS. The eIF4F and eIFiso4F complexes of plants: an evolutionary perspective. Comp Funct Genomics. 2012;2012:287814.
15. Rhoads RE, Dinkova TD, Jagus R. Approaches for analyzing the differential activities and functions of eIF4E family members. Methods Enzymol. 2007;429:261–97.
16. Browning KS. Plant translation initiation factors: it is not easy to be green. Biochem Soc Trans. 2004;32:589–91.
17. Carberry SE, Goss DJ. Wheat germ initiation factors 4F and (iso)4F interact differently with oligoribonucleotide analogues of rabbit α-globin mRNA. Biochemistry. 1991;30:4542–5.
18. Gallie DR, Browning KS. eIF4G functionally differs from eIFiso4G in promoting internal initiation, cap-independent translation, and translation of structured mRNAs. J Biol Chem. 2001;276:36951–60.
19. Dinkova TD. Sanchez de Jimenez E. Differential expression and regulation of translation initiation factors -4E and -iso4E during maize germination. Physiol Plant. 1999;107:419–25.
20. Bush MS, Hutchins AP, Jones AM, Naldrett MJ, Jarmolowski A, Lloyd CW, Doonan JH. Selective recruitment of proteins to 5' cap complexes during the growth cycle in Arabidopsis. Plant J. 2009;59:400–12.
21. Ruud KA, Kuhlow C, Goss DJ, Browning KS. Identification and characterization of a novel cap-binding protein from *Arabidopsis thaliana*. J Biol Chem. 1998;273:10325–30.
22. 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, Sonenberg N. A novel 4EHP-GIGYF2 translational repressor complex is essential for mammalian development. Mol Cell Biol. 2012;32:3585–93.
23. Patrick RM, Mayberry LK, Choy G, Woodard LE, Liu JS, White A, Mullen RA, Tanavin TM, Latz CA, Browning KS. Two Arabidopsis loci encode novel eukaryotic initiation factor 4E isoforms that are functionally distinct from the conserved plant eukaryotic initiation factor 4E. Plant Physiol. 2014;164:1820–30.
24. Duprat A, Caranta C, Revers F, Menand B, Browning KS, Robaglia C. The Arabidopsis eukaryotic initiation factor (iso)4E is dispensable for plant growth but required for susceptibility to potyviruses. Plant J. 2002;32:927–34.
25. Combe JP, Petracek ME, van Eldik G, Meulewaeter F, Twell D. Translation initiation factors eIF4E and eIFiso4E are required for polysome formation and regulate plant growth in tobacco. Plant Mol Biol. 2005;57:749–60.
26. Mayberry LK, Allen ML, Nitka KR, Campbell L, Murphy PA. Browning KS Plant cap-binding complexes eukaryotic initiation factors eIF4F and eIFiso4F: molecular specificity of subunit binding. J Biol Chem. 2011;286:42566–74.
27. Robaglia C, Caranta C. Translation initiation factors: a weak link in plant RNA virus infection. Trends Plant Sci. 2006;11:40–5.

28. Lellis AD, Allen ML, Aertker AW, Tran JK, Hillis DM, Harbin CR, Caldwell C, Gallie DR, Browning KS. Deletion of the eIFiso4G subunit of the Arabidopsis eIFiso4F translation initiation complex impairs health and viability. Plant Mol Biol. 2010;74:249–63.
29. Gallie DR. Cap-independent translation conferred by the 5' leader of tobacco etch virus is eukaryotic initiation factor 4G dependent. J Virol. 2001;75:12141–52.
30. Gallie DR. Eukaryotic initiation factor eIFiso4G1 and eIFiso4G2 are isoforms exhibiting distinct functional differences in supporting translation in Arabidopsis. J Biol Chem. 2016;291:1501–13.
31. Yoshii M, Nishikiori M, Tomita K, Yoshioka N, Kozuka R, Naito S, Ishikawa M. The Arabidopsis cucumovirus multiplication 1 and 2 loci encode translation initiation factors 4E and 4G. J Virol. 2004;78:6102–11.
32. Nicaise V, Gallois JL, Chafiai F, Allen LM, Schurdi-Levraud V, Browning KS, Candresse T, Caranta C, Le Gall O, German-Retana S. Coordinated and selective recruitment of eIF4E and eIF4G factors for potyvirus infection in Arabidopsis thaliana. FEBS Lett. 2007;581:1041–6.
33. Lee JH, Muhsin M, Atienza GA, Kwak DY, Kim SM, De Leon TB, Angeles ER, Coloquio E, Kondoh H, Satoh K, Cabunagan RC, Cabauatan PQ, Kikuchi S, Leung H, Choi IR. Single nucleotide polymorphisms in a gene for translation initiation factor (eIF4G) of rice (Oryza sativa) associated with resistance to Rice tungro spherical virus. Mol Plant Microbe Interact. 2010;23:29–38.
34. Martinez-Silva AV, Aguirre-Martinez C, Flores-Tinoco CE, Alejandri-Ramirez ND, Dinkova TD. Translation initiation factor AteIF(iso)4E is involved in selective mRNA translation in *Arabidopsis thaliana* seedlings. PLoS ONE. 2012;7:e31606.
35. Jackson RJ, Hellen CU, Pestova TV. The mechanism of eukaryotic translation initiation and principles of its regulation. Nat Rev. 2010;10:113–27.
36. Sonenberg N, Hinnebusch AG. New modes of translational control in development, behavior, and disease. Mol Cell. 2007;28:721–9.
37. Stebbins-Boaz B, Cao Q, de Moor CH, Mendez R, Richter JD. Maskin is a CPEB-associated factor that transiently interacts with elF-4E. Mol Cell. 1999;4:1017–27.
38. John F, Roffler S, Wicker T, Ringli C. Plant TOR signaling components. Plant Signal Behav. 2011;6:1700–5.
39. Robaglia C, Thomas M, Meyer C. Sensing nutrient and energy status by SnRK1 and TOR kinases. Curr Opin Plant Biol. 2012;15:301–7.
40. Xiong Y, Sheen J. The role of target of rapamycin signaling networks in plant growth and metabolism. Plant Physiol. 2014;164:499–512.
41. Dinkova TD, Aguilar R, Sánchez de Jiménez E. Expression of maize eukaryotic initiation factor (eIF) iso4E is regulated at the translational level. Biochem J. 2000;351:825–31.
42. Dinkova TD, Reyes de la Cruz H, Garcia-Flores C, Aguilar R, Jimenez-Garcia LF, Sanchez de Jimenez E. Dissecting the TOR-S6K signal transduction pathway in maize seedlings: relevance on cell growth regulation. Physiol Plant 2007;130:1–10.
43. Rodriguez CM, Freire MA, Camilleri C, Robaglia C. The *Arabidopsis thaliana* cDNAs coding for eIF4E and eIF(iso)4E are not functionally equivalent for yeast complementation and are differentially expressed during plant development. Plant J. 1998;13:465–73.
44. Freire MA, Tourneur C, Granier F, Camonis J, El Amrani A, Browning KS, Robaglia C. Plant lipoxygenase 2 is a translation initiation factor-4E-binding protein. Plant Mol Biol. 2002;44:129–40.
45. Freire MA. Translation initiation factor (iso) 4E interacts with BTF3, the β-subunit of the nascent polypeptide-associated complex. Gene. 2005;345:271–7.
46. Lazaro-Mixteco PE, Dinkova TD. Identification of proteins from Cap-binding complexes by mass spectrometry during maize (*Zea mays* L.) germination. J Mex Chem Soc. 2012;56:36–50.
47. Yeam I, Cavatorta JR, Ripoll DR, Kang BC, Jahn MM. Functional dissection of naturally occurring amino acid substitutions in eIF4E that confers recessive potyvirus resistance in plants. Plant Cell. 2007;19:2913–28.

48. Nicaise V, German-Retana S, Sanjuan R, Dubrana MP, Mazier M, Maisonneuve B, Candresse T, Caranta C, Le Gall O. The eukaryotic translation initiation factor 4E controls lettuce susceptibility to the potyvirus *Lettuce mosaic virus*. Plant Physiol. 2003;132:1272–82.
49. Gao Z, Johansen E, Eyers S, Thomas C, Ellis T, Maule A. The potyvirus recessive resistance gene, *sbm1*, identifies a novel role for translation initiation factor eIF4E in cell-to-cell trafficking. Plant J. 2004;40:376–85.
50. Lellis A, Kasschau K, Whitham S, Carrington J. Loss-of-susceptibility mutants of *Arabidopsis thaliana* reveal an essential role for eIF(iso)4E during potyvirus infection. Curr Biol. 2002;12:1046–51.
51. Hwang J, Li J, Liu WY, An SJ, Cho H, Her NH, Yeam I, Kim D, Kang BC. Double mutations in eIF4E and eIFiso4E confer recessive resistance to Chilli veinal mottle virus in pepper. Mol Cells. 2009;27:329–36.
52. Wang X, Kohalmi SE, Svircev A, Wang A, Sanfacon H, Tian L. Silencing of the host factor eIF(iso)4E gene confers plum pox virus resistance in plum. PLoS ONE. 2013;8:e50627.
53. Ruffel S, Gallois JL, Moury B, Robaglia C, Palloix A, Caranta C. Simultaneous mutations in translation initiation factors eIF4E and eIF(iso)4E are required to prevent Pepper veinal mottle virus infection of pepper. J Gen Virol. 2006;87:2089–98.
54. Walsh JA, Rusholme RL, Hughes SL, Jenner CE, Bambridge JM, Lydiate DJ, Green SK. Different classes of resistance to Turnip mosaic virus in *Brassica rapa*. Eur J Plant Pathol. 2002;108:15–20.
55. Rusholme RL, Higgins EE, Walsh JA, Lydiate DJ. Genetic control of broad-spectrum resistance to turnip mosaic virus in *Brassica rapa* (Chinese cabbage). J Gen Virol. 2007;88:3177–86.
56. Nellist CF, Qian W, Jenner CE, Moore JD, Zhang S, Wang X, Briggs WH, Barker GC, Sun R, Walsh JA. Multiple copies of eukaryotic translation initiation factors in *Brassica rapa* facilitate redundancy, enabling diversification through variation in splicing and broad-spectrum virus resistance. Plant J. 2014;77:261–8.
57. Ruffel S, Dussault MH, Palloix A, Moury B, Bendahmane A, Robaglia C, Caranta C. A natural recessive resistance gene against Potato virus Y in pepper corresponds to the eukaryotic initiation factor 4E (eIF4E). Plant J. 2002;32:1067–75.
58. Ruffel S, Caranta C, Palloix A, Lefebvre V, Caboche M, Bendahmane A. Structural analysis of the eukaryotic initiation factor 4E gene controlling potyvirus resistance in pepper: exploitation of a BAC library. Gene. 2004;338:209–16.
59. Kang BC, Yeam I, Frantz JD, Murphy JF, Jahn MM. The *pvr1* locus in Capsicum encodes a translation initiation factor eIF4E that interacts with Tobacco etch virus VPg. Plant J. 2005;42:392–405.
60. Charron C, Nicolai M, Gallois JL, Robaglia C, Moury B, Palloix A, Caranta C. Natural variation and functional analyses provide evidence for co-evolution between plant eIF4E and potyviral VPg. Plant J. 2008;54:56–68.
61. Rubio M, Nicolai M, Caranta C, Palloix A. Allele mining in the pepper gene pool provided new complementation effects between *pvr2*-eIF4E and *pvr6*-eIFISO4E alleles for resistance to pepper veinal mottle virus. J Gen Virol. 2009;90:2808–14.
62. Ling KS, Harris KR, Meyer JD, Levi A, Guner N, Wehner TC, Bendahmane A, Havey MJ. Non-synonymous single nucleotide polymorphisms in the watermelon eIF4E gene are closely associated with resistance to zucchini yellow mosaic virus. Theor Appl Genet. 2009;120:191–200.
63. Nieto C, Piron F, Dalmais M, Marco CF, Moriones E, Gomez-Guillamon ML, Truniger V, Gomez P, Garcia-Mas J, Aranda MA, Bendahmane A. EcoTILLING for the identification of allelic variants of melon eIF4E, a factor that controls virus susceptibility. BMC Plant Biol. 2007;7:34. doi:10.1186/1471-2229-7-34.
64. Truniger V, Nieto C, Gonzalez-Ibeas D, Aranda M. Mechanism of plant eIF4E-mediated resistance against a Carmovirus (*Tombusviridae*): Cap-independent translation of a viral RNA controlled in *cis* by an (a)virulence determinant. Plant J. 2008;56:716–27.

65. Kanyuka K, Druka A, Caldwell DG, Tymon A, McCallum N, Waugh R, Adams MJ. Evidence that the recessive bymovirus resistance locus *rym4* in barley corresponds to the eukaryotic translation initiation factor 4E gene. Mol Plant Pathol. 2005;6:449–58.
66. Stein N, Perovic D, Kumlehn J, Pellio B, Stracke S, Streng S, Ordon F, Graner A. The eukaryotic translation initiation factor 4E confers multiallelic recessive Bymovirus resistance in *Hordeum vulgare* (L.). Plant J. 2005;42:912–22.
67. Hofinger BJ, Russell JR, Bass CHG, Baldwin T, dos Reis M, Hedley PE, Li Y, Macaulay M, Waugh R, Hammond-Kosack KE, Kanyuka K. An exceptionally high nucleotide and haplotype diversity and a signature of positive selection for the eIF4E resistance gene in barley are revealed by allele mining and phylogenetic analyses of natural populations. Mol Ecol. 2011;20:3653–68.
68. German-Retana S, Walter J, Doublet B, Roudet-Tavert G, Nicaise V, Lecampion C, Houvenaghel MC, Robaglia C, Michon T, Le Gall O. Mutational analysis of plant cap-binding protein eIF4E reveals key amino acids involved in biochemical functions and potyvirus infection. J Virol. 2008;82:7601–12.
69. Ruffel S, Gallois JL, Lesage ML, Caranta C. The recessive potyvirus resistance gene *pot-1* is the tomato orthologue of the pepper *pvr2*-eIF4E gene. Mol Genet Genomics. 2005;274: 346–53.
70. Naderpour M, Lund OS, Larsen R, Johansen E. Potyviral resistance derived from cultivars of *Phaseolus vulgaris* carrying *bc-3* is associated with the homozygotic presence of a mutated eIF4E allele. Mol Plant Pathol. 2010;11:255–63.
71. Hart JP, Griffiths PD. A series of eIF4E alleles at the Bc-3 locus are associated with recessive resistance to Clover yellow vein virus in common bean. Theor Appl Genet. 2013;126: 2849–63.
72. Bruun-Rasmussen M, Moller IS, Tulinius G, Hansen JK, Lund OS, Johansen IE. The same allele of translation initiation factor 4E mediates resistance against two Potyvirus spp. in *Pisum sativum*. Mol Plant Microbe Interact. 2007;20:1075–82.
73. Andrade M, Abe Y, Nakahara KS, Uyeda I. The *cyv-2* resistance to clover yellow vein virus in pea is controlled by the eukaryotic initiation factor 4E. J Gen Plant Pathol. 2009;75:241–9.
74. Ashby JA, Stevenson CE, Jarvis GE, Lawson DM, Maule AJ. Structure-based mutational analysis of eIF4E in relation to *sbm1* resistance to pea seed-borne mosaic virus in pea. PLoS ONE. 2011;6:e15873.
75. Konecna E, Safarova D, Navratil M, Hanacek P, Coyne C, Flavell A, Vishnyakova M, Ambrose M, Redden R, Smykal P. Geographical gradient of the eIF4E alleles conferring resistance to potyviruses in pea (Pisum) germplasm. PLoS ONE. 2014;9:e90394.
76. Marandel G, Salava J, Abbott A, Candresse T, Decroocq V. Quantitative trait loci meta-analysis of Plum pox virus resistance in apricot (*Prunus armeniaca* L.): new insights on the organization and the identification of genomic resistance factors. Mol. Plant Pathol. 2009;10:347–60.
77. Monzingo AF, Dhaliwal S, Dutt-Chaudhuri A, Lyon A, Sadow JH, Hoffman DW, Robertus JD, Browning KS. The structure of eukaryotic translation initiation factor-4E from wheat reveals a novel disulfide bond. Plant Physiol. 2007;143:1504–18.
78. Michon T, Estevez Y, Walter J, German-Retana S, Le Gall O. The potyviral virus genome-linked protein VPg forms a ternary complex with the eukaryotic initiation factors eIF4E and eIF4G and reduces eIF4E affinity for a mRNA cap analogue. FEBS J. 2006;273:1312–22.
79. Miyoshi H, Suehiro N, Tomoo K, Muto S, Takahashi T, Tsukamoto T, Ohmori T, Natsuaki T. Binding analyses for the interaction between plant virus genome-linked protein (VPg) and plant translational initiation factors. Biochimie. 2006;88:329–40.
80. Roudet-Tavert G, Michon T, Walter J, Delaunay T, Redondo E, Le Gall O. Central domain of a potyvirus VPg is involved in the interaction with the host translation initiation factor eIF4E and the viral protein HcPro. J Gen Virol. 2007;88:1029–33.

81. Contreras-Paredes CA, Silva-Rosales L, Daros JA, Alejandri-Ramirez ND, Dinkova TD. The absence of eukaryotic initiation factor eIFISO4E affects the systemic spread of a Tobacco etch virus isolate in *Arabidopsis thaliana*. Mol Plant Microbe Interact. 2013;26:461–70.
82. Calvo M, Martinez-Turino S, Garcia JA. Resistance to Plum pox virus Strain C in *Arabidopsis thaliana* and *Chenopodium foetidum* involves genome-linked viral protein and other viral determinants and might depend on compatibility with host translation initiation factors. Mol Plant Microbe Interact. 2014;27:1291–301.
83. Mazier M, Flamain F, Nicolai M, Sarnette V, Caranta C. Knock-down of both eIF4E1 and eIF4E2 genes confers broad-spectrum resistance against potyviruses in tomato. PLoS ONE. 2011;6:e29595.
84. Albar L, Bangratz-Reyser M, Hebrard E, Ndjiondjop MN, Jones M, Ghesquiere A. Mutations in the eIFISO4G translation initiation factor confer high resistance of rice to Rice yellow mottle virus. Plant J. 2006;47:417–26.
85. Hebrard E, Poulicard N, Gerard C, Traore O, Wu HC, Albar L, Fargette D, Bessin Y, Vignols F. Direct interaction between the Rice yellow mottle virus (RYMV) VPg and the central domain of the rice eIFISO4G1 factor correlates with rice susceptibility and RYMV virulence. Mol Plant Microbe Interact. 2010;23:1506–13.
86. Leonard S, Plante D, Wittmann S, Daigneault N, Fortin MG, Laliberte JF. Complex formation between potyvirus VPg and translation eukaryotic initiation factor 4E correlates with virus infectivity. J Virol. 2000;74:7730–7.
87. Khan MA, Miyoshi H, Gallie DR, Goss DJ. Potyvirus genome-linked protein, VPg, directly affects wheat germ in vitro translation: Interactions with translation initiation factors eIF4F and eIFiso4F. J Biol Chem. 2008;283:1340–9.
88. Wang A, Krishnaswamy S. Eukaryotic translation initiation factor 4E-mediated recessive resistance to plant viruses and its utility in crop improvement. Mol Plant Pathol. 2012;13:795–803.
89. Beauchemin C, Boutet N, Laliberte JF. Visualisation of the interaction between the precursors of the viral protein linked to the genome (VPg) of Turnip mosaic virus and the translation eukaryotic initiation factor iso 4E in planta. J Virol. 2007;81:775–82.
90. Beauchemin C, Laliberte JF. The poly(A) binding protein is internalized in virus-induced vesicles or redistributed to the nucleolus during Turnip mosaic virus infection. J Virol. 2007;81:10905–13.
91. Thivierge K, Cotton S, Dufresne PJ, Mathieu I, Beauchemin C, Ide C, Fortin MG, Laliberte JF. Eukaryotic elongation factor 1A interacts with Turnip mosaic virus RNA-dependent RNA polymerase and VPg-Pro in virus-induced vesicles. Virology. 2008;377:216–25.
92. Ala-Poikela M, Goytia E, Haikonen T, Rajamaki ML, Valkonen JPT. Helper component proteinase of genus *Potyvirus* is an interaction partner of translation initiation factors eIF(iso) 4E and eIF4E that contains a 4E binding motif. J Virol. 2011;85:6784–94.
93. Estevan J, Marena A, Callot C, Lacombe S, Moretti A, Caranta C, Gallois JL. Specific requirement for translation initiation factor 4E or its isoform drives plant host susceptibility to Tobacco etch virus. BMC Plant Biol. 2014;14:67.
94. Cavatorta JR, Savage AE, Yeam I, Gray SM, Jahn MM. Positive Darwinian selection at single amino acid sites conferring plant virus resistance. J Mol Evol. 2008;67:551–9.
95. Moury B, Charron C, Janzac B, Simon V, Gallois JL, Palloix A, Caranta C. Evolution of plant eukaryotic initiation factor 4E (eIF4E) and potyvirus genome-linked protein (VPg): a game of mirrors impacting resistance spectrum and durability. Infect Genet Evol. 2014;27:472–80.
96. Moury B, Morel C, Johansen E, Guilbaud L, Souche S, Ayme V, Caranta C, Palloix A, Jacquemond M. Mutations in Potato virus Y genome-linked protein determine virulence toward recessive resistances in *Capsicum annuum* and *Lycopersicon hirsutum*. Mol Plant Microbe Interact. 2004;17:322–9.

97. Quenouille J, Saint-Felix L, Moury B, Palloix A. Diversity of genetic backgrounds modulating the durability of a major resistance gene. Analysis of a core collection of pepper landraces resistant to Potato virus Y. Mol. Plant Pathol. 2016;17:296–302.
98. Ayme V, Souche S, Caranta C, Jacquemond M, Chadoeuf J, Palloix A, Moury B. Different mutations in the genome-linked protein VPg of potato virus Y confer virulence on the *pvr2 (3)* resistance in pepper. Mol Plant Microbe Interact. 2006;19:557–63.
99. Ayme V, Petit-Pierre J, Souche S, Palloix A, Moury B. Molecular dissection of the potato virus Y VPg virulence factor reveals complex adaptations to the *pvr2* resistance allelic series in pepper. J Gen Virol. 2007;88:1594–601.
100. Quenouille J, Montarry J, Palloix A, Moury B. Farther, slower, stronger: how the plant genetic background protects a major resistance gene from breakdown. Mol Plant Pathol. 2012;14:109–18.
101. Stamatakis A. RAxML Version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014;30:1312–3.
102. Miller MA, Pfeiffer W, Schwartz T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In: Proceedings of the Gateway Computing Environments Workshop (GCE), 2010;New Orleans, LA:1–8.

Evolution of eIF4E-Interacting Proteins

Greco Hernández, Kathleen M. Gillespie, Tsvetan R. Bachvaroff, Rosemary Jagus, Cátia Igreja, Daniel Peter, Manuel Bulfoni and Bertrand Cosson

1 Introduction

Regulation of gene expression at the translation level is fundamental for many cellular and developmental processes. In eukaryotes, the vast majority of mRNAs are translated in a cap-dependent manner. During the initiation step of this mechanism, a preassembled 43S preinitiation complex (PIC) is targeted to the capped 5′-end of the mRNA through the interaction of the cap-binding protein eIF4E, with the scaffold protein eIF4G, the DEAD-box RNA helicase eIF4A and eIF4B (also eIF4H in mammals) [1–4]. Helicase eIF4A, bound to eIF4G, is thought to expose a single-stranded region in mRNA for interaction with the ribosome. Binding sites in eIF4G for either eIF3 (in mammals) or eIF5 and eIF1 (in yeast) facilitate recruitment of the 43S PIC to eIF4G bound at the cap structure [1, 5]. In many eukaryotes, eIF4G also harbors a binding site for the poly(A)-binding protein (PABP) that, together with an RNA-binding domain in the middle region of mammalian eIF4G, increases the stability of the assembly of eIF4E, eIF4G and eIF4A to the 5′ end of mRNA and promotes circularization of mRNA to improve initiation efficiency in

G. Hernández (✉)
Division of Basic Research, National Institute of Cancer (INCan), 22 San Fernando Avenue, Tlalpan, 14080 Mexico City, Mexico
e-mail: greco.hernandez@gmail.com

K.M. Gillespie · T.R. Bachvaroff · R. Jagus
Institute of Marine and Environmental Technology, University of Maryland Center for Environmental Science, 701 E. Pratt Street, Baltimore, MD 21202, USA

C. Igreja · D. Peter
Department of Biochemistry, Max Planck Institute for Developmental Biology, Spemannstrasse 35, 72076 Tübingen, Germany

M. Bulfoni · B. Cosson
Université Paris Diderot, Sorbonne Paris Cité, Epigenetics and Cell Fate, UMR7216, CNRS, 75013 Paris, France

G. Hernández and R. Jagus (eds.), *Evolution of the Protein Synthesis Machinery and Its Regulation*, DOI 10.1007/978-3-319-39468-8_10

repeated rounds of translation [1–4]. This 'closed-loop' model of translation initiation hypothesizes that mutual interactions of 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. The ribosomal complex thus formed is termed a 48S pre-initiation complex. This complex scans in a 5′ ➔ 3′ direction along the 5′-UTR to reach the start codon, usually an AUG. The interaction eIF4E-eIF4G is one of the most critical events regulating mRNA recruitment and offers a major target for regulation of gene expression [1–4].

In the 1990s, it was discovered that mammalian eIF4E is regulated by three phylogenetically related proteins, termed eIF4E-binding proteins (4E-BPs) 1, 2 and 3 [6–8]. 4E-BPs share with eIF4G the eIF4E-binding motif (4E-BM) $YX_4L\phi$ (where X is any amino acid and ϕ is a hydrophobic residue), which interacts with the dorsal surface of eIF4E. Thus, the 4E-BPs act as competitive inhibitors of the interaction of eIF4E with eIF4G, thereby repressing cap-dependent translation [9–12]. In mammals and *Drosophila*, the 4E-BPs are able to respond to changes in nutrient status and diverse stress stimuli, growth factors and hormones, through signaling via the mTOR pathway downstream of the P(I)3k signaling pathway. The activity of 4E-BPs is controlled by its phosphorylation state: hypophosphorylated 4E-BP shows high affinity for eIF4E; in contrast, phosphorylation results in reduced affinity for eIF4E and prevents the ability of 4E-BP to competitively inhibit eIF4G binding to eIF4E [13–18]. Hypophosphorylated 4E-BP selectively regulates certain classes of mRNAs such as those possessing terminal oligopyrimidine tracts (5′-TOP and 5′-TOP-like) [19]. Following the discovery of mammalian 4E-BPs, additional eIF4E interacting proteins have been identified in diverse organisms (Tables 1 and 2 [20–23]). Since these are not phylogenetically related to the mammalian 4E-BPs, in this chapter they are termed "4E-interacting proteins (4E-IPs)" as has been proposed previously to distinguish them from the 4E-BPs [22]. Recent examples of 4E-IPs also include proteins that enhance translation [24] or initiate mRNA decay [25].

With the rapid advance of genome-wide sequencing of hundreds of species from disparate lineages, it has become apparent that most organisms possess not just the cap binding translational initiation factor eIF4E, but also a family of eIF4Es that serve a range of functions [80–86]. These eIF4E cognates often exhibit differential expression and varying abilities to interact with the cap, eIF4G or 4E-BP/4E-IPs. Thus, diverse 4E-IPs may interact differently with the many eIF4E cognates across eukaryotes. In this chapter, we discuss recent research showing that, in contrast to the monophyletic origin of eIF4Es, 4E-IPs have arisen independently multiple times, often co-opted by preexisting molecules with a wide range of original functions to bind to eIF4E to regulate mRNA recruitment in eukaryotes [22].

1.1 eIF4E Diversification

While the general process of translation is well conserved across all forms of life, in eukaryotes the initiation step and the number of mechanisms to regulate it have

Table 1 4E-IPs with known role in translation

Protein	Binding partner	Organism	Biological process	Reference
4E-BP1 4E-BP2 4E-BP3	eIF4E-1A	Mouse, human, zebrafish	Cell cycle progression, cell growth and proliferation; synaptic plasticity and memory formation	[6–8, 11, 26]
4E-BP/Thor	eIF4E-1	*Drosophila*	Resistance to viral and microbial infections; fat metabolism and response to nutrient starvation	[27, 28]
4E-BP	eIF4E-1	Echinoderm	Embryonic development	[29, 30]
Neuroguidin	eIF4E-1	Mouse	Neurogenesis	[31]
Angel1	eIF4E-1	Human	Endo-/exo-nuclease-phosphatase domain-containing protein; no known biological role	[32]
Brain tumor complex	4E-HP*	*Drosophila*	Embryo patterning and neurogenesis	[33]
p20	eIF4E	*S. cerevisiae*	Pseudohyphal growth under lack of nitrogen	[34–37]
Mextli	eIF4E-1** eIF4E-2** eIF4E-3** eIF4E-4** eIF4E-7**	*Drosophila*	Oogenesis; germ line stem cells maintenance and early embryogenesis; promotes translation	[24]
Leish4E-IP	LeishIF4E-1	*Leishmania*	Promastigote-specific protein	[38]

** Class I eIF4E

undergone a substantial increase in complexity as compared to prokaryotes. Only three initiation factors, IF1, IF2 and IF3, are required to place the small ribosomal subunit at the start codon in eubacteria [87]. Archaea possess a larger number of translation initiation factors: at least six translation initiation factors, which contain up to three subunits [88–90]. In contrast to mRNA recruitment in the prokaryotic life domains, in eukaryotes there is a need for mechanisms to shuttle the transcripts into the cytoplasm and provide for their protection against degradation. This has complicated the process of mRNA recruitment; eukaryotic translation initiation relies on a scanning mechanism to locate the start codon in mRNAs with 5′-protected caps and involves 13 core initiation factors, some of which are large, multimeric complexes [1, 4]. Accordingly, with the exception of eIF5, all the eukaryotic-specific initiation factors, eIF4E, eIF4G, eIF4B, eIF4H (some lineages), eIF3 and poly(A) binding protein, PABP, are involved in the 5′-cap-binding and scanning processes [91–93]. The increase in complexity of the translation initiation step during eukaryotic evolution [22, 88, 89, 92, 94–96] promoted the increase of the number of initiation factors, such as eIF4E, as well as the evolution of new regulatory mechanisms, such as the arousal of a plethora of 4E-IPs in different lineages [22, 91, 92].

eIF4E is a highly conserved protein across eukaryotes. It is defined by the cupped hand structure within which the mRNA cap is bound. This novel fold is

Table 2 4E-IPs involved in additional biological processes independent of translation

Protein	Binding partner	Organism	Biological process	Reference
GEMIN5	eIF4E-1A	Human	WD repeat domain-containing protein; RNP assembly; RNA transport; RNA alternative splicing; apoptosis; motor neuron system development	[39, 40]
PML	eIF4E-1A	Human	Nuclear mRNA export and DNA repair; cell growth and apoptosis	[41, 42]
GYGYF2[a]	4E-HP/eIF4E-2	Human	Glycine-Tyrosine-phenylalanine (GYF) domain-containing protein; insulin metabolism	[43]
CYFIP1	eIF4E-1A	Human	FMRP-interaction factor during neuronal activity; actin polymerization	[44]
LRPPRC	eIF4E-1A	Human	Leucine-rich pentatricopeptide repeat containing protein; Mitochondrial RNA transport and expression; nuclear mRNA metabolism; neurogenesis; mitochondrial unfolded protein response	[45, 46]
PRH[b]	eIF4E-1A	Human	Homeobox transcription factor; hematopoiesis	[47]
HOXA9	eIF4E-1A	Human	Homeobox transcription factor; hematopoiesis.	[48]
Belle/DDX3	eIF4E-1-1A eIF4E-1*	Human; *Drosophila*	DEAD box RNA helicase; Transcription; RNA splicing and transport; development	[49–52]
EMX2[c]	eIF4E-1A	Mouse	Homeobox transcription factor; neurogenesis	[53]
PREP1	4E-HP/eIF4E-2	Mouse	Homeobox transcription factor; embryo development; hematopoietic stem cell biology	[54, 55]
4E-T[d]/Cup	*Drosophila* eIF4E-1; *Xenopus* eIF4E-1B; human eIF4E-1A, -1B, 4E-HP/eIF4E-2	Human; *Drosophila; Xenopus*	Nucleo-cytoplasmic shuttling protein; RNA decay and P-body formation; repression of translation; important for oogenesis and embryogenesis in *Drosophila and Xenopus*	[21, 25, 56–61]
Bicoid	4E-HP*	*Drosophila*	Homeobox transcription factor; embryogenesis	[33]
Ago2–RISC[e] complex	eIF4E-1*	*Drosophila*	RNA interference machinery	[62]
Diap1[f]	eIF4E-1*	*Drosophila*	Inhibitor of apoptosis; epithelial tracheal tube morphogenesis	[63–65]
Eap1[g]	eIF4E	*S. cerevisiae*	Walker A motif-containing protein; genetic stability	[37, 66–68]

(continued)

Table 2 (continued)

Protein	Binding partner	Organism	Biological process	Reference
PGL-1[h]	IFE-1**	*C. elegans*	RGG-box protein; Component of P-granules; germline development.	[69, 70]
SPN-2[i]	IFE-1** IFE-2** IFE-3** IFE-5**	*C. elegans*	Protein with partial similarity to human 4E-T and *Drosophila* Cup; Spindle formation during meiosis	[71]
Maskin	eIF4E-1A	*Xenopus*	Transforming acidic coiled-coil (TACC3)-motif-containing protein; oogenesis; cell division; Mitotic spindle assembly and microtubule growth during mitosis	[72–76]
CPEB[j]	eIF4E-1B	*Xenopus*	Regulates polyadenylation of mRNAs important for oogenesis and neurogenesis	[56]
Z protein	eIF4E-1A	Human arenavirus	Viral life cycle	[42]
VPg[k]	eIF4E*** eIF(iso)4E***	Plant potyviruses, vertebrate caliciviruses, and picornaviruses	Viral life cycle	[77–79]

*Class II eIF4E
**Class I eIF4E (binds mono- and tri-methylated caps
***Class I eIF4Es
[a]Grb10-interacting GYF protein 2;
[b]Proline-rich homeodomain;
[c]Vertebrate homologue of *Drosophila* gene *empty spiracles*;
[d]eIF4E-transporter;
[e]Argonaute2-RNA-induced silencing complex;
[f]*Drosophila* inhibitor of apoptosis protein;
[g]eIF4E-associated protein
[h]P-granule protein;
[i]Spindle orientation defective;
[j]Cytoplasmic polyadenylation element-binding protein;
[k]Viral genome-linked protein

characteristic of the eIF4E family, as exemplified by the prototypical mouse eIF4E [PDB:1L8B] [97, 98]. The mRNA cap-binding region is found within a conserved core of 160–170 amino acids containing eight aromatic residues with conserved spacing [81]. The secondary structure consists of eight antiparallel beta sheets and three major alpha helices [97, 98]. The beta sheets line the binding pocket, and recognition of the 7-methylguanosine moiety is mediated by base sandwich-stacking between conserved aromatic (usually tryptophan) residues [97, 98]. Alpha helix one, containing the recognition motif of S/TVXXW, interacts with eIF4G and 4E-IPs [10]. All eIF4E cognates for which there is a structure or which

have been modeled correspond to this three-dimensional core structure, although mammalian eIF4E-3 seems to recruit additional contacts to offset the decline in binding energies caused by the Trp to Cys substitution [99]. These findings argue for a monophyletic origin of eIF4E [93].

Expansion of the eIF4E family has occurred across the eukaryotic domain, from excavates (such as the trypanosomes) and alveolates (such as the dinoflagellates) to the different multicellular lineages. Multiple eIF4E family members (between 1 and 15) have been identified in a wide range of organisms that include plants, flies, mammals, frogs, fish, birds, nematodes and various protist lineages [80–84, 86, 100]. It seems likely that the wide eukaryotic radiation into many ecological niches, whether of unicellular protists or multicellular organisms, as well as the development of multicellular organisms from a single cell, and the physiological specialization of cells into different cell types and tissues in multicellular organisms, represented the driving force for expansion of the eIF4E family. Thus, a single early eIF4E gene underwent a series of gene duplications throughout evolution, generating multiple structural classes and in some cases subclasses. Some eIF4Es support general translational initiation. Others promote or inhibit translation of specific mRNAs or classes of mRNAs. Some are involved in translation in response to stress; others show variable biochemical properties or affinities for different cap structures, or may not be involved in translation at all [80].

It has been noted that the nomenclature for eIF4E family members has evolved with confusion [99]. One proposed classification has divided the different members of the eIF4E family from metazoans and fungi into structural classes: class I, class II and class III, to give eIF4E-1, eIF4E-2 and eIF4E-3 and subclasses of these [81]. Three structural classes of eIF4E have been recognized in plants, fungi and metazoans, most easily apparent by variations in the residues equivalent to Trp-43 and Trp-56 (as per the numbering in human eIF4E-1) [81]. Class I members contain both Trp residues; class II members (also named 4E-HP) contain Tyr, Phe or Leu at the first position and Tyr or Phe at the second position; class III proteins contain Trp at the first position and Cys or Tyr at the second position. Subclasses of eIF4Es in each class can be given additional letter designations, after the class designation, but these would only be equivalent within taxa. 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* [101, 102] or eIF4Es that fulfill regulatory functions such as the vertebrate eIF4E-1Bs [26, 56, 103] and the class I eIF4E-3 of *Drosophila* [104].

Phylogenies of the eIF4E family are in general poorly resolved, but strongly support monophyly of metazoan and fungal classes I–III. Increasingly, this nomenclature is in use, but has not been universally adopted by investigators accustomed to an earlier nomenclature or because the gene name is in a public database and its adoption would be confusing in the respective field. *Drosophila* has seven different cognates of class I eIF4Es, termed eIF4E-1 to eIF4E-7 [84, 105] in FlyBase. The class II eIF4E of *Drosophila* was recognized to be related to contemporaneously discovered class II eIF4Es in vertebrates and plants termed 4EHP (for eIF4E homologous protein [106]) and nCBP (for novel cap-binding

protein [107]), respectively. Therefore this *Drosophila* eIF4E was given the same non-systematic name, 4EHP or eIF4E-8 [105, 108]. In *C. elegans*, eIF4E family members have been termed IFE-1, -2, -3, -4 and -5 (with ife standing for initiation factor of elegans) [102] in WormBase. Four of the five *C. elegans* eIF4E family members, IFE-1, -2 and -3 and -5, are class I eIF4Es. IFE-3 corresponds to mammalian eIF4E-1 and binds to monomethylated cap structures [102]. IFE-1, -2 and -5 are also class I members but interact with monomethylated and the trimethylated cap structures found in *trans*spliced mRNAs. Here we will keep the FlyBase and WormBase nomenclature, indicating the class each eIF4E belongs to as appropriate in the text and tables. Any deuterostome eIF4E will be given its systematic name.

1.2 The Ancestral 4E-BP

4E-BPs are small, acidic, heat-stable proteins first described in vertebrates as a three-member protein family, 4E-BP1, 4E-BP2 and 4E-BP3, that share 55 % amino acid identity [12, 109]. 4E-BPs are involved in disparate cellular processes, including cell cycle progression, cell growth and proliferation, longevity, synaptic plasticity and fat metabolism, resistance to nutrient starvation and oxidative stress, and the responses against viral and microbial infections [14–18]. Thus, 4E-BPs have been proposed to act as metabolic "brakes" that can rapidly shut down cap-dependent translation in response to various challenges [17, 110, 111].

In mammalian 4E-BP1 and 4E-BP3, the 4E-BM is YDRKFLM; in 4E-BP2, the 4E-BM is YDRKFLL [7]. The consensus 4E-BM is Y(D/E)RXFL(L/M) in all 4E-BPs except basidiomycetes, in which it is YSRDXLL. Similarly, the non-canonical 4E-BM [112–116] is recognizable in all 4E-BPs except those from basidiomycetes. In addition to the 4E-BM, each 4E-BP harbors two canonical TOR phosphorylation sites, TPGGT and several *trans*regulatory phosphorylation sites. Once activated, mTORC1 phosphorylates Thr37 and Thr46 in human 4E-BP1, which are priming sites for subsequent phosphorylation at Ser65 and Thr70 [17, 109, 117, 118]. A comparison of 4E-BPs across the eukaryotes shows that the second TOR phosphorylation site (TPGGT) is absolutely conserved. The first TOR phosphorylation site (TPGGT) is conserved in all 4E-BPs found except for those from basidiomycetes in which it is missing. The residues equivalent to Ser65 and Thr70 are universally conserved. Whether TOR phosphorylates 4E-BP in other, so-far uninvestigated species, such as Basidiomycetes or different protists, remains an open and exiting question.

Figures 1 and 2 show a phylogenetic analysis of the 4E-BPs across the eukaryotes. 4E-BP is widely distributed in metazoan lineages, including sponges, Placozoa, *Hydra*, jellyfish and sea anemones, but is absent in nematodes. In general,

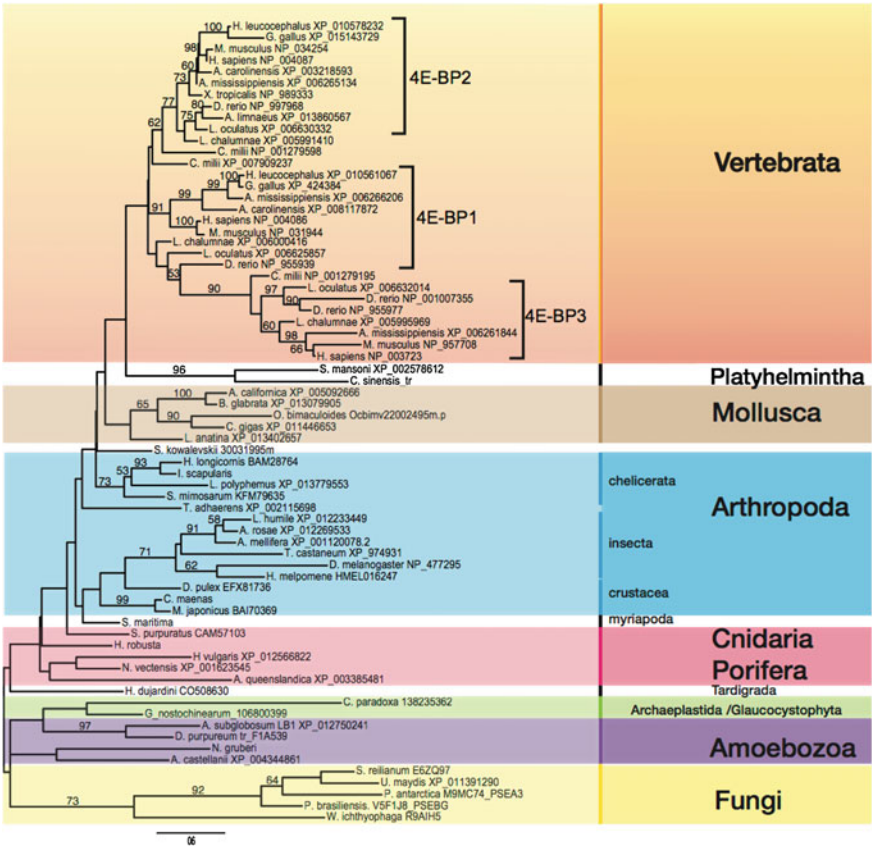


Fig. 1 Phylogenetic analysis of 4E-BPs. A maximum likelihood phylogeny of the 4E-BP family. The tree was calculated using RAxML with the Jones Taylor Thornton amino acid substitution matrix with gamma rate correction. There were 70 sequences in the alignment of 16 sites. Bootstrap proportions >50 % of 100 replicates are shown above branches. The individual species are given only a single letter for genus followed by the species name

the 4E-BPs cluster according to the phylogenetic relationships of the species; vertebrate 4E-BPs cluster together; mollusk 4E-BPs cluster together. Tick, spider and Chelicerata 4E-BPs cluster together. 4E-BPs from insects cluster together, close to those from Crustacea. 4E-BPs from *Hydra*, sea anemone and sponge cluster together; those from Amoebozoa cluster together, as do 4E-BPs from basidiomycetes. The 4E-BPs from basidiomycetes are the most derived. Multiple cognates of 4E-BPs are only found in vertebrates in which they mirror to some extent the vertebrate expansion of eIF4Es (Gillespie et al., this book). 4E-BP has not been found so far in choanoflagellates, a group of free-living, unicellular and colonial flagellate eukaryotes considered to be the closest living relatives of the metazoans. Choanoflagellates are part of the Supergroup referred to as Unikonta [119],

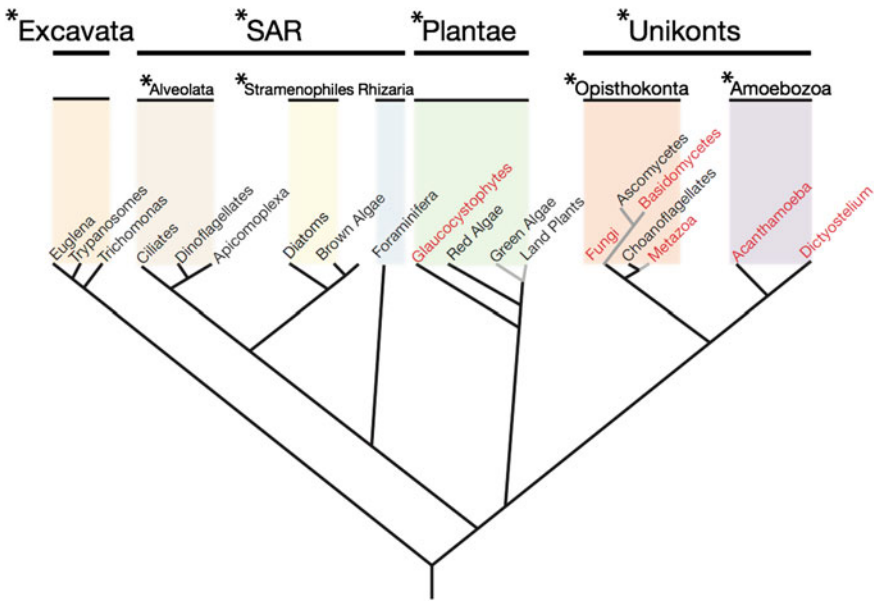


Fig. 2 Distribution of 4E-BP and neuroguidin in eukaryotic lineages. The distribution of 4E-BPs and neuroguidins in the eukaryotic superfamilies are shown based on the sequences found. Seven major lineages are shown grouped into five supergroups. Multicellular lineages are shown as *gray branches*. Lineages with 4E-BP are shown in *red*. *Lineages in which neuroguidin is found are marked with an *asterisk*

Amorphea [120] or Podiata [121], composed of the Opisthokonts (Metazoa, Fungi, Choanoflagellata) and Amoebozoa, which are all thought to have arisen from a common protist ancestor. Among the protist lineages, the distribution of 4E-BP is limited. Although it is found in the Amoebozoa and Glaucocystophyta, it is not found in the major protist lineages Alveolata, Stramenopiles, Rhizaria and Excavata (for example, *Trypanosoma* and *Euglena*). 4E-BP is also not found in the Jakobids, which contain the most deeply rooted eukaryotes such as *Reclinomonas*. In fact, with the exception of the glaucocystophytes, 4E-BP is only found in the eukaryotic Supergroup Unikonta/Amorphea/Podiata. The finding of 4E-BP in glaucocystophytes was a surprise, since they lie outside this group and are generally considered to represent a form close to the original algal type that gave rise to green algae, red algae and land plants. 4E-BP is not found in plants and has a restricted distribution in fungi, being found only in basidiomycetes as well as some glomeromycetes and zygomycetes. If the glaucocystophytes are basal to the green algae, red algae and plants, then they have been lost in these lineages. Figure 2 shows the distribution of 4E-BP across all eukaryotes. This distribution does not support the idea that a single ancestral 4E-BP gene might have emerged along with eIF4E in the last eukaryotic common ancestor (LECA).

1.3 *Molecular Tinkering: A Recurrent Phenomenon for the Emergence of Novel 4E-IPs*

Besides 4E-BPs, an increasing number of proteins that interact with eIF4E have been discovered in distantly related species (Tables 1 and 2; [20–23]). They control translation in disparate biological processes such as development, response to different stresses or neurogenesis, sometimes in an mRNA-specific manner [12, 20–23]. Most 4E-IPs are phylogenetically unrelated to each other or to the 4E-BPs. Surprisingly, whereas some have a unique role in translation (Table 1), most 4E-IPs play other or additional roles in different cellular processes such as RNP assembly, RNA transport, alternate splicing or mRNA turnover. Many of them are homeodomain proteins (Table 2) [22, 122]. Unique lineage-specific 4E-IPs have evolved independently in some taxonomic groups such as p20 [34] and Eap1p [66] in yeast, Mextli in higher dipteran and some nematodes [24], and SPN-2 in *C. elegans* [71]. Most 4E-IPs use the consensus 4E-BM YXXXXLϕ, or a similar one, for eIF4E interaction. However, human promyelocytic leukemia (PML) [41], cytoplasmic fragile X mental retardation protein (FMRP)-interacting protein (CYFIP1) [44] and arenavirus Z protein [42] do not contain this motif. Instead, PML and Z protein bind eIF4E via a RING motif [42] and CYFIP1 uses a peptide with a tertiary structure that mimics the consensus 4E-BM [44].

Considering that (1) the ranges of roles that many 4E-IPs play are unrelated to translation, (2) the lack of a phylogenetic relationship to each other and to the 4E-BPs, (3) the identification of some 4E-IPs only in specific lineages, (4) the variability in the 4E-BM used by different 4E-IPs and (5) the wide spectrum of protein architecture, it seems likely that the ability to bind eIF4E evolved independently multiple times. During evolution, genes are often subject to duplication events that can affect single genes, a stretch of several genes, whole chromosomes or even whole genomes (WGD). After duplication events, particularly after WGD, there is extensive gene loss and genomic rearrangements. Duplicate gene pairs can undergo different fates; one can be lost or change to provide a new function as has happened with the eIF4Es (Gillespie et al., this book). These processes are likely to account for the plethora of 4E-IPs. This phenomenon was famously anticipated by Francois Monod [123] who referred to it as "molecular tinkering" in the pre-genomic era. Overall, *the emerging view is that the ability to bind eIF4E has been a recurrent phenomenon that has facilitated control of mRNA translation in different tissues, developmental stages or conditions, throughout eukaryotic evolution.* 4E-IPs have been found in all major eukaryotic lineages although none has been identified so far in plants. The notion that different 4E-IPs have independently emerged either independently or by co-option of the 4E-BM by existing proteins is illustrated in this section by the following five different examples of unique eIF4E-4E-IP interactions that may have evolved only in specific lineages.

(1) **Bicoid**: The transcriptional factor Bicoid (Bcd) is a Hox protein that interacts with class II eIF4E (4EHP/eIF4E-2), but not with eIF4E-1, during *Drosophila* embryogenesis and oogenesis [33, 49, 124]. *Drosophila* 4EHP/eIF4E-2 is not able

to bind eIF4G [105]; hence, it represses translation mRNAs [33, 49, 124]. During early embryogenesis, binding of Bcd to the 3′-UTR of *caudal* mRNA recruits 4EHP, which associates with the cap structure, thereby blocking *caudal* translation. Although *caudal* mRNA is evenly distributed all over the embryo, Bcd is anteriorly localized, so 4EHP/eIF4E-2 causes specific repression of *caudal* mRNA translation to give a posterior distribution of Caudal [124].

Hox proteins are transcription factors that specify morphological identities and embryo patterning along the anterior-posterior axis across metazoan [125]. The sister genes *bcd* and *zerknüllt* (*zen*) are Hox genes that emerged from a duplication of the *Hox3* gene when higher dipteran insects originated [126–128]. After this duplication, *zen* and *bcd* retained their original function in embryo patterning but *bcd* gained the novel features that gave rise to a protein with the ability to bind 4EHP/eIF4E-2 for translational inhibition of *caudal* mRNA. Thus, Bcd and its function in embryo patterning are unique to close relatives of *Drosophila* and are absent in all other insects [126–128]. Accordingly, non-dipteran insects evolved a different mean to block Caudal accumulation in the anterior pole that is independent from a Bcd-4E-HP interaction, namely through localization of *caudal* mRNA in the posterior pole [127, 129]. Moreover, the 4EHP/eIF4E-2-binding motif of Bcd [124] is absent in Zen and Hox3 [128, 130]. Thus, the Bcd-4EHP/eIF4E-2 interaction is a feature that evolved only in higher dipterans by molecular tinkering, i.e., by transforming Bcd from the transcription machinery into a translational regulator to play a novel, different role. In another study, it was found that translation of *hunchback* mRNA is regulated by the same 4EHP/eIF4E-2, but in this case, the eIF4E-binding partner is Brat [33].

In a related example, from mouse oocytes, 4EHP/eIF4E-2 co-localizes with Prep1, a homeodomain transcription factor, which contains an eIF4E-binding motif [54]. The Prep1-4EHP/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. In order to determine the role of eIF4E-2 in mouse, a 4EHP/eIF4E-2 knockout was created that leads to increased translation but perinatal lethality in mice [43]. Overexpression of Hoxb4 in mouse zygotes in vitro resulted in the slowing of development. Since homeodomain proteins are widely distributed in angiosperms, fungi and metazoa, including the early branching metazoan phyla, Cnidaria, Placozoa and Porifera, more instances of translational regulation through 4E-BM-containing homeodomain proteins seem likely.

(2) **Mextli**: Mextli (Mxt) is a *Drosophila* 4E-IP that binds some eIF4E paralogs, promotes translation, and plays a role during oogenesis and early embryogenesis. Mextli is present in all Drosophilidae species and in *C. elegans*, but has no counterpart in other eukaryotes [24]. Comparison of the Mxt primary sequence and its predicted secondary structure strongly suggests that Mxt might have evolved from ancestral subtilase-like proteases containing KH and HEAT domains. Although these proteases are present across all eukaryotic phyla, the 4E-BM was acquired only two times (in Drosophilidae and Nematoda), most likely by convergent evolution.

(3) **Leish4E-IP**: In the lineage of parasitic protists *Leishmania spp*, a 4E-IP termed Leish4E-IP has been found that interacts only with LeishIF4E-1 among the several eIF4E cognates in the promastigote life stage [38]. Although it is 985 amino acids, secondary structure prediction suggests that the protein is mostly unstructured with a small-coiled coil region between amino acids 187–217. It has a YTREELL motif close to the N terminal. Although LeishIF4E1 has been reported not to play a role in translation [131], Leish4E-1P is found associated with Leish4E-1 in promastigotes. Leish4E-IP is conserved in Trypanosomatidae species, but does not have a counterpart in other eukaryote lineages and may represent a novel 4E-IP.

(4) **Angel1**: The CCR4-Not complex is a central regulator of mRNA metabolism. Angel (also called CCR4d/e) belongs to the family of the CCR4-like proteins that all possess a conserved deadenylase domain. Interestingly, in vertebrates, the Angel1 paralog acquired a 4E-BM. An ancestral gene coding for Angel, but lacking the 4E-BM, is present in the chordates *Branchiostoma floridae and Ciona intestinalis,* representatives of the two primary deuterostome subphyla *Cephalochordates and Urochordates.* Early during vertebrate divergence, the ancestral gene coding for Angel gave rise to two paralogs coding for Angel1 and Angel2 from a duplication of the *Angel* locus [132]. An extra exon coding for the 4E-BM appeared in the *Angel1* gene that has no equivalent in the *Angel2* gene [32]. This motif was acquired simultaneously or immediately after the duplication of the ancestral gene *Angel* and has been conserved throughout the evolution of vertebrates. Thus, the eIF4E-Angel1 interaction is a vertebrate-specific feature that appeared by neofunctionalization of a duplicated copy of the *Angel1* gene through the insertion of an extra exon.

(5) **Neuroguidin**: Neuroguidin (Ngd) was first detected in the mouse central nervous system in association with CPEB and is able to promote neural development by regulating the translation of CPE-containing mRNAs [31]. Ngd is also found in the embryonic nervous system and neural crest of *Xenopus* embryos. Ngd belongs to a protein family that includes Utp3 and LCP5, which are components of the U3 RNP. It also includes the human C1D protein and *Saccharomyces cerevisiae* YHR081 W (rrp47), an exosome-associated protein required for the 3′ processing of stable RNAs, and Sas10, which has been identified as a regulator of chromatin silencing. Ngd contains three 4E-BMs contained in a disordered region. In addition to binding to eIF4E, Ngd also interacts with CPEB and inhibits mRNA translation in a CPE-dependent manner when injected into *Xenopus* oocytes. However, whether Ngd orthologs from different *phyla* bind eIF4E is not know.

Surprisingly, little work has been done on neuroguidin outside of *Xenopus* and *Drosophila*. Notwithstanding, it is currently the only identified 4E-IP that is represented in all eukaryotic lineages. It is found in all deuterostomes, all metazoan phyla, sponges, placozoa, jellyfish and sea anemones, *Arabidopsis* and rice, but also in the major protist lineages, namely excavates (trypanosomes and *Naeglaria*) and alveolates (dinoflagellates and apicomplexans). Figure 2 shows the distribution of 4E-BP and Ngd across the eukaryotic tree of life. Thus, Ngd might be the 4E-IP candidate that arose with eIF4E in early eukaryotes, but that has been superseded in

significance as additional 4E-IPs evolved. It would be worth experimentally testing whether Ngd interacts with eIF4E in different lineages.

1.4 Conservation and Plasticity of the Molecular Architecture of 4E-IP and EIF4E Interaction Across Eukaryotes

The function of eIF4E in translation initiation depends on the binding of mRNA and proteins to its conserved and specific surfaces (Fig. 3). The cap-binding cavity of eIF4E resides in the concave and ventral surface of the protein formed by the strongly bent β-sheet containing the eight antiparallel β-strands [97, 98]. At the convex and dorsal surface, composed by the three α-helices, a conserved patch of hydrophobic residues interacts with the canonical eIF4E-binding motif, 4E-BM, present in eIF4G and 4E-IPs (Fig. 3, top) [9, 10, 97, 98]. A second and conserved hydrophobic cavity on the lateral surface of eIF4E is also of crucial importance for the association of 4E-IPs with eIF4E. At this surface, eIF4E is contacted by sequences of 4E-IPs located usually 15–30 residues after the canonical motif. In contrast to the canonical motifs, these sequences differ highly among all proteins and might reflect the absence of phylogenetic relation among eIF4E partners; thus, these were termed as non-canonical 4E-BMs (Fig. 3, bottom). These motifs contribute to the binding of 4E-IPs to eIF4E and are absolutely required for their ability to compete with eIF4G [112–116, 133, 135, 136].

Despite the diversity in function, sequence and origin, recent studies have revealed that *4E-IPs have common binding principles when they are in complex with eIF4E* (Fig. 3). All exhibit a bipartite binding mode defined by three common structural elements: (1) a canonical motif docked at the dorsal surface of eIF4E, (2) an elbow loop following the canonical α-helix that the bends the protein backbone 90 degrees downward and (3) a non-canonical motif that addresses the conserved lateral surface of eIF4E (Fig. 4b) [112, 113, 133]. These eIF4E-binding regions are usually located within intrinsically disordered regions (IDRs) or present in small intrinsically disordered proteins (IDPs), such as 4E-BPs, i.e., protein sequences that are mainly unstructured in the absence of a binding partner and undergo a disorder-to-order transition upon binding to it [113, 138–142]. Inside these IDRs, the 4E-BMs represent short sequence elements (also called short linear motifs or SLiMs) that sample alternative conformations and become ordered upon binding to eIF4E. Single motif binding is usually weak and transient, but as in the case of 4E-IPs, flanking motifs act in a concerted and cooperative manner, providing increased affinity and specificity and enabling the reversibility of the binding mechanism [114–116, 143]. These disordered regions evolved rapidly and have the capability to adapt to new demands as they provided an evolutionary neutral platform in which considerable variation can be accommodated. As such, new functions can be gained and different proteins can be rewired to novel pathways

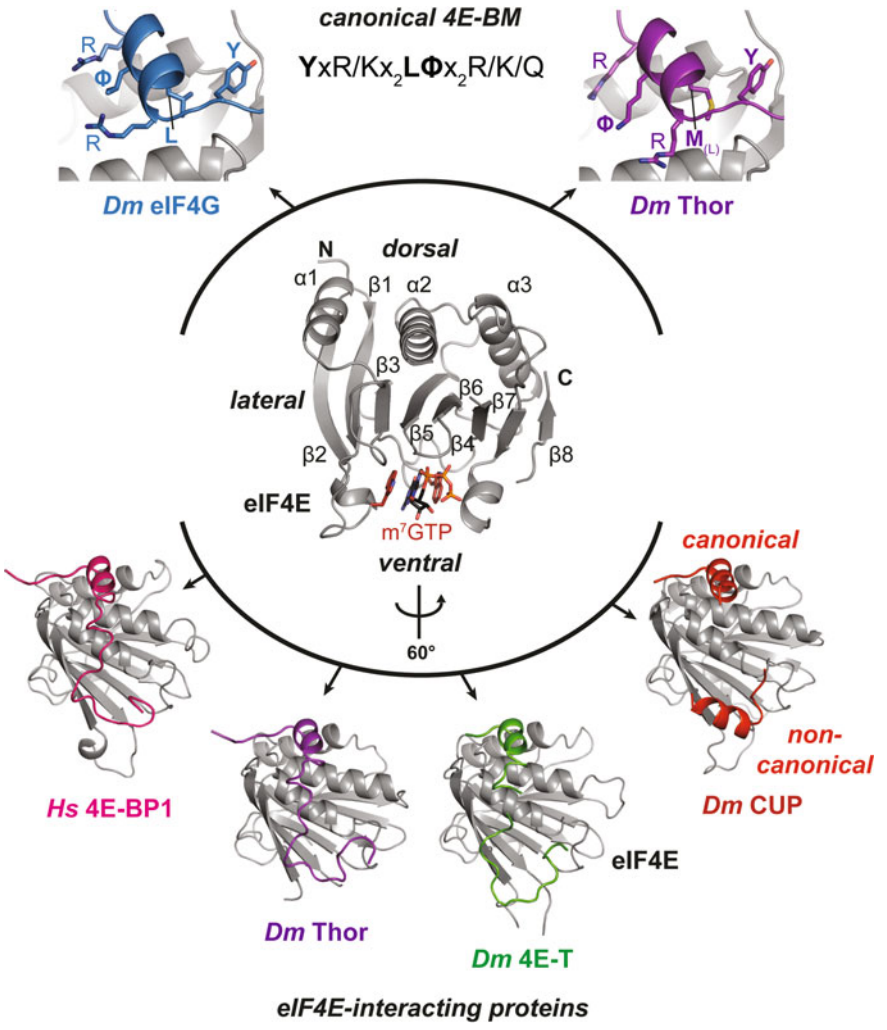


Fig. 3 Molecular architecture of 4E-IPs in complex with eIF4E. (*Top*) View of the canonical eIF4E-binding motifs bound to the dorsal surface of eIF4E. Close-up views of the canonical helix of *Drosophila melanogaster* (*Dm*) eIF4G (*blue*) and *Dm* Thor (*purple*) in complex with eIF4E. The consensus residues involved in the interaction with eIF4E are shown as *sticks*. The Arg residues at positions 2 and 9 of the motif cover a hydrophobic patch at the dorsal surface of eIF4E. The Leu residue in the canonical helix was replaced in *Dm* Thor by a Met. (*Middle*) Surfaces and structure of cap-bound eIF4E. The secondary structural elements (α-helices and β-strands) are labeled in *black*. The bound m^7GTP cap is shown in *sticks* and labeled in *red*, and the tryptophan residues involved in cap binding are highlighted in *red*. N (N terminus), C (C terminus). Protein databank (PDB): 4UEC [112]. (*Bottom*) Structures of different 4E-IPs bound to eIF4E. These structural overviews highlight the common bipartite arrangement that 4E-IPs adopt upon binding to eIF4E. The 4E-IPs are as follows: *Hs* 4E-BP1 (PDB: 4UED and 5BXV), *Dm* Thor (PDB: 4UE8), *Dm* 4E-Transporter (4E-T, PDB: 4UE9) and *Dm* CUP (PDB: 4AXG) [112, 133, 134]. eIF4E is colored in *gray* and the 4E-IPs peptides in *pink* (4E-BP1), *purple* (Thor), *dark green* (4E-T) or *red* (CUP)

[141, 143]. The diversity of 4E-IPs, their involvement in distinct cellular pathways, their different molecular functions and the fact that they cover a whole bandwidth of protein complexity suggest that the IDRs of the different 4E-IPs independently evolved the ability to interact with eIF4E.

Since the IDRs of 4E-IPs are engaged in the interaction with eIF4E, their interfaces offer new opportunities to develop structure-guided small molecules. Such molecules would bind to eIF4E and mimic the function of 4E-IPs halting translation in a multitude of biological contexts. To date three small molecules have been described to inhibit the eIF4E-eIF4G interaction. Interestingly, one of these molecules, the 4EGI-1 inhibitor, binds like the non-canonical motif of 4E-IPs to the lateral surface of eIF4E and displaces eIF4G from the dorsal surface by an allosteric mechanism while it promotes binding of 4E-BP1 [133, 144, 145]. The two other molecules, 4E1RCat and 4E2RCat, block the association of eIF4E with eIF4G and 4E-BPs, but their inhibitory mechanism is currently unknown [146, 147]. Despite their anti-cancer activity in different cell culture systems and animal tumor models [145, 147, 148], the selectivity and potency of these translational inhibitors require further improvement and highlight the need to develop structure-based drugs.

1.4.1 The Canonical 4E-BM and the Dorsal Surface of EIF4E

The canonical 4E-BM is the best-conserved feature among 4E-IPs. Its consensus sequence, $YX_4L\Phi$, folds into an α-helix upon recognition of invariant residues on the dorsal surface of eIF4E. In detail, the LΦ residues hold the position of the canonical motif through hydrophobic contacts with conserved Val and Trp residues in helix α1 of eIF4E. Moreover, the His-Pro-Leu motif present at the N terminus of strand β1 of eIF4E proteins is in contact with the Tyr residue of the canonical motifs [10, 112, 134, 149]. In the majority of the proteins, the canonical motif is also flanked by conserved Arg/Lys/Gln residues that shield the hydrophobic patches on the dorsal surface of eIF4E from solvent exposure [10, 112]. These residues adopt a similar arrangement in all structures of 4E-IPs bound to eIF4E available so far and suggest that the canonical motif is better defined by the following sequence YX[R/K]X_2LΦX_2[R/K/Q] (Fig. 3, top). Interestingly, exceptions to this extended motif lead to a significantly different mode of binding of the 4E-IP, causing structural variability and altered properties, and will be discussed below.

1.4.2 The Non-canonical 4E-BMs and the Lateral Surface of eIF4E

4E-IPs utilize non-conserved regions contiguous to the canonical 4E-BMs to extend their interaction with eIF4E. These sequences include variable linker regions and versatile non-canonical 4E-BMs that associate with the lateral binding site on eIF4E [112, 115, 133] (Fig. 4, bottom). Although not conserved at the sequence level, the non-canonical motifs are enriched in hydrophobic residues and use similar mechanisms to contact eIF4E [112, 113]. In essence, the non-canonical motifs adopt to a

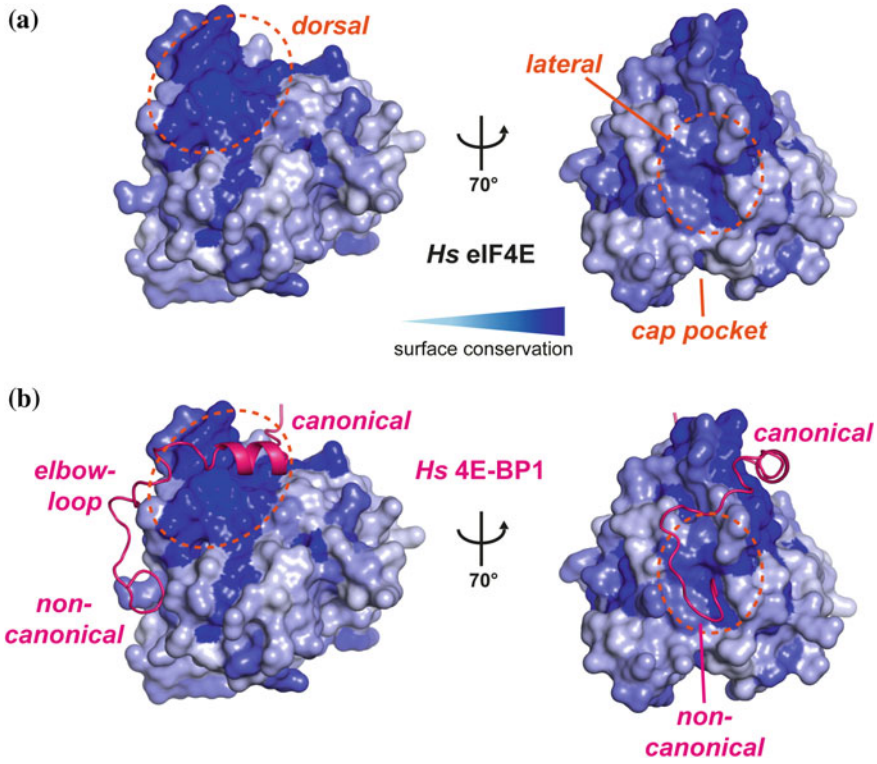


Fig. 4 Representation of sequence conservation at the dorsal and lateral surfaces of eIF4E. **a**, **b** eIF4E is represented as a surface and 4E-BP1 as a *pink cartoon*. The surface of eIF4E is colored from *light* to *dark blue* according to sequence conservation obtained using the Protskin software with an alignment of eIF4E sequences from yeast to human [137]. eIF4E is shown in two orientations to highlight the differences on sequence conservation between the dorsal and lateral surfaces of eIF4E

small hydrophobic pocket of eIF4E, lined by conserved Ile and aromatic amino acids (Phe in *Hs* eIF4E, Tyr in *Dm* and *Ce* eIF4E), using different arrangements (unstructured vs. helical) and residues (smaller hydrophobic residues vs. aromatic residues) to maintain the major interactions (Fig. 5) [112, 113].

The variety of the molecular arrangements employed by the non-canonical motifs to adapt to eIF4E is in striking contrast to the conserved binding mode of the canonical motifs and suggests that the lateral surface of eIF4E imposes lower molecular constraints than the dorsal surface onto 4E-BPs. Such molecular constraints are associated with the degree of sequence conservation on the different surfaces of eIF4E (Fig. 4) and might have represented the evolutionary pressure that determined the way the 4E-BMs interact with eIF4E. On the one hand, a highly similar molecular signature present in the canonical motifs was optimized to bind to the strictly conserved and dorsal surface of eIF4E. Such a binding mode provided

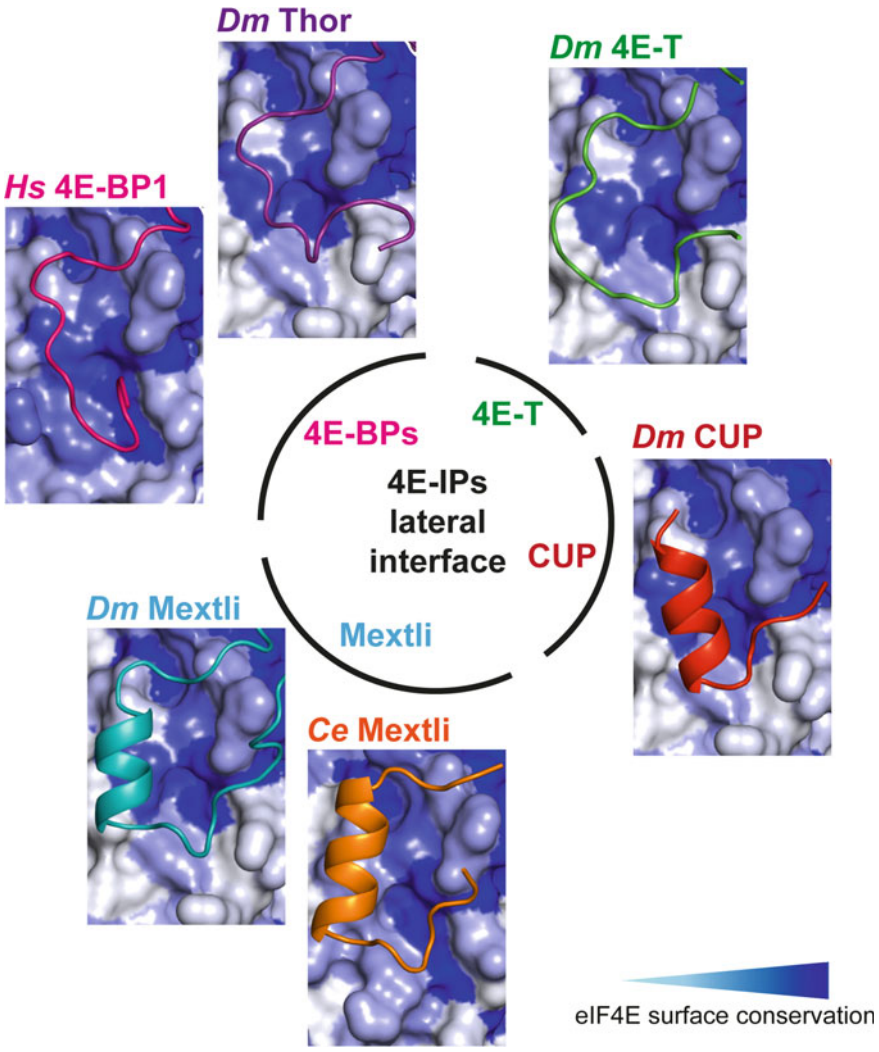


Fig. 5 The multiple molecular arrangements of the non-canonical motifs of 4E-IPs at the lateral surface of eIF4E. Non-canonical 4E-BMs dock at the lateral surface of eIF4E and get ordered into loops or helical structures. In this figure eIF4E is represented as a surface color coded as described in Fig. 2. The non-canonical motifs of 4E-IPs are depicted as a cartoon. The 4E-IPs are as follow: *Hs* 4E-BP1, *Dm* Thor, *Dm* 4E-T, *Dm* CUP, *Dm* Mextli and *Ce* Mextli

high specificity toward eIF4E. On the other hand, variable non-canonical sequences co-evolved in different proteins to associate with the lateral and less conserved surface of eIF4E (Fig. 5). A binding site with lower complexity might have emerged to recruit and select potential eIF4E partners that later in evolution adapted to the higher molecular constraints present on the dorsal surface of eIF4E through

the acquisition of a canonical motif. Alternatively, the additional interactions provided by the non-canonical motifs evolved to increase the affinity of 4E-IPs to eIF4E and provide novel properties to the different eIF4E complexes during the regulation of translation. Independently of the evolutionary scenario, the presence of canonical and non-canonical motifs in 4E-IPs can be explained by the rapid evolution of IDRs toward 4E-BMs.

1.4.3 Plasticity of the eIF4E-Binding Mode: Bipartite Versus Tripartite 4E-IPs

Although variations in the arrangement of the non-canonical motifs occurred frequently, most of the eIF4E partners have evolved a bipartite mode of binding. Yet, deviations of this evolutionary conserved binding strategy exist in nature and resulted, for instance, from minor changes in the consensus sequence of the canonical motifs. Strikingly, these changes lead to an even higher plasticity and complexity in the overall interaction mode with eIF4E. The *Drosophila* Mextli protein is a remarkable example of the molecular diversity of 4E-IPs. In contrast to other 4E-IPs, Mxt is reported to promote translation, as it not only binds several eIF4Es but also recruits eIF3 components to the mRNA, ultimately driving translation initiation [24]. Mxt is an invertebrate specific protein, and in *Drosophila* it regulates germ stem cell maintenance and early embryogenesis [24]. Moreover, Mxt is also exceptional in its way of binding to eIF4E (Fig. 6a). Recent structural insights have revealed that the classical bipartite molecular architecture of 4E-IPs is extended in Mxt by an additional linker and an auxiliary 4E-BM, which lead back and address again the dorsal surface of eIF4E (Fig. 6a). Such a unique tripartite binding mode is only possible because of specific changes on the sequence of canonical 4E-BM of Mxt, namely, the long aliphatic Arg/Lys/Gln residues that flank the canonical helix in the majority of 4E-IPs were replaced by shorter amino acids (Ile and Ser, Fig. 6c). Due to such modifications, an auxiliary helical motif can accommodate on the dorsal surface of eIF4E, antiparallel to the canonical helix, replacing the interactions and covering the surface usually occupied by the side chains of the aliphatic residues present in the majority of 4E-IPs (Fig. 6c) [113]. Interestingly, this tripartite binding mode seems to be only conserved in dipteran and some non-dipteran insects, as the sequences of the auxiliary linker and helix are more divergent or absent in Mxt proteins from other non-dipteran species, arthropods and nematodes. Curiously, in Mxt homologs lacking the auxiliary sequences, like *C. elegans* (*Ce*) Mxt, the canonical motif contains Arg residues at positions 2 and 9 (Fig. 6c). Accordingly, the structure of the *Ce* Mxt-eIF4E complex shows that in nematodes Mxt arranges in a classical bipartite manner and lacks the auxiliary sequences of the fruit fly protein. As such, the replacement of the long aliphatic residues within the canonical motif co-evolved with the acquisition of an auxiliary binding region (Fig. 6).

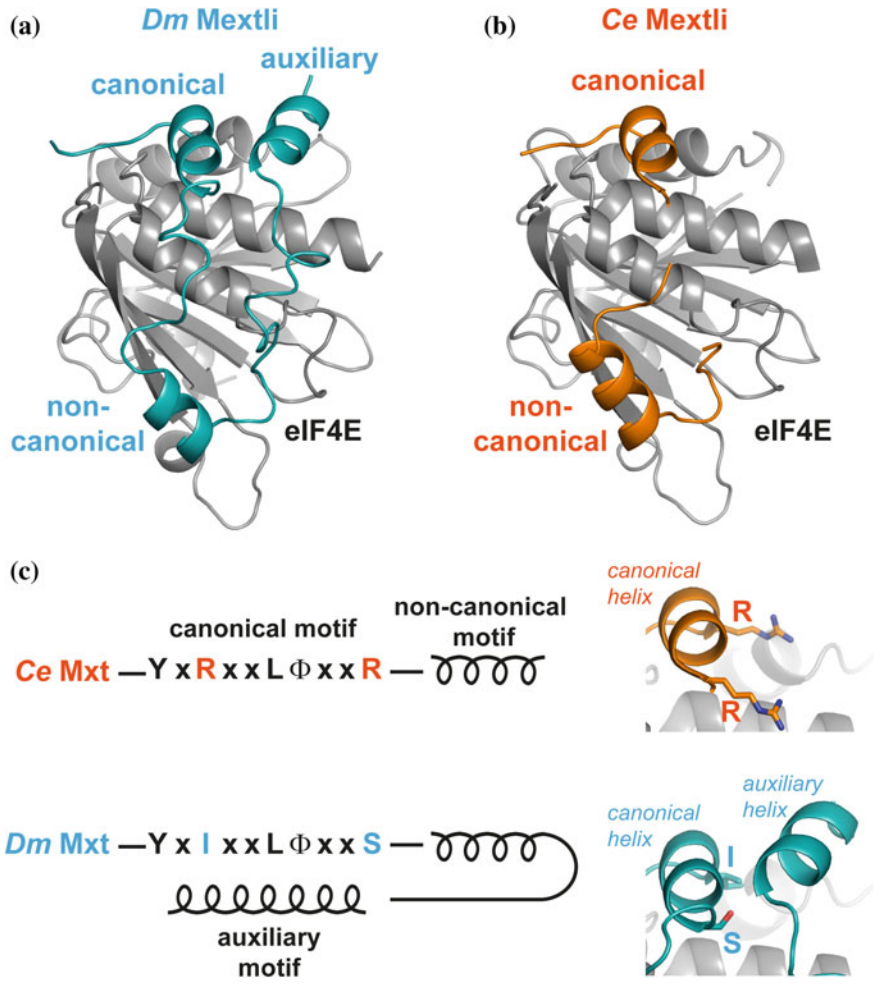


Fig. 6 The tripartite and bipartite binding modes of Mextli proteins. **a**, **b** Overview of the structures of *Dm* (**a**) and *Ce* (**b**) Mxt proteins bound to eIF4E. PDBs: 5ABV and 5ABY [113]. eIF4E is colored in *gray*, *Dm* Mxt in *cyan* and *Ce* Mxt in *orange*. (**c**, *Right*) Schematic representation of the bipartite (*Ce* Mxt) and tripartite (*Dm* Mxt) mode of binding to eIF4E. In *Dm* Mxt, the absence of Arg/Lys/Gln residues at positions 2 and 9 of the sequence of the canonical was compensated by the presence of an additional auxiliary 4E-BM that replaces the interactions the long aliphatic side chains perform at the dorsal surface of eIF4E. (**c**, *Left*) Close-up views of the canonical helices and the auxiliary helix of *Ce* (*orange*) and *Dm* (*cyan*) Mxt proteins bound to the dorsal surface of eIF4E. The Arg residues at positions 2 and 9 of the canonical motif of *Ce* Mxt, and the corresponding replacements (Ile and Ser) in the *Dm* canonical helix are *highlighted* in the figures

The distinct binding strategies present in Mxt proteins evolved to form complexes with eIF4E that display distinct functional properties. As a consequence of a more elaborate mode of binding, the tripartite Mxt (in flies) has a reduced ability to compete with eIF4G for binding to eIF4E than the bipartite Mxt (other species). However, in flies the eIF4E-Mxt complex is more stable and highly resistant to regulation by other 4E-IPs, allowing the function of Mxt in translation to be maintained in conditions in which the association of eIF4G with eIF4E is blocked [113]. Thus, Mxt proteins are an excellent example of how the plasticity of the mode of binding to eIF4E conferred unique properties to the proteins or to the assembled complexes. Most of the 4E-IPs so-far known adopt a bipartite arrangement when bound to eIF4E; it is less complex, easier to evolve, provides specificity and high affinity for eIF4E and an ideal competitive behavior to regulate translation initiation. In the case of *Drosophila* Mxt, specific cellular requirements, absent in other animal species, might have generated the evolutionary driving forces underlying the development of a novel binding mode. The evolution of a tripartite binding mode might also be associated with the repertoire of 4E-IPs present in different tissues or species in which the Mxt proteins have to fulfill their function, but more insight into the diversity of 4E-IPs and the function of Mxt proteins is required to draw further conclusions.

1.5 Concluding Remarks

One of the most creative forces of evolution is the transformation of a gene to perform a new function. Francois Jacob first proposed the concept of "molecular tinkering" 40 years ago [123] to account for such a fundamental law of nature at the molecular level. In the post-genomic era, such changes can be accounted for as arising from the consequences of gene or genome duplications that gave rise to duplicate copies that adopted new functions along with genome rearrangements.

In this chapter, we have discussed the lack of phylogenetic kinship among most 4E-IPs; the existence of some 4E-IPs only in specific lineages; their additional roles in distinct cellular processes such as transcription, RNA transport, splicing and degradation; their different molecular mechanisms; the wide spectrum of protein architecture; and the variability in the motif used to interact with eIF4E. In this way, evolution has provided a myriad of 4E-IPs from a plethora of different proteins, including several transcriptional factors. However, underlying all of this is the common theme of using intrinsically disordered proteins to fit the 4E-IP to the eIF4E.

The successive radiations of eukaryotes into many lineages led to the emergence of an amazing spectrum of ecological niches, body plans, organismal complexity, metabolic requirements, developmental programs and behavioral patterns. These events most probably constituted both the driving causes and effects of a parallel diversification, to different degrees in different taxa, of components and

mechanisms of the translation apparatus such as eIF4E and 4E-IPs. These molecules, in turn, expanded the translational regulatory capabilities of eukaryotes.

Acknowledgments G.H. thanks the National Institute of Cancer (INCan), Mexico, and the National Council of Science and Technology (CONACyT, Mexico, grant no. 168154) for funding this work. B.C and M.B. acknowledge Université Sorbonne Paris Cité (USPC) for the Research Project 2014 grant and 2015 International Fellowship. R.J and T.B are supported by NIH R01ES021949-01 and NSF OCE1313888 to R.J. and Allen R. Place. K.G. was supported by a graduate fellowship from the NOAA-EPP-funded Living Marine Sciences Cooperative Science Center (LMRCSC), NA11SEC4810002.

References

1. Hinnebusch AG. Molecular mechanism of scanning and start codon selection in eukaryotes. Microbiol Mol Biol Rev. 2011;75:434–67.
2. Aitken CE, Lorsch JR. A mechanistic overview of translation initiation in eukaryotes. Nat Struct Mol Biol. 2012;19:568–76.
3. Sonenberg N, Hinnebusch A. Regulation of translation initiation in eukaryotes: mechanisms and biological targets. Cell. 2009;136:731–45.
4. Jackson RJ, Hellen CU, Pestova TV. The mechanism of eukaryotic translation initiation and principles of its regulation. Nat Rev Molec Cell Biol. 2010;11:113–27.
5. Yanagiya A, Svitkin YV, Shibata S, Mikami S, Imataka H, Sonenberg N. Requirement of RNA binding of mammalian eukaryotic translation initiation factor 4GI (eIF4GI) for efficient interaction of eIF4E with the mRNA cap. Mol Cell Biol. 2009;29:1661–9.
6. Lin TA, Kong X, Haystead TA, Pause A, Belsham G, Sonenberg N, Lawrence JC. PHAS-I as a link between mitogen-activated protein kinase and translation initiation. Science. 1994;266:653–6.
7. Pause A, Belsham GJ, Gingras AC, Donzé O, Lin TA, Lawrence JC, Sonenberg N. Insulin-dependent stimulation of protein synthesis by phosphorylation of a regulator of 5'-cap function. Nature. 1994;371:762–7.
8. Poulin F, Gingras AC, Olsen H, Chevalier S, Sonenberg N. 4E-BP3, a new member of the eukaryotic initiation factor 4E-binding protein family. J Biol Chem. 1998;273:14002–7.
9. Mader S, Lee H, Pause A, Sonenberg N. The translation initiation factor eIF-4E binds to a common motif shared by the translation factor eIF-4 gammaand the translational repressors 4E-binding proteins. Mol Cell Biol. 1995;15:4990–7.
10. Marcotrigiano J, Gingras AC, Sonenberg N, Burley SK. Cap-dependent translation initiation in eukaryotes is regulated by a molecular mimic of eIF4G. Mol Cell. 1999;3:707–16.
11. Haghighat A, Mader S, Pause A, Sonenberg N. Repression of cap-dependent translation by 4E-binding protein 1: competition with p220 for binding to eukaryotic initiation factor-4E. EMBO J. 1995;14:5701–9.
12. Richter JD, Sonenberg N. Regulation of cap-dependent translation by eIF4E inhibitory proteins. Nature. 2005;433:477–80.
13. Bah A, Vernon RM, Siddiqui Z, Krzeminski M, Muhandiram R, Zhao C, Sonenberg N, Kay LE, Forman-Kay JD. Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. Nature. 2015;519:106–9.
14. Dowling RJ, Topisirovic I, Fonseca BD, Sonenberg N. Dissecting the role of mTOR: lessons from mTOR inhibitors. Biochim Biophys Acta. 2010;1804:433–9.
15. Ma XM, Blenis J. Molecular mechanisms of mTOR-mediated translational control. Nat Rev Mol Cell Biol. 2009;10:307–18.

16. Fonseca BD, Smith EM, Yelle N, Alain T, Bushell M, Pause A. The ever-evolving role of mTOR in translation. Semin Cell Dev Biol. 2014;36:102–12.
17. Raught B, Gingras AC. Signaling to translation initiation. In: Mathews MB, Sonenberg N, Hershey JWB, editors. Translational control in biology and medicine. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press; 2007. p. 369–400.
18. Sonenberg N. eIF4E, the mRNA cap-binding protein: from basic discovery to translational research. Biochem Cell Biol. 2008;86:178–83.
19. Thoreen CC, Chantranupong L, Keys HR, Wang T, Gray NS, Sabatini DM. A unifying model for mTORC1-mediated regulation of mRNA translation. Nature. 2012;485:109–13.
20. Rhoads RE. eIF4E: new family members, new binding partners, new roles. J Biol Chem. 2009;284:16711–5.
21. Kamenska A, Simpson C, Standart N. eIF4E-binding proteins: new factors, new locations, new roles. Biochem Soc Trans. 2014;42:1238–45.
22. Hernández G, Altmann M, Lasko P. Origins and evolution of the mechanisms regulating translation initiation in eukaryotes. Trends Biochem Sci. 2010;35:63–73.
23. Sonenberg N, Hinnebusch AG. New modes of translation control in development, behavior, and disease. Mol Cell. 2007;28:721–9.
24. Hernández G, Miron M, Han H, Liu N, Magescas J, Tettweiler G, Frank F, Siddiqui N, Sonenberg N, Lasko P. Mextli is a novel eukaryotic translation initiation factor 4E-binding protein that promotes translation in *Drosophila melanogaster*. Mol Cell Biol. 2013;33:2854–64.
25. Kamenska A, Lu WT, Kubacka D, Broomhead H, Minshall N, Bushell M, Standart N. Human 4E-T represses translation of bound mRNAs and enhances microRNA-mediated silencing. Nucleic Acids Res. 2014;42:3298–313.
26. Robalino J, Joshi B, Fahrenkrug SC, Jagus R. Two zebrafish eIF4E family members are differentially expressed and functionally divergent. J Biol Chem. 2004;279:10532–41.
27. Bernal A, Kimbrell DA. *Drosophila* Thor participates in host immune defence and conects a translational regulator with innate immunity. Proc Natl Acad Sci USA. 2000;97:6019–24.
28. Miron M, Verdu J, Lachance PE, Birnbaum MJ, Lasko P, Sonenberg N. The translational inhibitor 4E-BP is an effector of PI(3)K/Akt signaling and cell growth. Nat Cell Biol. 2001;3:596–601.
29. Oulhen N, Boulben S, Bidinosti M, Morales J, Cormier P, Cosson B. A variant mimicking hyperphosphorylated 4E-BP inhibits protein synthesis in a sea urchin cell-free, cap-dependent translation system. PLoS ONE. 2009;4:e5070.
30. Salaün P, Pyronnet S, Morales J, Mulner-Lorillon O, Bellé R, Sonenberg N, Cormier P. eIF4E/4E-BP dissociation and 4E-BP degradation in the first mitotic division of the sea urchin embryo. Dev Biol. 2003;255:428–39.
31. Jung MY, Lorenz L, Richter JD. Translational control by neuroguidin, a eukaryotic initiation factor 4E and CPEB binding protein. Mol Cell Biol. 2006;26:4277–87.
32. Gosselin P, Martineau Y, Morales J, Czjzek M, Glippa V, Gauffeny I, Morin E, Le Corguillé G, Pyronnet S, Cormier P, Cosson B. Tracking a refined eIF4E-binding motif reveals Angel1 as a new partner of eIF4E. Nucleic Acid Res. 2013;41:7783–92.
33. Cho PF, Gamberi C, Cho-Park YA, Cho-Park IB, Lasko P, Sonenberg N. Cap-dependent translational inhibition establishes two opposing morphogen gradients in *Drosophila* embryos. Curr Biol. 2006;16:2031–41.
34. Altmann M, Schmitz N, Berset C, Trachsel H. A novel inhibitor of cap-dependent translation initiation in yeast: p20 competes with eIF4G for binding to eIF4E. EMBO J. 1997;16:1114–21.
35. de la Cruz J, Iost I, Kressler D, Linder P. The p20 and Ded1 proteins have antagonistic roles in eIF4E-dependent translation in Saccharomyces cerevisiae. Proc Natl Acad Sci USA. 1997;94:5201–6.
36. Park YU, Hur H, Ka M, Kim J. Identification of translation regulation target genes during filamentous growth in S. cerevisiae: regulatory role of Caf20 and Dhh1. Eukaryotic Cell. 5:2120–7.

37. Ibrahimo S, Holmes LE, Ashe MP. Regulation of translation initiation by the yeast eIF4E binding proteins is required for the pseudohyphal response. Yeast. 2006;23:1075–88.
38. Zinoviev A, Leger M, Wagner G, Shapira M. A novel 4E-interacting protein in *Leishmania* is involved in stage-specific translation pathways. Nucleic Acids Res. 2011;39:8404–15.
39. Fierro-Monti I, Mohammed S, Matthiesen R, Santoro R, Burns JS, Williams DJ, Proud CG, Kassem M, Jensen ON, Roepstorff P. Quantitative proteomics identifies Gemin5, a scaffolding protein involved in ribonucleoprotein assembly, as a novel partner for eukaryotic initiation factor 4E. J Proteome Res. 2006;5:1367–78.
40. Piñeiro D, Fernandez-Chamorro J, Francisco-Velilla R, Martinez-Salas E. Gemin5: a multitasking RNA-binding protein involved in translation control. Biomolecules. 2015;5:528–44.
41. Cohen N, Sharma M, Kentsis A, Perez JM, Strudwick S, Borden KL. PML RING suppresses oncogenic transformation by reducing the affinity of eIF4E for mRNA. EMBO J. 2001;20:4547–59.
42. Kentsis A, Dwyer EC, Perez JM, Sharma M, Chen A, Pan ZQ, Borden KL. The RING domains of the promyelocytic leukemia protein PML and the arenaviral protein Z repress translation by directly inhibiting translation initiation factor eIF4E. J Mol Biol. 2001;312:609–23.
43. Morita M, Ler LW, Fabian MR, Siddiqui N, Mullin M, Henderson VC, Alain TF. B.D., Karashchuk G, Bennett CF, Kabuta T, Higashi S, Larsson O, Topisirovic, I., Smith RJ, Gingras AC, Sonenberg N. A novel 4EHP-GIGYF2 translational repressor complex is essential for mammalian development. Mol Cell Biol. 2012;32:3585–93.
44. Napoli I, Mercaldo V, Boyl PP, Eleuteri B, Zalfa F, De Rubeis S, Di Marino D, Mohr E, Massimi M, Falconi M, Witke W, Costa-Mattioli M, Sonenberg N, Achsel TB. C. The fragile X syndrome protein represses activity-dependent transation through CYFIP1, a new 4E-BP. Cell. 2008;134:1042–54.
45. Topisirovic I, Siddiqui N, Lapointe VL, Trost M, Thibault P, Bangeranye C, Piñol-Roma S, Borden KL. Molecular dissection of the eukaryotic initiation factor 4E export-competent RNP. EMBO J. 2009;28:1087–98.
46. Köhler F, Müller-Rischart AK, Conradt BR. S.G. The loss of LRPPRC function induces the mitochondrial unfolded protein response. Aging. 2015;7:701–17.
47. Topisirovic I, Culjkovic B, Cohen N, Perez JM, Skrabanek L, Borden KL. The prolin-rich homeodomain protein, PRH, is a tissue-specific inhibitor of eIF4E-dependent cyclin D1 mRNA transport and growth. EMBO J. 2003;22:689–703.
48. Topisirovic I, Kentsis A, Perez JM, Guzman ML, Jordan CT, Borden KL. Eukaryotic translation initiation factor 4E activity is modulated by HOXA9 at multiple levels. Mol Cell Biol. 2005;25:1100–12.
49. Yarunin A, Harris RE, Ashe MP, Ashe HL. Patterning of the Drosophila oocyte by a sequential translation repression program involving the d4EHP and Belle translational repressors. RNA Biol. 2011;8:904–12.
50. Shih JW, Tsai TY, Chao CH, Lee YH. Candidate tumor suppressor DDX3 RNA helicase specifically represses cap-dependent translation by acting as an eIF4E inhibitory protein. Oncogene. 2008;24:700–14.
51. Rosnerm A, Rinkevich B. The DDX3 subfamily of the DEAD box helicases: divergent roles as unveiled by studying different organisms and in vitro assays. Curr Med Chem. 2007;14:2517–25.
52. Tarn WY, Chang TH. The current understanding of Ded1p/DDX3 homologs from yeast to human. RNA Biol. 2009;6:17–20.
53. Nédélec S, Nédélec S, Foucher I, Brunet I, Bouillot C, Prochiantz A, Trembleau A. Emx2 homeodomain transcription factor interacts with eukaryotic translation initiation factor 4E (eIF4E) in the axons of olfactory sensory neurons. Proc Natl Acad Sci USA. 2004;101:10815–20.

54. Villaescusa JC, Buratti C, Penkov D, Mathiasen L, Planagumà J, Ferretti E, Blasi F. Cytoplasmic Prep1 interacts with 4EHP inhibiting Hoxb4 translation. PLoS ONE. 2009;4: e5213.
55. Longobardi E, Penkov D, Mateos D, De Florian G, Torres M, Blasi F. Biochemistry of the tale transcription factors PREP, MEIS, and PBX in vertebrates. Dev Dyn. 2014;243:59–75.
56. Minshall N, Reiter MH, Weil D, Standart N. CPEB interacts with an ovary-specific eIF4E and 4E-T in early *Xenopus* oocytes. J Biol Chem. 2007;282:37389–401.
57. Piccioni F, Zappavigna V, Verrotti AC. A cup full of functions. RNA Biol. 2005;2:125–8.
58. Keyes LN, Spradling AC. The Drosophila gene fs(2)cup interacts with otu to define a cytoplasmic pathway required for the structure and function of germ-line chromosomes. Development. 1997;124:1419–31.
59. Dostie J, Ferraiuolo M, Pause A, Adam SA, Sonenberg N. A novel shuttling protein, 4E-T, mediates the nuclear import of the mRNA 5' cap-binding protein, eIF4E. EMBO J. 2000;19:3142–56.
60. Ferraiuolo MA, Basak S, Dostie J, Murray EL, Schoenberg DR, Sonenberg N. A role for the eIF4E-binding protein 4E-T in P-body formation and mRNA decay. J Cell Biol. 2005;170:913–24.
61. Kubacka D, Kamenska A, Broomhead H, Minshall N, Darzynkiewicz E, Standart N. Investigating the consequences of eIF4E2 (4EHP) interaction with 4E-transporter on its cellular distribution in HeLa cells. PLoS ONE. 2013;8:e72761.
62. Iwasaki S, Kawamata T, Tomari Y. *Drosophila* Argonaute1 and Argonaute2 employ distinct mechanisms for translational repression. Mol Cell. 2009;34:58–67.
63. Lee SK, Lee JS, Shin KS, Yoo SJ. Translation initiation factor 4E is regulated by cell death inhibitor, Diap1. Mol Cells. 2007;24:445–51.
64. Vasudevan D. Don Ryoo H. Regulation of cell death by IAPs and their Antagonists. Curr Top Dev Biol. 2015;114:185–208.
65. Robbins RM, Gbur SC, Beitel GJ. Non-canonical roles for Yorkie and Drosophila Inhibitor of Apoptosis 1 in epithelial tube size control. PLoS ONE. 2014;9:e101609.
66. Cosentino GP, Schmelzle T, Haghighat A, Helliwell SB, Hall MN, Sonenberg N. Eap1p, a novel eukaryotic translation initiation factor 4E-associated protein in *Saccharomyces cerevisiae*. Mol Cell Biol. 2000;20:4604–13.
67. Chial HJ, Stemm-Wolf AJ, McBratney S, Winey M. Yeast Eap1, an eIF4E-associated protein, has a separate function involving genetic stability. Curr Biol. 2000;10:1519–22.
68. Deloche O, de la Cruz J, Kressler D, Doère M, Linder P. A membrane transport defect leads to a rapid attenuation of translation initiation in *Saccharomyces cerevisiae*. Mol Cell. 2004;13:357–66.
69. Amiri A, Keiper BD, Kawasaki I, Fan Y, Kohara Y, Rhoads RE, Strome S. An isoform of eIF4E is a component of germ granules and is required for spermatogenesis in *C. elegans*. Development. 2001;128:3899–912.
70. Wang JT, Seydoux G. P granules. Curr Biol. 2014;24:R637–8.
71. Li W, DeBella LR, Guven-Ozkan T, Lin R, Rose L. An eIF4E-binding protein regulates katanin protein levels in *C. elegans* embryos. J Cell Biol. 2009;187:33–42.
72. Stebbins-Boaz B, Cao Q, de Moor CH, Mendez R, Richter JD. Maskin is a CPEB-associated factor that transiently interacts with eIF4E. Mol Cell. 1999;4:1017–27.
73. O'Brien LL, Albee AJ, Liu L, Tao W, Dobrzyn P, Lizarraga SB, Wiese C. The *Xenopus* TACC homologue, maskin, functions in mitotic spindle assembly. Mol Biol Cell. 2005;16:2836–47.
74. Peset I, Seiler J, Sardon T, Bejarano LA, Rybina S, Vernos I. Function and regulation of Maskin, a TACC family protein, in microtubule growth during mitosis. J Cell Biol. 2005;170:1057–66.
75. Albee AJ, Wiese C. *Xenopus* TACC3/maskin is not required for microtubule stability but is required for anchoring microtubules at the centrosome. Mol Biol Cell. 2008;19:3347–56.

76. Kinoshita K, Noetzel TL, Pelletier L, Mechtler K, Drechsel DN, Schwager A, Lee MR. J.W., Hyman AA. Aurora A phosphorylation of TACC3/maskin is required for centrosome-dependent microtubule assembly in mitosis. J Cell Biol. 2005;170:1047–55.
77. Moury B, Charron C, Janzac B, Simon V, Gallois JL, Palloix A, Caranta C. Evolution of plant eukaryotic initiation factor 4E (eIF4E) and potyvirus genome-linked protein (VPg): a game of mirrors impacting resistance spectrum and durability. Infect Genet Evol. 2014;27:472–80.
78. Jiang J, Laliberté JF. The genome-linked protein VPg of plant viruses-a protein with many partners. Curr Opin Virol. 2011;1:347–54.
79. Goodfellow I. The genome-linked protein VPg of vertebrate viruses - a multifaceted protein. Curr Opin Virol. 2011;1:355–62.
80. Hernández G, Vazquez-Pianzola P. Functional diversity of the eukaryotic translation initiation factors belonging to eIF4 families. Mech Dev. 2005;122:865–76.
81. Joshi B, Lee K, Maeder DL, Jagus R. Phylogenetic analysis of eIF4E-family members. BMC Evol Biol. 2005;5:48.
82. Patrick RM, Browning KS. The eIF4F and eIF(iso)4F complexes of plants: an evolutionary perspective. Comp Funct Genomics. 2012;2012:287814.
83. Jones GD, Williams EP, Place AR, Jagus R, Bachvaroff TR. The alveolate translation initiation factor 4E family reveals a custom toolkit for translational control in core dinoflagellates. BMC Evol Biol. 2015;15:14.
84. Tettweiler G, Kowanda M, Lasko P, Sonenberg N, Hernández G. The distribution of eIF4E-family members across insecta. Comp Funct Genom. 2012;2012:960420.
85. Evsikov AV. Marín de Evsikova C. Evolutionary origin and phylogenetic analysis of the novel oocyte-specific eukaryotic translation initiation factor 4E in Tetrapoda. Dev Genes Evol. 2009;219:111–8.
86. Jagus R, Bachvaroff TR, Joshi B, Place AR. Diversity of eukaryotic translational initiation factor eIF4E in protists. Comp Funct Genom. 2012;2012:134839.
87. Laursen BS, Sørensen HP, Mortensen KK, Sperling-Petersen HU. Initiation of protein synthesis in bacteria. Microbiol Mol Biol Rev. 2005;69:101–23.
88. Benelli D, Londei P. Begin at the beginning: evolution of translational initiation Res Microbiol. 2009;160:493–501.
89. Benelli D, Londei P. Translation initiation in Archaea: conserved and domain-specific features. Biochem Soc Trans. 2011;19:89–93.
90. Londei P. Evolution of translational initiation: news insights from the archaea. FEMS Microbiol Rev. 2005;29:185–200.
91. Hernández G. Was the initiation of translation in early eukaryotes IRES-driven? Trends Biochem Sci. 2008;33:58–64.
92. Hernández G. On the origin of the cap-dependent initiation of translation in eukaryotes. Trends Biochem Sci. 2009;34:166–75.
93. Aravind L, Koonin EV. Eukaryotic-specific domains in translation initiation factors: implications for translation regulation and evolution of the translation system. Genome Res. 2000;10:1172–84.
94. Kyrpides NC, Woese CR. Universally conserved translation initiation factors. Proc Natl Acad Sci USA. 1998;95:224–8.
95. Hernández G. On the emergence and evolution of the eukaryotic translation apparatus. In: Biyani M, editor. Cell-free protein synthesis. Rijekra, Croatia: InTech; 2012. p. 31–50.
96. Hernández G, Proud CG, Preiss T, Parsyan A. On the diversification of the translation apparatus across eukaryotes. Comp Funct Genom. 2012;2012:256848.
97. Marcotrigiano J, Gingras AC, Sonenberg N, Burley SK. Cocrystal structure of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP. Cell. 1997;89:951–61.
98. Matsuo H, Li H, McGuire AM, Fletcher CM, Gingras AC, Sonenberg N, Wagner G. Structure of translation factor eIF4E bound to m7GDP and interaction with 4E-binding protein. Nature Structural Biology. 1997;4:717–24.

99. Osborne MJ, Volpon L, Kornblatt JA, Culjkovic-Kraljacic B, Baguet A, Borden KL. eIF4E3 acts as a tumor suppressor by utilizing an atypical mode of methyl-7-guanosine cap recognition. Proc Natl Acad Sci USA. 2013;110:3877–82.
100. Li L, Wang CC. Identification in the ancient protist *Giardia lamblia* of two eukaryotic translation initiation factor 4E homologues with distinctive functions. Eukaryot Cell. 2005;4:948–59.
101. Jankowska-Anyszka M, Lamphear BJ, Aamodt EJ, Harrington T, Darzynkiewicz E, Stolarski R, Rhoads RE. Multiple isoforms of eukaryotic protein synthesis initiation factor 4E in *Caenorhabditis elegans* can distinguish between mono- and trimethylated mRNA cap structures. J Biol Chem. 1998;273:10538–42.
102. Keiper BD, Lamphear BJ, Deshpande AM, Jankowska-Anyszka M, Aamodt EJ, Blumenthal T, Rhoads RE. Functional characterization of five eIF4E isoforms in *Caenorhabditis elegans*. J Biol Chem. 2000;275:10590–6.
103. Standart N, Minshall N. Translational control in early development: CPEB, P-bodies and germinal granules. Biochem Soc Trans. 2008;36:671–6.
104. Hernández G, Gandin V, Han H, Ferreira T, Sonenberg N, Lasko P. Translational control by *Drosophila* eIF4E-3 is essential for cell differentiation during spermiogenesis. Development. 2012;139:3211–20.
105. Hernández G, Altmann M, Sierra JM, Urlaub H, Corral RD, Schwartz P, Rivera-Pomar R. Functional analysis of seven genes encoding eight translation initiation factor 4E (eIF4E) isoforms in *Drosophila*. Mech Dev. 2005;122:529–43.
106. Rom E, Kim HC, Gingras AC, Marcotrigiano J, Favre D, Olsen H, Burley SK, Sonenberg N. Cloning and characterization of 4E-HP, a novel mammalian eIF4E-related cap-binding protein. J Biol Chem. 1998;273:13104–9.
107. Ruud KA, Kuhlow C, Goss DJ, Browning KS. Identification and characterization of a novel cap-binding protein from *Arabidopsis thaliana*. J Biol Chem. 1998;273:10325–30.
108. Lasko P. The *Drosophila melanogaster* genome: translational factors and RNA binding proteins. J Cell Biol. 2000;150:F51–6.
109. Gingras AC, Raught B, Sonenberg N. eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. Annu Rev Biochem. 1999;68:913–63.
110. Teleman AA, Chen YW, Cohen SM. 4E-BP functions as a metabolic brake used under stress conditions but not during norml growth. Genes Dev. 2005;19:1844–8.
111. Tettweiler G, Miron M, Jenkins M, Sonenberg N, Lasko P. Starvation and oxidative stress resistance in *Drosophila* are mediated through the eIF4E-binding protein, d4E-BP. Genes Dev. 2005;19:1840–3.
112. Peter D, Igreja C, Weber R, Wohlbold L, Weiler C, Ebertsch L, Weichenrieder O, Izaurralde E. Molecular architecture of 4E-BP translational inhibitors bound to eIF4E. Mol Cell. 2015;57:1074–87.
113. Peter D, Weber R, Kone C, Chung MY, Ebertsch L, Truffault V, Weichenrieder O, Igreja C, Izaurralde E. Mextli proteins use both canonical bipartite and novel tripartite binding modes to form eIF4E complexes that display differential sensitivity to 4E-BP regulation. Genes & Dev. 2015;29:1835–49.
114. Igreja C, Peter D, Weiler C, Izaurralde E. 4E-BPs require non-canonical 4E-binding motifs and a lateral surface of eIF4E to repress translation. Nature Commun. 2014;5:4790.
115. Lukhele S, Bah A. H. L, Sonenberg N, Forman-Kay JD. Interaction of the eukaryotic initiation factor 4E with 4E-BP2 at a dynamic bipartite interface. Structure. 2013;21:2186–96.
116. Paku KS, Umenaga Y, Usui T, Fukuyo A, Mizuno A, In Y, Ishida T, Tomoo K. 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. 2012;441:237–45.
117. Gingras AC, Kennedy SG, O'Leary MA, Sonenberg N, Hay N. 4E-BP1, a repressor of mRNA translation, is phosphorylated and inactivated by the Akt(PKB) signaling pathway. Genes Dev. 1998;12:502–13.

118. Gingras AC, Raught B, Gygi SP, Niedzwiecka A, Miron M, Burley SK, Polakiewicz RD, Wyslouch-Cieszynska A, Aebersold R, Sonenberg N. Hierarchical phosphorylation of the translation inhibitor 4E-BP1. Genes Dev. 2001;15:2852–64.
119. Cavalier-Smith T. The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa. Int J Syst Evol Microbiol. 2002;52:297–354.
120. He D, Fiz-Palacios O, Fu CJ, Fehling J, Tsai CC, Baldauf SL. An alternative root for the eukaryote tree of life. Curr Biol. 2014;24:465–70.
121. Cavalier-Smith T, Chao EE, Snell EA, Berney C, Fiore-Donno AM, Lewis R. Multigene eukaryote phylogeny reveals the likely protozoan ancestors of opisthokonts (animals, fungi, choanozoans) and Amoebozoa. Mol Phylogenet Evol. 2014;81:71–85.
122. Topisirovic I, Borden KL. Homeodomain proteins and eukaryotic translation initiation factor 4E (eIF4E): an unexpected relationship. Histol Histopathol. 2005;20:1275–84.
123. Jacob F. Evolution and Tinkering. Science. 1977;196:1161–6.
124. Cho PF, Poulin F, Cho-Park YA, Cho-Park IB, Chicoine JD, Lasko P, Sonenberg N. A new paradigm for translational control: inhibition via 5'-3' mRNA tethering by Bicoid and the eIF4E cognate 4E-HP. Cell. 2005;121:411–23.
125. Garcia-Fernandez J. The genesis and evolution of homeobox gene clusters. Nat Rev Genet. 2005;6:881–92.
126. McGregor AP. How to get ahead: the origin, evolution and function of bicoid. BioEssays. 2005;27:904–13.
127. Rosenberg MI, Lynch JA, Desplan C. Heads and tails: evolution of antero-posterior patterning in insects. Biochim Biophys Acta. 2009;1789:333–42.
128. Stauber M, Jäckle H, Schmidt-Ott U. The anterior determinant bicoid of Drosophila is a derived Hox class 3 gene. Proc Natl Acad Sci USA. 1999;96:3786–9.
129. Olesnicky EC, Brent AE, Tonnes L, Walker M, Pultz MA, Leaf D, Desplan C. A caudal mRNA gradient controls posterior development in the wasp *Nasonia*. Development. 2006;133:3973–82.
130. Stauber M, Prell A, Schmidt-Ott U. A single Hox3 gene with composite bicoid and zerknüllt expression characteristics in non-Cyclorrhaphan flies. Proc Natl Acad Sci USA. 2002;99:274–9.
131. Freire ER, Dhalia R, Moura DM, Lima TD, Lima RP, Reis CR, Hughes K, Figueiredo RC, Standart N, Carrington M, de Melo NOP. The four trypanosomatid eIF4E homologues fall into two separate groups, with distinct features in primary sequence and biological properties. Mol Biochem Parasitol. 2011;176:25–36.
132. Lipovich L, Hughes AL, King MC, Abkowitz JL, Quigley JG. Genomic structure and evolutionary context of the human feline leukemia virus subgroup C receptor (hFLVCR) gene: evidence for block duplications and de novo gene formation within duplicons of the hFLVCR locus. Gene. 2002;286:203–13.
133. Sekiyama N, Arthanari H, Papadopoulos E, Rodriguez-Mias RA, Wagner G, Leger-Abraham M. Molecular mechanism of the dual activity of 4EGI-1: Dissociating eIF4G from eIF4E but stabilizing the binding of unphosphorylated 4E-BP1. Proc Natl Acad Sci USA. 2015;112:E4036–45.
134. Kinkelin K, Veith K, Grunwald M, Bono F. Crystal structure of a minimal eIF4E-Cup complex reveals a general mechanism of eIF4E regulation in translational repression. RNA. 2012;18:1624–34.
135. Nelson MR, Leidal AM, Smibert CA. *Drosophila* Cup is an eIF4E-binding protein that functions in Smaug-mediated translational repression. EMBO J. 2004;23:150–9.
136. Mizuno A, In Y, Fujita Y, Abiko F, Miyagawa H, Kitamura K, Tomoo K, Ishida T. mportance of C-terminal flexible region of 4E-binding protein in binding with eukaryotic initiation factor 4E. FEBS Lett. 2008;582:3439–44.
137. Ritter B, Denisov AY, Philie J, Deprez C, Tung EC, Gehring K, McPherson PS. Two WXXF-based motifs in NECAPs define the specificity of accessory protein binding to AP-1 and AP-2. EMBO J. 2004;23:3701–10.

138. Gosselin P, Oulhen N, Jam M, Ronzca J, Cormier P, Czjzek M, Cosson B. The translational repressor 4E-BP called to order by eIF4E: new structural insights by SAXS. Nucleic Acid Res. 2011;39:3496–503.
139. Fletcher CM, Wagner G. The interaction of eIF4E with 4E-BP1 is an induced fit to a completely disordered protein. Protein Sci. 1998;7:1639–42.
140. Fletcher CM, McGuire AM, Gingras AC, Li H, Matsuo H, Sonenberg N, Wagner G. 4E binding proteins inhibit the translation factor eIF4E without folded structure. Biochemistry. 1998;37:9–15.
141. Tompa P. Intrinsically disordered proteins: a 10-year recap. Trends Biochem Sci. 2012;37:509–16.
142. Dyson HJ, Wright PE. Intrinsically unstructured proteins and their functions. Nat Rev Mol Cell Biol. 2005;6:197–208.
143. Stein A, Pache RA, Bernado P, Pons M, Aloy P. Dynamic interactions of proteins in complex networks: a more structured view. FEBS J. 2009;276:5390–405.
144. 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, Wagner G. Structure of the eukaryotic translation initiation factor eIF4E in complex with 4EGI-1 reveals an allosteric mechanism for dissociating eIF4G. Proc Natl Acad Sci USA. 2014;111:E3187–95.
145. Moerke NJ, Aktas H, Chen H, Cantel S, Reibarkh MY, Fahmy A, Gross JD, Degterev A, Yuan J, Chorev MH. J.A., Wagner G. Small-molecule inhibition of the interaction between the translation initiation factors eIF4E and eIF4G. Cell. 2007;128:257–67.
146. Cencic R, Desforges M, Hall DR, Kozakov D, Du Y, Min J, Dingledine R, Fu H, Vajda S, Talbot PJ, Pelletier J. Blocking eIF4E-eIF4G interaction as a strategy to impair coronavirus replication. J Virol. 2011;85:6381–9.
147. Cencic R, Hall DR, Robert F, Du Y, Min J, Li L, Qui M, Lewis I, Kurtkaya S, Dingledine R, Fu H, Kozakov D, Vajda S, Pelletier J. Reversing chemoresistance by small molecule inhibition of the translation initiation complex eIF4F. Proc Natl Acad Sci USA. 2011;108:1046–51.
148. Lee T, Pelletier J. Eukaryotic initiation factor 4F: a vulnerability of tumor cells. Future Med Chem. 2012;4:19–31.
149. Gross JD, Moerke NJ, von der Haar T, Lugovskoy AA, Sachs AB, McCarthy JE, Wagner G. Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E. Cell. 2003;115:739–50.

Evolution of eIF2α Kinases: Adapting Translational Control to Diverse Stresses

Stefan Rothenburg, Millie M. Georgiadis and Ronald C. Wek

1 Overview

Protein synthesis is dynamic and is modulated rapidly in response to diverse physiological and environmental stresses. The focus of this review is a family of protein kinases that control translation by phosphorylation of the α subunit of eukaryotic initiation factor-2 (eIF2) during diverse stresses, a process that regulates the initiation phase of translation. A central theme of eIF2α kinases is that each family member is activated by different stress conditions, triggering global and gene-specific translation, which provides for cell adaptation to the underlying stress. To respond to different stress conditions, each eIF2α kinase contains unique regulatory regions that recognize perturbations in cells. The juxtaposition of a related eIF2α kinase domain to unique regulatory sequences suggests that members of this protein kinase family arose by a process of exon shuffling, which culminated in adjoining new combinations of stress sensing regions to the catalytic domain. This review will begin with an overview of eIF2 and its role in the initiation of protein synthesis and translation control, followed by a description of the phylogenetic relationships between eIF2α kinases and the contributions of diverse regulatory domains in the mechanisms activating each eIF2α kinase.

S. Rothenburg
Laboratory for Host-Specific Virology, Division of Biology, Kansas State University, Manhattan, KS 66506, USA
e-mail: sr1hsv@ksu.edu

M.M. Georgiadis · R.C. Wek (✉)
Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, IN 46202, USA
e-mail: rwek@iu.edu

G. Hernández and R. Jagus (eds.), *Evolution of the Protein Synthesis Machinery and Its Regulation*, DOI 10.1007/978-3-319-39468-8_11

2 Eukaryotic Translation and eIF2

Regulation of protein synthesis in eukaryotes occurs predominantly during the initiation phase, which features multiple associated proteins that enable assembly of competent 80S ribosomes onto mRNAs. These initiation factors, each designated as eukaryotic initiation factors (eIFs), help recruit the small 40S ribosomal subunit coupled with initiator methionyl tRNAs to mRNAs [1, 2]. Upon binding to the 5'-end of the gene transcript, the small ribosomal subunit processively scans in search of an appropriate initiation codon. Once the initiator tRNA is bound to the start codon in the P site of the ribosomal subunit, the large 60S ribosome joins, which allows ribosomes to proceed to the elongation phase. One of the initiation factors, eIF2, plays an integral role in the delivery of the initiator tRNA to ribosomes. This delivery process requires eIF2 to be in an active conformation that features association with GTP. Ribosome recognition of the mRNA initiation codon and subsequent joining of the small and large ribosomal subunits require hydrolysis of the GTP associated with eIF2 producing GDP, followed by release of eIF2-GDP from the ribosome machinery (Fig. 1). To facilitate subsequent rounds of translation initiation, the eIF2-GDP is recycled to eIF2-GTP by a mechanism aided by a

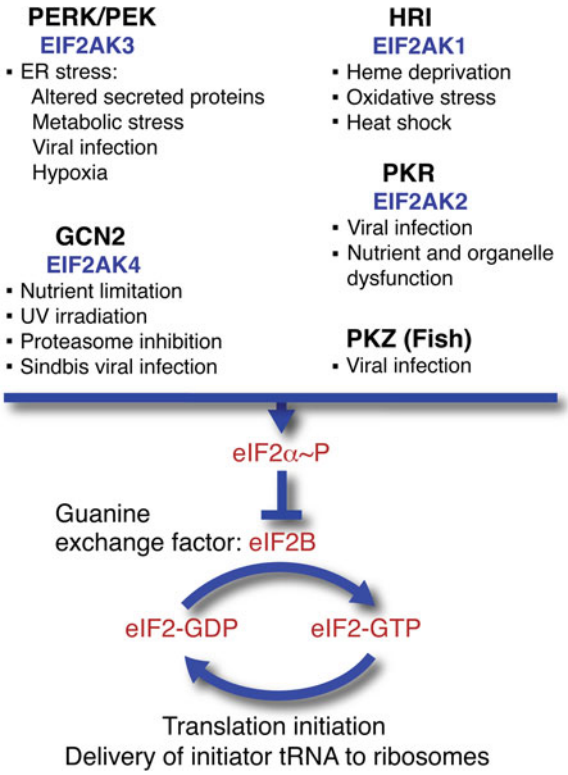


Fig. 1 The family of eIF2α kinases regulates translation in response to different stresses. Each of the eIF2α kinases is activated by distinct sets of stress arrangements, leading to phosphorylation of the α subunit of eIF2. Phosphorylation of eIF2α converts the translation initiation factor to an inhibitor of the guanine nucleotide exchange factor eIF2B. As a consequence there are lowered levels of eIF2-GTP, which reduces the delivery of initiator tRNA to ribosomes and represses global protein synthesis. Four eIF2α kinases, designated EIF2AK1-4, are expressed in humans, whereas PKZ is present in fish

guanine nucleotide exchange factor, eIF2B. Hence eIF2-GTP can again combine with initiator methionyl tRNAs and serve to expedite another round of translation initiation.

An important step in the regulation of translation initiation features control of recycling of eIF2-GDP to the active GTP-bound form of the initiation factor. The eIF2 consists of three distinct subunits, α, β, and γ, and phosphorylation of eIF2α at a residue designated as serine-51 can dramatically lower the efficiency of this guanine nucleotide exchange. Phosphorylation of eIF2α converts the initiation factor from a substrate to an inhibitor of its guanine exchange factor eIF2B, resulting in reduction in eIF2-GTP that sharply deters eIF2 binding to initiator tRNA and delivery to ribosomes (Fig. 1) [1, 2]. As a consequence, translation initiation is rapidly lowered, which allows cells to conserve energy and resources and reconfigure gene expression to adapt to the stress conditions.

It is important to note that while eIF2α phosphorylation can repress global protein synthesis, translation of many mRNAs is largely indifferent and some are in fact preferentially translated upon eIF2α phosphorylation. An underlying reason for this differential translation control among mRNAs involves the presence of short upstream ORFs (uORFs) located in the 5'-end of the gene transcripts [3]. Among the preferentially translated genes are those encoding transcription factors that help cells to alleviate damage from stress. For example, yeast *Saccharomyces cerevisiae GCN4* and mammalian *ATF4* contain multiple uORFs that allow for the coding sequences of these gene transcripts to be translated predominantly during eIF2α phosphorylation and stress [3–7]. This translation control scheme has remarkable conservation among virtually all eukaryotes, including plants, fungi, protists, and animals. Another class of mRNAs that can be translated with low eIF2-GTP levels includes those that show eIF2-independent translation initiation through internal ribosome entry sites (IRES) [2, 8]. A stress response can also be the result of the different turnover rates of proteins and mRNAs. An example for this regulation of gene expression is the activation of the transcription factor NF-κB because the levels of its labile inhibitory protein IκBα are rapidly lowered during eIF2α phosphorylation and stress because of its repressed translation of IκBα mRNA [9, 10].

3 Family of eIF2α Kinases Responds to Different Stresses

In vertebrates, up to five distinct members of the eIF2α kinase family, designated GCN2 (EIF2AK4), PERK (EIF2AK3/PEK), HRI (EIF2AK1), PKR (EIF2AK2/PKRK), and PKZ, phosphorylate the same serine-51 residue in eIF2α and mediate translation control. An important distinguishing difference between these eIF2α kinases is that each is activated in response to different stress conditions and stimuli (Fig. 1). As illustrated in Fig. 2 and discussed in depth below, while each of the family members possesses a related kinase catalytic domain, there are flanking regulatory sequences by which each eIF2α kinase can recognize distinct stress

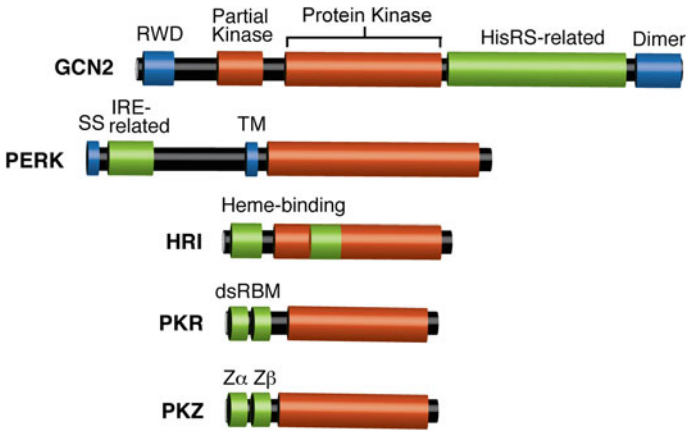


Fig. 2 The eIF2α kinases contain different stress responsive domains that function to coordinate translational control upon diverse stresses. The eIF2α kinases share a related protein kinase domain that is flanked by distinct regulatory regions that function to monitor stress conditions. The size of the eIF2α kinase regions is variable among family members because of the different lengths of the signature insert situated in the N-terminal lobe of the protein kinase domains

signals. These regulatory regions function to bind to different cellular and viral molecules whose levels are modulated during stress, serving to control activation of the eIF2α kinases by a process involving autophosphorylation in the kinase domain. The juxtaposition of related eIF2α kinase domains to unique regulatory sequences suggests that members of this protein kinase family arose by a process of exon shuffling or retrotransposition of partial mRNAs, which culminated in new combinations of stress sensing motifs to the kinase catalytic domain. In addition to the five eIF2α kinases with well-defined regulatory domains, a plethora of proven or putative eIF2α kinases are found in many protist species. While some of the protist eIF2α kinases contain some regulatory domains that are found in GCN2, others do not contain any well-defined modulatory sequences and may be activated by unique stimuli that reflect the life cycle of the protist and its environmental conditions.

4 Phylogenetic Relationships Between eIF2α Kinases

eIF2α kinases are present uniformly among eukaryotic organisms. Among the eIF2α kinases that are expressed in plants, fungi, and animals, GCN2 shows the widest distribution and is found in organisms of all three kingdoms (Table 1). By comparison, HRI is absent in plants but is found in some but not all fungi and animals, while PERK is only present in animals. PKR has only been identified in vertebrates, whereas PKZ is only present in some fish species. Since the eIF2α kinase domains are more closely related to one another than other protein kinases, the different eIF2α kinases likely emerged through gene duplications. One copy of

Table 1 Distribution of eIF2α kinases in model organisms

		GCN2	HRI	PERK	PKR	PKZ
A. thaliana	(Plant)	1	–	–	–	–
S. cerevisiae	(Fungus)	1	–	–	–	–
S. pombe	(Fungus)	1	2	–	–	–
C. elegans	(Nematode)	1	–	1	–	–
D. melanogaster	*(Insect)*	1	–	1	–	–
A. mellifera	*(Insect)*	1	1	1	–	–
S. purpuratus	(Echinoderm)	1	1	1	–	–
D. rerio	(Fish)	1	1	1	1	1
T. nigroviridis	(Fish)	1	1	1	3	–
X. tropicalis	(Amphibian)	1	1	1	3	–
H. sapiens	(Mammal)	1	1	1	1	–

Numbers of genes present in the indicated genomes are shown
Dashes indicate absence of genes

the eIF2α kinase genes subsequently acquired different regulatory domains, which provided selective advantages, whereas one copy of the duplicated gene was maintained. Drawing from the distribution of eIF2α kinases among different organisms, phylogenetic analyses of the kinase domains suggest GCN2 to be ancestral to the other eIF2α kinases (Fig. 3) [11, 12]. A duplication of the GCN2 kinase domain in a common ancestor of fungi and animals likely led to the emergence of HRI, which itself is ancestral to PERK and PKR. Thus, a duplication of HRI probably occurred in an early ancestor of extant animals, which is suggested to have resulted in the evolution of PERK. A duplication of PERK likely occurred in a common ancestor of teleost fish and tetrapods, which resulted in the emergence of PKR [11, 12]. Within teleost fish, a duplication of PKR led to the evolution of PKZ [13]. Other eIF2α kinases may have evolved during the evolution of plants, fungi, and animals, which either existed temporarily or might be still present in some extant organisms and have so far eluded identification. In the latter case, inclusion of newly identified eIF2α kinases would allow the construction of a more refined eIF2α kinase family tree.

An interesting observation in the phylogenetic relationships among eIF2α kinases is that family members were not only gained during evolution, but also lost, as observed for HRI genes. While HRI likely evolved in an ancestor of fungi and animals, it is absent in some organisms belonging to these kingdoms. For example, HRI is found in many fungi, including *Schizosaccharomyces pombe* (fission yeast) and *Aspergillus* species, but is absent from *S. cerevisiae* (budding yeast) and other *Saccharomycotina* fungi. Furthermore, not all animals possess *HRI* genes. HRI is found in many insects including mosquitoes, bees, and silkworm but is absent from the genomes of *Drosophilidae*, including *Drosophila melanogaster* (Table 1). HRI is also absent in *Caenorhabditis elegans* and other *Caenorhabditis* species, but present in other nematodes including *Trichinella* species. In each of these cases, *HRI* became lost independently in different lineages.

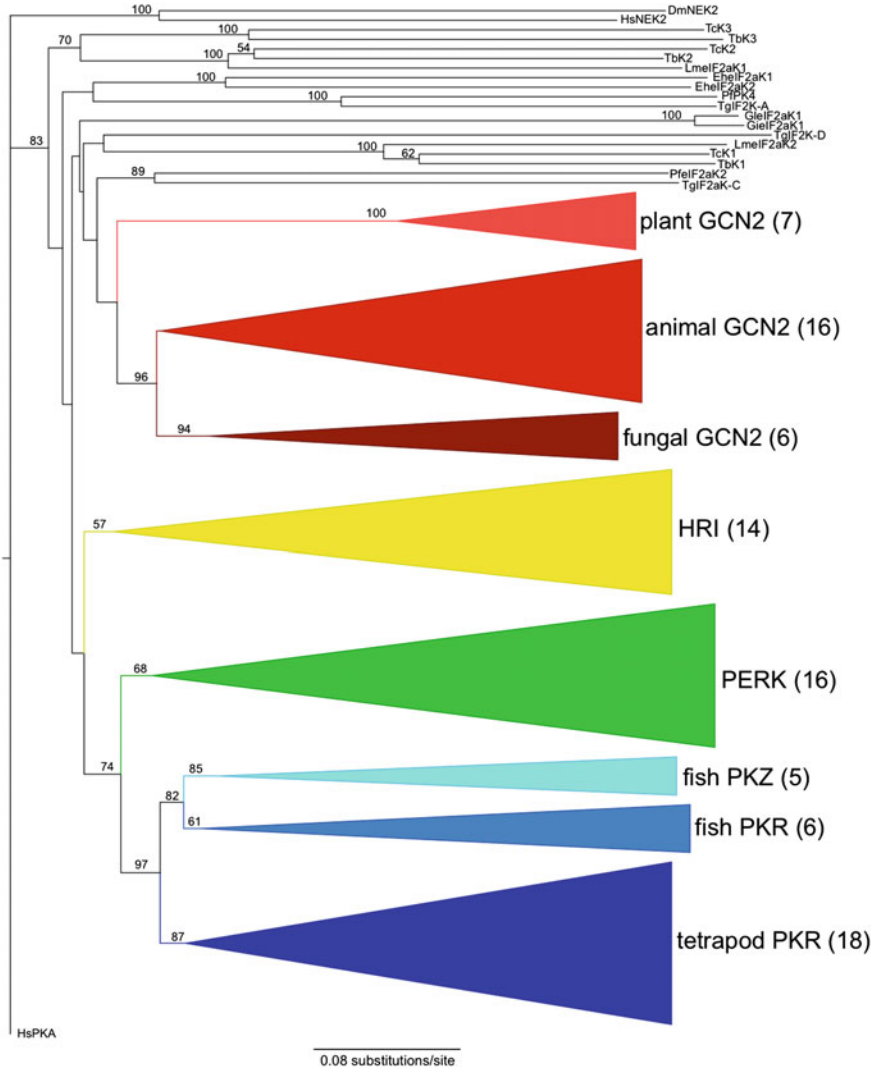


Fig. 3 Phylogenetic analysis of eIF2α kinases. The sequences of the kinase domains of 105 eIF2α kinases, deleted for the kinase insert between the N- and C-lobe, along with dmNEK2 (NIMA-related kinase 2), as a closely related kinase and hsPKA (cAMP-dependent protein kinase) as a distantly related kinase, were aligned using the MUSCLE program [147]. The phylogenetic tree was generated using the NJ method with nodal support assessed via bootstrapping (10,000 replicates) as implemented in PAUP [148]. The tree was rooted to PKA. Bootstrap values above 50 are indicated above the branch nodes. For a clearer representation, the branches for major eIF2α kinase groups are shown as cartoons, and the number of sequences is indicated. Abbreviations are as follows: Dm *Drosophila melanogaster*; Eh *Entamoeba histolytica*; Gi *Giardia intestinalis*; Gl *Giardia lamblia*; Hs *Homo sapiens*; Lm *Leishmania major*; Pf *Plasmodium falciparum*; Tb *Trypanosoma brucei*; Tc *Trypanosoma cruzi*; Tg *Toxoplasma gondii*

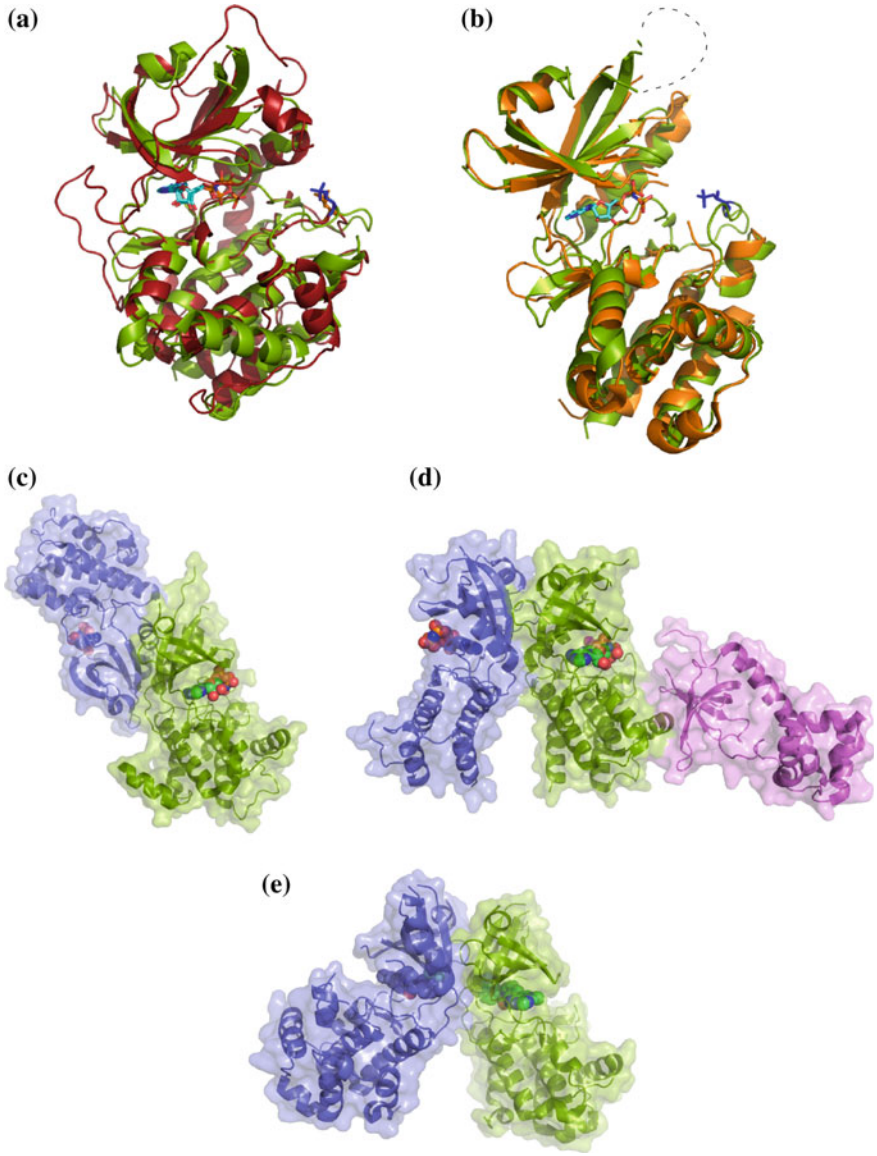
Another phenomenon suggested in the evolution of eIF2α kinases is that some gene duplications occurred in some lineages, with duplicated genes maintaining their regulatory domains, for example, two HRI copies in *S. pombe* and three independently evolved PKR copies in some fish and amphibian lineages (Table 1). The duplication of PKR genes and emergence of PKZ in fish will be discussed further below in the presentation of the mechanisms regulating these eIF2α kinases. Additionally, we will highlight below the expansion of the eIF2α kinase family; PKR has been subject to rapid evolution to alter sensitivity of this eIF2α kinase to viral inhibitors [12, 14].

The evolutionary relationship of protist eIF2α kinases to those of plants, fungi, and animals has not yet been studied in detail. Our phylogenetic analysis shows some protist eIF2α kinases to be ancestral to the family members from plants, fungi, and animals (Fig. 3). Consistent with GCN2 being broadly expressed among eukaryotes, a subgroup of eIF2α kinases actually clusters as a sister clade to the GCN2 clade. As will be discussed below, while some protists clearly possess GCN2-related eIF2α kinases, they appear to lack eIF2α kinases that are directly related to HRI, PERK, and PKR. Some protist eIF2α kinases are found in the lowest branches of the tree and therefore appear to be more ancestral (Table 1).

5 Catalytic Domain of eIF2α Kinases

Like other eukaryotic protein kinases, the catalytic domain of the eIF2α kinase family members have a bi-lobal structure, with the ATP-binding region positioned in the hinge region between the two lobes (Fig. 4a). The N-terminal lobe features a largely antiparallel β-sheet structure that binds nucleotides through an unusual motif. A defining feature of the eIF2α kinases is inclusion of insert sequences ranging from ~15 amino acid residues to well over 100 residues, which are situated at the top surface of the N-terminal lobe between β-strands (Fig. 4b). While the position of this insert is conserved among eIF2α kinases, the sequences even between orthologs of a specific eIF2α kinase are variable but can be enriched for acidic amino acid residues [11]. Structural analyses of PKR, GCN2, and PERK required the partial deletion of this insert sequence for protein crystallization [15–18]; hence, the structures of the inserts are not yet known, and there are suggestions that the insert regions may be largely unstructured or assume different structural arrangements depending on the activation state of the eIF2α kinase. Nonetheless, residue substitutions and small deletion analyses indicate that the inserts are required for activity of the eIF2α kinase family members [16, 18–21].

The larger α-helical C-terminal lobe is connected to the N-terminal lobe by a hinge region and contains residues that are critical for substrate recognition, activation, and catalysis. Autophosphorylation of a Thr residue in the activation loop facilitates interactions with the N-terminal lobe resulting in a closed conformation of the enzyme and formation of the substrate binding site [22]. This feature is conserved across Ser/Thr kinase family members as shown in a comparison of the


(a)
(b)
(c)
(d)
(e)

◀ **Fig. 4** Comparison of structures of the eIF2α kinase domains. **a** An activated kinase domain from PKR (PDB:2A19, A chain) is shown as a *green cartoon* rendering superimposed on cAPK, cyclic AMP-dependent protein kinase, (PDB:1ATP), shown as a *red cartoon* rendering. The phosphorylated Thr residues are shown in a *stick model* for each structure in *orange* for PKR and *blue* for cAPK. Bound ATPs for cAPK or phosphoaminophosphonic acid-adenylate ester (ANP) are shown in *stick models* with C, *cyan*, O, *red*, N, *blue*, and P, *orange*. **b** PKR (PDB:2A19, A chain) and PERK (PDB:4G31) kinase domains are shown superimposed. PKR is shown in *green* and PERK in an *orange cartoon* rendering. The location for the deleted polypeptide corresponding to the insertion found in eIF2α kinases is indicated as a *dashed black line*. **c** The dimeric structure of the kinase domain from GCN2 (PDB:1ZY5) is shown with the two polypeptide chains as cartoon renderings in *green* and *blue*. A semi-transparent molecular surface is shown for each chain; AMP-PNP is shown as van der Waals spheres with C, *green* or *blue*, O, *red*, N, *blue*, and P, *orange*. **d** The dimeric structure of the PKR kinase domain (PDB:2A19) is shown as a cartoon/surface rendering as described in **c**. Bound ANP molecules are shown as van der Waals spheres as in panel C. eIF2α is shown in a *pink cartoon/surface rendering* interacting with the activated subunit of the PKR dimer. **e** The dimeric structure of the PERK kinase dimer (PDB:4G31) is shown in a similar orientation and rendering as in panels C and D with the small molecule inhibitor 7-methyl-5-(1-{[3-(trifluoromethyl)phenyl]acetyl}-2,3-dihydro-1H-indol-5-yl)-7H-pyrrolo[2,3-d] pyrimidin-4-amine (GSK2606414) bound to each polypeptide chain. In this case, a single polypeptide crystallized in the asymmetric unit; the dimer is generated by crystallographic symmetry

cyclic AMP-dependent protein kinase [23] with PKR [17] (Fig. 4a). The S1 subdomain comprising a 5-stranded β-barrel of eIF2α interacts with a single alpha helix, αG, in the C-terminal lobe of PKR in the PKR-eIF2α complex structure (Fig. 4d). Interaction of eIF2α with the PKR kinase domain is suggested to allow for a conformational transition in the substrate that positions the adjacent serine-51 phosphorylation site into the phospho-acceptor binding site of PKR [17].

PKR function can be inhibited by a number of different viral proteins and RNAs. One protein designated K3L from vaccinia virus has sequence conservation with eIF2α, allowing for K3L to mimic the structure of the translation factor [24–27]. It is important to note that the K3L sequence diverges from that of eIF2α at the site of phosphorylation by PKR. Hence, K3L is thought to compete with eIF2α for binding to PKR, thereby preventing eIF2α phosphorylation during viral infection. As a consequence, translational control in infected cells is abrogated, rendering the host more sensitive to viral disease.

6 Activation of eIF2α Kinases upon Sensing of Stress Conditions

Determining the mechanisms of activation of the eIF2α kinases remains an active area of investigation. Although the catalytic domains of eIF2α kinases are closely related to those of other mammalian Ser/Thr kinases and Tyr kinases, some members differ in oligomeric state, and all have insertions within their N-terminal lobes (Fig. 4). By analogy, initial models of activation for eIF2α kinases were

based on the activation of receptor tyrosine kinases, which dimerize in response to ligand binding resulting in activation through autophosphorylation.

The structural basis for receptor dimerization upon ligand binding was first reported for the human growth hormone-receptor complex [28]. Extending this concept of dimerization upon ligand binding, the crystal structure of the first receptor tyrosine kinase-ligand complex was reported for the fibroblast growth factor-receptor tyrosine kinase complex [29]. This work in addition to numerous other studies led to a unifying model for activation for receptor tyrosine kinases, which was subsequently proposed as a mechanistic basis for the activation of eIF2α kinases in response to stress. Thus, the crystal structures of the isolated eIF2α kinase domains from these enzymes were interpreted in light of a model in which the inactive kinase existed as a monomer and was activated through dimerization.

The eIF2α kinase domains dimerize through interactions of their N-terminal lobes in the crystal structures of PKR, GCN2, and PERK kinases (Fig. 3). In the structure of the kinase domain of GCN2 with bound AMP-PNP (PDB:1ZY5) [18], the N-terminal lobes are juxtaposed in a head-to-head interaction with the C-terminal lobes positioned on opposite ends of the dimer. This dimeric arrangement differs significantly from that of other eIF2α kinase structures. While it is possible that it represents an inactive dimeric form of the enzyme that would undergo a significant conformational change upon activation through binding of uncharged tRNA to the histidyl tRNA synthetase-related domain [30], it is also possible that this dimer forms as a consequence of crystal packing forces within the lattice.

Dimeric interactions within the PERK kinase domain structure with the small molecule GSK6515 bound (PDB:4G31) [15] arise through crystallographic symmetry and also involve interactions between the N-terminal lobes. However, in this case, the C-terminal lobes are proximally positioned in a back-to-back arrangement, albeit with their active sites and activation loops pointing away from one another. Given that the lumenal domain of PERK forms dimeric or tetrameric structures [31], the dimeric kinase may phosphorylate a second PERK dimer in trans. Back-to back dimeric kinase domain interactions in both the apo form of the mutant K296R PKR kinase domain (PDB:3UIU) and a complex of a PKR kinase dimer with eIF2α and AMP-PNP (PDB:2A19) (Fig. 4d) [17] are very similar to that of the crystallographic PERK dimer (Fig. 4c). In the latter structure (PDB:2A19), one kinase domain within the dimer interacts with eIF2α, and its activation loop is ordered through interactions of the N- and C-terminal lobes with phosphorylated Thr-446 revealing an active conformation for this molecule (Fig. 4d). In the other kinase domain comprising the PKR dimer, the activation loop is disordered. The activation loops in PERK and GCN2 kinase domains are also disordered in the crystal structures. In a second crystal form including one PKR kinase domain and one eIF2α (PDB:2A1A), the kinase dimer generated by crystallographic symmetry, while still in a back-to-back conformation with interacting N-terminal lobes, has C-terminal lobes that are splayed much farther apart than in the other structure.

In contrast to GCN2 and PERK, the sensing domain in PKR, a dsRNA-binding motif (dsRBM), is monomeric in the absence of viral dsRNA but dimerizes upon

binding to specific dsRNA sequences [32–35]. Thus, PKR may be more similar to the receptor tyrosine kinases in being monomeric in the absence of stress. Activation of inactive monomeric PKR requires both dimerization and autophosphorylation. One model suggests that each kinase domain phosphorylates its own activation loop following conformation changes induced by dimerization [22].

7 Stress Regulation of eIF2α Kinases

Each of the eIF2α kinases consists of a combination of stress-sensing regions adjoined with the kinase catalytic domain. The above-described phylogenetic relationships indicate that the eIF2α kinase family expanded with the evolution of metazoic animals. The physiological traits of the organism and nature of environmental exposures, along with the activating properties of each eIF2α kinase, helped to determine which of the eIF2α kinases is expressed in a given organism. Highlighted below is a summary of the mechanisms modulating each eIF2α kinase and the critical roles that the regulatory domains unique to each family member play in translation control during stress.

GCN2: The eIF2α kinase GCN2 is represented among virtually all eukaryotic phyla, suggesting an integral function in stress resistance that was implemented early in evolution. GCN2 consists of five defined domains that each play integral parts for GCN2 activation by amino acid starvation (Fig. 2). In addition to the above-described protein kinase domain, GCN2 includes the RWD domain, a partial kinase or pseudokinase domain, a histidyl-tRNA synthetase (HisRS)-related domain, and C-terminal dimerization domain. Among these GCN2 regulatory regions, crystal structures are available for the murine RWD [36] and yeast and murine C-terminal domains [37] (Fig. 5). Mutations or deletions in each of these regions can inactivate GCN2, indicating that each is required for GCN2 function [21, 38–41].

Foremost among these regulatory domains is the HisRS-related domain that is suggested to be important for GCN2 sensing depletion of amino acid in cells. Nutrient starvation enhances the levels of uncharged tRNAs, which can bind directly to the HisRS-related domain and trigger activation of the adjacent eIF2α kinase domain. It is important to note that starvation for many different amino acids, not just histidine, can activate GCN2; thus, GCN2 is suggested to bind multiple different uncharged tRNAs that can accumulate during diverse starvation conditions. Although activation of GCN2 was initially proposed to require dimerization, crystal structures of the C-terminal domain (CTD) of GCN2 feature an interdigitated dimeric structure [37] (Fig. 5b, c). Thus, in the absence of stress, GCN2 is already a dimer, suggesting that activation involves either a conformational change or formation of a higher oligomeric state. A role for localizing GCN2 to the ribosome through interaction of the CTD with the ribosome is not conserved; yeast CTD is ribosome associated, whereas the murine CTD is not [37].

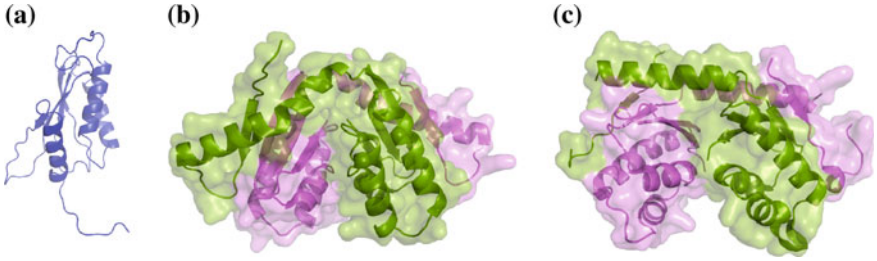


Fig. 5 Regulatory domains of GCN2. **a** A single solution structure from an ensemble of 20 structures for the RWD domain from murine GCN2 (PDB:1UKX) is shown as a *blue cartoon rendering*. **b** The crystal structure of the murine C-terminal domain (CTD) of GCN2 (PDB:4OTN) is shown as *green and pink cartoon renderings* with semi-transparent surfaces. **c** A similar rendering is shown for the yeast C-terminal domain (PDB:4OTM). Although the murine and yeast CTDs have similar overall folds for each polypeptide chain, the dimeric structures differ significantly

Featured at the N-terminus of GCN2 is a domain found in RING finger- and WD repeat-containing proteins and DEAD-like helicases, prompting the designation RWD domain. The RWD of GCN2 consists of a β-sheet sandwiched between α-helices (Fig. 5a), which is proposed to serve in protein-protein interactions that facilitate GCN2 binding of uncharged tRNAs in the context of the translational machinery [36, 38, 42–44]. There are GCN2 inhibitory proteins, such as IMPACT, which also possess an RWD domain and can compete with GCN2 for binding to a common docking protein designated GCN1, which is required for activation of the eIF2α kinase [45, 46]. Finally, GCN2 contains a partial kinase domain that consists of sequences central to the C-terminal lobe, but is without catalytic function [30, 40, 41, 47, 48]. This partial kinase region is thought to directly interact with the adjoining kinase catalytic domain, and this engagement is proposed to expedite conformational changes that activate GCN2 upon binding to uncharged tRNAs [30].

GCN2 is activated by other stresses not directly linked to amino acid availability, including high salinity, glucose depletion, and UV irradiation [10, 39, 49–51]. In the case of high sodium stress, there is reported lowering of charged tRNAs, indicating that the mechanism of GCN2 activation also is likely to involve binding to accumulating uncharged tRNAs [52]. It remains to be determined whether other stresses, such as UV irradiation, also lower tRNA charging or trigger alternative modes of GCN2 activation.

Loss of *GCN2* has significant consequences for adaptation to nutrient depletion. In the case of *S. cerevisiae*, deletion of *GCN2* renders the yeast more sensitive to limiting amino acids or glucose [51, 53]. In plants such as *Arabidopsis thaliana*, GCN2 regulates seed germination in diverse environmental conditions, and loss of GCN2 can adversely affect leaf morphology and plant growth and development [54, 55]. Deletion of *GCN2* in mice impairs translational control and appropriate adaptation to starvation for essential amino acids [56]. In this case, loss of *GCN2* diminishes the ability of liver to appropriately attenuate translation upon nutrient

starvation, leading to accelerated turnover of protein in muscle. GCN2 also plays a significant role in adaptation to dietary imbalances for essential amino acids in mice by altering feeding behavior [57–59]. Finally, mutations in *GCN2* in humans were reported to lead to pulmonary veno-occlusive disease (PVOD) and pulmonary capillary hemangiomatosis (PCH) [60, 61]. These lung disorders feature extensive fibrous intimal proliferation of septal veins and preseptal venules, along with pulmonary capillary dilatation and proliferation. Currently, the rationale for *GCN2* deficiencies afflicting lung tissues is not yet understood, and treatments often require bilateral lung transplants.

PERK: Phosphorylation of eIF2α by PERK is induced in response to accumulation of unfolded protein in the endoplasmic reticulum (ER), which is suggested to occur with different perturbations in this organelle. PERK is a type 1 ER transmembrane protein, which features an N-terminal signal sequence that is suggested to facilitate transport of the N-terminal portion of the eIF2α kinase into the lumen of this organelle (Fig. 6) [62, 63]. An α-helical hydrophobic segment serves as a stop-transfer segment during translocation, as well as the subsequent transmembrane region, which results in the C-terminal eIF2α kinase domain of PERK residing in the cytosol on the exterior of the ER. The luminal portion of PERK is suggested to sense accumulating unfolded proteins in the ER, triggering phosphorylation of eIF2α, which serves to repress translation and lower the influx of nascent polypeptides into the stressed ER [62, 64–67]. In addition to translational control, PERK induces transcriptional regulation in concert with the unfolded protein response (UPR), which collectively serves to expand the processing capacity of the ER to better manage the incoming protein load on this organelle [68–73]. Along with PERK, there are two other ER transmembrane proteins, IRE1 and ATF6, which function to recognize unfolded proteins in the ER and induce the transcription regulation of the UPR.

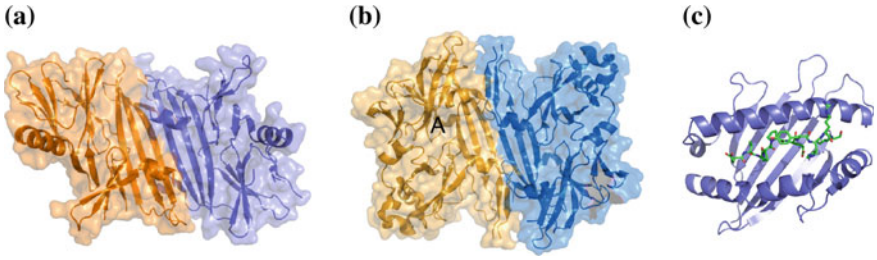


Fig. 6 Lumenal sensing domain of PERK. **a** The structure of the dimeric lumenal domain of PERK (PDB:4YZS) is shown as a cartoon rendering with a semi-transparent surface. The two polypeptide chains that comprise the dimer are shown in *blue* and *orange*. **b** A similar rendering of the yeast IRE lumenal domain (PDB:2BE1) is shown with polypeptide chains shown in *blue* and *orange*. **c** A *blue cartoon rendering* is shown for a class I major histocompatibility complex (MHC) with bound peptide (PDB:1TMC) as a *stick model*, C, *green*, O, *red*, and N, *blue*. Both PERK and IRE dimeric structures present a β-sheet platform with α-helices defining a potential peptide binding groove reminiscent of the peptide binding site observed in the MHC molecules

There are two prevailing models for how PERK senses accumulating unfolded protein in the ER. Both models center on the portion of PERK that resides in the lumen of the ER. This portion of PERK is homologous to IRE1 and exists as both dimeric and tetrameric species in crystal structures; a tetramer is generated through crystallographic symmetry (4YZS) [31]. The dimeric structure of both PERK and yeast IRE1 lumenal domains features a central groove formed by α-helices, one from each subunit of the dimer, and a β-sheet base, again formed by the dimeric interaction, which is likened to the peptide-binding domains of major histocompatibility complexes (Fig. 6). The ER lumenal structure of yeast IRE1 is suggested to bind directly to unfolded protein, triggering its activation in the UPR [74, 75]. Given the similarities in structures between IRE1 and human PERK, it is also plausible that PERK can bind to accumulating unfolded protein in the ER, leading to conformational changes that induce PERK phosphorylation of eIF2α. An alternative model for PERK and IRE1 sensing of ER stress is that both can bind to BiP (GRP78/HSPA5), an ER molecular chaperone that binds to nascent polypeptides and aids protein folding [64, 67, 75]. BiP binds to the lumenal portion of PERK. In one study, the interaction is non-canonical, involving the folded lumenal domain of PERK and the ATPase domain of BiP [76]. Other studies suggest a canonical interaction involving the substrate-binding domain of BiP and an unstructured region of the PERK lumenal domain near the transmembrane portion of the protein [67]. Independent of the nature of the interactions, BiP binding may lead to repression of PERK activation. Upon accumulation of unfolded protein, BiP would be titrated from PERK, allowing for appropriate conformation changes that may involve oligomerization. The indirect BiP binding model may also function as a coarse regulatory system in conjunction with more direct PERK binding to unfolded protein.

Loss of *PERK* function is thought to adversely affect all tissues that specialize in protein secretion. This is illustrated by Wolcott-Rallison syndrome, which results from *PERK* (*EIFAK3*) mutations and features neonatal diabetes, digestive dysfunction, bone abnormalities, growth retardation, and episodic liver dysfunction [77–79]. Mouse models featuring deletions of *PERK* recapitulate the key phenotypes associated with human patients, providing for a model for disease progression and treatment [80, 81].

HRI: The eIF2α kinase HRI contains two reversible heme binding sites, one in the N-terminus and another within the insert region of the kinase domain [82]. These so-called heme-responsive motifs are present in other proteins associating with heme and are central to the mechanisms regulating HRI [83]. In mammals, HRI is predominantly expressed in erythroid tissues, and heme binding is suggested to repress eIF2α kinase function [82, 84]. However, upon iron depletion and lowered heme levels, heme is released from HRI, leading to eIF2α phosphorylation that would lower the synthesis of proteins, which in erythroid tissues is predominantly globin. Hence, globin synthesis is linked to iron availability. Loss of *HRI* in mouse models leads to microcytic hypochromic anemia, which is linked with inclusions consisting

of excess heme-free α- and β- globins in heme-deficiency states [85, 86]. There is mounting evidence that HRI can be activated by other stresses, including those involving oxidative damage, osmotic stress, and heat shock [87]. Oxidative stress has been shown to induce HRI in *Schizosaccharomyces pombe*, and depletion of two HRI orthologs from this yeast sensitizes cells to oxidizing agents, such as arsenite [88, 89]. It is likely that heme binding is integrated with HRI sensing of oxidation stress, and this mode of regulation is central for translational control in the lower eukaryotes expressing this eIF2α kinase.

PKR: The eIF2α kinase PKR is transcriptionally expressed at moderate levels and is induced by interferon and functions to thwart viral infections in vertebrate organisms [90]. Regulation of PKR centers on the two dsRBMs situated in the N-terminus that serve to bind dsRNA expressed during viral infections [22, 91–94]. The dsRBMs are each ~70 residues in length and consist of an αββββα fold (Fig. 7a) that are present in a number of different proteins that function in RNA editing, processing, transport, and silencing [92, 95]. PKR is mainly present in the cytosol and to a lesser extent in the nucleus of cells in a conformation that transitions between monomeric and dimeric forms [22, 33, 96–98]. Upon viral infection, PKR would bind to dsRNA produced by viruses during infection and replication. PKR binding to the viral dsRNA is suggested to fully induce PKR dimerization by bridging PKR molecules that help orientate specific interactions between the kinase domains of two PKR molecules. Emphasizing the importance of dimerization in the mechanism of PKR activation, the dsRBMs of PKR can be functionally substituted by fusing heterologous dimerization regions to its eIF2α kinase domain [99, 100]. Trans-intradimer autophosphorylation is not thought to occur, given the back-to-back orientation of the PKR dimer. This suggests that autophosphorylation occurs between PKR dimers, which would include

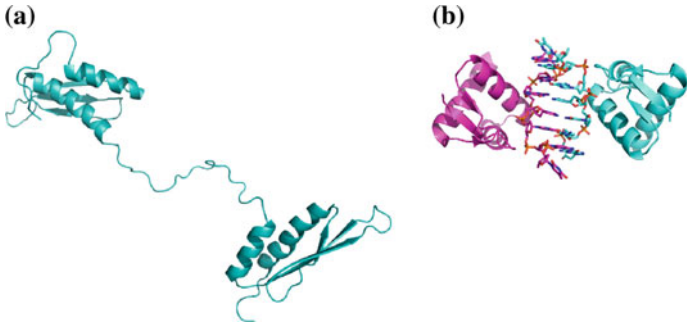


Fig. 7 Nucleic acid binding domains from PKR and PKZ. **a** A single solution structure from the ensemble of 21 structures for the dsRNA binding domain from human PKR (PDB:1QU6) is shown as a *cyan cartoon rendering*. **b** The dimeric winged helix-turn-helix Z-DNA binding domain bound to the Z-DNA structure is shown with one polypeptide chain in *magenta* and the second in *cyan*. The Z-DNA is shown as a *stick rendering* with C, *magenta* or *cyan*, O, *red*, N, *blue*, and P, *orange*. In this case, the dimer is generated by crystallographic symmetry. One polypeptide and a single strand of DNA are in the asymmetric unit

phosphorylation in the activation loop region that would then aid eIF2α substrate phosphorylation and lowered global translation in the infected host cells. Viral replication is dependent upon host translation, which is repressed following activation of PKR, and eIF2α phosphorylation is also suggested to trigger apoptosis, which would further lower the ability of viruses to infect neighboring cells [101, 102].

The mechanism of PKR activation may provide a rationale for why this eIF2α kinase is induced by viral dsRNAs but not those encoded by the mammalian host. For example, the mammalian genome is composed of noncoding retrotransposons, such as ALU sequences, and expresses precursors of miRNAs and lncRNAs that can contain dsRNA. It is known that optimal activation of PKR requires dsRNA segments that are at least 30 bp in length. Furthermore, some secondary structures in RNA may bind to PKR but are not optimal for induction of dimerization arrangements that promote PKR activation. In fact, this is a strategy used by some viruses to bar activation of PKR. For example, adenovirus VA RNAi and Epstein-Barr virus EBER1 and 2 can bind to PKR and inhibit this eIF2α kinase, and these RNAs are important for these viruses to overcome interferon-induced apoptosis [103–108]. Viruses can also inhibit PKR through expression of proteins that possess dsRBMs. Vaccinia virus protein E3L contains a C-terminal dsRBM that is suggested to inhibit PKR by sequestration of activator dsRNAs [109–114]. Additionally, there is mounting evidence that E3L can also inhibit PKR by formation of heterodimers. Regulation of PKR by dsRBM-containing proteins also extends to those encoded by mammalian cells. The mammalian protein PACT contains three dsRBMs and is suggested to engage with PKR and trigger activation through a conformational change that enhances eIF2α phosphorylation [115, 116].

Mice deficient for PKR show enhanced susceptibility to infection by viruses, such as vesicular stomatitis virus [117, 118]. Furthermore, deletion of PKR inhibitory genes in vaccinia virus attenuates infection [111, 119]. Viruses encode a plethora of PKR inhibitors that can inhibit all steps in the PKR activation cycle [120]. PKR has been subject to rapid evolution because of an ongoing molecular arms race between PKR and viral PKR inhibitors [12, 14]. Within the vertebrate lineage, the PKR kinase domain evolved much faster than the kinase domains of the other eIF2α kinases. The rationale for this observation is that viral proteins that are homologs of eIF2α, including vaccinia virus K3L and ranavirus vIF2α can directly target the PKR kinase domain to prevent the interaction with eIF2α and act as pseudosubstrate inhibitors [26, 121]. As a consequence, variation in the PKR kinase domain among mammals resulted in differential sensitivity to vaccinia virus K3L [12, 14]. The interaction between PKR and pseudosubstrate inhibitors can be very species-specific. For example, it was shown that M156R, the K3L ortholog from myxoma virus, was only able to inhibit PKR from European rabbits but not human, sheep, mouse, rat, Guinea pig, Syrian hamster or Chinese hamster PKR. These findings correlate well with the host range of myxoma virus, which is a rabbit-specific pathogen, and emphasize the importance of PKR inhibition for virus replication [122].

PKZ, a PKR-related eIF2α kinase in fish: Because of the importance of PKR in the response to viruses in mammals, there was much interest in identifying PKR in non-mammalian species. For example, infections cause substantial economic and ecological problems in both farmed and wild fish, amphibian and reptile populations, and characterization of their innate immune response to pathogens is of great importance. In the fish rainbow trout, PKR-like activity was described after the infection with infectious pancreatic necrosis and poly(I:C) incubation, which mimics viral dsRNA [123]. The first PKR-related genes in fish were cloned and characterized from goldfish and zebrafish [11, 124]. Surprisingly, the deduced proteins lacked dsRBMs, but instead contained two Z-DNA/RNA-binding (Zα) domains and were hence named PKZ [11]. The PKZ-associated Zα domains are winged helix-turn-helix motifs, a common DNA-binding motif, which form a dimeric complex with bound DNA as seen in the goldfish Zα domain: Z-DNA complex (PDB:4KMF) [125] (Fig. 7b). Phylogenetic analyses showed that the eIF2α kinase domain in the PKZ orthologs was more closely related to that of PKR than other eIF2α kinases (Fig. 3) [11, 13]. Interestingly, Zα domains are found in other interferon-inducible proteins, ADAR1 and ZBP1 (Dlm-1), as well as in the poxvirus virulence and host range factor E3L and cyprinid herpes virus 3 ORF112 [126–129]. These domains are thought to induce a rapid transition from B- to Z-DNA through binding to dCdG repeat sequences [125]. PKZ was also described in other fish, including Atlantic salmon, rare minnow, and grass carp [130–132].

PKZ and PKR genes are arranged in tandem in the zebrafish genomes and are separated by approximately 8 kb, suggesting that the PKZ gene was created by a chromosomal duplication. Gene duplication events also amplified PKR genes. For example, spotted green pufferfish (*Tetraodon nigroviridis*) express three PKR genes that are contiguously arranged, which encode three, one, or no dsRBMs, respectively. The closely related pufferfish *Takifugu rubripes* contains two PKR genes, which each encode two dsRBMs. Three in tandem duplicated PKR genes are also found in the frog *Xenopus tropicalis* [13]. PKR and PKZ genes appear to be absent in non-bony fish, including sharks and rays. Taking the presence of PKR genes in all vertebrate lineages and PKZ in some teleost fish into account, it is most likely that PKR evolved in an ancestor of both extant bony fish and tetrapods. PKZ likely evolved after a duplication of a PKR gene occurred, followed by the substitution of the dsRBMs with Zα domains. Later PKR duplications arose independently in some fish and amphibian lineages [13]. Because the Zα domains bind to different forms of nucleic acids than the dsRBMs of PKR, the acquisition of Zα domains in combination with an eIF2α kinase domain likely extended the spectrum of viruses that could be recognized, enhancing their antiviral effects [11]. Both goldfish PKR and PKZ can inhibit replication of grass carp reovirus, and both eIF2α kinases are suggested to function cooperatively for their antiviral functions [133]. Another advantage of the presence of multiple copies of antiviral eIF2α kinases is that the differences could render certain antiviral molecules that directly target PKR or PKZ less effective, which as a consequence would better restrict viral replication.

8 Diversity of eIF2α Kinases in Protozoan Parasites

Recent research has described the diverse eIF2α kinases that are expressed in protozoan parasites and their roles in the stress resistance and differentiation processes. A striking feature of this collection is the large numbers of different eIF2α kinases in protozoan including new family members and their large size with protein kinase inserts of extended length. These ideas are illustrated in *Toxoplasma gondii*, which can infect cells from virtually any warm-blooded vertebrate and cause toxoplasmosis in humans. *Toxoplasma* expresses four different eIF2α kinases designated TgIF2 K-A to -D [134–137]. The eIF2α kinases TgIF2 K-C and -D are related to GCN2 and function to manage nutrient depletion and overcome exposure to the extracellular environment upon release or egress from the host cells, respectively [134, 135]. eIF2 K-A possesses a transmembrane segment and is situated at the ER where it is suggested to function analogously to PERK, providing *Toxoplasma* resistance to disruptions in this organelle (Sullivan et al. [137, 138]. The final variant eIF2 K-B is suggested to be a novel eIF2α kinase. The eIF2 K-B has a predicted molecular weight of 278 kDa, with an 866 residue insert predicted to be in the N-terminus and a second insert 700 residues in length situated in the predicted activation loop in the C-terminal lobe [136]. While the TgIF2 K-B has been shown to be an eIF2α kinase biochemically, there is no homology in the insert sequences or in the regions flanking the kinase domain to suggest how it is regulated by stress. *Toxoplasma* has two developmental stages, a proliferating tachyzoite and encysted bradyzoite, which is induced upon stress. Phosphorylation of the *Toxoplasma* eIF2α occurs upon differentiation into the quiescent bradyzoite, suggesting that one or more of these eIF2α kinases are critical for the developmental program for this parasite [136]. These findings suggest that small molecules that alter the status of *Toxoplasma* eIF2α phosphorylation may be effective treatment strategies for toxoplasmosis [139, 140].

Similar themes are also found in the related parasite, *Plasmodium falciparum*, which causes malaria. The Plasmodium eIF2α kinases eIK1, related to GCN2, and PK4, which shares features with PERK, are expressed in the asexual blood stages, with PK4 being required for the erythrocytic cycle that facilitates disease [141, 142]. By comparison, the *Plasmodium* eIK2 is largely expressed in salivary sporozoites, which are injected by the mosquito vector into the bloodstream of the infected host [143]. Three eIF2α kinases are also expressed in each of the pathogenic trypanosomatids, *Trypanosoma cruzi*, *Trypanosoma brucei,* and *Leishmania* major, which cause Chagas' disease, African sleeping sickness, and cutaneous leishmaniasis, respectively [144]. In *T. cruzi*, phosphorylation of the parasite eIF2α is suggested to be required for differentiation of the non-infective epimastigote into infective metacyclic trypomastigotes, a process suggested to involve parasite starvation for nutrients [145]. Nutrients are stored in the endosome of *T. cruzi*, and the eIF2α kinase TcK2 is suggested to be situated in the membranes of this compartment where it senses cytosolic heme, which is obtained by the parasite during blood

meals from the host [146]. If the heme levels are sufficiently depleted in the cytosol, activation of TcK2 and the attendant lowering of protein synthesis triggers differentiation to the infective form of the parasite. In this regard, each of the protozoan parasites has adapted the utility of eIF2α kinases for sensing diverse stresses to coordinate optimal patterns of parasite infection, growth, and differentiation.

Acknowledgments The authors acknowledge support by NIH grant GM049164 (RCW and MMG) and AI114851 (SR) and the Grace M. Showalter Research Trust Fund (RCW).

References

1. Hinnebusch AG. The scanning mechanism of eukaryotic translation initiation. Annu Rev Biochem. 2014;83:779–812.
2. Jackson RJ, Hellen CU, Pestova TV. The mechanism of eukaryotic translation initiation and principles of its regulation. Nat Rev Mol Cell Biol. 2010;11:113–27.
3. Baird TD, Wek RC. Eukaryotic initiation factor 2 phosphorylation and translational control in metabolism. Adv Nutr. 2012;3:307–21.
4. Abastado JP, Miller PF, Jackson BM, Hinnebusch AG. Suppression of ribosomal reinitiation at upstream open reading frames in amino acid-starved cells forms the basis of *GCN4* translational control. Mol Cell Biol. 1991;11:486–96.
5. Harding HP, Novoa I, Zhang Y, Zeng H, Wek R, Schapira M, Ron D. Regulated translation initiation controls stress-induced gene expression in mammalian cells. Mol Cell. 2000;6:1099–108.
6. Hinnebusch AG. Translational regulation of GCN4 and the general amino acid control of yeast. Annu Rev Microbiol. 2005;59:407–50.
7. Vattem KM, Wek RC. Reinitiation involving upstream open reading frames regulates *ATF4* mRNA translation in mammalian cells. Proc Natl Acad Sci USA. 2004;101:11269–74.
8. Gilbert WV. Alternative ways to think about cellular internal ribosome entry. J Biol Chem. 2010;285:29033–8.
9. Deng J, Lu PD, Zhang Y, Scheuner D, Kaufman RJ, Sonenberg N, Harding HP, Ron D. Translational repression mediates activation of nuclear factor kappa B by phosphorylated translation initiation factor 2. Mol Cell Biol. 2004;24:10161–8.
10. Jiang HY, Wek RC. Gcn2 phosphorylation of eIF2α activates NF-κB in response to UV irradiation. Biochem J. 2005;385:371–80.
11. Rothenburg S, Deigendesch N, Dittmar K, Koch-Nolte F, Haag F, Lowenhaupt K, Rich A. A PKR-like eukaryotic initiation factor 2alpha kinase from zebrafish contains Z-DNA binding domains instead of dsRNA binding domains. Proc Natl Acad Sci USA. 2005;102:1602–7.
12. Rothenburg S, Seo EJ, Gibbs JS, Dever TE, Dittmar K. Rapid evolution of protein kinase PKR alters sensitivity to viral inhibitors. Nat Struct Mol Biol. 2009;16:63–70.
13. Rothenburg S, Deigendesch N, Dey M, Dever TE, Tazi L. Double-stranded RNA-activated protein kinase PKR of fishes and amphibians: varying the number of double-stranded RNA binding domains and lineage-specific duplications. BMC Biol. 2008;6:12.
14. Elde NC, Child SJ, Geballe AP, Malik HS. Protein kinase R reveals an evolutionary model for defeating viral mimicry. Nature. 2009;457:485–9.
15. Axten JM, Medina JR, Feng Y, Shu A, Romeril SP, Grant SW, Li WH, Heerding DA, Minthorn E, Mencken T, Atkins C, Liu Q, Rabindran S, Kumar R, Hong X, Goetz A, Stanley T, Taylor JD, Sigethy SD, Tomberlin GH, Hassell AM, Kahler KM, Shewchuk LM,

Gampe RT. Discovery of 7-methyl-5-(1-{[3-(trifluoromethyl)phenyl]acetyl}-2,3-dihydro-1H-indol-5-yl)-7H-p yrrolo[2,3-d]pyrimidin-4-amine (GSK2606414), a potent and selective first-in-class inhibitor of protein kinase R (PKR)-like endoplasmic reticulum kinase (PERK). J Med Chem. 2012;55:7193–207.
16. Cui W, Li J, Ron D, Sha B. The structure of the PERK kinase domain suggests the mechanism for its activation. Acta Crystallogr D Biol Crystallogr. 2011;67:423–8.
17. Dar AC, Dever TE, Sicheri F. Higher-order substrate recognition of eIF2α by the RNA-dependent protein kinase PKR. Cell. 2005;122:887–900.
18. Padyana AK, Qiu H, Roll-Mecak A, Hinnebusch AG, Burley SK. Structural basis for autoinhibition and mutational activation of eukaryotic initiation factor 2alpha protein kinase GCN2. J Biol Chem. 2005;280:29289–99.
19. Cai R, Williams BR. Mutations in the double-stranded RNA-activated protein kinase insert region that uncouple catalysis from eIF2alpha binding. J Biol Chem. 1998;273:11274–80.
20. Craig AW, Cosentino GP, Donze O, Sonenberg N. The kinase insert domain of interferon-induced protein kinase PKR is required for activity but not for interaction with the pseudosubstrate K3L. J Biol Chem. 1996;271:24526–33.
21. Wek RC, Jackson BM, Hinnebusch AG. Juxtaposition of domains homologous to protein kinases and histidyl-tRNA synthetases in GCN2 protein suggests a mechanism for coupling *GCN4* expression to amino acid availability. Proc Natl Acad Sci USA. 1989;86:4579–83.
22. Dey M, Cao C, Dar AC, Tamura T, Ozato K, Sicheri F, Dever TE. Mechanistic link between PKR dimerization, autophosphorylation, and eIF2α substrate recognition. Cell. 2005;122:901–13.
23. Knighton DR, Zheng JH, Ten Eyck LF, Ashford VA, Xuong NH, Taylor SS, Sowadski JM. Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. Science. 1991;253:407–14.
24. Beattie E, Tartaglia J, Paoletti E. Vaccinia virus-encoded eIF-2 alpha homolog abrogates the antiviral effect of interferon. Virology. 1991;183:419–22.
25. Carroll K, Elroy-Stein O, Moss B, Jagus R. Recombinant vaccinia virus K3L gene product prevents activation of double-stranded RNA-dependent, initiation factor 2 alpha-specific protein kinase. J Biol Chem. 1993;268:12837–42.
26. Dar AC, Sicheri F. X-ray crystal structure and functional analysis of vaccinia virus K3L reveals molecular determinants for PKR subversion and substrate recognition. Mol Cell. 2002;10:295–305.
27. Davies MV, Elroy-Stein O, Jagus R, Moss B, Kaufman RJ. The vaccinia virus K3L gene product potentiates translation by inhibiting double-stranded-RNA-activated protein kinase and phosphorylation of the alpha subunit of eukaryotic initiation factor 2. J Virol. 1992;66:1943–50.
28. de Vos AM, Ultsch M, Kossiakoff AA. Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. Science. 1992;255:306–12.
29. Plotnikov AN, Schlessinger J, Hubbard SR, Mohammadi M. Structural basis for FGF receptor dimerization and activation. Cell. 1999;98:641–50.
30. Lageix S, Rothenburg S, Dever TE, Hinnebusch AG. Enhanced interaction between pseudokinase and kinase domains in Gcn2 stimulates eIF2alpha phosphorylation in starved cells. PLoS Genet. 2014;10:e1004326.
31. Carrara M, Prischi F, Nowak PR, Ali MM. Crystal structures reveal transient PERK luminal domain tetramerization in endoplasmic reticulum stress signaling. EMBO J. 2015;34:1589–600.
32. Cole JL. Activation of PKR: an open and shut case? Trends Biochem Sci. 2007;32(2):57–62.
33. Husain B, Hesler S, Cole JL. Regulation of PKR by RNA: formation of active and inactive dimers. Biochemistry. 2015;54:6663–72.
34. Proud CG. PKR: a new name and new roles. Trends Biochem Sci. 1995;20:217–56.
35. Robertson HD, Mathews MB. The regulation of the protein kinase PKR by RNA. Biochimie. 1996;78:909–14.

36. Nameki N, Yoneyama M, Koshiba S, Tochio N, Inoue M, Seki E, Matsuda T, Tomo Y, Harada T, Saito K, Kobayashi N, Yabuki T, Aoki M, Nunokawa E, Matsuda N, Sakagami N, Terada T, Shirouzu M, Yoshida M, Hirota H, Osanai T, Tanaka A, Arakawa T, Carninci P, Kawai J, Hayashizaki Y, Kinoshita K, Guntert P, Kigawa T, Yokoyama S. Solution structure of the RWD domain of the mouse GCN2 protein. Protein Sci. 2004;13:2089–100.
37. He H, Singh I, Wek SA, Dey S, Baird TD, Wek RC, Georgiadis MM. Crystal structures of GCN2 C-terminal domains suggest regulatory differences in yeast and mammals. J Biol Chem. 2014;289:15023–34.
38. Garcia-Barrio M, Dong J, Ufano S, Hinnebusch AG. Association of GCN1-GCN20 regulatory complex with the N-terminus of eIF2alpha kinase GCN2 is required for GCN2 activation. EMBO J. 2000;19:1887–99.
39. Narasimhan J, Staschke KA, Wek RC. Dimerization is required for activation of eIF2 kinase Gcn2 in response to diverse environmental stress conditions. J Biol Chem. 2004;279: 22820–32.
40. Qiu H, Garcia-Barrio MT, Hinnebusch AG. Dimerization by translation initiation factor 2 kinase GCN2 is mediated by interactions of the C-terminal ribosome binding region and the protein kinase domain. Mol Cell Biol. 1998;18:2697–711.
41. Wek RC, Ramirez M, Jackson BM, Hinnebusch AG. Identification of positive-acting domains in GCN2 protein kinase required for translational activation of *GCN4* expression. Mol Cell Biol. 1990;10:2820–31.
42. Marton MJ, Vazquez de Aldana CR, Qiu H, Chakraburtty K, Hinnebusch AG. Evidence that GCN1 and GCN20, translational regulators of GCN4, function on elongating ribosomes in activation of eIF2alpha kinase GCN2. Mol Cell Biol. 1997;17:4474–89.
43. Sattlegger E, Hinnebusch AG. Separate domains in GCN1 for binding protein kinase GCN2 and ribosomes are required for GCN2 activation in amino acid-starved cells. EMBO J. 2000;19:6622–33.
44. Vazquez de Aldana CR, Marton MJ, Hinnebusch AG. GCN20, a novel ATP binding cassette protein, and GCN1 reside in a complex that mediates activation of the eIF-2 alpha kinase GCN2 in amino acid-starved cells. EMBO J. 1995;14:3184–99.
45. Pereira CM, Sattlegger E, Jiang HY, Longo BM, Jaqueta CB, Hinnebusch AG, Wek RC, Mello LE, Castilho BA. IMPACT, a protein preferentially expressed in the mouse brain, binds GCN1 and inhibits GCN2 activation. J Biol Chem. 2005;280:28316–23.
46. Roffe M, Hajj GN, Azevedo HF, Alves VS, Castilho BA. IMPACT is a developmentally regulated protein in neurons that opposes the eukaryotic initiation factor 2alpha kinase GCN2 in the modulation of neurite outgrowth. J Biol Chem. 2013;288:10860–9.
47. Murphy JM, Zhang Q, Young SN, Reese ML, Bailey FP, Eyers PA, Ungureanu D, Hammaren H, Silvennoinen O, Varghese LN, Chen K, Tripaydonis A, Jura N, Fukuda K, Qin J, Nimchuk Z, Mudgett MB, Elowe S, Gee CL, Liu L, Daly RJ, Manning G, Babon JJ, Lucet IS. A robust methodology to subclassify pseudokinases based on their nucleotide-binding properties. Biochem J. 2014;457:323–34.
48. Zhu S, Sobolev AY, Wek RC. Histidyl-tRNA synthetase-related sequences in GCN2 protein kinase regulate in vitro phosphorylation of eIF-2. J Biol Chem. 1996;271:24989–94.
49. Deng J, Harding H, Raught B, Gingras A, Berlanga J, Scheuner D, Kaufman R, Ron D, Sonenberg N. Activation of GCN2 in UV-irradiated cells inhibits translation. Curr Biol. 2002;12:1279–86.
50. Goosens A, Dever TE, Pascual-Ahuir A, Serrano R. The protein kinase Gcn2p mediates sodium toxicity in yeast. J Biol Chem. 2001;276:30753–60.
51. Yang R, Wek SA, Wek RC. Glucose limitation induces *GCN4* translation by activation of Gcn2 protein kinase. Mol Cell Biol. 2000;20:2706–17.
52. Zaborske JM, Narasimhan J, Jiang L, Wek SA, Dittmar KA, Freimoser F, Pan T, Wek RC. Genome-wide analysis of tRNA charging and activation of the eIF2 kinase Gcn2p. J Biol Chem. 2009;284:25254–67.
53. Hinnebusch AG. A hierarchy of trans-acting factors modulates translation of an activator of amino acid biosynthetic genes in Saccharomyces cerevisiae. Mol Cell Biol. 1985;5:2349–60.

54. Lageix S, Lanet E, Pouch-Pelissier MN, Espagnol MC, Robaglia C, Deragon JM, Pelissier T. Arabidopsis eIF2alpha kinase GCN2 is essential for growth in stress conditions and is activated by wounding. BMC Plant Biol. 2008;8:134.
55. Liu X, Merchant A, Rockett KS, McCormack M, Pajerowska-Mukhtar KM. Characterization of Arabidopsis thaliana GCN2 kinase roles in seed germination and plant development. Plant Signal Behav. 2015;10:e992264.
56. Anthony TG, McDaniel BJ, McGrath BC, Cavener DR, McNurlan MA, Wek RC. Preservation of liver mass during dietary leucine starvation occurs at the expense of skeletal muscle mass in mice deleted for eIF2 kinase GCN2. J Biol Chem. 2004;279:36553–61.
57. Hao S, Sharp JW, Ross-Inta CM, McDaniel BJ, Anthony TG, Wek RC, Cavener DR, McGrath BC, Rudell JB, Koehnle TJ, Gietzen DW. Uncharged tRNA and sensing of amino acid deficiency in mammalian piriform cortex. Science. 2005;307:1776–8.
58. Maurin AC, Benani A, Lorsignol A, Brenachot X, Parry L, Carraro V, Guissard C, Averous J, Jousse C, Bruhat A, Chaveroux C, B'Chir W, Muranishi Y, Ron D, Penicaud L, Fafournoux P. Hypothalamic eIF2alpha signaling regulates food intake. Cell Rep. 2014;6:438–44.
59. Maurin AC, Jousse C, Averous J, Parry L, Bruhat A, Cherasse Y, Zeng H, Zhang Y, Harding HP, Ron D, Fafournoux P. The GCN2 kinase biases feeding behavior to maintain amino acid homeostasis in omnivores. Cell Metab. 2005;1:273–7.
60. Best DH, Sumner KL, Austin ED, Chung WK, Brown LM, Borczuk AC, Rosenzweig EB, Bayrak-Toydemir P, Mao R, Cahill BC, Tazelaar HD, Leslie KO, Hemnes AR, Robbins IM, Elliott CG. EIF2AK4 mutations in pulmonary capillary hemangiomatosis. Chest. 2014;145:231–6.
61. Eyries M, Montani D, Girerd B, Perret C, Leroy A, Lonjou C, Chelghoum N, Coulet F, Bonnet D, Dorfmuller P, Fadel E, Sitbon O, Simonneau G, Tregouet DA, Humbert M, Soubrier F. EIF2AK4 mutations cause pulmonary veno-occlusive disease, a recessive form of pulmonary hypertension. Nat Genet. 2014;46:65–9.
62. Harding HP, Zhang Y, Ron D. Protein translation and folding are coupled by an endoplasmic-reticulum-resident kinase. Nature. 1999;397:271–4.
63. Shi Y, Vattem KM, Sood R, An J, Liang J, Stramm L, Wek RC. Identification and characterization of pancreatic eukaryotic initiation factor 2 a-subunit kinase, PEK, involved in translation control. Mol Cell Biol. 1998;18:7499–509.
64. Bertolotti A, Zhang Y, Hendershot LM, Harding HP, Ron D. Dynamic interaction of BiP and ER stress transducers in the unfolded protein response. Nat Cell Biol. 2000;2:326–32.
65. Gardner BM, Pincus D, Gotthardt K, Gallagher CM, Walter P. Endoplasmic reticulum stress sensing in the unfolded protein response. Cold Spring Harb Perspect Biol. 2013;5:a013169.
66. Liu CY, Schroder M, Kaufman RJ. Ligand-independent dimerization activates the stress response kinases IRE1 and PERK in the lumen of the endoplasmic reticulum. J Biol Chem. 2000;275:24881–5.
67. Ma K, Vattem KM, Wek RC. Dimerization and release of molecular chaperone inhibition facilitate activation of eukaryotic initiation factor-2 kinase in response to endoplasmic reticulum stress. J Biol Chem. 2002;277:18728–35.
68. Han J, Back SH, Hur J, Lin YH, Gildersleeve R, Shan J, Yuan CL, Krokowski D, Wang S, Hatzoglou M, Kilberg MS, Sartor MA, Kaufman RJ. ER-stress-induced transcriptional regulation increases protein synthesis leading to cell death. Nat Cell Biol. 2013;15:481–90.
69. Harding HP, Zhang Y, Zeng H, Novoa I, Lu PD, Calfon M, Sadri N, Yun C, Popko B, Paules R, Stojdl DF, Bell JC, Hettmann T, Leiden JM, Ron D. An integrated stress response regulates amino acid metabolism and resistance to oxidative stress. Mol Cell. 2003;11:619–33.
70. Lu PD, Jousse C, Marciniak SJ, Zhang Y, Novoa I, Scheuner D, Kaufman RJ, Ron D, Harding HP. Cytoprotection by pre-emptive conditional phosphorylation of translation initiation factor 2. EMBO J. 2004;23:169–79.
71. Marciniak SJ, Yun CY, Oyadomari S, Novoa I, Zhang Y, Jungreis R, Nagata K, Harding HP, Ron D. CHOP induces death by promoting protein synthesis and oxidation in the stressed endoplasmic reticulum. Genes Dev. 2004;18:3066–77.

72. Teske BF, Fusakio ME, Zhou D, Shan J, McClintick JN, Kilberg MS, Wek RC. CHOP induces activating transcription factor 5 (ATF5) to trigger apoptosis in response to perturbations in protein homeostasis. Mol Biol Cell. 2013;24:2477–90.
73. Teske BF, Wek SA, Bunpo P, Cundiff JK, McClintick JN, Anthony TG, Wek RC. The eIF2 kinase PERK and the integrated stress response facilitate activation of ATF6 during endoplasmic reticulum stress. Mol Biol Cell. 2011;22:4390–405.
74. Gardner BM, Walter P. Unfolded proteins are Ire1-activating ligands that directly induce the unfolded protein response. Science. 2011;333:1891–4.
75. Korennykh A, Walter P. Structural basis of the unfolded protein response. Annu Rev Cell Dev Biol. 2012;28:251–77.
76. Carrara M, Prischi F, Nowak PR, Kopp MC, Ali MM. Noncanonical binding of BiP ATPase domain to Ire1 and Perk is dissociated by unfolded protein CH1 to initiate ER stress signaling. Elife. 2015;4:e03522.
77. Delepine M, Nicolino M, Barrett T, Golamaully M, Lathrop GM, Julier C. EIF2AK3, encoding translation initiation factor 2-a kinase 3, is mutated in patients with Wolcott-Rallison syndrome. Nat Genet. 2000;25:406–9.
78. Senée V, Vattem KM, Delépine M, Rainbow L, Haton C, Lecoq A, Shaw N, Robert J-J, Rooman R, Diatloff-Zito C, Michaud JL, Bin-Abbas B, Taha D, Zabel B, Franceschini P, Topaloglu AK, Lathrop M, Barrett T, Nicolino M, Wek RC, Julier C. Wolcott-Rallison syndrome: clinical, genetic, and functional study of *EIF2AK3* mutations, and suggestion of genetic heterogeneity. Diabetes. 2004;53:1876–83.
79. Wolcott CD, Rallison ML. Infancy-onset diabetes mellitus and multiple epiphyseal dysplasia. J Pediatr. 1972;80:292–7.
80. Harding H, Zeng H, Zhang Y, Jungreis R, Chung P, Plesken H, Sabatini DD, Ron D. Diabetes mellitus and exocrine pancreatic dysfunction in *Perk -/-* mice reveals a role for translational control in secretory cell survival. Mol Cell. 2001;7:1153–63.
81. Zhang P, McGrath BC, Reinert J, Olsen DS, Lei L, Gill S, Wek SA, Vattem KM, Wek RC, Kimball SR, Jefferson LS, Cavener DR. The GCN2 eIF2alpha kinase is required for adaptation to amino acid deprivation in mice. Mol Cell Biol. 2002;22:6681–8.
82. Rafie-Kolpin M, Chefalo PJ, Hussain Z, Hahn J, Uma S, Matts RL, Chen JJ. Two heme-binding domains of heme-regulated eukaryotic initiation factor-2alpha kinase. N terminus and kinase insertion. J Biol Chem. 2000;275(7):5171–8.
83. Girvan HM, Munro AW. Heme sensor proteins. J Biol Chem. 2013;288:13194–203.
84. Crosby JS, Lee K, London IM, Chen JJ. Erythroid expression of the heme-regulated eIF-2 alpha kinase. Mol Cell Biol. 1994;14:3906–14.
85. Chen JJ. Regulation of protein synthesis by the heme-regulated eIF2alpha kinase: relevance to anemias. Blood. 2007;109:2693–9.
86. Han A, Yu C, Lu L, Fujiwara Y, Browne C, Chin G, Fleming P, Leboulch P, Orkin SH, Chen J-J. Heme-regulated eIF2a kinase (HRI) is required for translational regulation and survival of erythroid precursors in iron deficiency. EMBO J. 2001;20:6909–18.
87. Lu L, Han AP, Chen J-J. Translation initiation control by heme-regulated eukaryotic initiation factor 2a kinase in erythroid cells under cytoplasmic stresses. Mol Cell Biol. 2001;21:7971–80.
88. Zhan K, Narasimhan J, Wek RC. Differential activation of eIF2 kinases in response to cellular stresses in *Schizosaccharomyces pombe*. Genetics. 2004;168:1867–75.
89. Zhan K, Vattem KM, Bauer BN, Dever TE, Chen J-J, Wek RC. Phosphorylation of eukaryotic initiation factor -2 by HRI-related protein kinases in *Schizosaccharomyces pombe* is important for resistance to environmental stresses. Mol Cell Biol. 2002;22:7134–46.
90. Meurs E, Chong K, Galabru J, Thomas NSB, Kerr IM, Williams BRG, Hovanessian AG. Molecular cloning and characterization of the human double-stranded RNA-activated protein kinase induced by interferon. Cell. 1990;62:379–90.
91. Green SR, Mathews MB. Two RNA-binding motifs in the double-stranded RNA-activated protein kinase, DAI. Genes Dev. 1992;6:2478–90.

92. Nanduri S, Carpick BW, Yang Y, Williams BR, Qin J. Structure of the double-stranded RNA-binding domain of the protein kinase PKR reveals the molecular basis of its dsRNA-mediated activation. EMBO J. 1998;17:5458–65.
93. Nanduri S, Rahman F, Williams BRG, Qin J. A dynamically tuned double-stranded RNA binding mechanism for the activation of antiviral kinase PKR. EMBO J. 2000;19:5567–74.
94. St Johnston D, Brown NH, Gall JG, Jantsch M. A conserved double-stranded RNA-binding domain. Proc Natl Acad Sci USA. 1992;89(22):10979–83.
95. Masliah G, Barraud P, Allain FH. RNA recognition by double-stranded RNA binding domains: a matter of shape and sequence. Cell Mol Life Sci. 2013;70:1875–95.
96. Besse S, Rebouillat D, Marie I, Puvion-Dutilleul F, Hovanessian AG. Ultrastructural localization of interferon-inducible double-stranded RNA-activated enzymes in human cells. Exp Cell Res. 1998;239:379–92.
97. Garcia MA, Meurs EF, Esteban M. The dsRNA protein kinase PKR: virus and cell control. Biochimie. 2007;89:799–811.
98. Jeffrey IW, Kadereit S, Meurs EF, Metzger T, Bachmann M, Schwemmle M, Hovanessian AG, Clemens MJ. Nuclear localization of the interferon-inducible protein kinase PKR in human cells and transfected mouse cells. Exp Cell Res. 1995;218:17–27.
99. Ung TL, Cao C, Lu J, Ozato K, Dever TE. Heterologous dimerization domains functionally substitute for the double-stranded RNA binding domains of the kinase PKR. EMBO J. 2001;20:3728–37.
100. Vattem K, Staschke KA, Wek RC. Mechanism of activation of the double-stranded-RNA-dependent protein kinase, PKR: Role of dimerization and cellular localization in the stimulation of PKR phosphorylation of eukaryotic initiation factor -2α (eIF2α). Eur J Biochem. 2001;268:3674–84.
101. Barber GN. The dsRNA-dependent protein kinase, PKR and cell death. Cell Death Differ. 2005;12(6):563–70.
102. Gil J, Alcami J, Esteban M. Induction of apoptosis by double-stranded-RNA-dependent protein kinase (PKR) involves the alpha subunit of eukaryotic translation initiation factor 2 and NF-kappaB. Mol Cell Biol. 1999;19:4653–63.
103. Clarke PA, Mathews MB. Interactions between the double-stranded RNA binding motif and RNA: definition of the binding site for the interferon-induced protein kinase DAI (PKR) on adenovirus VA RNA. RNA. 1995;1:7–20.
104. Laing KG, Elia A, Jeffrey I, Matys V, Tilleray VJ, Souberbielle B, Clemens MJ. In vivo effects of the Epstein-Barr virus small RNA EBER-1 on protein synthesis and cell growth regulation. Virology. 2002;297:253–69.
105. McKenna SA, Kim I, Liu CW, Puglisi JD. Uncoupling of RNA binding and PKR kinase activation by viral inhibitor RNAs. J Mol Biol. 2006;358:1270–85.
106. Nanbo A, Yoshiyama H, Takada K. Epstein-Barr virus-encoded poly(A)- RNA confers resistance to apoptosis mediated through Fas by blocking the PKR pathway in human epithelial intestine 407 cells. J Virol. 2005;79:12280–5.
107. Schneider RJ, Safer B, Munemitsu SM, Samuel CE, Shenk T. Adenovirus VAI RNA prevents phosphorylation of the eukaryotic initiation factor 2 alpha subunit subsequent to infection. Proc Natl Acad Sci USA. 1985;82:4321–5.
108. Sharp TV, Xiao Q, Jeffrey I, Gewert DR, Clemens MJ. Reversal of the double-stranded-RNA-induced inhibition of protein synthesis by a catalytically inactive mutant of the protein kinase PKR. Eur J Biochem. 1993;214:945–8.
109. Brandt TA, Jacobs BL. Both carboxy- and amino-terminal domains of the vaccinia virus interferon resistance gene, E3L, are required for pathogenesis in a mouse model. J Virol. 2001;75:850–6.
110. Davies MV, Chang HW, Jacobs BL, Kaufman RJ. The E3L and K3L vaccinia virus gene products stimulate translation through inhibition of the double-stranded RNA-dependent protein kinase by different mechanisms. J Virol. 1993;67:1688–92.
111. Langland JO, Jacobs BL. The role of the PKR-inhibitory genes, E3L and K3L, in determining vaccinia virus host range. Virology. 2002;299:133–41.

112. Langland JO, Jacobs BL. Inhibition of PKR by vaccinia virus: role of the N- and C-terminal domains of E3L. Virology. 2004;324:419–29.
113. Sharp TV, Moonan F, Romashko A, Joshi B, Barber GN, Jagus R. The vaccinia virus E3L gene product interacts with both the regulatory and the substrate binding regions of PKR: implications for PKR autoregulation. Virology. 1998;250:302–15.
114. Zhang P, Jacobs BL, Samuel CE. Loss of protein kinase PKR expression in human HeLa cells complements the vaccinia virus E3L deletion mutant phenotype by restoration of viral protein synthesis. J Virol. 2008;82:840–8.
115. Li S, Peters GA, Ding K, Zhang X, Qin J, Sen GC. Molecular basis for PKR activation by PACT or dsRNA. Proc Natl Acad Sci USA. 2006;103:10005–10.
116. Patel RC, Sen GC. PACT, a protein activator of the interferon-induced protein kinase, PKR. EMBO J. 1998;17:4379–90.
117. Balachandran S, Roberts PC, Brown LE, Truong H, Pattnaik AK, Archer DR, Barber GN. Essential role for the dsRNA-dependent protein kinase PKR in innate immunity to viral infection. Immunity. 2000;13:129–41.
118. Stojdl DF, Abraham N, Knowles S, Marius R, Brasey A, Lichty BD, Brown EG, Sonenberg N, Bell JC. The murine double-stranded RNA-dependent protein kinase PKR is required for resistance to vesicular stomatitis virus. J Virol. 2000;74:9580–5.
119. White SD, Jacobs BL. The amino terminus of the vaccinia virus E3 protein is necessary to inhibit the interferon response. J Virol. 2012;86:5895–904.
120. Langland JO, Cameron JM, Heck MC, Jancovich JK, Jacobs BL. Inhibition of PKR by RNA and DNA viruses. Virus Res. 2006;119:100–10.
121. Rothenburg S, Chinchar VG, Dever TE. Characterization of a ranavirus inhibitor of the antiviral protein kinase PKR. BMC Microbiol. 2011;11:56.
122. Peng C, Haller SL, Rahman MM, McFadden G, Rothenburg S. Myxoma virus M156 is a specific inhibitor of rabbit PKR but contains a loss-of-function mutation in Australian virus isolates. Proc Natl Acad Sci USA. 2016;38:55–60.
123. Garner JN, Joshi B, Jagus R. Characterization of rainbow trout and zebrafish eukaryotic initiation factor 2alpha and its response to endoplasmic reticulum stress and IPNV infection. Dev Comp Immunol. 2003;27:217–31.
124. Hu CY, Zhang YB, Huang GP, Zhang QY, Gui JF. Molecular cloning and characterisation of a fish PKR-like gene from cultured CAB cells induced by UV-inactivated virus. Fish Shellfish Immunol. 2004;17:353–66.
125. Kim D, Hur J, Park K, Bae S, Shin D, Ha SC, Hwang HY, Hohng S, Lee JH, Lee S, Kim YG, Kim KK. Distinct Z-DNA binding mode of a PKR-like protein kinase containing a Z-DNA binding domain (PKZ). Nucleic Acids Res. 2014;42:5937–48.
126. Herbert A, Alfken J, Kim YG, Mian IS, Nishikura K, Rich A. A Z-DNA binding domain present in the human editing enzyme, double-stranded RNA adenosine deaminase. Proc Natl Acad Sci USA. 1997;94:8421–6.
127. Kim YG, Muralinath M, Brandt T, Pearcy M, Hauns K, Lowenhaupt K, Jacobs BL, Rich A. A role for Z-DNA binding in vaccinia virus pathogenesis. Proc Natl Acad Sci USA. 2003;100:6974–9.
128. Schwartz T, Behlke J, Lowenhaupt K, Heinemann U, Rich A. Structure of the DLM-1-Z-DNA complex reveals a conserved family of Z-DNA-binding proteins. Nat Struct Biol. 2001;8:761–5.
129. Tome AR, Kus K, Correia S, Paulo LM, Zacarias S, de Rosa M, Figueiredo D, Parkhouse RM, Athanasiadis A. Crystal structure of a poxvirus-like zalpha domain from cyprinid herpesvirus 3. J Virol. 2013;87:3998–4004.
130. Bergan V, Jagus R, Lauksund S, Kileng O, Robertsen B. The Atlantic salmon Z-DNA binding protein kinase phosphorylates translation initiation factor 2 alpha and constitutes a unique orthologue to the mammalian dsRNA-activated protein kinase R. FEBS J. 2008;275:184–97.
131. Su J, Zhu Z, Wang Y. Molecular cloning, characterization and expression analysis of the PKZ gene in rare minnow Gobiocypris rarus. Fish Shellfish Immunol. 2008;25:106–13.

132. Yang PJ, Wu CX, Li W, Fan LH, Lin G, Hu CY. Cloning and functional analysis of PKZ (PKR-like) from grass carp (*Ctenopharyngodon idellus*). Fish Shellfish Immunol. 2011;31:1173–8.
133. Liu TK, Zhang YB, Liu Y, Sun F, Gui JF. Cooperative roles of fish protein kinase containing Z-DNA binding domains and double-stranded RNA-dependent protein kinase in interferon-mediated antiviral response. J Virol. 2011;85:12769–80.
134. Konrad C, Wek RC, Sullivan WJ Jr. A GCN2-like eukaryotic initiation factor 2 kinase increases the viability of extracellular Toxoplasma gondii parasites. Eukaryot Cell. 2011;10:1403–12.
135. Konrad C, Wek RC, Sullivan WJ Jr. GCN2-like eIF2alpha kinase manages the amino acid starvation response in Toxoplasma gondii. Int J Parasitol. 2014;44:139–46.
136. Narasimhan J, Joyce BR, Naguleswaran A, Smith AT, Livingston MR, Dixon SE, Coppens I, Wek RC, Sullivan WJ Jr. Translation regulation by eukaryotic initiation factor-2 kinases in the development of latent cysts in Toxoplasma gondii. J Biol Chem. 2008;283:16591–601.
137. Sullivan WJ Jr, Narasimhan J, Bhatti MM, Wek RC. Parasite-specific eukaryotic initiation factor -2 (eIF2) kinase required for stress-induced translation control. Biochem J. 2004; 380:523–31.
138. Joyce BR, Tampaki Z, Kim K, Wek RC, Sullivan WJ Jr. The unfolded protein response in the protozoan parasite Toxoplasma gondii features translational and transcriptional control. Eukaryot Cell. 2013;12:979–89.
139. Benmerzouga I, Checkley LA, Ferdig MT, Arrizabalaga G, Wek RC, Sullivan WJ Jr. Guanabenz repurposed as an antiparasitic with activity against acute and latent toxoplasmosis. Antimicrob Agents Chemother. 2015;59:6939–45.
140. Konrad C, Queener SF, Wek RC, Sullivan WJ Jr. Inhibitors of eIF2alpha dephosphorylation slow replication and stabilize latency in Toxoplasma gondii. Antimicrob Agents Chemother. 2013;57:1815–22.
141. Fennell C, Babbitt S, Russo I, Wilkes J, Ranford-Cartwright L, Goldberg DE, Doerig C. PfeIK1, a eukaryotic initiation factor 2alpha kinase of the human malaria parasite Plasmodium falciparum, regulates stress-response to amino-acid starvation. Malar J. 2009;8:99.
142. Zhang M, Mishra S, Sakthivel R, Rojas M, Ranjan R, Sullivan WJ Jr, Fontoura BM, Menard R, Dever TE, Nussenzweig V. PK4, a eukaryotic initiation factor 2alpha(eIF2alpha) kinase, is essential for the development of the erythrocytic cycle of Plasmodium. Proc Natl Acad Sci USA. 2012;109:3956–61.
143. Zhang M, Fennell C, Ranford-Cartwright L, Sakthivel R, Gueirard P, Meister S, Caspi A, Doerig C, Nussenzweig RS, Tuteja R, Sullivan WJ Jr, Roos DS, Fontoura BM, Menard R, Winzeler EA, Nussenzweig V. The *Plasmodium* eukaryotic initiation factor-2alpha kinase IK2 controls the latency of sporozoites in the mosquito salivary glands. J Exp Med. 2010;207:1465–74.
144. Moraes MC, Jesus TC, Hashimoto NN, Dey M, Schwartz KJ, Alves VS, Avila CC, Bangs JD, Dever TE, Schenkman S, Castilho BA. Novel membrane-bound eIF2alpha kinase in the flagellar pocket of *Trypanosoma brucei*. Eukaryot Cell. 2007;6:1979–91.
145. Tonelli RR, Augusto Lda S, Castilho BA, Schenkman S. Protein synthesis attenuation by phosphorylation of eIF2alpha is required for the differentiation of Trypanosoma cruzi into infective forms. PLoS ONE. 2011;6:e27904.
146. da Silva Augusto L, Moretti NS, Ramos TC, de Jesus TC, Zhang M, Castilho BA, Schenkman S. A membrane-bound eIF2 alpha kinase located in endosomes is regulated by heme and controls differentiation and ROS levels in Trypanosoma cruzi. PLoS Pathog. 2015;11:e1004618.
147. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004;32:1792–7.
148. Wilgenbusch JC, Swofford D. Inferring evolutionary trees with PAUP*. Curr Protoc Bioinformatics Chapter 6: Unit 6.4 (2003).

eIF2α Kinases and the Evolution of Stress Response in Eukaryotes

Juan José Berlanga, César de Haro, Miguel A. Rodríguez-Gabriel and Iván Ventoso

1 The Origin of eIF2

Translation initiates with the recognition of the initiation codon in mRNA. This is a rate-limiting step in translation that involves the transfer of Met-tRNAi to the small ribosomal subunit for codon-anticodon pairing. To achieve this recognition, the ribosomes require the participation of GTPase-bearing proteins that bring the Met-tRNAi (fMet-tRNA in bacteria) to the small ribosomal subunit (30S in bacteria and archaea, 40S in eukaryotes) [1–3]. Upon codon recognition, the hydrolysis of GTP triggers the release of initiation factors and the joining of large ribosomal subunit (50S in bacteria and archaea, 60S in eukaryotes). Bacteria and eukaryotes significantly differ in the way they transfer Met-tRNAi to the ribosome (Fig. 1). Whereas in bacteria this process requires a monomeric GTPase that binds first the ribosome and then accommodates the formyl-Met-tRNA in the P site of 30S subunit (IF2), eukaryotic cells use a complex of three non-identical subunits, α, β and γ (eIF2), that binds GTP and Met-tRNAi first to form a ternary complex [1, 4, 5]. The ternary complex (TC) then binds the 40S subunit in a process that is stimulated by other initiation factors (e.g., eIF3) that are specific to eukaryotic cells. Both IF2 and the γ subunit of eIF2 contain a G domain subtype (tr-type G domain) involved in GTP binding and GTPase activation upon large ribosomal subunit joining (Fig. 1). However, unlike IF2, the intrinsic GTPase activity of eIF2 is low and requires a GTPase activating protein (GAP) called eIF5 that associates with eIF2 in the pre-initiation 43S complex [6, 7]. After GTP hydrolysis, exchange of GTP for GDP in eIF2 requires eIF2B, an additional factor that replenishes the GTP necessary for eIF2 activity recycling [8]. In bacteria IF2 performs all these activities, a fact that probably makes the initiation step more efficient in these organisms. In eukaryotes,

J.J. Berlanga (✉) · C. de Haro · M.A. Rodríguez-Gabriel · I. Ventoso
Centro de Biología Molecular Severo Ochoa (CSIC-UAM),
Universidad Autónoma de Madrid, 28049 Campus Cantoblanco, Madrid, Spain
e-mail: stress-lab@cbm.csic.es

G. Hernández and R. Jagus (eds.), *Evolution of the Protein Synthesis Machinery and Its Regulation*, DOI 10.1007/978-3-319-39468-8_12

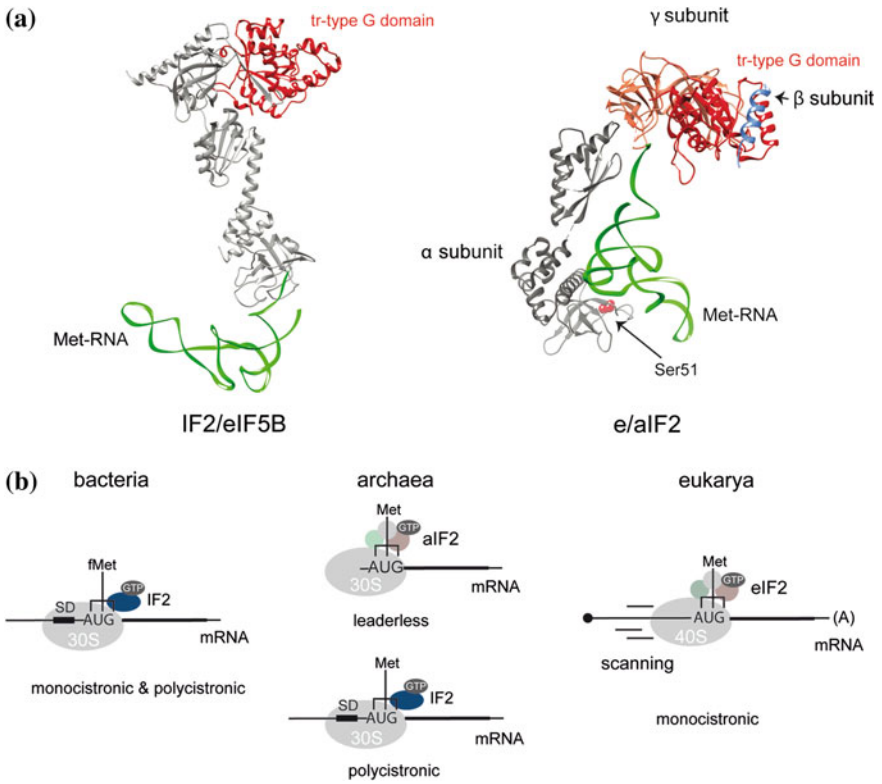


Fig. 1 Factors involved in Met-tRNAi binding and AUG recognition in bacteria, archaea and eukaryotes. **a** Structure of IF2/eIF5B and a/eIF2 factors bound to Met-tRNAi (*green*). IF2/eIF5B structure corresponds to mammalian eIF5B and a/eIF2 structure corresponds to yeast eIF2. The GTP binding-GTPase domain is shown in *red* (tr-type G domain). The three subunits of eIF2 are indicated, although the structure of the β-subunit is incomplete. Ser51 of the α subunit is labeled. **b** Mechanisms involved in AUG recognition in bacteria, archaea and eukaryotes. The Shine-Dalgarno sequence (SD) in mRNA promotes the recruitment of the 30s subunit in bacteria, but also in some polycistronic mRNAs of archaea. a/eIF2 binds Met-tRNAi and GTP to form the ternary complex (TC) which in turn binds archaeal 30S and eukaryotic 40S, promoting AUG recognition in most eukaryotic mRNAs and in leaderless archaeal mRNA

however, the splitting of eIF2 activities in two additional associated factors (eIF5 and eIF2B) resulted in a new layer of functional regulation and complexity associated to the evolution of eukaryotes [9].

The emergence of eIF2 was perhaps linked to a new mode of translation initiation that involved the attachment of ribosome to the 5′ termini of the mRNA. This process also required a new set of eukaryotic proteins (eIF4F and eIF3) that propelled the 40S subunit in the 5′–3′ direction (scanning) to reach the initiation codon [10]. This new mode of initiation probably replaced the existing mechanism that operated in primitive bacteria, where the small subunits possibly loaded the mRNAs

directly as occurs in modern prokaryotes [5] (Fig. 1). How did these primitive eukaryotes move from a prokaryotic-like mechanism of initiation toward the one conducted by eIF2 in modern eukaryotes? A look at the archaeal mode of initiation could shed some light on this question. An eIF2 homolog has been found in many species of archaea (aIF2), being extensively characterized in the prototypical thermophilic archaeon *Sulfolobus solfataricus* [11, 12]. aIF2 is structurally related to eIF2, although it still retains some functional resemblance with bacterial IF2. Like eIF2, aIF2 is a α-β-γ trimer that binds GTP, Met-tRNAi and the ribosomal small subunit (30S). Thus, aIF2 can replace eIF2 for delivering the Met-tRNAi to the eukaryotic 40S subunit in vitro [13]. However, unlike eIF2, aIF2's affinity for GTP is similar to its affinity for GDP and exhibits a significant GTPase activity in the absence of any GAP [14]. Consistently, no homolog of eIF5 has been found in archaea, whereas the existence of the eIF2B homolog in archaea is still controversial [15]. Moreover, as bacterial IF2, aIF2 seems to first associate with GTP and the 30S subunit before recruiting the Met-tRNAi [16]. Interestingly, besides Met-tRNAi binding activity, aIF2γ also binds the 5′ triphosphate of mRNAs and protects them from degradation [17, 18]. Thus, aIF2 has been implicated in the translation and stabilization of leaderless mRNA, which are very abundant in archaea (about 50 % of all mRNAs) [11, 16]. At the same time, archaea also use the homolog of bacterial IF2 to translate some mRNAs, especially for those that are included in operons (polycistronic) (Fig. 1). Eukaryotes also have a homolog of bacterial IF2, eIF5B, although its function has diverged from the former to promote 60S subunit joining after initiation codon recognition [19]. The fact that these two modes of translation initiation still coexist in archaea suggests that a/eIF2 perhaps emerged in the archaeal/eukaryotic branch as a new posttranscriptional regulator, involved in translation and stabilization of certain mRNAs with a distinctive 5′ end. Later, once in the eukaryotic branch, the machinery evolved so that eIF2 began to take over the translation of an increasing number of mRNAs. In this regard, it is interesting to note that eIF2 has been reported to interact with the 5′ UTR of some mammalian and viral mRNAs, an activity that could still remain in eIF2 as a relic of a primitive function [20].

2 The Origin and Diversification of eIF2α Kinases

In eukaryotes, the Ser51 of eIF2α can be phosphorylated by up to five different kinases with an exquisite degree of specificity, suggesting that a/eIF2 and its primitive kinase could have emerged as a working couple. In agreement, all eukaryotes studied to date contain at least one eIF2α kinase. In archaea, the sequence around the phosphorylatable S51 in aIF2α is not conserved. However, a nearby residue (Ser48) in aIF2 can be phosphorylated by human PKR in vitro [16, 21], consistent with the fact that the eIF2α motif involved in recognition by eIF2α kinases is quite conserved in aIF2α (KGYID in eIF2α vs. KGHID in aIF2α). However, such specific kinase activity on aIF2α has not been detected in any archaea analyzed so far [21]. Therefore,

at least in prototypical archaea, aIF2 could have emerged without an associated a/eIF2α kinase. Nonetheless, given the wide biodiversity found in the archaea domain, the possibility of finding a specific aIF2α kinase cannot be ruled out.

Phosphorylation at eIF2α Ser51 affects the activity of the entire α-β-γ complex, since it prevents the GDP-eIF2 to GTP-eIF2 recycling promoted by eIF2B. eIF2B binds phosphorylated eIF2α with 100-fold more affinity than to the non-phosphorylated form, causing the accumulation of a non-productive complex that sequesters eIF2B in the cell [22]. So, in order to impact eIF2 activity, primitive a/eIF2α kinase had to emerge together with (or after) eIF2B emergence in ancient eukaryotes. Interestingly, eIF2B-like activity was reported in extracts of thermophilic archaea [15], although it is not clear whether a similar activity could also be present in other archaeal groups.

eIF2α kinases (EIF2AKs) belong to the large protein-serine/threonine kinase family. Among EIF2AKs, the kinase domain (KD) shares approximately 27 % identity and 45 % similarity [23], which is much higher than the identity shared with other serine/threonine kinases (about 17 % identity). Besides the differences in amino acid sequence, the KD of eIF2α kinases contains two distinctive features that make them highly specific for eIF2α, their only well-characterized substrate. One of these features is the presence of an insert, variable in size, ranging from 14 to 248 residues and located between kinase subdomains IV and V (sheets β4 and β5) of KD that is non-conserved in the amino acid sequence (Fig. 2). Another characteristic hallmark is represented by a shorter linker connecting subdomains IX and X (helices αF and αG) together with a larger helix αG in the carboxy-terminal region of the kinase domain that is distinctly oriented towards the substrate [24].

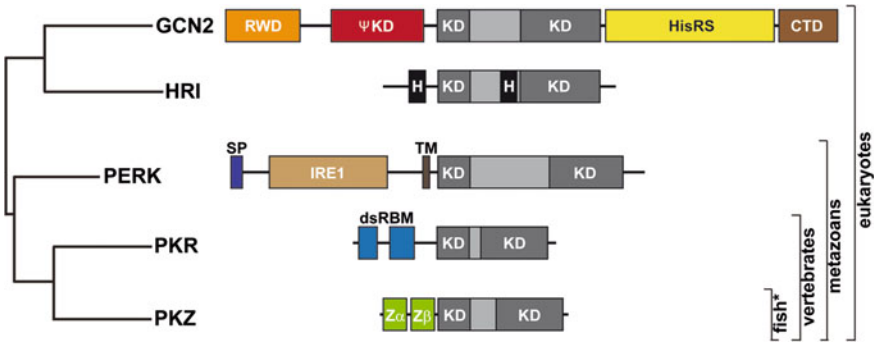


Fig. 2 Domain organization and phylogenetic relationships of eIF2α kinases. Phylogenetic tree was constructed with sequences of the kinase domains using the maximum likelihood method (PhyML). All branches gave a support higher than 0.8. Kinase domains (*KD*), including the internal insert (*lighter gray*), are represented together with the different regulatory domains found in eIF2α kinases (see text for details). The distribution of eIF2α kinases in eukaryotes is also indicated

GCN2 is the only eIF2α kinase that is present in virtually all eukaryotes (the anemone *N. vectensis* is the only exception to date), indicating that it may be the most ancient member of the EIF2AK family (Fig. 2). Furthermore, GCN2 is the only eIF2α kinase found in plants. HRI, the eIF2α kinase regulated by heme groups, is the second more widely distributed member of the family, and it is present in some unicellular eukaryotes such as fission yeast [25, 26]. Searching for HRI homologs in other available genomes revealed that HRI is much more widespread than previously thought, so that it is present in many species of invertebrates including insects, worms, bivalve mollusks and crustaceans (data not shown). Surprisingly, HRI genes are not present in some of the best studied model organisms, such as *D. melanogaster, C. elegans* and *S. cerevisiae*, a fact that led to the misleading assumption that HRI was only present in vertebrates. PERK is the third eIF2α kinase according to its lineage appearance, and it has been found in virtually all metazoans analyzed to date [27].

Finally, PKR, the eIF2α kinase involved in the innate cellular antiviral response, is only present in vertebrates. Additionally, its closely related eIF2α kinase, PKZ, has been found exclusively in fishes [28–32] (goldfish, zebrafish, Atlantic salmon, carp).

Of particular interest is the case of protozoan parasites such as *Plasmodium*, *Toxoplasma*, *Trypanosoma* or *Leishmania* in which at least three eIF2α kinases are found [33–36]. One of them, eIF2K1, is very similar to GCN2, but lacks the pseudokinase domain, which usually precedes the eIF2α KD. The second, eIF2K2, has a topology very similar to PERK, with a predicted N-terminal signal peptide and a potential transmembrane region, although lacking any known regulatory domain. Recently, a regulation by heme has been proposed for this kinase [37]. The third member of the family, eIF2K3, shows very little similarity with any other member of the eIF2α kinase family.

A comparison of the different kinase subdomains reveals certain phylogenetic relationships among the five members of this kinase family. The highest identity scores are found in subdomains VII and VIII, with 54 and 44 %, respectively, whereas subdomains II, X and XI are the most divergent with global identities of 15, 9 and 7 %, respectively [23]. Surprisingly, subdomain X (helix αG), which is directly involved in eIF2 recognition, shows a very limited conservation (9 % identity), although it contains 5 out of the 26 residues that are preferentially conserved only among the eIF2α kinases [24]. When comparing in pairs, the highest identity scores (49 %) correspond to GCN2-HRI (72 % similarity), HRI-PERK (68 % similarity) and GCN2-PERK (68 % similarity). PKR shares a relatively low identity with HRI and GCN2 (40 and 42 %, respectively) and higher with PERK (49 %) and PKZ (52 %).

The ability of each eIF2α kinase to respond to different stimuli resides in the presence of unique regulatory regions, most of them located in the N-terminal part of the protein (PERK, PKR and PKZ), but also in the C-terminus (GCN2) or even inside the kinase domain (HRI) (Fig. 2). The diversity of stress-sensing regulatory domains found in eIF2α kinases represents a good example of domain shuffling by recruitment and ligation of non-kinase domains to a primitive kinase domain (KD).

Non-homologous recombination and gene fusion events have been proposed as sources of genetic diversification of multimodular proteins, including kinases [38, 39]. In eIF2α kinases, the addition of regulatory domains occurred preferentially at the N-terminal of the KD, whereas GCN2 is the only member with an extended regulatory domain at the C-terminal of the KD. The length and organization of the regulatory domains in eIF2α kinases are very diverse, and no sequence homology has been detected among them, suggesting that they might have been added to the primitive KD by independent events of recombination.

GCN2, first identified in budding yeast, is mainly activated by binding of deacylated tRNA, which accumulates in cells starved for any amino acid [40]. Beside the KD, this kinase carries distinct well-defined domains: (1) an amino terminal charged region (RWD), responsible for the binding of the regulatory protein GCN1/GCN20; (2) a region similar to a kinase domain, but lacking essential residues (pseudokinase, ψKD); (3) a histidyl-tRNA synthetase (HisRS)-like domain; (4) a carboxy-terminal domain (CTD) with no obvious similarity to any known functional domain, but responsible for dimerization, binding to the ribosome and kinase activation by tRNA [41–43]. GCN2 activation requires the m2 motif of the HisRS domain, which binds deacylated tRNA and is responsible for GCN2 activation in response to amino acid starvation. This domain is also necessary for the activation of the kinase by viral RNA [44, 45]. The kinase is thought to be a constitutive dimer where the CTD interacts with the HisRS and kinase domains blocking its interaction with eIF2, until the binding of tRNA to HisRS releases these interdomain interactions allowing the interaction of the substrate with the kinase domain [46].

HRI was discovered as a modulator of globin synthesis in reticulocytes. HRI coordinates hemoglobin production to iron availability because of its activation when heme is scarce [47]. The presence of this protein in many other organisms, including fission yeast *Schizosaccharomyces pombe*, was later reported, extending its implication in the response to other stress forms such as exposure to heavy metals, heat shock, osmotic stress, oxidative stress, nutrient starvation or nitric oxide [47–49]. HRI seems to be a constitutively active homodimer, which binds heme through a region located in the amino-terminal portion of the protein. A second heme-binding site inside the kinase domain also regulates its activity. It has been proposed that kinase activity is inhibited in the presence of heme as a result of the amino-terminal domain interaction with the KD as well as the formation of intermolecular disulfide bonds. Under heme deprivation, heme dissociates from the kinase domain-binding site, releasing the intra- and intermolecular interactions that result in kinase activation [50, 51].

PERK is an endoplasmic reticulum (ER) resident kinase with an amino-terminal signal peptide (SP), a luminal amino-terminal regulatory region (IRE1-like) and a transmembrane domain (TM) preceding the kinase domain. This protein is activated upon ER stress because of its oligomerization through the IRE1-like regulatory domain [52, 53]. In unstressed cells PERK interacts with ER chaperones such as BiP (GRP78), which binds to the amino-terminal IRE1-like regulatory domain, preventing the dimerization and keeping the kinase in an inactive state.

Accumulation of unfolded proteins during ER stress leads to the dissociation of the chaperones, promoting kinase oligomerization and activation [54, 55].

Expression of PKR is stimulated by interferon, and it is the main eIF2α kinase involved in the response to viral infection. The kinase is activated upon binding to double-stranded RNA (dsRNA) generated during viral replication [56]. It contains two dsRNA-binding motifs (dsRBM) in the amino-terminal region that are responsible for the activation of the kinase. It has been reported that the folding of these dsRBMs over the kinase domain keeps the kinase inactive in its latent state. Upon dsRNA binding to dsRBM, the kinase domain unlocks. Furthermore, the dsRNA binding promotes the dimerization of the protein, which is also necessary for its activation [24, 57, 58]. There is not a dsRNA sequence requirement, but the length must be at least 16 bp to allow the binding of two dsRBMs, present in two PKR molecules, to the same dsRNA molecule and the formation of PKR dimers, the active form of the kinase [59].

PKZ is an eIF2α kinase very closely related to PKR, only present in fish. PKZ differs from PKR because it contains two Z-DNA binding domains instead of the dsRBM present in PKR. The Z conformation is an alternative form that can be adopted by dsRNA or dsDNA, consisting in a left-handed helix. In cells, Z-DNA can be generated by RNA polymerases during transcription. PKZ can provide advantages in the innate response against virus infection, because it could make the spectrum of viruses sensitive to eIF2α kinase activity wider by responding to distinct nucleic acid structures and by altering the sensitivity of cells to viral PKR inhibitors [30].

The diversification of eIF2α kinases exemplifies the specialization of stress response found in vertebrates. However, not all eIF2α kinases exhibit the same degree of specialization. GCN2 and HRI, the most primitive members of the family, exhibit some degree of promiscuity so that they can be activated by several different types of stress (Fig. 3). However, PERK and PKR, the newest members, show a much higher degree of specialization so that they respond to just one type of stress (ER stress/protein unfolding or virus, respectively). GCN2 and HRI are activated by stresses that can be considered as universal to all eukaryotic cells (nutrient deprivation, oxidation and heat shock). However, the ER stress that activates PERK mainly affects some specialized cells that are involved in the secretory function of metazoans and vertebrate tissues [52]. Similarly, the viral infections that activate PKR are more prevalent in vertebrate hosts.

No significant overlapping in eIF2α kinase activation was found when comparing different types of stress in human and mouse cells. For example, drugs that perturb ER homeostasis or that induce protein misfolding activated PERK, but no other eIF2α kinases. Similarly, viral infection can induce the activation of PKR exclusively, although a weak secondary activation of PERK has been documented for some viruses because of the massive accumulation of viral proteins, which resulted in ER overloading [60]. Exceptionally, some degree of overlapping has been found for PKR and GCN2 in cells infected with alphavirus, where an additional activation of GCN2 in addition to primary PKR activation has been reported

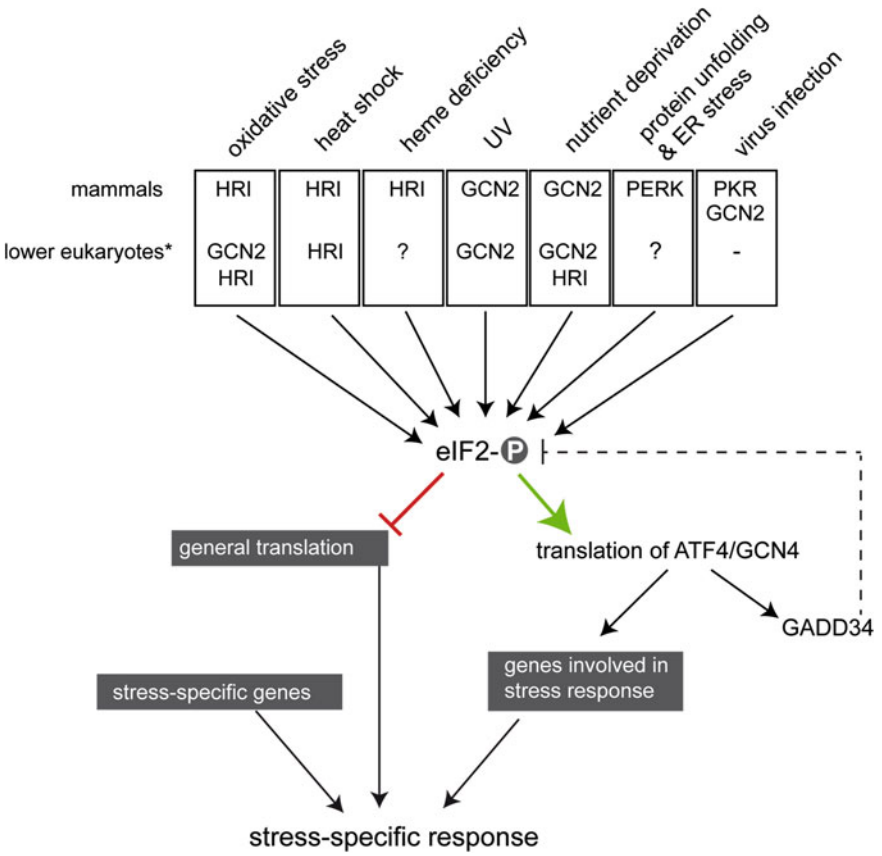


Fig. 3 eIF2α kinases and the integrated stress response (ISR). Cells sense environmental stresses and differentially activate eIF2α kinases, in both lower and higher eukaryotes. These signals funnel into eIF2α phosphorylation leading to attenuation of general translation and activating translation of specific mRNAs (e.g., ATF4). ATF4 activates transcription of a basic set of genes involved in stress response and recovery (e.g., GADD34). This translational reprogramming, along with stress-specific transcription, generates a tailored response to different stresses

[44]. On the contrary, in lower eukaryotes such as fission yeast, an overlapping of Gcn2 and Hri1 and Hri2 kinases is more evident. For example, nitrogen source deprivation activated Gcn2 and Hri1 in fission yeast, whereas Hri2 activated early during glucose deprivation [48]. Although some degree of specialization is already apparent in lower eukaryote eIF2α kinases, stress response in these organisms seems to be less diversified when compared to higher eukaryotes, so that different stresses can activate a similar set of genes, a feature much less evident in higher eukaryotes [49, 61, 62].

3 eIF2α Phosphorylation and Stress Response

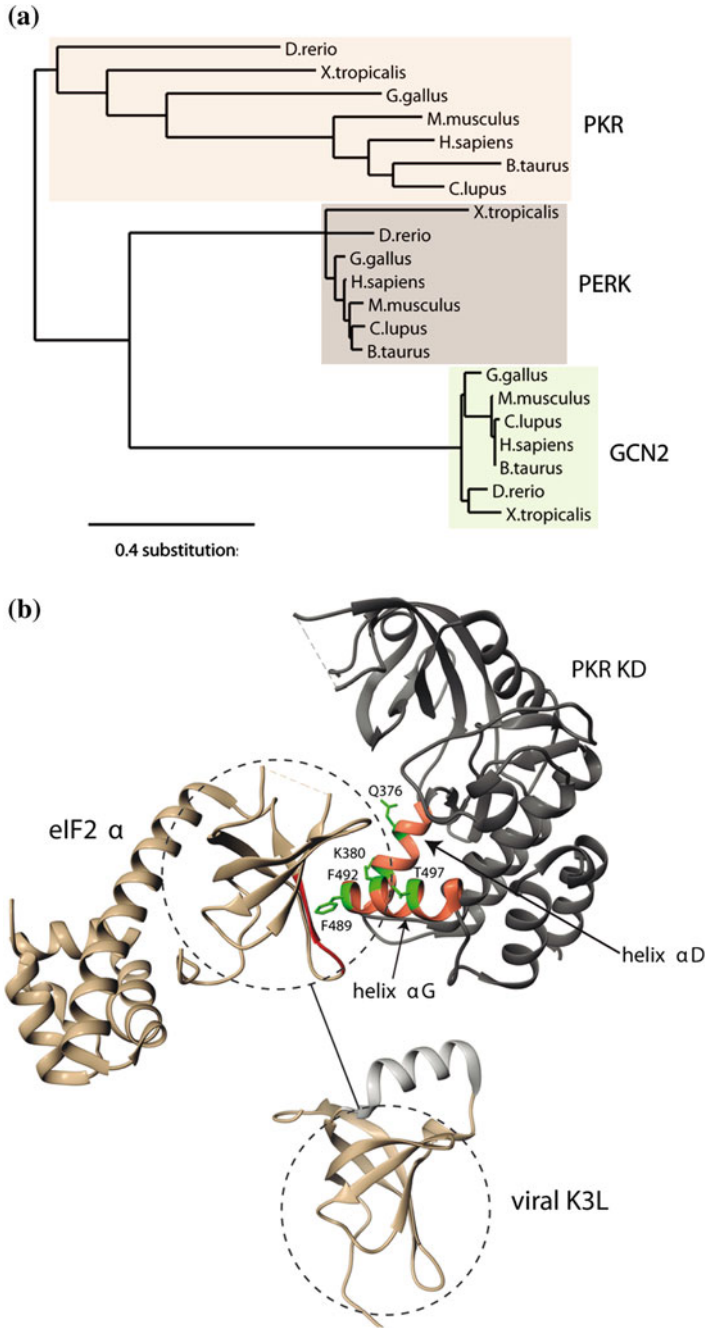
In eukaryotes, the responses to amino acid deprivation, viral infection, iron deficiency and protein misfolding constitute the integrated stress response (ISR), which is activated by eIF2α phosphorylation [63]. All these stresses trigger kinase activation and eIF2α phosphorylation in order to attenuate the synthesis of new proteins and, at the same time, to activate a transcriptional program that includes the induction of a set of common and stress-specific genes [64]. Therefore, the successful execution of the stress response requires a tight coordination between translation reprogramming and transcription. Given that eIF2B is at limiting amounts in the cell, and that phosphorylated eIF2α showed an increased affinity for eIF2B, a relatively low level of eIF2α phosphorylation causes a disproportioned inhibition of general translation, especially in mammalian cells [4]. This blockage in protein synthesis by eIF2α phosphorylation has been interpreted as an attempt of the cell to prevent stress-induced damage of proteins that are being synthesized and also to save energy that could be further necessary for recovering from stress. Reducing the bulk of protein synthesis is particularly relevant during the unfolded protein response (UPR) that occurs in the ER. Translation halt due to eIF2 phosphorylation alleviates the protein load of the ER almost instantly, minimizing the chance for co-translational misfolding [65, 66]. For GCN2, the response to amino acid deprivation couples the cell metabolism and proliferation to nutrient availability, an elemental response for all eukaryotic cells.

The seminal studies with GCN4 mRNA in yeast, coupled with modern high-throughput approaches (polysome and ribosome profiling) have revealed the prevalence of a set of mRNAs whose translation resisted or was even activated upon eIF2α phosphorylation [67]. This paradoxical effect relies on the presence of distinctive regulatory elements in these mRNAs. For translation-activated mRNA it consists in the presence of one or several short upstream open reading frames (uORFs), which control the re-initiation of 40S subunits on the true AUG of these mRNAs. ATF4, the transcription factor homolog of GCN4 in higher eukaryotes, is a paradigm of this mode of translation regulation that tightly controls the level of this protein in the cell [63, 68]. ATF4 is a master regulator of the ISR, which activates the transcription of many genes involved in antioxidant functions, protein synthesis, amino acid transport and chaperons, which are necessary for recovering from stress [63]. Interestingly, ATF4 also controls the status of eIF2α phosphorylation by inducing the expression of the specific regulatory subunit of the eIF2α phosphatase (GADD34) via ATF3 and CHOP induction, a circuit that restores translation once the stress stimulus has disappeared [63]. Similar to ATF4, the mRNAs of ATF3, CHOP and GADD34 genes also contain uORFs, which regulate their translation during the stress response. Remarkably, all these genes are also activated at the transcriptional level upon stress, a fact that potentiates the stress response. Whereas translation reprogramming due to eIF2α phosphorylation is predicted to be very similar for all types of stress, transcriptional activation profiles are clearly stress specific in mammalian cells [62]. For example, UPR activates the

transcription of a number of genes that, in combination with translation reprogramming by PERK/eIF2α phosphorylation, generates a more specific response that clearly differentiates from amino acid deprivation, UV or virus infection (Fig. 3). The coupling of these two levels of regulation (translation and transcription) generates the specialization in the stress response observed in mammals.

4 Evolution of PKR and Virus Defeating

Phylogenetic analysis revealed that PKR is the eIF2α kinase family member that accumulated more changes (diversification) in the recent evolution of this kinase family (last 30–50 million years). Signs of intense episodes of positive evolution became apparent when ratios of synonymous and non-synonymous changes (dN/dS) were compared among PKR genes from fish to primates [69, 70]. Thus, PKR genes have been described as one of the fastest evolving genes identified to date [70]. These rates of evolution are not observed in PERK and GCN2, which are highly conserved among vertebrates, whereas HRI genes evolved at rates similar to that observed in other kinase families (Fig. 4). Since PKR is directly involved in defeating viruses, the accelerated rates of evolution found in PKR genes have been explained by a sustained "arms race" between the kinase and viral products that antagonize with it [69, 70]. This is exemplified by the poxvirus K3L gene, a pseudo-substrate of PKR that mimics the eIF2 moiety involved in the interaction with the enzymatic pocket of eIF2α kinases (Fig. 4). Thus, the combination of adaptive changes in the αG, αD and αE helices of the PKR kinase domain (KD) of the hominoid lineage (including human, chimpanzee, gorilla and orangutan) resulted in an almost complete resistance to the vaccinia K3L product [69]. This illustrates a remarkable case of evolutionary flexibility at protein interaction interfaces, since the resistance of hominoid PKR to K3L binding was acquired without affecting the interaction with the natural substrate (eIF2). Interestingly, tandem duplication and triplication of PKR genes have been detected in the genome of some amphibians and fishes, a fact that has not been observed for the rest of eIF2α kinases. This raises the possibility that intense episodes of arms race with viruses or other pathogens could have promoted duplications in the PKR locus of some vertebrate clades, a source of diversity that is also suggestive for the primary diversification of the eIF2 kinases that may have occurred in primitive eukaryotes.


(a)
D.rerio
X.tropicalis
G.gallus
M.musculus
H.sapiens
B.taurus
C.lupus
PKR
X.tropicalis
D.rerio
G.gallus
H.sapiens
M.musculus
C.lupus
B.taurus
PERK
G.gallus
M.musculus
C.lupus
H.sapiens
B.taurus
D.rerio
X.tropicalis
GCN2
0.4 substitution
(b)
PKR KD
eIF2 α
Q376
K380
F492
T497
F489
helix αD
helix αG
viral K3L

◀ **Fig. 4** Accelerated evolution of PKR. **a** Maximum-likelihood phylogram of PKR, GCN2 and PERK kinases in seven vertebrates (human, mouse, cow, dog, chicken, frog and fish). **b** Positively selected sites in the PKR kinase domain as resulting from the arms race with viral PKR inhibitors (*K3L*). The crystal structure of PKR KD bound to eIF2α is shown (pdb2a1a), including the residues of αG and αD helices subjected to positive evolution (*numbered residues*). The structural domains of eIF2α and vaccinia K3L (pdb1luz) showing molecular mimicry are encircled in *dashed lines*

References

1. Hinnebusch AG. Mechanism and regulation of initiatior methionyl-tRNA binding to ribosome. In: Sonenberg, N; Hershey, JWB and Mathews, MB, editors. Translational control of gene expression. NY: Cold Spring Harbor Laboratory Press, Cold Spring Harbor; 2000. pp. 185–243.
2. Lee JH, Choi SK, Roll-Mecak A, Burley SK, Dever TE. Universal conservation in translation initiation revealed by human and archaeal homologs of bacterial translation initiation factor IF2. Proc Natl Acad Sci USA. 1999;96:4342–7.
3. Roll-Mecak A, Shin BS, Dever TE, Burley SK. Engaging the ribosome: universal IFs of translation. Trends Biochem Sci. 2001;26:705–9.
4. Hinnebusch AG. The scanning mechanism of eukaryotic translation initiation. Annu Rev Biochem. 2014;83:779–812. doi:10.1146/annurev-biochem-060713-035802.
5. Gualerzi CO, Pon CL. Initiation of mRNA translation in bacteria: structural and dynamic aspects. Cell Mol Life Sci. 2015;72:4341–67. doi:10.1007/s00018-015-2010-3.
6. Asano K, Shalev A, Phan L, Nielsen K, Clayton J, Valasek L, Donahue TF, Hinnebusch AG. Multiple roles for the C-terminal domain of eIF5 in translation initiation complex assembly and GTPase activation. EMBO J. 2001;20:2326–37. doi:10.1093/emboj/20.9.2326.
7. Paulin FE, Campbell LE, O'Brien K, Loughlin J, Proud CG. Eukaryotic translation initiation factor 5 (eIF5) acts as a classical GTPase-activator protein. Curr Biol. 2001;11:55–9.
8. Konieczny A, Safer B. Purification of the eukaryotic initiation factor 2-eukaryotic initiation factor 2B complex and characterization of its guanine nucleotide exchange activity during protein synthesis initiation. J Biol Chem. 1983;258:3402–8.
9. Alone PV, Dever TE. Direct binding of translation initiation factor eIF2gamma-G domain to its GTPase-activating and GDP-GTP exchange factors eIF5 and eIF2B epsilon. J Biol Chem. 2006;281:12636–44. doi:10.1074/jbc.M511700200.
10. Jackson RJ, Hellen CU, Pestova TV. The mechanism of eukaryotic translation initiation and principles of its regulation. Nat Rev Mol Cell Biol. 2010;11:113–27. doi:10.1038/nrm2838.
11. La Teana A, Benelli D, Londei P, Blasi U. Translation initiation in the crenarchaeon Sulfolobus solfataricus: eukaryotic features but bacterial route. Biochem Soc Trans. 2013;41:350–5. doi:10.1042/BST20120300.
12. Pedulla N, Palermo R, Hasenohrl D, Blasi U, Cammarano P, Londei P. The archaeal eIF2 homologue: functional properties of an ancient translation initiation factor. Nucleic Acids Res. 2005;33:1804–12. doi:10.1093/nar/gki321.
13. Dmitriev SE, Stolboushkina EA, Terenin IM, Andreev DE, Garber MB, Shatsky IN. Archaeal translation initiation factor aIF2 can substitute for eukaryotic eIF2 in ribosomal scanning during mammalian 48S complex formation. J Mol Biol. 2011;413:106–14. doi:10.1016/j.jmb.2011.08.026.
14. Dubiez E, Aleksandrov A, Lazennec-Schurdevin C, Mechulam Y, Schmitt E. Identification of a second GTP-bound magnesium ion in archaeal initiation factor 2. Nucleic Acids Res. 2015;43:2946–57. doi:10.1093/nar/gkv053.

15. Dev K, Santangelo TJ, Rothenburg S, Neculai D, Dey M, Sicheri F, Dever TE, Reeve JN, Hinnebusch AG. Archaeal aIF2B interacts with eukaryotic translation initiation factors eIF2alpha and eIF2Balpha: Implications for aIF2B function and eIF2B regulation. J Mol Biol. 2009;392:701–22. doi:10.1016/j.jmb.2009.07.030.
16. Benelli D, Londei P. Translation initiation in Archaea: conserved and domain-specific features. Biochem Soc Trans. 2011;39:89–93. doi:10.1042/BST0390089.
17. Arkhipova V, Stolboushkina E, Kravchenko O, Kljashtorny V, Gabdulkhakov A, Garber M, Nikonov S, Martens B, Blasi U, Nikonov O. Binding of the 5′-Triphosphate End of mRNA to the gamma-Subunit of Translation Initiation Factor 2 of the Crenarchaeon Sulfolobus solfataricus. J Mol Biol. 2015;427:3086–95. doi:10.1016/j.jmb.2015.07.020.
18. Hasenohrl D, Lombo T, Kaberdin V, Londei P, Blasi U. Translation initiation factor a/eIF2(-gamma) counteracts 5′ to 3′ mRNA decay in the archaeon Sulfolobus solfataricus. Proc Natl Acad Sci USA. 2008;105:2146–50. doi:10.1073/pnas.0708894105.
19. Pestova TV, Lomakin IB, Lee JH, Choi SK, Dever TE, Hellen CU. The joining of ribosomal subunits in eukaryotes requires eIF5B. Nature. 2000;403:332–5. doi:10.1038/35002118.
20. Ben-Asouli Y, Banai Y, Hauser H, Kaempfer R. Recognition of 5′-terminal TAR structure in human immunodeficiency virus-1 mRNA by eukaryotic translation initiation factor 2. Nucleic Acids Res. 2000;28:1011–8.
21. Tahara M, Ohsawa A, Saito S, Kimura M. In vitro phosphorylation of initiation factor 2 alpha (aIF2 alpha) from hyperthermophilic archaeon Pyrococcus horikoshii OT3. J Biochem. 2004;135:479–85.
22. Krishnamoorthy T, Pavitt GD, Zhang F, Dever TE, Hinnebusch AG. Tight binding of the phosphorylated alpha subunit of initiation factor 2 (eIF2alpha) to the regulatory subunits of guanine nucleotide exchange factor eIF2B is required for inhibition of translation initiation. Mol Cell Biol. 2001;21:5018–30. doi:10.1128/MCB.21.15.5018-5030.2001.
23. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, Thompson JD, Higgins DG. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011;7:539. doi:10.1038/msb.2011.75.
24. Dar AC, Dever TE, Sicheri F. Higher-order substrate recognition of eIF2alpha by the RNA-dependent protein kinase PKR. Cell. 2005;122:887–900. doi:10.1016/j.cell.2005.06.044.
25. Zhan K, Vattem KM, Bauer BN, Dever TE, Chen JJ, Wek RC. Phosphorylation of eukaryotic initiation factor 2 by heme-regulated inhibitor kinase-related protein kinases in Schizosaccharomyces pombe is important for fesistance to environmental stresses. Mol Cell Biol. 2002;22:7134–46.
26. Dever TE DA, Sicheri F. The eIF2α kinases. In: Mathews MBSN, Hershey JWB, editors. Translational control in biology and medicine. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press; 2007. pp. 319–344.
27. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL. GenBank. Nucleic Acids Res. 2005;33:D34–8. doi:10.1093/nar/gki063.
28. Rothenburg S, Deigendesch N, Dittmar K, Koch-Nolte F, Haag F, Lowenhaupt K, Rich A. A PKR-like eukaryotic initiation factor 2alpha kinase from zebrafish contains Z-DNA binding domains instead of dsRNA binding domains. Proc Natl Acad Sci USA. 2005;102:1602–7. doi:10.1073/pnas.0408714102.
29. Bergan V, Jagus R, Lauksund S, Kileng O, Robertsen B. The Atlantic salmon Z-DNA binding protein kinase phosphorylates translation initiation factor 2 alpha and constitutes a unique orthologue to the mammalian dsRNA-activated protein kinase R. FEBS J. 2008;275:184–97. doi:10.1111/j.1742-4658.2007.06188.x.
30. Rothenburg S, Deigendesch N, Dey M, Dever TE, Tazi L. Double-stranded RNA-activated protein kinase PKR of fishes and amphibians: varying the number of double-stranded RNA binding domains and lineage-specific duplications. BMC Biol. 2008;6:12. doi:10.1186/1741-7007-6-12.

31. Su J, Zhu Z, Wang Y. Molecular cloning, characterization and expression analysis of the PKZ gene in rare minnow Gobiocypris rarus. Fish Shellfish Immunol. 2008;25:106–13. doi:10.1016/j.fsi.2008.03.006.
32. Yang PJ, Wu CX, Li W, Fan LH, Lin G, Hu CY. Cloning and functional analysis of PKZ (PKR-like) from grass carp (Ctenopharyngodon idellus). Fish Shellfish Immunol. 2011;31:1173–8. doi:10.1016/j.fsi.2011.10.012.
33. Mohrle JJ, Zhao Y, Wernli B, Franklin RM, Kappes B. Molecular cloning, characterization and localization of PfPK4, an eIF-2alpha kinase-related enzyme from the malarial parasite Plasmodium falciparum. Biochem J. 1997;328(Pt 2):677–87.
34. Narasimhan J, Joyce BR, Naguleswaran A, Smith AT, Livingston MR, Dixon SE, Coppens I, Wek RC, Sullivan WJ Jr. Translation regulation by eukaryotic initiation factor-2 kinases in the development of latent cysts in Toxoplasma gondii. J Biol Chem. 2008;283:16591–601. doi:10.1074/jbc.M800681200.
35. Konrad C, Wek RC, Sullivan WJ Jr. GCN2-like eIF2alpha kinase manages the amino acid starvation response in Toxoplasma gondii. Int J Parasitol. 2014;44:139–46. doi:10.1016/j.ijpara.2013.08.005.
36. Moraes MC, Jesus TC, Hashimoto NN, Dey M, Schwartz KJ, Alves VS, Avila CC, Bangs JD, Dever TE, Schenkman S, Castilho BA. Novel membrane-bound eIF2alpha kinase in the flagellar pocket of Trypanosoma brucei. Eukaryot Cell. 2007;6:1979–91. doi:10.1128/EC.00249-07.
37. da Silva Augusto L, Moretti NS, Ramos TC, de Jesus TC, Zhang M, Castilho BA, Schenkman S. A membrane-bound eIF2 alpha kinase located in endosomes is regulated by heme and controls differentiation and ROS levels in Trypanosoma cruzi. PLoS Pathog. 2015;11:e1004618. doi:10.1371/journal.ppat.1004618.
38. Deshmukh K, Anamika K, Srinivasan N. Evolution of domain combinations in protein kinases and its implications for functional diversity. Prog Biophys Mol Biol. 2010;102:1–15. doi:10.1016/j.pbiomolbio.2009.12.009.
39. Trifonov EN, Frenkel ZM. Evolution of protein modularity. Curr Opin Struct Biol. 2009;19:335–40. doi:10.1016/j.sbi.2009.03.007.
40. Hinnebusch AG. Evidence for translational regulation of the activator of general amino acid control in yeast. Proc Natl Acad Sci USA. 1984;81:6442–6.
41. Dong J, Qiu H, Garcia-Barrio M, Anderson J, Hinnebusch AG. Uncharged tRNA activates GCN2 by displacing the protein kinase moiety from a bipartite tRNA-binding domain. Mol Cell. 2000;6:269–79.
42. Garcia-Barrio M, Dong J, Ufano S, Hinnebusch AG. Association of GCN1-GCN20 regulatory complex with the N-terminus of eIF2alpha kinase GCN2 is required for GCN2 activation. EMBO J. 2000;19:1887–99.
43. Wek SA, Zhu S, Wek RC. The histidyl-tRNA synthetase-related sequence in the eIF-2 alpha protein kinase GCN2 interacts with tRNA and is required for activation in response to starvation for different amino acids. Mol Cell Biol. 1995;15:4497–506.
44. Berlanga JJ, Ventoso I, Harding HP, Deng J, Ron D, Sonenberg N, Carrasco L, de Haro C. Antiviral effect of the mammalian translation initiation factor 2alpha kinase GCN2 against RNA viruses. EMBO J. 2006;25:1730–40.
45. del Pino J, Jimenez JL, Ventoso I, Castello A, Munoz-Fernandez MA, de Haro C, Berlanga JJ. GCN2 has inhibitory effect on human immunodeficiency virus-1 protein synthesis and is cleaved upon viral infection. PLoS ONE. 2012;7:e47272. doi:10.1371/journal.pone.0047272.
46. Qiu H, Dong J, Hu C, Francklyn CS, Hinnebusch AG. The tRNA-binding moiety in GCN2 contains a dimerization domain that interacts with the kinase domain and is required for tRNA binding and kinase activation. EMBO J. 2001;20:1425–38.
47. Lu L, Han AP, Chen JJ. Translation initiation control by heme-regulated eukaryotic initiation factor 2alpha kinase in erythroid cells under cytoplasmic stresses. Mol Cell Biol. 2001;21:7971–80. doi:10.1128/MCB.21.23.7971-7980.2001.

48. Martin R, Berlanga JJ, de Haro C. New roles of the fission yeast eIF2alpha kinases Hri1 and Gcn2 in response to nutritional stress. J Cell Sci. 2013;126:3010–20. doi:10.1242/jcs.118067.
49. Zhan K, Narasimhan J, Wek RC. Differential activation of eIF2 kinases in response to cellular stresses in Schizosaccharomyces pombe. Genetics. 2004;168:1867–75.
50. Yang JM, London IM, Chen JJ. Effects of hemin and porphyrin compounds on intersubunit disulfide formation of heme-regulated eIF-2 alpha kinase and the regulation of protein synthesis in reticulocyte lysates. J Biol Chem. 1992;267:20519–24.
51. Yun BG, Matts JA, Matts RL. Interdomain interactions regulate the activation of the heme-regulated eIF 2 alpha kinase. Biochim Biophys Acta. 2005;1725:174–81. doi:10.1016/j.bbagen.2005.07.011.
52. Harding HP, Zhang Y, Ron D. Protein translation and folding are coupled by an endoplasmic-reticulum-resident kinase. Nature. 1999;397:271–4.
53. Shi Y, Vattem KM, Sood R, An J, Liang J, Stramm L, Wek RC. Identification and characterization of pancreatic eukaryotic initiation factor 2 alpha-subunit kinase, PEK, involved in translational control. Mol Cell Biol. 1998;18:7499–509.
54. Bertolotti A, Zhang Y, Hendershot LM, Harding HP, Ron D. Dynamic interaction of BiP and ER stress transducers in the unfolded-protein response. Nat Cell Biol. 2000;2:326–32. doi:10.1038/35014014.
55. Ma K, Vattem KM, Wek RC. Dimerization and release of molecular chaperone inhibition facilitate activation of eukaryotic initiation factor-2 kinase in response to endoplasmic reticulum stress. J Biol Chem. 2002;277:18728–35. doi:10.1074/jbc.M200903200.
56. Garcia MA, Gil J, Ventoso I, Guerra S, Domingo E, Rivas C, Esteban M. Impact of protein kinase PKR in cell biology: from antiviral to antiproliferative action. Microbiol Mol Biol Rev. 2006;70:1032–60. doi:10.1128/MMBR.00027-06.
57. Green SR, Mathews MB. Two RNA-binding motifs in the double-stranded RNA-activated protein kinase. DAI. Genes Dev. 1992;6:2478–90.
58. Manche L, Green SR, Schmedt C, Mathews MB. Interactions between double-stranded RNA regulators and the protein kinase DAI. Mol Cell Biol. 1992;12:5238–48.
59. Schmedt C, Green SR, Manche L, Taylor DR, Ma Y, Mathews MB. Functional characterization of the RNA-binding domain and motif of the double-stranded RNA-dependent protein kinase DAI (PKR). J Mol Biol. 1995;249:29–44. doi:10.1006/jmbi.1995.0278.
60. Zhang L, Wang A. Virus-induced ER stress and the unfolded protein response. Front Plant Sci. 2012;3:293. doi:10.3389/fpls.2012.00293.
61. Gasch AP, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, Storz G, Botstein D, Brown PO. Genomic expression programs in the response of yeast cells to environmental changes. Mol Biol Cell. 2000;11:4241–57.
62. Murray JI, Whitfield ML, Trinklein ND, Myers RM, Brown PO, Botstein D. Diverse and specific gene expression responses to stresses in cultured human cells. Mol Biol Cell. 2004;15:2361–74. doi:10.1091/mbc.E03-11-0799.
63. Harding HP, Zhang Y, Zeng H, Novoa I, Lu PD, Calfon M, Sadri N, Yun C, Popko B, Paules R, Stojdl DF, Bell JC, Hettmann T, Leiden JM, Ron D. An integrated stress response regulates amino acid metabolism and resistance to oxidative stress. Mol Cell. 2003;11:619–33.
64. Dever TE. Gene-specific regulation by general translation factors. Cell. 2002;108:545–56.
65. Harding HP, Calfon M, Urano F, Novoa I, Ron D. Transcriptional and translational control in the Mammalian unfolded protein response. Annu Rev Cell Dev Biol. 2002;18:575–99. doi:10.1146/annurev.cellbio.18.011402.160624.
66. Scheuner D, Song B, McEwen E, Liu C, Laybutt R, Gillespie P, Saunders T, Bonner-Weir S, Kaufman RJ. Translational control is required for the unfolded protein response and in vivo glucose homeostasis. Mol Cell. 2001;7:1165–76.
67. Hinnebusch AG. Translational regulation of yeast GCN4. A window on factors that control initiator-trna binding to the ribosome. J Biol Chem. 1997;272:21661–4.
68. Vattem KM, Wek RC. Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. Proc Natl Acad Sci USA. 2004;101:11269–74.

69. Elde NC, Child SJ, Geballe AP, Malik HS. Protein kinase R reveals an evolutionary model for defeating viral mimicry. Nature. 2009;457:485–9. doi:10.1038/nature07529.
70. Rothenburg S, Seo EJ, Gibbs JS, Dever TE, Dittmar K. Rapid evolution of protein kinase PKR alters sensitivity to viral inhibitors. Nat Struct Mol Biol. 2009;16:63–70. doi:10.1038/nsmb.1529.

Translation Elongation and Termination: Are They Conserved Processes?

Sandra Eltschinger, Peter Bütikofer and Michael Altmann

1 Introduction

Translation initiation is followed by a process in which sequential addition of amino acid residues enables peptide chain formation, called translation elongation. Elongation decodes the codons on an mRNA and depends on elongation factors (EFs). In a first step, a ternary complex consisting of EF1A/EF-Tu-GTP-aminoacyl-tRNA (aa-tRNA) is formed and the elongator aa-tRNA is recruited to the ribosomal acceptor (A-) site. Hydrolysis of EF1A/EF-Tu-GTP is activated upon codon-anticodon decoding at the A-site and mRNA-tRNA interaction with the ribosome. Subsequently, peptide bond formation occurs between the aa moiety of the A-site aa-tRNA and the peptidyl-tRNA located at the ribosomal peptidyl (P-) site. Thereafter, the deacylated tRNA is moved to the exit (E-) site [1–4]. This elongating process is repeated until a stop codon (UAA, UAG and UGA) is encountered. Exposing a stop codon at the A-site initiates the process of translation termination. Polypeptide release factors mediate the release of the polypeptide chain from the ribosome by hydrolysis of the ester bond between the polypeptide chain and the tRNA at the P-site [1–3, 5].

S. Eltschinger · P. Bütikofer · M. Altmann (✉)
Institute of Biochemistry and Molecular Medicine (IBMM),
Bühlstrasse 28, 3012 Bern, Switzerland
e-mail: michael.altmann@ibmm.unibe.ch

S. Eltschinger
e-mail: Sandra.Eltschinger@unige.ch

G. Hernández and R. Jagus (eds.), *Evolution of the Protein Synthesis Machinery and Its Regulation*, DOI 10.1007/978-3-319-39468-8_13

2 Elongation Factors

2.1 Mechanism of Translation Elongation

2.1.1 The EF1 Complex and Bacterial EF-Tu/EF-Ts

Translation elongation is a conserved process across the three kingdoms of life—eukaryotes, bacteria and archaea [6]. Once the eukaryotic 80S ribosome—or the 70S ribosome in prokaryotes—is positioned on the start AUG codon of the messenger RNA (mRNA), a polypeptide chain is formed by sequential addition of amino acids. Eukaryotic translation elongation requires several elongation factors: eukaryotic elongation factor complex 1 [7]—consisting of the G-protein eEF1A, its GEF (guanine nucleotide exchange factor) eEF1Bα and a stimulatory factor eEF1Bγ in fungi and additionally eEF1Bβ in plants and animals [8]—eEF2 and fungal-specific eEF3 [9–13]. Analogously, archaea and bacteria require aEF1 or EF-Tu (Elongation Factor Thermo-unstable [14]), respectively, for aa-tRNA binding and peptide elongation. These proteins are orthologs of eEF1A. Furthermore, aEF1 and EF-Tu nucleotide exchange is performed by EF-Ts (Elongation Factor thermostable [14]) and aEF1β, respectively, and translocation of peptidyl-tRNAs is catalyzed by aEF2 and EF-G, respectively (reviewed in [2]).

tRNAs are specific for a particular amino acid. A given tRNA recognizes its cognate mRNA codon by its anticodon. The genetic code is—with a few exceptions—universal in all domains of life and degenerated, allowing correspondence of several mRNA triplets to one amino acid residue, as recently reviewed in [15]. tRNAs bind to three different binding sites within a ribosome: Incoming aminoacyl-tRNAs enter the A-site. The P-site serves as entry site of the initiator tRNA (Met-tRNA$_i$) during translation initiation. Furthermore, stable binding of the peptidyl-tRNA during elongation also occurs at the P-site. Lastly, the E-site is bound by deacylated tRNAs before they leave the ribosome (reviewed in [16, 17]).

Peptide elongation starts with an initiator methionyl-tRNA placed at the ribosomal P-site next to the A-site. eEF1A, aEF1 and bacterial EF-Tu, respectively, are among the most abundant proteins within a cell [18] and consist of three different domains that perform a specific function. Domain I—also called G-domain—is responsible for the guanosine-triphosphatase (GTPase) activity of eEF1A/EF-Tu. When interacting with GTP, eEF1A/EF-Tu binds to an aminoacyl-tRNA (aa-tRNA) with very high affinity and forms a ternary complex. Domains II and III assist in binding the aa-tRNA. The aa-tRNA is directed to the ribosomal A site [1]. Hydrolysis of GTP occurs upon codon-anticodon match between the A-site codon of the ribosome-bound mRNA and the aa-tRNA, enabling accommodation of the aminoacyl-tRNA, release and recycling of eEF1A. After accommodation of the 3′ end of an aa-tRNA at the A-site, the ribosome—a complex ribozyme—catalyzes peptide bond formation with the P-site located peptidyl tRNA at the catalytic core of the ribosome—the peptidyl transferase center (PTC), located on the large ribosomal subunit rRNA. Nucleotide exchange of eEF1A has been proposed to be the

rate-controlling step of translation in eukaryotes [19]. In eukaryotes, this function is performed by eEF1B. In contrast to eukaryotes where exchange factors form a high macromolecular complex, bacteria have a single exchange factor, EF-Ts. The composition of the nucleotide exchange factor eEF1B differs between eukaryotes depending on the organism [7]. Yeast eEF1B contains eEF1Bα and eEF1Bγ while metazoans contain a heteromer containing at least four subunits: eEF1Bα, eEF1Bγ, eEF1Bδ and the valine-tRNA synthetase (Table 1) ([20]; reviewed in [7]). Only the catalytic α-subunit is required for nucleotide exchange, and eEF1Bγ may stimulate eEF1Bα-activity [21–23]. Unlike eEF1A and EF-Tu, the catalytic eEF1Bα subunit is entirely different from bacterial EF-Ts. Therefore, it is not surprising that EF-Ts interacts with EF-Tu in a manner distinct from the binding of eEF1Bα to eEF1A. In fact, prokaryotic EF-Tu binds EF-Ts through domains I and III (Fig. 1b) [24, 25], whereas eEF1A interacts with eEF1Bα during GDP-GTP exchange through domains I and II [26] (Fig. 1a). Furthermore, eEF1Bα and EF-Ts promote

Table 1 Translation elongation and release factors in eukaryotes, archaea and prokaryotes

Protein					Function
Eukaryotes			Archaea	Bacteria (and organelles)	
eEF1A			aEF1	EF-Tu	Binding of aa-tRNA and recruitment to ribosomal A-site, GTP-binding and hydrolysis
eEFSec			SelB	SelB	Binding of Sec-tRNASec and recruitment to ribosomal A-site, GTP-binding and hydrolysis
eEF1B	eEF1Bα	eEF1Bα	aEF1β	EF-Ts	Nucleotide exchange
	eEF1Bβ	eEF1Bβ			Nucleotide exchange
		eEF1Bδ			
	eEF1Bγ	eEF1Bγ			Structural protein
	–	Val-RS			Aminoacyl-tRNA synthetase
eEF2			aEF2	EF-G	Translocation of peptidyl-tRNA to P-site
Fungal eEF3			–	–	Release of deacylated tRNA from E-site
–			–	LepA/EF4	Back-translocase
eIF5A/4D			aIF5A/aEF5	EF-P	Positioning of Met-tRNAi at P-site, alleviation of ribosomal stalling upon consecutive Pro residues
eRF1			aRF1	RF1 and RF2	Class I RFs
eRF3			–	RF3	Class II RFs

Eukaryotic nucleotide exchange factor eEF1B composition varies between kingdoms. Valyl-tRNA synthetase is abbreviated “Val-RS”

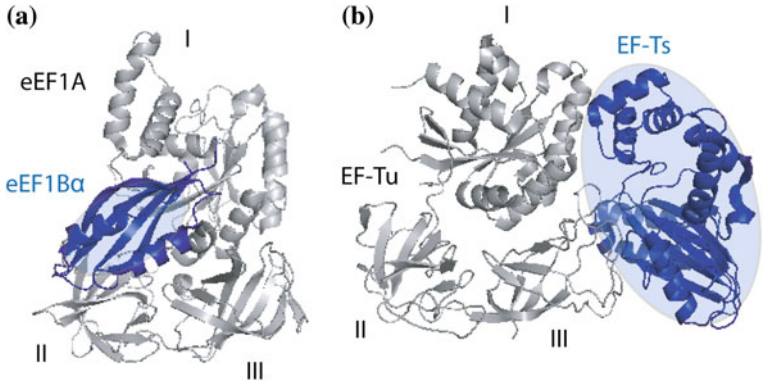


Fig. 1 eEF1A/EF-Tu GDP-GTP exchange and GEFs are not conserved. **a** *S. cerevisiae* eEF1A in complex with its guanine nucleotide exchange factor (GEF) eEF1Bα (*highlighted*). The 3D structure was obtained from PDB, accession number 1F60. **b** *E. coli* EF-Tu in complex with its GEF EF-Ts (*highlighted*). The 3D structure was obtained from PDB, accession number 4PC7. **a** and **b** eEF1A and EF-Tu consist of three domains

nucleotide exchange through mechanisms distinct from each other [6]. eEF1Bα recycles eEF1A-GDP to eEF1A-GTP by inserting a lysine residue into the γ-phosphate and Mg^{2+} binding site to destabilize binding of Mg^{2+} [27], allowing eEF1A to enter a new round of elongation and forming a new aa-tRNA-eEF1A-GTP ternary complex [28]. In contrast, EF-Ts binds to the G-domain of EF-Tu, thereby indirectly destabilizing Mg^{2+} binding and inducing GDP release [27]. Why eukaryotes use a different mechanism than prokaryotes is not clear. A recent report claims that Mg^{2+} is not absolutely required in higher eukaryotes for nucleotide exchange [29], explaining the similar affinity of eEF1A for GDP and GTP in contrast to EF-Tu [30, 31], which binds GDP approximately 100-fold tighter than GTP [32].

Eukaryotic elongation factor 2 (eEF2)—which is very similar to its bacterial counterpart EF-G (Table 1) [4]—is responsible for translocation of tRNAs to the E- and P-sites by binding to the ribosome [33] and was proposed to prevent backward movements of tRNAs (reviewed in [2]). As only two tRNAs can occupy the ribosome at the same time, deacylated tRNAs have to exit to ensure delivery of the next aa-tRNA to the ribosomal A-site. In yeast, eEF3 catalyzes the release of deacylated tRNA from the E-site of the ribosome, thereby acting as an "Exit-(E)-site factor" (Fig. 4) [11, 13] (discussed in the text below).

Some translation elongation factors carry unique post-translational modifications (Fig. 2). Generally, the range of activity of a specific protein can increase by introducing post-translational modifications which change the properties of an amino acid or by impinging on structural features. Furthermore, post-translational modifications may affect interactions with other proteins, change the subcellular organization or alter turnover rates of proteins. In some cases, the significance of post-translational modifications is not completely clear. eEF1A of the eukaryotic

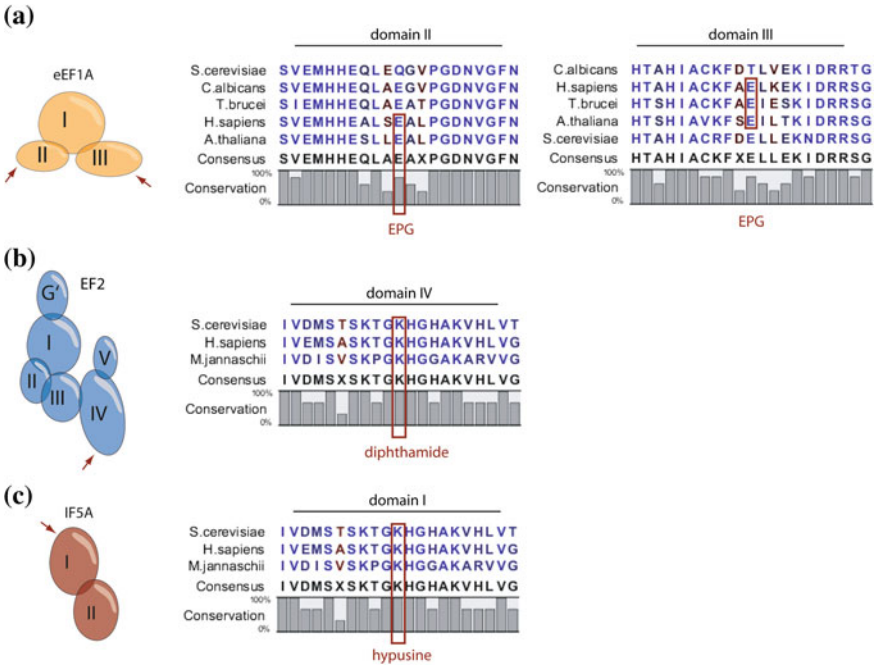


Fig. 2 Unique post-translational modifications in translation elongation factors. **a** eEF1A consists of three domains and is post-translationally modified with ethanolamine phosphoglycerol (EPG) on glutamate residues. This modification is unique to eukaryotic cells. Human eEF1A carries an EPG moiety in domain II and III, whereas *T. brucei* is EPG-modified on domain III only. *S. cerevisiae* (and possibly *C. albicans*) eEF1A lacks EPG modifications. A specific motif requirement for EPG attachment is not known. **b** eEF2 and aEF2 are diphthamide-modified on a conserved lysine residue in domain IV. Despite strict conservation of the lysine residue, bacterial EF-G is not modified. **c** Hypusine is attached to a conserved lysine residue of IF5A in eukaryotes and archaea on domain I of the protein. Some bacteria carry a lysylation or a rhamnosylation modification on the conserved arginine residue. **a**–**c** *Arrows* indicate the sites of modification. Modified residues are framed. Residue conservation is indicated, alignments were created using CLC Sequence Viewer

eEF1 complex undergoes different post-translational modifications such as phosphorylation [34], methylation [35, 36], C-terminal methyl-esterification [36] and ethanolamine phosphoglycerol (EPG) attachment [37, 38]. In addition, a lysine trimethylation conserved between yeast and human has been reported recently [39]. Interestingly, the methyl transferase responsible for methylation of this conserved residue is overexpressed in many cancers indicating that it might have an enhancing effect on protein synthesis [40].

The most conserved post-translational modification on eEF1A is ethanolamine phosphoglycerol (EPG) (Fig. 2a). So far, no other protein has been found to have an EPG attachment. This unique modification was discovered in eEF1A from a human erythroleukemia cell line [37] and a murine lymphocyte cell line [38]. Since then, it has been detected in eEF1A from many eukaryotic species. In plants [41]

and mammals [35, 37, 38], EPG was found to be attached to two glutamic acid side chains within eEF1A domains II and III (Fig. 2a). Notably, residues in close proximity to the EPG attachment site show only little sequence similarity indicating that EPG attachment may be independent of a consensus motif but strictly specific for glutamate. Despite its early discovery, the function of EPG modification of eEF1A has remained elusive. The ancient eukaryote *Trypanosoma brucei* carries a single EPG moiety on domain III only [42], although both glutamate residues in domains II and III are conserved. It has been reported that an 80-amino-acid peptide of domain III attached to a reporter protein is sufficient to recruit the EPG-attachment machinery as long as the critical Glu362 residue is present, indicating that changes around the linkage site have no effect on EPG attachment but may rather depend on the three-dimensional structure of domain III [43]. The three-dimensional structure prediction of *T. brucei* eEF1A domain III indicates that EPG is positioned at the surface of a β-sheet. Surprisingly and unlike other eukaryotic species, *S. cerevisiae* eEF1A is not EPG-modified [44, 45], although the glutamate residue at amino acid position 372 within eEF1A domain III is conserved. In contrast, *Homo sapiens* eEF1A is EPG-modified when expressed in *T. brucei*, suggesting that EPG attachment is conserved between both species [45]. It is possible that—even though the protein is remarkably conserved from its primary to its tertiary structure—a structural requirement in yeast eEF1A may be missing for EPG attachment to occur. Surprisingly, *S. cerevisiae* eEF1A is not EPG-modified when it is expressed in *T. brucei* (E. Greganova and P. Bütikofer, unpublished results). A lack of genes encoding for enzymes required for EPG attachment and/or synthesis in *S. cerevisiae* cannot be excluded. Remarkably, in *Candida albicans* eEF1A the glutamate residue in the putative EPG modification site is replaced by a threonine residue, suggesting that other yeast species are lacking EPG as well. In contrast to most eukaryotes, where the glutamate residues in domains II and III of eEF1A are conserved, EF-Tu lacks those residues, and experiments have shown that *E. coli* EF-Tu is not EPG-modified [38]. Additionally, aEF1A from *Halobacterium salinarum* and *Haloquadratum walsbyi* also lacks EPG [46]. Thus, EPG modification seems to be a feature found in most but not all eEF1As.

Not only the function, but also the mechanism of EPG attachment is unclear. According to one hypothesis, individual EPG components are sequentially added to the glutamic acid residues of eEF1A. Another hypothesis favors the idea that fully synthesized EPG is attached to eEF1A. It seems more likely that a larger structure containing EPG is attached to eEF1A with modification reactions taking place at the protein itself [42]. The aminophospholipid phosphatidylethanolamine (PE) was shown to act as donor of the ethanolamine moiety for EPG. Although a PE-modified version of eEF1A has not been detected yet, the possibility remains that eEF1A is first modified with PE, followed by subsequent deacylation to EPG [42]. It is noteworthy that EPG attachment is not essential for normal growth of *T. brucei* in culture [43, 46, 47]. EPG might be critical in the natural environment of the organism, e.g., under stress conditions or—in the case of *T. brucei*—during the parasite life cycle. Furthermore, it may play a role in protein-protein interactions.

2.1.2 eEFSec/SelB: A Specialized Elongation Factor for Incorporation of the 21st Amino Acid

Selenocysteine occurs in all three domains of life and is co-translationally incorporated as part of the active site of selenoproteins. This specific class consists of proteins with peroxidase and reductase activity and is of relevance for redox reactions, during selenocysteine (Sec) synthesis, selenium transport, hormone metabolism and protein folding (reviewed in [48]). As this "21st" amino acid is essential, a complex molecular machinery to re-define an internal UGA triplet as selenocysteine instead of a stop codon has been developed. UGA is decoded as selenocysteine when the codon is located in a specific context, namely a downstream RNA structure. This structured cis sequence named SECIS (Sec insertion sequence) harbors the shape of a stem loop, lies downstream of UGA and directs insertion of selenocysteine [49]. In bacteria, SECIS lies in close proximity to UGA [49], whereas in eukaryotes, the cis-sequence resides 500-5300 nucleotides downstream of the UGA codon within the 3′ untranslated region (UTR) (Fig. 3) [50–52]. Selenocysteine is synthesized via multiple reactions and requires the action of several enzymes. In eukaryotes and archaea, transcribed tRNASec is in a first step serylated by the conventional SerRS (seryl-tRNA synthetase). In a next step, Sep (O-phosphoserine) is generated by phosphorylation of serine by a phosphoseryl tRNA kinase (PSTK). Finally, Sep is modified to Sec by a Sec synthetase (SepSecS) (reviewed in [53]). In bacteria, conventional SerRS charges tRNASec with serine. Subsequently, the seryl residue is converted to selenocysteine by the selenocysteine synthase SelA [54]. *SELC* encodes tRNASec [55], and the selenium donor selenophosphate is synthesized by the selenophosphate synthetase SelD [56]. Sec-tRNASec (selenocysteyl-tRNASec) is the largest tRNA, and its incorporation requires a specialized elongation factor, SelB (Fig. 3), which cannot be replaced by EF-Tu/EF1A [57–59] and specifically interacts with both downstream sequences and Sec-tRNASec [60–63]. Depending on the UGA codon and the SECIS element, prokaryotic SelB and eukaryotic eEFSec bring tRNASec to the ribosomal A site in a GTP-dependent manner, which is bound with a higher affinity than GDP [49, 57, 64], explaining the observation that—unlike eEF1A—SelB and eEFSec do not require GEF to function [65]. This has led to the hypothesis that there might be no physiological role for eEFSec:GDP. In contrast, GDP-bound eEF1A was suggested to interact with actin [66]. In a yeast two-hybrid assay with *Drosophila melanogaster* eEFSec, dGAPSec—a GTPase-activating protein—was identified as eEFSec interactor and proposed to support UGA read-through in a SECIS-dependent manner [67]. Furthermore, in bacteria SelB directly binds to the SECIS element through its C-terminal domain (Fig. 3a). In contrast, UGA-approaching eEFSec:GTP:tRNASec is recruited to the ribosome by an essential additional protein, SBP2 (SECIS binding protein 2), which is found in a complex with SECIS (Fig. 3b) [68–70]. Other SECIS-binding proteins have been identified, including nucleolin, the DEAD-box helicase eIF4A3, SBP2L and the ribosomal protein L30. But only SBP2 seems to be essential for efficient Sec incorporation, as reviewed in [53].

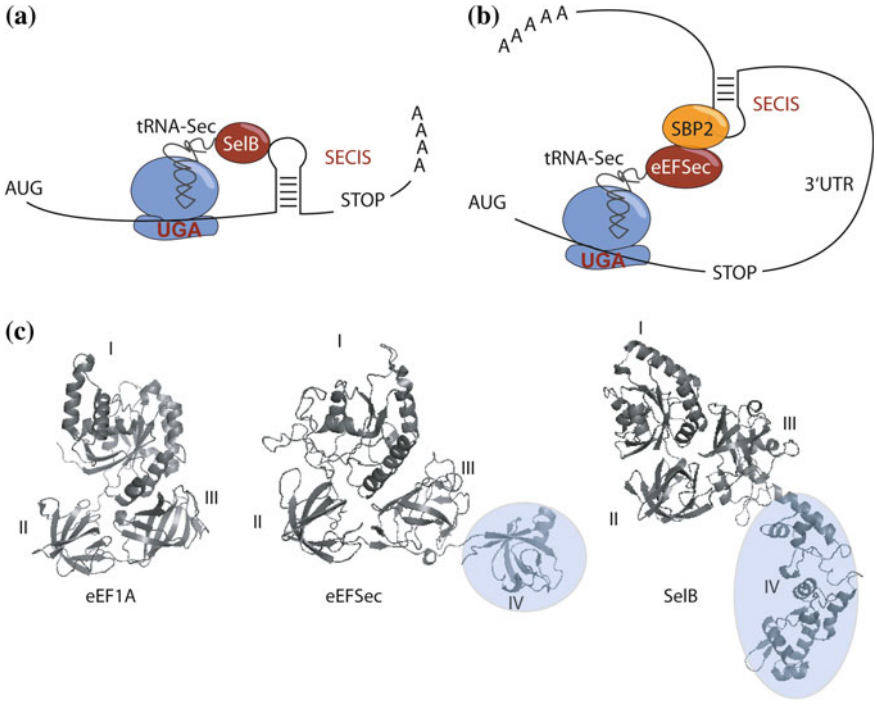


Fig. 3 Incorporation of selenocysteine in bacteria and eukaryotes. **a** In bacteria, UGA is decoded as selenocysteine in the presence of SECIS in the coding region in close proximity to UGA. The *cis*-acting SECIS stem-loop RNA directs the insertion of selenocysteine. SelB, a Sec-tRNA^{Sec}—specific elongation factor delivers Sec-tRNA^{Sec} to the ribosomal A-site. SelB recognizes SECIS via its C-terminal extension (domain IV). **b** In eukaryotes and archaea, the SECIS element is placed in the 3′UTR, 500-5300 nucleotides downstream of the UGA codon. In certain archaea, SECIS resides in the 5′UTR (not shown). Analogously to eEF1A, eEFSec recruits Sec-tRNA^{Sec} to the ribosomal A-site and recognizes the SECIS element by binding to SECIS binding protein 2 (SBP2) found in a complex with SECIS. **c** 3D structures of eEF1A, eEFSec and SelB. In contrast to eEF1A and EF-Tu (not shown), eEFSec and SelB contain a C-terminal extension (domain IV). eEFSec domain IV and SelB domain IV are evolutionarily unrelated. **d** The 3D structure of *S. cerevisiae* eEF1A was obtained from PDB, accession number 1F60. Predicted 3D structures of *D. melanogaster* eEFSec (UniProt Q9W2H0) and *E. coli* SelB (P14081) were obtained using PHYRE^2. 88 % (452 residues) of eEFSec were modelled with 100 % confidence, 94 % (577 residues) of SelB were modelled with 100 % confidence

eEF1A and eEFSec display structural domain similarity. They highly diverge in their primary amino acid sequence with only around 35 % pairwise identity between Drosophila and human [71]. In contrast to eEF1A consisting of three domains (Figs. 1 and 3c), eEFSec contains a C-terminal extension (domain IV), which is also found in prokaryotic SelB (Fig. 3c). Notably, the fourth domains of eEFSec and SelB are not evolutionarily related to each other, which is likely a result of the diversity of tRNASec incorporation. Bacterial and archaeal SelB domain IVs are structurally unrelated as well [72], proposing that domain IV evolved in a

convergent manner. In eukaryotes, it was proposed that domain IV is required for SBP2 and tRNA binding and might be involved in regulating hydrolysis of GTP [73]. The inability of eEFSec to directly bind to the SECIS element and the requirement of SBP2 is the result of an adaptation as SECIS is located far away from the UGA codon in eukaryotes.

Intriguingly, the crystal structure of SelB from the archaeon *Methanococcus maripaludis* revealed that its overall domain arrangement resembles translation initiation factor IF2/eIF5B [74]. An early report proposed that SelB is related to archaebacterial and eukaryotic eIF2γ. They hypothesized that eIF2γ/SelB represent an ancient GTPase subfamily [75], raising the possibility of mechanistic similarities between translation initiation and selenocysteine incorporation and/or the interaction with the ribosome [74, 76].

The evolution and origin of selenoproteins are still a matter of debate. Is incorporation of selenocysteine an ancient trait mostly lost or a new property acquired along evolution? Why should cells maintain selenoproteins instead of using cysteine? Why should a separate machinery for selenoprotein biosynthesis be required? In fact, some organisms such as higher plants and fungi do not have selenoproteins. Instead, activities carried out by selenoproteins are performed by cysteine-containing homologs [77, 78]. However, Sec is advantageous over Cys (cysteine) as it prevents enzymatic inactivation by irreversible oxidation. Oxidized Sec (SeOH) can be recycled to Se, while Cys is not recycled once it is over-oxidized to SO_2^- or SO_3^- [71]. Therefore, selenocysteine catalyzes peroxide scavenging faster than cysteine. Furthermore, selenocysteine can force a substrate to bind differently because of a specific protein folding initiating its incorporation [79].

2.1.3 EF2 and EF-G Mediate Translocation of tRNAs on the Ribosome

A new round of peptide elongation requires translocation of a complex consisting of mRNA, deacylated tRNA in the P-site and the peptidyl-tRNA at the A-site by one codon, positioning the deacylated tRNA at the E-site and the peptidyl-tRNA at the P-site. This pre- to post-translocational state is dependent on EF2 or EF-G, respectively (Table 1). EF2 in eukaryotes and archaea and EF-G in prokaryotes display homologous structures and function [46]. Characteristically, they consist of six domains (termed the I-V and G' domains) with domain I acting as GDP/GTP binding pocket [80]. In mammalian cells, eEF2 function is regulated through phosphorylation on Thr56 by the Ca^{2+}—activated protein kinase eEF2 K resulting in impaired binding to the ribosome and thereby blocking translation [81–83].

By binding to the ribosome, translocation during translation elongation is promoted by eEF2 and EF-G. eEF2 and EF-G insert domain IV into the decoding center of the small ribosomal subunit. Outstandingly, in archaea and eukaryotes a/eEF2 is post-translationally modified with diphthamide at a conserved histidine residue at the structural tip of domain IV mimicking a tRNA-anticodon loop (Fig. 2b) [84]. In contrast to EPG on eEF1A (Fig. 2a), the amino acid residues

around the site of modification are highly conserved and required for formation of diphthamide on eEF2 [84, 85]. Diphthamide is attached to eEF2 in two steps and requires five factors encoded by the genes *DPH1-5*. A well-known consequence of diphthamide modification is to serve as an ADP-ribosylation site for diphtheria toxin from *Corynebacterium diphtheria* (reviewed in [86, 87]), cholix toxin from *Vibrio cholera* [88] and exotoxin A from *Pseudomonas aeruginosa* [89]. ADP ribosylation is an ancient process and serves to attenuate protein synthesis in prokaryotes and eukaryotes as well as viruses [90–92]. Exotoxin A interacts with diphthamide-modified EF2 leading to its ADP-ribosylation by mimicking part of the 80S ribosomal unit, thereby rendering eEF2 inactive and resulting in a block of protein synthesis [93]. Although the diphthamide modification was detected more than 30 years ago [94], its role is not fully understood. Notably, it is not essential for cell viability [95, 96], and the function of eEF2 seems to not directly depend on diphthamide [97]. Therefore, the question remains whether cells have retained this modification to enable pathogen inhibition of protein synthesis. Even more striking, this modification is broadly conserved between eukaryotes and archaea, but absent in bacteria. In light of the fact that it is positioned at the tip of domain IV (Fig. 2b), one might speculate that diphthamide contacts the tRNA, mRNA or rRNA of the ribosomal decoding center to enhance translocation. Indeed some reports have demonstrated that this modification may promote eEF2 function. In mutant mammalian cells, diphthamide was proposed to protect the ribosome from RIPs (ribosome-inactivating proteins) such as ricin, probably by sterically occluding RIP's access to the ribosome [98]. Moreover, frameshifting was demonstrated to increase in diphthamide-deficient eEF2 yeast mutants impairing cell growth [84, 95]. One report suggests that diphthamide may play a role at specific time points during eukaryote development [98]. These reports all propose an enhancing role for diphthamide in translation. Its exact contribution remains to be determined.

2.1.4 eEF3, An Unusual ATPase Promoting tRNA Exit From the Ribosome

Generally, protein synthesis requires two canonical GTPases, eEF1 (EF-Tu in prokaryotes) and eEF2 (EF-G in prokaryotes). In ascomycete yeast species such as *Saccharomyces cerevisiae*, *Candida albicans*, *Cryptococcus neoformans* and *Pneumocystis carinii* [99–103], another factor, eEF3, is found (Fig. 4). eEF3 is an ATPase with two ATP-binding cassettes and—unlike other proteins of this class that are involved in transport across membranes—binds to and functions at the ribosome [104]. Its ATPase activity is stimulated by the ribosome, and ATP hydrolysis is required for dissociation of eEF3 from the ribosome. Unlike eEF1A and eEF2, which bind to the ribosomal A-site, eEF3 contacts the 60S and the head of the 40S ribosomal subunit [105–107] and competes with eEF2 for binding to the ribosome. During translation elongation, eEF3 functions in stimulating eEF1A-dependent binding of cognate aminoacyl-tRNA to the ribosomal A-site [108]. Furthermore, eEF3 acts as an "E-site factor." eEF3-ATP hydrolysis

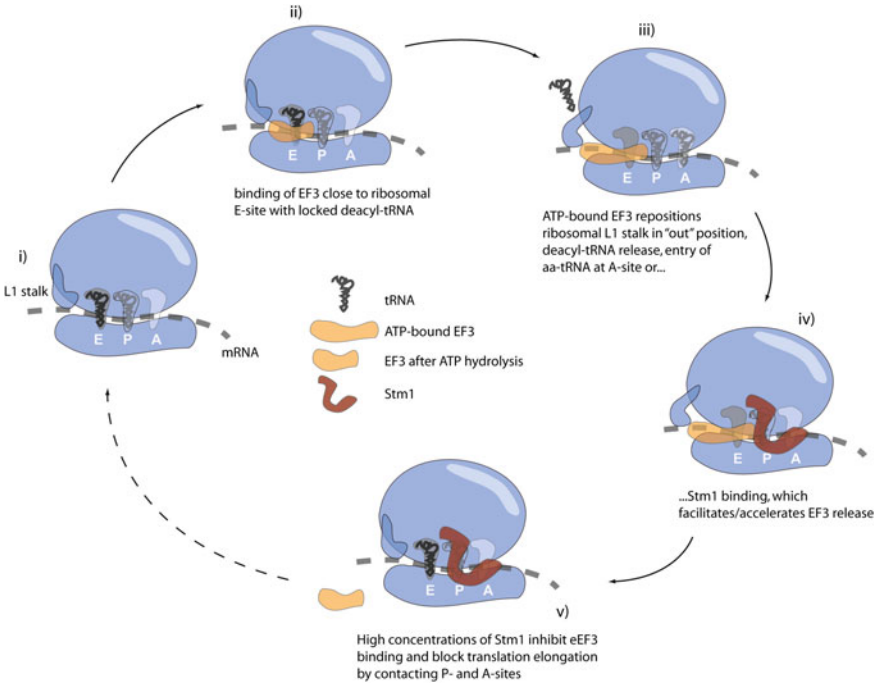


Fig. 4 eEF3-Stm1 antagonism. **i** Before the next aa-tRNA can enter the ribosomal A-site, the E-site tRNA needs to leave. **ii** The ATPase eEF3 acts as a translocase in yeast, lower fungi and certain algae and promotes as an "E-site factor" the release of deacylated-tRNA from the ribosome. **iii** In its ATP-bound form, eEF3 undergoes a conformational change, leading to a stabilization of the ribosomal L1 stalk in its "out position", thereby facilitating the exit of a deacylated-tRNA. **iv** The non-ribosomal protein Stm1 has been proposed to facilitate or accelerate eEF3 release. **v** High concentrations of Stm1, which contacts the ribosomal A- and P-sites can inhibit translation and binding of eEF3

facilitates the release of deacylated tRNA from the ribosomal E-site [11]. The opening of the E-site is a prerequisite for the release of the tRNA [33, 109]. It was proposed that the chromodomain of eEF3 stabilizes the ribosomal L1 stalk in its "out" position (Fig. 4). In yeast, eEF3 seems to be required for rearranging the L1 stalk [110].

In *S. cerevisiae*, eEF3 has a paralog encoded by the non-essential HEF3 gene that arose from whole genome duplication [111, 112]. In other organisms, eEF3 homologs do not exist. Regarding its requirement during translation elongation, it is not surprising that eEF3 is essential for yeast viability [108, 113]. It is not clear yet why yeast translation depends on eEF3, while translation in bacteria and most eukaryotic organisms is eEF3-independent. Furthermore, unique features of yeast ribosomes that would explain the need for an additional elongation factor have not been found yet [2]. Different ribosomal protein composition might explain an eEF3-requirement as discussed below. Very recently, eEF3 has been demonstrated

to bind to the 3′UTR of a specific mRNA in a ribosome-independent manner proposing a new regulatory function [114].

2.1.5 LepA/EF4 Enables tRNA Back-Translocation

Bacterial EF4 is a GTPase originally called LepA [115] since its coding sequence is the first cistron of the bicistronic lep operon of the signal peptidase I in *E. coli* [116]. Due to its remarkable feature catalyzing a backward movement of both peptidyl-tRNA and deacylated tRNA from the post- to the pre-translational state (in the opposite direction as the one catalyzed by EF-G)—i.e., from the ribosomal P- and E-sites to the A- and P-sites—it has been renamed elongation factor 4 (EF4). Hence, EF4 is thought to function as a back-translocase on the elongating ribosome providing a possibility for a proper translocation reaction catalyzed by EF-G to increase translational fidelity. A report has demonstrated that overexpression of EF4 in vivo is toxic and leads to non-productive translation [117]. EF4 is found in bacteria and mitochondria and chloroplasts of eukaryotes—but is absent in archaea and the cytoplasm of eukaryotes—and displays a similar structure to the elongation factor EF-G [118–120], resembling the aa-tRNA-EF-Tu-GTP ternary complex [121]. With EF-Tu and EF-G, EF4 is the third most highly conserved bacterial protein displaying 55-68 % amino acid identity among different species ([117]; reviewed in [121]). EF4 consists of six domains of which four are similar to EF-G. Domain IV of EF-G is lacking a corresponding domain in EF4. EF-G and EF4 bind to the ribosomal A-site by contacting the ribosome with domains I, II, III and V. The C-terminal domain of EF-4 contacts the ribosome in a manner different from domain IV of EF-G [122]. It is able to remodel the ribosomal decoding center in a way opposite to that induced by the accommodation of a tRNA in the A-site, leading to an A-site re-opening [123]. EF4 is also thought to compete with EF-G for pre-translocational ribosomes. Recent data propose that the last 44 C-terminal amino acids of EF4 play a role in the GTP-dependent function on the ribosome. Furthermore, the C-terminal domain is universally conserved among EF4 homologs and represents an identity element of EF4, i.e., if an EF4-CTD is found within a protein, it is classified as EF4 (reviewed in [124]).

Eukaryotic cells have an EF4 homolog, named Guf1 (GTPase of unknown function) [125]. Guf1 is conserved in all eukaryotic genomes and is found in mitochondria where it interacts with translating mitochondrial ribosomes [117]. Notably, Guf1 is not essential; in fact, yeast deletion mutants display faster growth at 14 °C compared to wild-type cells [125].

2.1.6 IF5/EF-P Promote Translation Elongation

Structurally, e/aIF5A is related to the bacterial protein EF-P (reviewed in [2]). Eukaryotic translation initiation factor eIF5A was formerly called IF-M2Bα and later eIF4D [126–128]. Initially, IF5A was thought to be involved in translation

initiation as it was able to enhance the yield of methionyl-puromycin synthesis in an in vitro translation system for *E. coli*, an assay supposed to monitor translation initiation [129–134]. More recent studies in yeast propose that eIF5A is involved in translation elongation rather than initiation as depletion or inactivation of eIF5A resulted in an increased ribosomal transit time, accumulation of polysomes, mimicking the effects of an eEF2 inhibitor [135–137]. It has been suggested that eIF5A promotes translation of a subset of mRNAs only [138]. Furthermore, IF5 and EF-P have been assigned to alleviate the ribosome stalling that often occurs during synthesis of proteins with consecutive Pro residues [139–142] by stimulating the formation of new peptide bonds [143]. The crystal structure of EF-P bound to the 70S ribosome of *Thermus thermophilus* demonstrated that EF-P binds between the P- and E-sites of the ribosome and mimics with its size and shape a tRNA, similar to eIF5A (Fig. 5) [143, 144]. For its association with the ribosome, eIF5A requires a specific and unique post-translational modification, hypusine (Fig. 2c). This modification is conserved in archaea and eukaryotes—but absent in bacteria—and is attached to a Lys residue at the top of domain I [145] in a two-step reaction requiring spermidine [146]. Although the mode of action of eIF5A during translation is not fully elucidated, early studies demonstrate that the hypusine modification on eIF5A is required to stimulate the formation of methionyl-puromycin [134, 147]. Noticeably and consistent with a stimulatory role of the hypusine modification, the deoxyhypusine synthase gene is essential in yeast (reviewed in [2]). Furthermore, hypusine was proposed to determine the localization of eIF5A in the cytoplasm [148].

Notably, the residue in EF-P corresponding to the hypusine attachment site in IF5 is also Lys or Arg, and some bacteria, e.g., *E. coli* and *Salmonella enterica*, modify the Lys residue by addition of β-lysine, resulting in a modification that resembles the hypusine side chain [149–151]. This modification was found to be essential for stimulation of methionyl-puromycin synthesis [152]. In bacteria

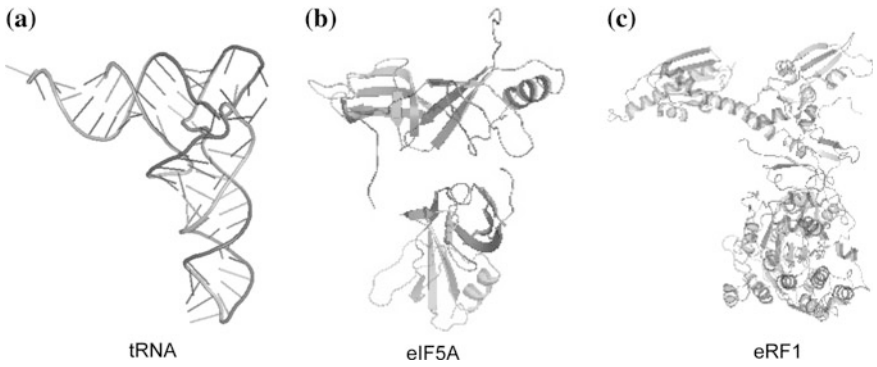


Fig. 5 IF5 and eRF1 structurally mimic tRNA. **a** 3D structure of a tRNA (PDB accession number 1EHZ). **b** 3D structure of *S. cerevisiae* eIF5A (PDB accession number 3ERO (UniProt P23301)). **c** 3D structure of *S. cerevisiae* eRF1 (PDB accession number 4CRM (UniProt P12385))

carrying an Arg residue at this position, e.g., *Shewanella oneidensis* and *Pseudomonas aeruginosa*, Arg-rhamnosylation was detected. In all cases, these special post-translational modifications are essential for EF-P activity. A lack of these modifications in bacteria impairs their virulence in mammalian hosts and decreases bacterial fitness [153, 154].

While aIF5A and eIF5A possess a two-domain structure (Fig. 5), bacterial EF-P consists of three domains [144, 155] with EF-P domains I and II corresponding to the N- and C-terminal domains of a/eIF5A [144]. It is unclear whether bacteria incorporated an additional EF-P domain III during evolution, i.e., whether the last common ancestor had an a/eIF5A-structure with two domains or whether the last common ancestor exhibited a three-domain architecture that was reduced to two domains in archaea and eukaryotes (reviewed in [156]). Remarkably, the structure of bacterial EF-P resembles an L-shaped tRNA, with the lysylation site corresponding to the 3′ end of the tRNA [151]. During translation, EF-P domain I is positioned close to the P-site-bound acceptor stem of the initiator tRNA and domain III of EF-P localizes to the anticodon stem loop of the P-site tRNA [143].

The genes encoding aIF5A, eIF5A and EF-P are ubiquitous [156]. Possibly all eukaryotes have two eIF5A isoforms while some plants, e.g., *Arabidopsis thaliana*, have four [157]. Eukaryotes and archaea depend on e/aIF5A for survival. However, the EF-P coding gene is non-essential in many [158, 159] but not all bacteria [160].

2.2 Conservation of Elongation Factors

Genes encoding elongation factors are suitable tools for phylogenetic analyses [161] as they are conserved and ubiquitous among all species. Their functions are essential for translation elongation. *S. cerevisiae* eEF1A displays approximately 80 % sequence identity with *H. sapiens* eEF1A and approximately 74 % with *L. major*, *T. brucei* and *A. thaliana* eEF1A. This raises the question whether eEF1A from one eukaryotic species may cross-complement in other eukaryotes. Complementation studies indirectly help to draw conclusions on how proteins or protein domains have adapted to their environment during evolution and may provide a basis to study functional differences and common features of orthologous factors that are not evident from their amino acid sequences. Functional complementation across certain species has been demonstrated for eukaryotic translation initiation factor 4E (eIF4E) [162]. Similarly, cross-species complementation experiments have been performed with EF1A. Despite very high sequence conservation and very similar predicted three-dimensional structures, eEF1A orthologs were not able to replace endogenous proteins. Nevertheless, inter-species chimeric forms of eEF1A expressed in *S. cerevisiae* were functional in vivo as long as domain I of yeast eEF1A was present [45]. Similarly, *E. coli* EF-Tu domain I is strictly required for protein translation and cannot be replaced by aEF1A domain I. This result may be explained by the fact that the GTPase activities of aEF1A and EF-Tu differ in a number of properties such as different ligands and effector

affinities [163]. The failure to cross-complement individual paralogs in different organisms may be related to the co-evolution of eEF1A, aEF1A and/or EF-Tu with their interacting partners resulting in differences that do not allow the protein to fulfill its function in a different cellular environment. GEFs of eEF1A/EF-Tu are much less conserved—eEF1Bα from yeast and mammals shows only 43 % sequence identity [164]—which might hold true for other interactors of eEF1A as well. Notably, beyond binding to tRNA, eEF1A was demonstrated to fulfill various functions such as nuclear export, viral propagation, apoptosis and proteolysis. Its most common moonlighting function is linked to actin organization (reviewed in [165]), suggesting that eEF1A has multiple interaction partners to which it has adapted and that are most likely much less conserved than eEF1A itself. eEF1A interacts with actin on different levels. It was demonstrated to bind the mRNA of F-actin and β-actin, thereby affecting their cellular localization [166]. This function of eEF1A seems to be conserved between eukaryotes and bacteria. MreB is the bacterial ortholog of actin, and they share a common ancestor [167]. Analogously to eEF1A, EF-Tu interacts with MreB in different bacterial species such as *B. subtilis* and *E. coli* affecting the dynamics and localization of MreB filaments [168–170]. Besides components of the cell wall and the cell membrane, actin is an important factor contributing to the shape of the cell for which eEF1A/EF-Tu interaction seems to be relevant.

Higher vertebrates have two eEF1A isoforms that are 97 % similar and encoded by different genes. These isoforms are mutually exclusively expressed and display tissue specific localization. While eEF1A2 is present in cardiomyocytes, skeletal myocytes and neurons, eEF1A1 is present in the remaining cell types. In some types of cancer eEF1A2 can have oncogene-like properties [171, 172] and is overexpressed [173, 174]. Intriguingly, some eukaryotes, mostly unicellular organisms, such as green and red algae, diatoms, euglenozoans, foraminifera, dinoflagellates and others—as mentioned in [3]—lack eEF1A, which is replaced by an elongation factor-like (EFL) protein [175], which most likely arose by a duplication event [3]. eEF1A and EFL appear to be functionally equivalent [175–181].

As mentioned above and in contrast to other eukaryotes, archaea and bacteria, yeast species require eEF3. From a ribosome-structural point of view, this requirement is not obvious. Mammalian eEF1 and eEF2 are not sufficient to promote in vitro protein biosynthesis in yeast [182], indicating that they do not take over the role of eEF3. Moreover, no eEF3-like factor has been detected in bacteria, archaea and other eukaryotes. Yet, a striking property of yeast is the unique non-ribosomal protein Stm1 that binds to the 40S subunit. Stm1 has been proposed to follow the mRNA path contacting ribosomal RNA residues at the A and P sites [183]. Apparently, Stm1 seems to antagonize the function of eEF3. Consistently, a lack of Stm1 enhances binding of eEF3 to the ribosome, and eEF3 overexpression results in impaired growth of Stm1-depleted cells (Fig. 4) [184]. Therefore, it has been proposed that Stm1 binding could facilitate release of eEF3. Stm1 could act as a translation inhibitor at high concentrations, e.g., under glucose starvation conditions or during quiescence, clamping the two ribosomal subunits until nutrient abundance increases [183–186].

Why is EF4 not found in eukaryotes and archaea? In fact, lepA was found among the genes essential for survival of *H. pylori* in the acidic environment, i.e., the stomach mucosa of its host [187]. Therefore, EF4 could be of physiological significance under conditions that increase the error rate in protein synthesis as well as the number of stalled ribosomes such as high ionic strength. Fast ionic changes may also occur in chloroplasts and mitochondria depending on the rates of photosynthesis and respiration (reviewed in [121]).

3 Termination/Release Factors

In contrast to translation elongation, factors involved in translation termination are not universally conserved at the level of structure and number. Although (release factors) in bacteria and eukaryotes display entirely different protein topologies [188–190], the mechanistic principles of translation termination are similar, and some common features are exploited by both prokaryotes and eukaryotes. The lack of homology between release factors raises the question whether they share a common origin or whether a convergent evolutionary process is responsible for the occurrence of different RFs.

3.1 Mechanism of Translation Termination

When a stop codon is positioned at the A-site of the decoding center, termination of protein synthesis is initiated [191]. This brings up the question of how the three stop codons—UAG, UGA and UAA—are efficiently and correctly recognized. In contrast to the previous steps during translation, termination does not depend on tRNAs but instead relies on codon-specific class-I release factors. These release factors recognize the stop codon presented at the A-site to induce hydrolysis of the peptidyl-tRNA linkage at the peptidyl-transferase center [192, 193]. Similar to elongation, recognition of stop codons is very accurate, and release factors performing on sense codons are rare [194]. Notably, while eukaryotes and archaea have an omnipotent decoding class-I release factor, e/aRF1, prokaryotes encode two evolutionarily unrelated codon-specific class-I factors with partial overlapping specificity, RF1 for UAG and UAA and RF2 for UGA and UAA [195, 196] (Fig. 6). Regardless of which stop codon has been engaged by a class-I release factor, the GTPase activity of a class-II release factor ((e)RF3) facilitates recycling of class-I release factors. In contrast to bacteria—as discussed below—termination in eukaryotes depends only on two factors, eRF1 and eRF3. eRF1 recognizes the stop codon and is further responsible for peptidyl-tRNA hydrolysis. eRF3, which is related to EF-Tu and eEF1A, collaborates during this process [197–199] and binds to eRF1 in the absence of the ribosome [197, 198]. Only in the presence of eRF1, eRF3 is able to bind GTP [200, 201]. Moreover, eRF3 is also proposed to promote

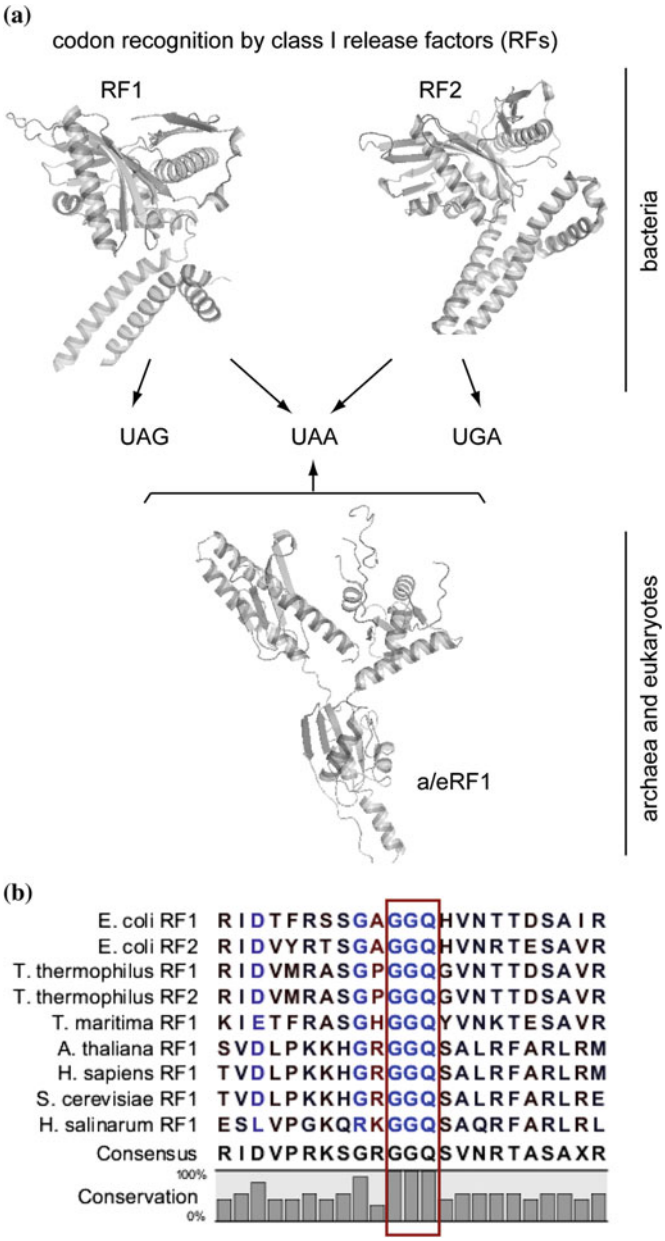


Fig. 6 Codon recognition by class I release factors. **a** Archaeal and eukaryotic class I release factor (a/eRF1) recognizes all three stop codons, UAG, UAA and UGA. In contrast, bacteria encode for two codon-specific class I RFs, RF1 and RF2. UAA is recognized by both factors, while UAG requires RF1 and UGA RF2, respectively. Bacterial RFs and a/eRF1 are evolutionarily not conserved. **b** The GGQ motif (*framed*) is highly conserved between prokaryotes and eukaryotes and required for accommodation of class I RFs to the ribosomal A-site

eRF1 binding to an A-site stop codon containing ribosomal pre-termination complexes (analogous to the function of EF-Tu helping to increase the aminoacyl-tRNA affinity for the A-site) (reviewed in [202]). Once an eRF1-eRF3-GTP ternary complex is formed and binds to a stop codon, an eRF3-mediated conformational change is induced by hydrolysis of GTP leading to eRF3 dissociation from eRF1. This conformational change permits eRF1 to fully accommodate to the A-site and brings its exceptionally well-conserved GGQ primary sequence motif (Fig. 6) in proximity to the ester bond between the tRNA and the nascent polypeptide [2, 5, 203]. eRF1 harbors the shape of a tRNA (Fig. 5). Notably, the role of eRF3 in bringing a tRNA-like molecule to the ribosome appears similar to the one performed by EF1A/EF-Tu. eRF1 is composed of three domains [188], N, M, and C [204]. Each of them is crucial for the termination process (see also below). The amino-terminal domain (domain N) recognizes the termination codon through a codon:anticodon-like interaction with a conserved NIKS peptide motif localized within the distal loop of the N-domain. Other motifs such as YxCxxxF and the GTS consensus motif also seem to be crucial for proper interaction with the stop codon [205–207]. Unlike in bacteria (see below), it has been proposed that eRF1 recognizes stop codons through a three-dimensional structural network formed by several residues of these motifs rather than by simple codon-peptide interactions [204]. The tRNA acceptor stem is mimicked by the middle (M) domain of eRF1 that extends into the peptidyl transferase center and promotes peptide release with the help of the highly conserved GGQ motif [188] to correctly position a water molecule for nucleophilic attack and to discriminate it from other potential nucleophiles [188, 208, 209]. The carboxy-terminal (C) domain of eRF1 promotes the interaction with eRF3 [210–212]. Archaeal translation termination seems to be similar to that found for eukaryotes. The main difference is that elongation factor aEF1A replaces a class-II release factor [213].

Stop codons are conserved between eukaryotes and bacteria. However, release factors involved in termination are significantly different. In contrast to eRF1, bacterial RF1 and RF2 consist of four domains. Domain I interacts with the 50S subunit during translation termination. Conserved P(A/V)T and SPF motifs found within RF1 and RF2, respectively, act as “tripeptide anticodons” [214] and are positioned to favor a codon:anticodon-like interaction [215–218] on domain 2 of RF1/2. These tripeptide anticodons also define the specificity for the second nucleotide of the stop codon, and they functionally correspond to the eRF1 NIKS motif found in the N-domain [188, 219–221]. As these primary sequences are completely unrelated, stop codon recognition is different and appears much simpler in bacteria than in eukaryotes. RF1/2 domain 3 contains the conserved GGQ motif involved in peptidyl-tRNA hydrolysis [222]. Together with domain 2, domain 4 is involved in stop-codon recognition by forming a compact superdomain (reviewed in [223]). Unlike in eukaryotic cells, prokaryotic RF3 is found in a subset of bacteria only and is not orthologous to eRF3 [120]. Its GTPase activity is ribosome-dependent and stimulated by RF1 and RF2 [188, 224, 225]. Bacterial RF3 resembles EF-G more than EF-Tu and plays a role as a dissociation factor for RF1 and RF2 from the ribosome upon hydrolysis of the peptidyl-tRNA [226].

Unlike eRF3, RF3 does not form a stable complex with class-I RFs, and it does not couple stop codon recognition and hydrolysis of peptidyl-tRNA [199, 227–229]. Reports have demonstrated that RF3 assists during a retrospective editing reaction initiated by mistakes occurring during elongation, leading to premature translation termination [230]. Recently, it has been postulated that maintaining high-fidelity protein synthesis is the main function of RF3 rather than supporting termination per se [231].

In contrast to prokaryotes, eukaryotic cells require the translational GTPase class-II eRF3 for viability. It is essential for stop codon decoding in a GTP-hydrolysis-dependent manner while prokaryotic RF3 is not [232–235].

3.2 Translation Termination in Prokaryotes and Eukaryotes: Independent Origins?

It has been suggested that release factors have evolved to replace an RNA-based machinery from the RNA world [223]. Strikingly and consistent with a former RNA-based release machinery, deacylated tRNAs are—although at lower rates than release factors—able to promote peptide release in a codon-dependent manner [236, 237], supporting the idea of deacylated tRNAs being "release RNAs" in an ancient RNA world. Low accuracy and efficiency could have favored the faster protein-based termination found nowadays. The existence of an ancient mechanism is further supported by the fact that a common ancient release factor ancestor is missing. Due to lack of sequence homologies between eukaryotic and bacterial release factors, independent origins of translation termination that have functionally converged during evolution are proposed [198, 238]. This hypothesis is also reflected by the occurrence of two class-I RFs in bacteria, while there is only one in archaea and eukaryotes. Furthermore, not all bacteria encode orthologs of class-II RF3. For example, RF3 was neither detected in small-genome bacteria nor in organelle translation systems (reviewed in [202]), whereas eRF3 is essential in eukaryotes [239]. Anyway, later studies reported that eRF1 overexpression is able to rescue translation termination activity of an eRF3 temperature-sensitive yeast mutant [240]. Bacterial RF1 and RF2 would probably result from gene duplication in the lineage of bacteria [221]. While bacterial RF3 might have arisen from the EF2/EF-G lineage [241], eukaryotic eRF3 could have originated from the eEF1 family [242]. In fact, it has been suggested that at an early stage in eukaryotic evolution an eEF1A gene duplication may have occurred with one paralog resulting in eRF3 [243]. Consistently and analogously to eEF1A binding to aa-tRNAs, eRF3 binds and transports tRNA-mimicking eRF1 [188] to the A-site of the ribosome. It has been demonstrated that eRF1 contacts the P-site tRNA to structurally resemble a tRNA in the A-site [204]. Thus, both proteins—eEF1A and eRF3—fulfill their function in the same molecular environment at the ribosomal A-site. So far, class-II RFs have been found in eukaryotes and bacteria but not in archaea, where aEF1A

fulfills the role of a class-II RF [242, 244], further supporting the idea that eRF3 originates from the eEF1 family.

a/eRF1 and RFs from bacteria carry out the same function at the same ribosomal locus. Nevertheless, their structures show no similarities [189]. Are these differences a result of divergence from a common origin? As mentioned above, a common RF ancestor is absent. Eukaryotic and archaeal class-I RFs are homologous (Fig. 6); therefore, it is thought that the eukaryotic termination system has evolved from an archaeal-based machinery [245]. Indeed, aRF1 is able to catalyze an in vitro release of peptides from ribosomes with eukaryotic origin [246]. As discussed above, eRF3 may have evolved from eEF1, which replaces eRF3 in archaea. Sequence similarity between a/eRF1 and bacterial class-I RFs is absent, and structural [247] and functional differences are pronounced as well [198, 210, 238, 240, 248–250], arguing against a common origin of bacterial and archaeal/eukaryotic RFs. Although release factors clearly differ between eukaryotes and bacteria, the GGQ motif in the primary sequence is conserved among all species (Fig. 6) and was proposed to act as a molecular mimicry of the aminoacyl-tRNA-CCA3′ acceptor stem [196, 205]. In fact, the Gln residue is essential for viability in RFs of prokaryotes and eukaryotes [188, 251], indicating that efficient hydrolysis of the peptidyl-tRNA provides a very strong constraint. Interestingly, the Gln side chain of this motif is post-translationally modified by methylation in both bacteria and eukarya [252, 253]. While the role of this RF methylation in eukaryotes is unclear, RF methylation in bacteria stimulates its function (reviewed in [254]). Intriguingly, the functional equivalency to the anticodon of a tRNA is mimicked in both bacteria and eukaryotes by the consensus motifs P(A/V)T and SPF or YxCxxxF, NIKS and GTS, respectively. The later ones —in contrast to RF1/2—form a three-dimensional network characteristic for eukaryotic eRF1 codon recognition. In bacteria, class-I RFs are codon-specific, and it is suggested that the first and the third amino acids of the P(A/V)T motif in RF1 and the SPF motif in RF2, respectively, discriminate the second and the third bases of the stop codon. Therefore, these motifs have been termed “tripeptide anticodons” (reviewed in [202]). The different motifs for codon recognition in eukaryotes and prokaryotes suggest independent mechanisms unlikely to share a common origin.

The ribosome core is formed by ribosomal RNA. Its three-dimensional structure as well as rRNA nucleotide sequences is broadly conserved. Therefore, what might be the reason for such a big difference between RFs fulfilling the same function and acting at the same site on the ribosome? The different factors may reflect the fact that ribosomal proteins decorating the periphery of the ribosome are much more divergent than the rRNA. Additionally, eRF1 and eRF3 family proteins are also involved in mRNA quality control surveillance mechanisms, such as NMD (nonsense mediated decay) and NGD (no-go decay) [240, 255–257]. Thus, RFs could have adapted to their differing protein environments and their potential interactors. On the other hand, eRF1 and eRF3 gene duplications were suggested to have driven the evolution of mRNA decay mechanisms in eukaryotes [243]. An additional level of complexity is brought in by deviations from the universal genetic code. For example, ciliates show codon reassignments that might be the result of a complex

interplay between eRF1 and tRNAs [258]. In vitro and in vivo studies have shown that eRF1 competes with nonsense suppressor tRNAs when recognizing a stop codon [259, 260]. These nonsense suppressor tRNAs are mutated in their anticodon to suppress the termination of translation upon a stop codon. Generally, changes in one of the translation factors may evolutionarily affect the rate of the entire translational machinery [258].

4 Conclusions

The ribosome is a highly conserved ribozyme, and protein synthesis is a process occurring in all domains of life that use the same genetic code for deciphering amino acid sequences. But yet, as evidenced above, specific differences are found within organisms. In particular, the number, composition and function of translation elongation and termination factors vary between organisms. Eukaryotic eEF1B, for instance, consists of multiple protein subunits, while bacteria rely on the single-protein nucleotide exchange factor EF-Ts (Table 1). Furthermore, these exchange factors exploit differing mechanisms for nucleotide exchange. A nice example of species-specific adaptation is given by the existence of eEF3 and its antagonizing factor Stm1 (Fig. 4). Why they only arise in fungi is unclear. The assumed interplay between these proteins seems to provide a strategy for fungi to quickly adjust translation elongation to fast-changing conditions such as nutrient deprivation. EF4 in certain bacteria and organelles is thought to enable a proof-reading mechanism crucial to maintain translational fidelity under conditions when significant changes in ion concentrations are occurring. Some factors—such as eIF5A or EF-P, respectively—able to promote translation under conditions that do not favor efficient translation are conserved between all domains of life (Table 1). Probably all organisms utilize e/aIF5A or EF-P for peptide bond formation when polyproline stretches are encountered to avoid stalling of ribosomes. More significant differences between factors are found within RFs. Currently, it is not clear whether they share a common origin or not. Remarkably, although there is no homology between RFs from bacteria and eukaryotes/archaea, they all share an extremely conserved GGQ motif required for peptide release (Fig. 6b). Another very outstanding feature is the highly recurring molecular mimicry. Class-I RFs imitate a tRNA, while class-II RF3 mimics a/eEF1A or EF-Tu recruiting an aa-tRNA to the ribosomal A-site. Further, a structural tRNA mimic is crucial for the role of a/eIF5A and EF-P during the enhancement of peptide formation between the ribosomal P- and E-sites. As a matter of particular interest, even pathogen toxins such as endotoxin A exploit the strategy of structurally mimicking a translational component. Surprisingly, unique modifications on translation elongation factors are also maintained among different forms of life. Nevertheless, although pronounced differences in translation elongation and termination factors exist among bacteria, archaea and eukaryotes, the principles of elongation and termination are conserved.

Acknowledgments This work was supported by Swiss National Foundation grants 31003A_146722/1 (M. Altmann) and 31003A_149353/1 (P. Bütikofer).

References

1. Kapp LD, Lorsch JR. The molecular mechanics of eukaryotic translation. Annu Rev Biochem. 2004;73:657–704. doi:10.1146/annurev.biochem.73.030403.080419.
2. Dever TE, Green R. The elongation, termination, and recycling phases of translation in eukaryotes. Cold Spring Harb Perspect Biol. 2012;4:a013706. doi:10.1101/cshperspect.a013706.
3. Hernandez G, Proud CG, Preiss T, Parsyan A. On the Diversification of the Translation Apparatus across Eukaryotes. Comp Funct Genomics. 2012;2012:256848. doi:10.1155/2012/256848.
4. Andersen GR, Nissen P, Nyborg J. Elongation factors in protein biosynthesis. Trends Biochem Sci. 2003;28:434–41. doi:10.1016/S0968-0004(03)00162-2.
5. Jackson RJ, Hellen CU, Pestova TV. Termination and post-termination events in eukaryotic translation. Adv Protein Chem Struct Biol. 2012;86:45–93. doi:10.1016/B978-0-12-386497-0.00002-5.
6. Rodnina MV, Wintermeyer W. Recent mechanistic insights into eukaryotic ribosomes. Curr Opin Cell Biol. 2009;21:435–43. doi:10.1016/j.ceb.2009.01.023.
7. Le Sourd F, Boulben S, Le Bouffant R, Cormier P, Morales J, Belle R, Mulner-Lorillon O. eEF1B: At the dawn of the 21st century. Biochim Biophys Acta. 2006;1759:13–31. doi:10.1016/j.bbaexp.2006.02.003.
8. Bec G, Kerjan P, Waller JP. Reconstitution in vitro of the valyl-tRNA synthetase-elongation factor (EF) 1 beta gamma delta complex. Essential roles of the NH2-terminal extension of valyl-tRNA synthetase and of the EF-1 delta subunit in complex formation. J Biol Chem. 1994;269:2086–92.
9. Belfield GP, Bauer M, Ross-Smith N, Tan P, Colthurst DR, Tuite MF. EF-3: a novel fungal elongation factor with homology to E. coli ribosomal protein S5. Biochem Soc Trans. 1993;21:331S.
10. Belfield GP, Tuite MF. Translation elongation factor 3: a fungus-specific translation factor? Mol Microbiol. 1993;9:411–8.
11. Triana-Alonso FJ, Chakraburtty K, Nierhaus KH. The elongation factor 3 unique in higher fungi and essential for protein biosynthesis is an E site factor. J Biol Chem. 1995;270:20473–8.
12. Hericourt F, Jupin I. Molecular cloning and characterization of the Arabidopsis thaliana alpha-subunit of elongation factor 1B. FEBS Lett. 1999;464:148–52.
13. Anand M, Balar B, Ulloque R, Gross SR, Kinzy TG. Domain and nucleotide dependence of the interaction between Saccharomyces cerevisiae translation elongation factors 3 and 1A. J Biol Chem. 2006;281:32318–26. doi:10.1074/jbc.M601899200.
14. Lucas-Lenard J. Protein biosynthesis. Annu Rev Biochem. 1971;40:409–48. doi:10.1146/annurev.bi.40.070171.002205.
15. Sengupta S, Higgs PG. Pathways of Genetic Code Evolution in Ancient and Modern Organisms. J Mol Evol. 2015;80:229–43. doi:10.1007/s00239-015-9686-8.
16. Graifer D, Karpova G. Interaction of tRNA with eukaryotic ribosome. Int J Mol Sci. 2015;16:7173–94. doi:10.3390/ijms16047173.
17. Wilson DN, Doudna Cate JH. The structure and function of the eukaryotic ribosome. Cold Spring Harbor perspectives in biology. 2012; 4. doi:10.1101/cshperspect.a011536.
18. Khalyfa A, Bourbeau D, Chen E, Petroulakis E, Pan J, Xu S, Wang E. Characterization of elongation factor-1A (eEF1A-1) and eEF1A-2/S1 protein expression in normal and wasted mice. J Biol Chem. 2001;276:22915–22. doi:10.1074/jbc.M101011200.

19. Janssen GM, Moller W. Kinetic studies on the role of elongation factors 1 beta and 1 gamma in protein synthesis. J Biol Chem. 1988;263:1773–8.
20. Minella O, Mulner-Lorillon O, Bec G, Cormier P, Belle R. Multiple phosphorylation sites and quaternary organization of guanine-nucleotide exchange complex of elongation factor-1 (EF-1betagammadelta/ValRS) control the various functions of EF-1alpha. Biosci Rep. 1998;18:119–27.
21. Motoyoshi K, Iwasaki K. Resolution of the polypeptide chain elongation factor-1 beta gamma into subunits and some properties of the subunits. J Biochem. 1977;82:703–8.
22. Motoyoshi K, Iwasaki K, Kaziro Y. Purification and properties of polypeptide chain elongation factor-1 beta gamma from pig liver. J Biochem. 1977;82:145–55.
23. Janssen GM, Moller W. Elongation factor 1 beta gamma from Artemia. Purification and properties of its subunits. European J Biochem/ FEBS. 1988; 171:119–129.
24. Kawashima T, Berthet-Colominas C, Wulff M, Cusack S, Leberman R. The structure of the Escherichia coli EF-Tu.EF-Ts complex at 2.5 A resolution. Nature. 1996;379:511–8. doi:10.1038/379511a0.
25. Wang Y, Jiang Y, Meyering-Voss M, Sprinzl M, Sigler PB. Crystal structure of the EF-Tu. EF-Ts complex from Thermus thermophilus. Nat Struct Biol. 1997;4:650–6.
26. Andersen GR, Pedersen L, Valente L, Chatterjee I, Kinzy TG, Kjeldgaard M, Nyborg J. Structural basis for nucleotide exchange and competition with tRNA in the yeast elongation factor complex eEF1A:eEF1Balpha. Mol Cell. 2000;6:1261–6.
27. Sivan G, Aviner R, Elroy-Stein O. Mitotic modulation of translation elongation factor 1 leads to hindered tRNA delivery to ribosomes. J Biol Chem. 2011;286:27927–35. doi:10.1074/jbc.M111.255810.
28. Andersen GR, Nyborg J. Structural studies of eukaryotic elongation factors. Cold Spring Harb Symp Quant Biol. 2001;66:425–37.
29. Crepin T, Shalak VF, Yaremchuk AD, Vlasenko DO, McCarthy A, Negrutskii BS, Tukalo MA, El'skaya AV. Mammalian translation elongation factor eEF1A2: X-ray structure and new features of GDP/GTP exchange mechanism in higher eukaryotes. Nucleic Acids Res. 2014;42:12939–48. doi:10.1093/nar/gku974.
30. Miller DL, Weissbach H. Interactions between the elongation factors: the displacement of GPD from the TU-GDP complex by factor Ts. Biochem Biophys Res Commun. 1970;38:1016–22.
31. Saha SK, Chakraburtty K. Protein synthesis in yeast. Isolation of variant forms of elongation factor 1 from the yeast Saccharomyces cerevisiae. J Biol Chem. 1986;261:12599–603.
32. Gaucher EA, Das UK, Miyamoto MM, Benner SA. The crystal structure of eEF1A refines the functional predictions of an evolutionary analysis of rate changes among elongation factors. Mol Biol Evol. 2002;19:569–73.
33. Spahn CM, Gomez-Lorenzo MG, Grassucci RA, Jorgensen R, Andersen GR, Beckmann R, Penczek PA, Ballesta JP, Frank J. Domain movements of elongation factor eEF2 and the eukaryotic 80S ribosome facilitate tRNA translocation. The EMBO J. 2004;23:1008–19. doi:10.1038/sj.emboj.7600102.
34. Venema RC, Peters HI, Traugh JA. Phosphorylation of elongation factor 1 (EF-1) and valyl-tRNA synthetase by protein kinase C and stimulation of EF-1 activity. J Biol Chem. 1991;266:12574–80.
35. Dever TE, Costello CE, Owens CL, Rosenberry TL, Merrick WC. Location of seven post-translational modifications in rabbit elongation factor 1 alpha including dimethyllysine, trimethyllysine, and glycerylphosphorylethanolamine. J Biol Chem. 1989;264:20518–25.
36. Zobel-Thropp P, Yang MC, Machado L, Clarke S. A novel post-translational modification of yeast elongation factor 1A. Methylesterification at the C terminus. J Biol Chem. 2000; 275:37150–37158. doi:10.1074/jbc.M001005200.
37. Rosenberry TL, Krall JA, Dever TE, Haas R, Louvard D, Merrick WC. Biosynthetic incorporation of [3H]ethanolamine into protein synthesis elongation factor 1 alpha reveals a new post-translational protein modification. J Biol Chem. 1989;264:7096–9.

38. Whiteheart SW, Shenbagamurthi P, Chen L, Cotter RJ, Hart GW. Murine elongation factor 1 alpha (EF-1 alpha) is posttranslationally modified by novel amide-linked ethanolamine-phosphoglycerol moieties. Addition of ethanolamine-phosphoglycerol to specific glutamic acid residues on EF-1 alpha. J Biol Chem. 1989;264:14334–41.
39. Hamey JJ, Winter DL, Yagoub D, Overall CM, Hart-Smith G, Wilkins MR. Novel N-terminal and lysine methyltransferases that target translation elongation factor 1A in yeast and human. Mol Cell Proteomics: MCP. 2015;. doi:10.1074/mcp.M115.052449.
40. Uhlen M, Bjorling E, Agaton C, Szigyarto CA, Amini B, Andersen E, Andersson AC, Angelidou P, Asplund A, Asplund C, Berglund L, Bergstrom K, Brumer H, Cerjan D, Ekstrom M, Elobeid A, Eriksson C, Fagerberg L, Falk R, Fall J, Forsberg M, Bjorklund MG, Gumbel K, Halimi A, Hallin I, Hamsten C, Hansson M, Hedhammar M, Hercules G, Kampf C, Larsson K, Lindskog M, Lodewyckx W, Lund J, Lundeberg J, Magnusson K, Malm E, Nilsson P, Odling J, Oksvold P, Olsson I, Oster E, Ottosson J, Paavilainen L, Persson A, Rimini R, Rockberg J, Runeson M, Sivertsson A, Skollermo A, Steen J, Stenvall M, Sterky F, Stromberg S, Sundberg M, Tegel H, Tourle S, Wahlund E, Walden A, Wan J, Wernerus H, Westberg J, Wester K, Wrethagen U, Xu LL, Hober S, Ponten F. A human protein atlas for normal and cancer tissues based on antibody proteomics. Mol Cell Proteomics: MCP. 2005;4:1920–32. doi:10.1074/mcp.M500279-MCP200.
41. Ransom WD, Lao PC, Gage DA, Boss WF. Phosphoglycerylethanolamine posttranslational modification of plant eukaryotic elongation factor 1alpha. Plant Physiol. 1998;117:949–60.
42. Signorell A, Jelk J, Rauch M, Bütikofer P. Phosphatidylethanolamine is the precursor of the ethanolamine phosphoglycerol moiety bound to eukaryotic elongation factor 1A. J Biol Chem. 2008;283:20320–9. doi:10.1074/jbc.M802430200.
43. Greganova E, Heller M, Bütikofer P. A structural domain mediates attachment of ethanolamine phosphoglycerol to eukaryotic elongation factor 1A in Trypanosoma brucei. PLoS ONE. 2010;5:e9486. doi:10.1371/journal.pone.0009486.
44. Cavallius J, Zoll W, Chakraburtty K, Merrick WC. Characterization of yeast EF-1 alpha: non-conservation of post-translational modifications. Biochim Biophys Acta. 1993;1163: 75–80.
45. Eltschinger S, Greganova E, Heller M, Bütikofer P, Altmann M. Eukaryotic translation elongation factor 1A (eEF1A) domain I from S. cerevisiae is required but not sufficient for inter-species complementation. PLoS ONE. 2012;7:e42338. doi:10.1371/journal.pone.0042338.
46. Greganova E, Altmann M, Bütikofer P. Unique modifications of translation elongation factors. The FEBS J. 2011;278:2613–24. doi:10.1111/j.1742-4658.2011.08199.x.
47. Greganova E, Bütikofer P. Ethanolamine phosphoglycerol attachment to eEF1A is not essential for normal growth of Trypanosoma brucei. Scientific reports. 2012;2:254. doi:10.1038/srep00254.
48. Burk RF, Hill KE. Orphan selenoproteins. BioEssays: News Rev Mol Cell Dev Biol. 1999;21:231–7. doi:10.1002/(SICI)1521-1878(199903)21:3<231:AID-BIES7>3.0.CO;2-D.
49. Huttenhofer A, Bock A. Selenocysteine inserting RNA elements modulate GTP hydrolysis of elongation factor SelB. Biochemistry. 1998;37:885–90. doi:10.1021/bi972298k.
50. Berry MJ, Banu L, Harney JW, Larsen PR. Functional characterization of the eukaryotic SECIS elements which direct selenocysteine insertion at UGA codons. The EMBO J. 1993;12:3315–22.
51. Shen Q, Chu FF, Newburger PE. Sequences in the 3′-untranslated region of the human cellular glutathione peroxidase gene are necessary and sufficient for selenocysteine incorporation at the UGA codon. J Biol Chem. 1993;268:11463–9.
52. Berry MJ, Banu L, Chen YY, Mandel SJ, Kieffer JD, Harney JW, Larsen PR. Recognition of UGA as a selenocysteine codon in type I deiodinase requires sequences in the 3′ untranslated region. Nature. 1991;353:273–6. doi:10.1038/353273a0.
53. Shetty SP, Copeland PR. Selenocysteine incorporation: A trump card in the game of mRNA decay. Biochimie. 2015;114:97–101. doi:10.1016/j.biochi.2015.01.007.

54. Forchhammer K, Leinfelder W, Boesmiller K, Veprek B, Bock A. Selenocysteine synthase from Escherichia coli. Nucleotide sequence of the gene (selA) and purification of the protein. J Biol Chem. 1991;266:6318–23.
55. Leinfelder W, Zehelein E, Mandrand-Berthelot MA, Bock A. Gene for a novel tRNA species that accepts L-serine and cotranslationally inserts selenocysteine. Nature. 1988;331:723–5. doi:10.1038/331723a0.
56. Ehrenreich A, Forchhammer K, Tormay P, Veprek B, Bock A. Selenoprotein synthesis in E. coli. Purification and characterisation of the enzyme catalysing selenium activation. European J Biochem/FEBS. 1992; 206:767–773.
57. Fagegaltier D, Hubert N, Yamada K, Mizutani T, Carbon P, Krol A. Characterization of mSelB, a novel mammalian elongation factor for selenoprotein translation. The EMBO J. 2000;19:4796–805. doi:10.1093/emboj/19.17.4796.
58. Rother M, Wilting R, Commans S, Bock A. Identification and characterisation of the selenocysteine-specific translation factor SelB from the archaeon Methanococcus jannaschii. J Mol Biol. 2000;299:351–8. doi:10.1006/jmbi.2000.3756.
59. Forchhammer K, Leinfelder W, Bock A. Identification of a novel translation factor necessary for the incorporation of selenocysteine into protein. Nature. 1989;342:453–6. doi:10.1038/342453a0.
60. Zinoni F, Heider J, Bock A. Features of the formate dehydrogenase mRNA necessary for decoding of the UGA codon as selenocysteine. Proc Natl Acad Sci USA. 1990;87:4660–4.
61. Heider J, Bock A. Targeted insertion of selenocysteine into the alpha subunit of formate dehydrogenase from Methanobacterium formicicum. J Bacteriol. 1992;174:659–63.
62. Baron C, Westhof E, Bock A, Giege R. Solution structure of selenocysteine-inserting tRNA (Sec) from Escherichia coli. Comparison with canonical tRNA(Ser). J Mol Biol. 1993;231:274–92. doi:10.1006/jmbi.1993.1282.
63. Chen GF, Fang L, Inouye M. Effect of the relative position of the UGA codon to the unique secondary structure in the fdhF mRNA on its decoding by selenocysteinyl tRNA in Escherichia coli. J Biol Chem. 1993;268:23128–31.
64. Baron C, Heider J, Bock A. Interaction of translation factor SELB with the formate dehydrogenase H selenopolypeptide mRNA. Proc Natl Acad Sci USA. 1993;90:4181–5.
65. Tujebajeva RM, Copeland PR, Xu XM, Carlson BA, Harney JW, Driscoll DM, Hatfield DL, Berry MJ. Decoding apparatus for eukaryotic selenocysteine insertion. EMBO Rep. 2000;1:158–63. doi:10.1038/sj.embor.embor604.
66. Pittman YR, Kandl K, Lewis M, Valente L, Kinzy TG. Coordination of eukaryotic translation elongation factor 1A (eEF1A) function in actin organization and translation elongation by the guanine nucleotide exchange factor eEF1Balpha. J Biol Chem. 2009;284:4739–47. doi:10.1074/jbc.M807945200.
67. Hirosawa-Takamori M, Ossipov D, Novoselov SV, Turanov AA, Zhang Y, Gladyshev VN, Krol A, Vorbruggen G, Jackle H. A novel stem loop control element-dependent UGA read-through system without translational selenocysteine incorporation in Drosophila. FASEB J: Official publication of the Federation of American Societies for Experimental Biology. 2009;23:107–13. doi:10.1096/fj.08-116640.
68. Copeland PR, Fletcher JE, Carlson BA, Hatfield DL, Driscoll DM. A novel RNA binding protein, SBP2, is required for the translation of mammalian selenoprotein mRNAs. The EMBO J. 2000;19:306–14. doi:10.1093/emboj/19.2.306.
69. Lesoon A, Mehta A, Singh R, Chisolm GM, Driscoll DM. An RNA-binding protein recognizes a mammalian selenocysteine insertion sequence element required for cotranslational incorporation of selenocysteine. Mol Cell Biol. 1997;17:1977–85.
70. Copeland PR, Driscoll DM. Purification, redox sensitivity, and RNA binding properties of SECIS-binding protein 2, a protein involved in selenoprotein biosynthesis. J Biol Chem. 1999;274:25447–54.
71. Gonzalez-Flores JN, Shetty SP, Dubey A, Copeland PR. The molecular biology of selenocysteine. Biomolecular concepts. 2013;4:349–65. doi:10.1515/bmc-2013-0007.

72. Selmer M, Su XD. Crystal structure of an mRNA-binding fragment of Moorella thermoacetica elongation factor SelB. The EMBO J. 2002;21:4145–53.
73. Gonzalez-Flores JN, Gupta N, DeMong LW, Copeland PR. The selenocysteine-specific elongation factor contains a novel and multi-functional domain. J Biol Chem. 2012;287:38936–45. doi:10.1074/jbc.M112.415463.
74. Leibundgut M, Frick C, Thanbichler M, Bock A, Ban N. Selenocysteine tRNA-specific elongation factor SelB is a structural chimaera of elongation and initiation factors. The EMBO J. 2005;24:11–22. doi:10.1038/sj.emboj.7600505.
75. Keeling PJ, Fast NM, McFadden GI. Evolutionary relationship between translation initiation factor eIF-2gamma and selenocysteine-specific elongation factor SELB: change of function in translation factors. J Mol Evol. 1998;47:649–55.
76. Allmang C, Krol A. Selenoprotein synthesis: UGA does not end the story. Biochimie. 2006;88:1561–71. doi:10.1016/j.biochi.2006.04.015.
77. Lobanov AV, Fomenko DE, Zhang Y, Sengupta A, Hatfield DL, Gladyshev VN. Evolutionary dynamics of eukaryotic selenoproteomes: large selenoproteomes may associate with aquatic life and small with terrestrial life. Genome Biol. 2007;8:R198. doi:10.1186/gb-2007-8-9-r198.
78. Lobanov AV, Hatfield DL, Gladyshev VN. Eukaryotic selenoproteins and selenoproteomes. Biochim Biophys Acta. 2009;1790:1424–8. doi:10.1016/j.bbagen.2009.05.014.
79. Nauser T, Steinmann D, Koppenol WH. Why do proteins use selenocysteine instead of cysteine? Amino Acids. 2012;42:39–44. doi:10.1007/s00726-010-0602-7.
80. Yates SP, Jorgensen R, Andersen GR, Merrill AR. Stealth and mimicry by deadly bacterial toxins. Trends Biochem Sci. 2006;31:123–33. doi:10.1016/j.tibs.2005.12.007.
81. Carlberg U, Nilsson A, Nygard O. Functional properties of phosphorylated elongation factor 2. European J Biochem/FEBS. 1990;191:639–45.
82. Ryazanov AG, Davydova EK. Mechanism of elongation factor 2 (EF-2) inactivation upon phosphorylation. Phosphorylated EF-2 is unable to catalyze translocation. FEBS Lett. 1989;251:187–90.
83. Price NT, Redpath NT, Severinov KV, Campbell DG, Russell JM, Proud CG. Identification of the phosphorylation sites in elongation factor-2 from rabbit reticulocytes. FEBS Lett. 1991;282:253–8.
84. Ortiz PA, Ulloque R, Kihara GK, Zheng H, Kinzy TG. Translation elongation factor 2 anticodon mimicry domain mutants affect fidelity and diphtheria toxin resistance. J Biol Chem. 2006;281:32639–48. doi:10.1074/jbc.M607076200.
85. Zhang Y, Liu S, Lajoie G, Merrill AR. The role of the diphthamide-containing loop within eukaryotic elongation factor 2 in ADP-ribosylation by Pseudomonas aeruginosa exotoxin A. Biochem J. 2008;413:163–74. doi:10.1042/BJ20071083.
86. Pappenheimer AM Jr. Diphtheria toxin. Annu Rev Biochem. 1977;46:69–94. doi:10.1146/annurev.bi.46.070177.000441.
87. Oppenheimer NJ, Bodley JW. Diphtheria toxin. Site and configuration of ADP-ribosylation of diphthamide in elongation factor 2. J Biol Chem. 1981;256:8579–81.
88. Jorgensen R, Purdy AE, Fieldhouse RJ, Kimber MS, Bartlett DH, Merrill AR. Cholix toxin, a novel ADP-ribosylating factor from Vibrio cholerae. J Biol Chem. 2008;283:10671–8. doi:10.1074/jbc.M710008200.
89. Chen JY, Bodley JW. Biosynthesis of diphthamide in Saccharomyces cerevisiae. Partial purification and characterization of a specific S-adenosylmethionine:elongation factor 2 methyltransferase. J Biol Chem. 1988;263:11692–6.
90. Pallen MJ, Lam AC, Loman NJ, McBride A. An abundance of bacterial ADP-ribosyltransferases–implications for the origin of exotoxins and their human homologues. Trends in microbiology. 2001; 9:302–307; discussion 308.
91. Krueger KM, Barbieri JT. The family of bacterial ADP-ribosylating exotoxins. Clin Microbiol Rev. 1995;8:34–47.
92. Corda D, Di Girolamo M. Functional aspects of protein mono-ADP-ribosylation. The EMBO J. 2003;22:1953–8. doi:10.1093/emboj/cdg209.

93. Jorgensen R, Merrill AR, Yates SP, Marquez VE, Schwan AL, Boesen T, Andersen GR. Exotoxin A-eEF2 complex structure indicates ADP ribosylation by ribosome mimicry. Nature. 2005;436:979–84. doi:10.1038/nature03871.
94. Van Ness BG, Howard JB, Bodley JW. ADP-ribosylation of elongation factor 2 by diphtheria toxin. NMR spectra and proposed structures of ribosyl-diphthamide and its hydrolysis products. J Biol Chem. 1980;255:10710–6.
95. Kimata Y, Kohno K. Elongation factor 2 mutants deficient in diphthamide formation show temperature-sensitive cell growth. J Biol Chem. 1994;269:13497–501.
96. Phan LD, Perentesis JP, Bodley JW. Saccharomyces cerevisiae elongation factor 2. Mutagenesis of the histidine precursor of diphthamide yields a functional protein that is resistant to diphtheria toxin. J Biol Chem. 1993;268:8665–8.
97. Jorgensen R, Merrill AR, Andersen GR. The life and death of translation elongation factor 2. Biochem Soc Trans. 2006;34:1–6. doi:10.1042/BST20060001.
98. Gupta PK, Liu S, Batavia MP, Leppla SH. The diphthamide modification on elongation factor-2 renders mammalian cells resistant to ricin. Cell Microbiol. 2008;10:1687–94. doi:10.1111/j.1462-5822.2008.01159.x.
99. Qin SL, Xie AG, Bonato MC, McLaughlin CS. Sequence analysis of the translational elongation factor 3 from Saccharomyces cerevisiae. J Biol Chem. 1990;265:1903–12.
100. Myers KK, Fonzi WA, Sypherd PS. Isolation and sequence analysis of the gene for translation elongation factor 3 from Candida albicans. Nucleic Acids Res. 1992;20:1705–10.
101. Di Domenico BJ, Lupisella J, Sandbaken M, Chakraburtty K. Isolation and sequence analysis of the gene encoding translation elongation factor 3 from Candida albicans. Yeast. 1992;8:337–52. doi:10.1002/yea.320080502.
102. Blakely G, Hekman J, Chakraburtty K, Williamson PR. Evolutionary divergence of an elongation factor 3 from Cryptococcus neoformans. J Bacteriol. 2001;183:2241–8. doi:10.1128/JB.183.7.2241-2248.2001.
103. Ypma-Wong MF, Fonzi WA, Sypherd PS. Fungus-specific translation elongation factor 3 gene present in Pneumocystis carinii. Infect Immun. 1992;60:4140–5.
104. Jones PM, George AM. Mechanism of ABC transporters: a molecular dynamics simulation of a well characterized nucleotide-binding subunit. Proc Natl Acad Sci USA. 2002;99:12639–44. doi:10.1073/pnas.152439599.
105. Kambampati R, Pellegrino C, Paiva A, Huang L, Mende-Mueller L, Chakraburtty K. Limited proteolysis of yeast elongation factor 3. Sequence and location of the subdomains. J Biol Chem. 2000;275:16963–8. doi:10.1074/jbc.M001157200.
106. Kovalchuke O, Kambampati R, Pladies E, Chakraburtty K. Competition and cooperation amongst yeast elongation factors. European J Biochem/FEBS. 1998;258:986–93.
107. Gontarek RR, Li H, Nurse K, Prescott CD. The N terminus of eukaryotic translation elongation factor 3 interacts with 18 S rRNA and 80 S ribosomes. J Biol Chem. 1998;273:10249–52.
108. Kamath A, Chakraburtty K. Role of yeast elongation factor 3 in the elongation cycle. J Biol Chem. 1989;264:15423–8.
109. Harms J, Schluenzen F, Zarivach R, Bashan A, Gat S, Agmon I, Bartels H, Franceschi F, Yonath A. High resolution structure of the large ribosomal subunit from a mesophilic eubacterium. Cell. 2001;107:679–88.
110. Andersen CB, Becker T, Blau M, Anand M, Halic M, Balar B, Mielke T, Boesen T, Pedersen JS, Spahn CM, Kinzy TG, Andersen GR, Beckmann R. Structure of eEF3 and the mechanism of transfer RNA release from the E-site. Nature. 2006;443:663–8. doi:10.1038/nature05126.
111. Sarthy AV, McGonigal T, Capobianco JO, Schmidt M, Green SR, Moehle CM, Goldman RC. Identification and kinetic analysis of a functional homolog of elongation factor 3, YEF3 in Saccharomyces cerevisiae. Yeast. 1998;14:239–53. doi:10.1002/(SICI)1097-0061(199802)14:3<239:AID-YEA219>3.0.CO;2-B.

112. Maurice TC, Mazzucco CE, Ramanathan CS, Ryan BM, Warr GA, Puziss JW. A highly conserved intraspecies homolog of the Saccharomyces cerevisiae elongation factor-3 encoded by the HEF3 gene. Yeast. 1998;14:1105–13. doi:10.1002/(SICI)1097-0061(19980915)14:12<1105:AID-YEA313>3.0.CO;2-Y.
113. Herrera F, Martinez JA, Moreno N, Sadnik I, McLaughlin CS, Feinberg B, Moldave K. Identification of an altered elongation factor in temperature-sensitive mutant ts 7′-14 of Saccharomyces cerevisiae. J Biol Chem. 1984;259:14347–9.
114. Samra N, Atir-Lande A, Pnueli L, Arava Y. The elongation factor eEF3 (Yef3) interacts with mRNA in a translation independent manner. BMC Mol Biol. 2015;16:17. doi:10.1186/s12867-015-0045-5.
115. March PE, Inouye M. Characterization of the lep operon of Escherichia coli. Identification of the promoter and the gene upstream of the signal peptidase I gene. J Biol Chem. 1985;260:7206–13.
116. Zwizinski C, Wickner W. Purification and characterization of leader (signal) peptidase from Escherichia coli. J Biol Chem. 1980;255:7973–7.
117. Qin Y, Polacek N, Vesper O, Staub E, Einfeldt E, Wilson DN, Nierhaus KH. The highly conserved LepA is a ribosomal elongation factor that back-translocates the ribosome. Cell. 2006;127:721–33. doi:10.1016/j.cell.2006.09.037.
118. Bauerschmitt H, Funes S, Herrmann JM. The membrane-bound GTPase Guf1 promotes mitochondrial protein synthesis under suboptimal conditions. J Biol Chem. 2008;283:17139–46. doi:10.1074/jbc.M710037200.
119. Ji DL, Lin H, Chi W, Zhang LX. CpLEPA is critical for chloroplast protein synthesis under suboptimal conditions in Arabidopsis thaliana. PLoS ONE. 2012;7:e49746. doi:10.1371/journal.pone.0049746.
120. Margus T, Remm M, Tenson T. Phylogenetic distribution of translational GTPases in bacteria. BMC Genom. 2007;8:15. doi:10.1186/1471-2164-8-15.
121. Yamamoto H, Qin Y, Achenbach J, Li C, Kijek J, Spahn CM, Nierhaus KH. EF-G and EF4: translocation and back-translocation on the bacterial ribosome. Nat Rev Microbiol. 2014;12:89–100. doi:10.1038/nrmicro3176.
122. Evans RN, Blaha G, Bailey S, Steitz TA. The structure of LepA, the ribosomal back translocase. Proc Natl Acad Sci USA. 2008;105:4673–8. doi:10.1073/pnas.0801308105.
123. Gagnon MG, Lin J, Bulkley D, Steitz TA. Crystal structure of elongation factor 4 bound to a clockwise ratcheted ribosome. Science. 2014;345:684–7. doi:10.1126/science.1253525.
124. Zhang D, Qin Y. The paradox of elongation factor 4: highly conserved, yet of no physiological significance? Biochem J. 2013;452:173–81. doi:10.1042/BJ20121792.
125. Kiser GL, Weinert TA. GUF1, a gene encoding a novel evolutionarily conserved GTPase in budding yeast. Yeast. 1995;11:1311–6. doi:10.1002/yea.320111312.
126. Thomas A, Goumans H, Amesz H, Benne R, Voorma HO. A comparison of the initiation factors of eukaryotic protein synthesis from ribosomes and from the postribosomal supernatant. European J Biochem/FEBS. 1979;98:329–37.
127. Gordon ED, Mora R, Meredith SC, Lee C, Lindquist SL. Eukaryotic initiation factor 4D, the hypusine-containing protein, is conserved among eukaryotes. J Biol Chem. 1987;262:16585–9.
128. Gordon ED, Mora R, Meredith SC, Lindquist SL. Hypusine formation in eukaryotic initiation factor 4D is not reversed when rates or specificity of protein synthesis is altered. J Biol Chem. 1987;262:16590–5.
129. Kemper WM, Berry KW, Merrick WC. Purification and properties of rabbit reticulocyte protein synthesis initiation factors M2Balpha and M2Bbeta. J Biol Chem. 1976;251:5551–7.
130. Schreier MH, Erni B, Staehelin T. Initiation of mammalian protein synthesis. I. Purification and characterization of seven initiation factors. J Mol Biol. 1977;116:727–53.
131. Jao DL, Chen KY. Tandem affinity purification revealed the hypusine-dependent binding of eukaryotic initiation factor 5A to the translating 80S ribosomal complex. J Cell Biochem. 2006;97:583–98. doi:10.1002/jcb.20658.

132. Benne R, Brown-Luedi ML, Hershey JW. Purification and characterization of protein synthesis initiation factors eIF-1, eIF-4C, eIF-4D, and eIF-5 from rabbit reticulocytes. J Biol Chem. 1978;253:3070–7.
133. Smit-McBride Z, Schnier J, Kaufman RJ, Hershey JW. Protein synthesis initiation factor eIF-4D. Functional comparison of native and unhypusinated forms of the protein. J Biol Chem. 1989;264:18527–30.
134. Park MH. The essential role of hypusine in eukaryotic translation initiation factor 4D (eIF-4D). Purification of eIF-4D and its precursors and comparison of their activities. J Biol Chem. 1989;264:18531–5.
135. Gregio AP, Cano VP, Avaca JS, Valentini SR, Zanelli CF. eIF5A has a function in the elongation step of translation in yeast. Biochem Biophys Res Commun. 2009;380:785–90. doi:10.1016/j.bbrc.2009.01.148.
136. Saini P, Eyler DE, Green R, Dever TE. Hypusine-containing protein eIF5A promotes translation elongation. Nature. 2009;459:118–21. doi:10.1038/nature08034.
137. Zanelli CF, Maragno AL, Gregio AP, Komili S, Pandolfi JR, Mestriner CA, Lustri WR, Valentini SR. eIF5A binds to translational machinery components and affects translation in yeast. Biochem Biophys Res Commun. 2006;348:1358–66. doi:10.1016/j.bbrc.2006.07.195.
138. Kang HA, Hershey JW. Effect of initiation factor eIF-5A depletion on protein synthesis and proliferation of Saccharomyces cerevisiae. J Biol Chem. 1994;269:3934–40.
139. Doerfel LK, Rodnina MV. Elongation factor P: Function and effects on bacterial fitness. Biopolymers. 2013;99:837–45. doi:10.1002/bip.22341.
140. Doerfel LK, Wohlgemuth I, Kothe C, Peske F, Urlaub H, Rodnina MV. EF-P is essential for rapid synthesis of proteins containing consecutive proline residues. Science. 2013;339:85–8. doi:10.1126/science.1229017.
141. Gutierrez E, Shin BS, Woolstenhulme CJ, Kim JR, Saini P, Buskirk AR, Dever TE. eIF5A promotes translation of polyproline motifs. Mol Cell. 2013;51:35–45. doi:10.1016/j.molcel.2013.04.021.
142. Ude S, Lassak J, Starosta AL, Kraxenberger T, Wilson DN, Jung K. Translation elongation factor EF-P alleviates ribosome stalling at polyproline stretches. Science. 2013;339:82–5. doi:10.1126/science.1228985.
143. Blaha G, Stanley RE, Steitz TA. Formation of the first peptide bond: the structure of EF-P bound to the 70S ribosome. Science. 2009;325:966–70. doi:10.1126/science.1175800.
144. Hanawa-Suetsugu K, Sekine S, Sakai H, Hori-Takemoto C, Terada T, Unzai S, Tame JR, Kuramitsu S, Shirouzu M, Yokoyama S. Crystal structure of elongation factor P from Thermus thermophilus HB8. Proc Natl Acad Sci USA. 2004;101:9595–600. doi:10.1073/pnas.0308667101.
145. Kim KK, Hung LW, Yokota H, Kim R, Kim SH. Crystal structures of eukaryotic translation initiation factor 5A from Methanococcus jannaschii at 1.8 A resolution. Proc Natl Acad Sci USA. 1998;95:10419–24.
146. Park MH, Nishimura K, Zanelli CF, Valentini SR. Functional significance of eIF5A and its hypusine modification in eukaryotes. Amino Acids. 2010;38:491–500. doi:10.1007/s00726-009-0408-7.
147. Park MH, Wolff EC, Smit-McBride Z, Hershey JW, Folk JE. Comparison of the activities of variant forms of eIF-4D. The requirement for hypusine or deoxyhypusine. J Biol Chem. 1991;266:7988–94.
148. Lee SB, Park JH, Kaevel J, Sramkova M, Weigert R, Park MH. The effect of hypusine modification on the intracellular localization of eIF5A. Biochem Biophys Res Commun. 2009;383:497–502. doi:10.1016/j.bbrc.2009.04.049.
149. Bailly M, de Crecy-Lagard V. Predicting the pathway involved in post-translational modification of elongation factor P in a subset of bacterial species. Biol Direct. 2010;5:3. doi:10.1186/1745-6150-5-3.

150. Navarre WW, Zou SB, Roy H, Xie JL, Savchenko A, Singer A, Edvokimova E, Prost LR, Kumar R, Ibba M, Fang FC. PoxA, yjeK, and elongation factor P coordinately modulate virulence and drug resistance in Salmonella enterica. Mol Cell. 2010;39:209–21. doi:10.1016/j.molcel.2010.06.021.
151. Yanagisawa T, Sumida T, Ishii R, Takemoto C, Yokoyama S. A paralog of lysyl-tRNA synthetase aminoacylates a conserved lysine residue in translation elongation factor P. Nat Struct Mol Biol. 2010;17:1136–43. doi:10.1038/nsmb.1889.
152. Park JH, Johansson HE, Aoki H, Huang BX, Kim HY, Ganoza MC, Park MH. Post-translational modification by beta-lysylation is required for activity of Escherichia coli elongation factor P (EF-P). J Biol Chem. 2012;287:2579–90. doi:10.1074/jbc.M111.309633.
153. Lassak J, Keilhauer EC, Furst M, Wuichet K, Godeke J, Starosta AL, Chen JM, Sogaard-Andersen L, Rohr J, Wilson DN, Haussler S, Mann M, Jung K. Corrigendum: Arginine-rhamnosylation as new strategy to activate translation elongation factor P. Nat Chem Biol. 2015;11:299. doi:10.1038/nchembio0415-299d.
154. Lassak J, Keilhauer EC, Furst M, Wuichet K, Godeke J, Starosta AL, Chen JM, Sogaard-Andersen L, Rohr J, Wilson DN, Haussler S, Mann M, Jung K. Arginine-rhamnosylation as new strategy to activate translation elongation factor P. Nat Chem Biol. 2015;11:266–70. doi:10.1038/nchembio.1751.
155. Choi S, Choe J. Crystal structure of elongation factor P from Pseudomonas aeruginosa at 1.75 A resolution. Proteins. 2011;79:1688–93. doi:10.1002/prot.22992.
156. Lassak J, Wilson DN, Jung K. Stall no more at polyproline stretches with the translation elongation factors EF-P and IF-5A. Mol Microbiol. 2015;. doi:10.1111/mmi.13233.
157. Jenkins ZA, Haag PG, Johansson HE. Human eIF5A2 on chromosome 3q25-q27 is a phylogenetically conserved vertebrate variant of eukaryotic translation initiation factor 5A with tissue-specific expression. Genomics. 2001;71:101–9. doi:10.1006/geno.2000.6418.
158. Ohashi Y, Inaoka T, Kasai K, Ito Y, Okamoto S, Satsu H, Tozawa Y, Kawamura F, Ochi K. Expression profiling of translation-associated genes in sporulating Bacillus subtilis and consequence of sporulation by gene inactivation. Biosci Biotechnol Biochem. 2003;67:2245–53. doi:10.1271/bbb.67.2245.
159. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol. 2006;2006(2):0008. doi:10.1038/msb4100050.
160. Sassetti CM, Boyd DH, Rubin EJ. Genes required for mycobacterial growth defined by high density mutagenesis. Mol Microbiol. 2003;48:77–84.
161. Ejiri S. Moonlighting functions of polypeptide elongation factor 1: from actin bundling to zinc finger protein R1-associated nuclear localization. Biosci Biotechnol Biochem. 2002;66:1–21. doi:10.1271/bbb.66.1.
162. Altmann M, Muller PP, Pelletier J, Sonenberg N, Trachsel H. A mammalian translation initiation factor can substitute for its yeast homologue in vivo. J Biol Chem. 1989;264:12145–7.
163. Arcari P, Masullo M, Arcucci A, Ianniciello G, de Paola B, Bocchini V. A chimeric elongation factor containing the putative guanine nucleotide binding domain of archaeal EF-1 alpha and the M and C domains of eubacterial EF-Tu. Biochemistry. 1999;38:12288–95.
164. Anand M, Chakraburtty K, Marton MJ, Hinnebusch AG, Kinzy TG. Functional interactions between yeast translation eukaryotic elongation factor (eEF) 1A and eEF3. J Biol Chem. 2003;278:6985–91. doi:10.1074/jbc.M209224200.
165. Mateyak MK, Kinzy TG. eEF1A: thinking outside the ribosome. J Biol Chem. 2010;285:21209–13. doi:10.1074/jbc.R110.113795.
166. Liu G, Grant WM, Persky D, Latham VM Jr, Singer RH, Condeelis J. Interactions of elongation factor 1alpha with F-actin and beta-actin mRNA: implications for anchoring mRNA in cell protrusions. Mol Biol Cell. 2002;13:579–92. doi:10.1091/mbc.01-03-0140.
167. van den Ent F, Amos LA, Lowe J. Prokaryotic origin of the actin cytoskeleton. Nature. 2001;413:39–44. doi:10.1038/35092500.

168. Defeu Soufo HJ. Reimold C, Linne U, Knust T, Gescher J, Graumann PL. Bacterial translation elongation factor EF-Tu interacts and colocalizes with actin-like MreB protein. Proc Natl Acad Sci USA. 2010;107:3163–8. doi:10.1073/pnas.0911979107.
169. Liu Z, Xing D, Su QP, Zhu Y, Zhang J, Kong X, Xue B, Wang S, Sun H, Tao Y, Sun Y. Super-resolution imaging and tracking of protein-protein interactions in sub-diffraction cellular space. Nat Commun. 2014;5:4443. doi:10.1038/ncomms5443.
170. Defeu Soufo HJ. Reimold C, Breddermann H, Mannherz HG, Graumann PL. Translation elongation factor EF-Tu modulates filament formation of actin-like MreB protein in vitro. J Mol Biol. 2015;427:1715–27. doi:10.1016/j.jmb.2015.01.025.
171. Anand N, Murthy S, Amann G, Wernick M, Porter LA, Cukier IH, Collins C, Gray JW, Diebold J, Demetrick DJ, Lee JM. Protein elongation factor EEF1A2 is a putative oncogene in ovarian cancer. Nat Genet. 2002;31:301–5. doi:10.1038/ng904.
172. Tomlinson VA, Newbery HJ, Wray NR, Jackson J, Larionov A, Miller WR, Dixon JM, Abbott CM. Translation elongation factor eEF1A2 is a potential oncoprotein that is overexpressed in two-thirds of breast tumours. BMC Cancer. 2005;5:113. doi:10.1186/1471-2407-5-113.
173. Pinke DE, Kalloger SE, Francetic T, Huntsman DG, Lee JM. The prognostic significance of elongation factor eEF1A2 in ovarian cancer. Gynecol Oncol. 2008;108:561–8. doi:10.1016/j.ygyno.2007.11.019.
174. Scaggiante B, Dapas B, Bonin S, Grassi M, Zennaro C, Farra R, Cristiano L, Siracusano S, Zanconati F, Giansante C, Grassi G. Dissecting the expression of EEF1A1/2 genes in human prostate cancer cells: the potential of EEF1A2 as a hallmark for prostate transformation and progression. Br J Cancer. 2012;106:166–73. doi:10.1038/bjc.2011.500.
175. Keeling PJ, Inagaki Y. A class of eukaryotic GTPase with a punctate distribution suggesting multiple functional replacements of translation elongation factor 1alpha. Proc Natl Acad Sci USA. 2004;101:15380–5. doi:10.1073/pnas.0404505101.
176. Sakaguchi M, Takishita K, Matsumoto T, Hashimoto T, Inagaki Y. Tracing back EFL gene evolution in the cryptomonads-haptophytes assemblage: separate origins of EFL genes in haptophytes, photosynthetic cryptomonads, and goniomonads. Gene. 2009;441:126–31. doi:10.1016/j.gene.2008.05.010.
177. Noble GP, Rogers MB, Keeling PJ. Complex distribution of EFL and EF-1alpha proteins in the green algal lineage. BMC Evol Biol. 2007;7:82. doi:10.1186/1471-2148-7-82.
178. Kamikawa R, Inagaki Y, Sako Y. Direct phylogenetic evidence for lateral transfer of elongation factor-like gene. Proc Natl Acad Sci USA. 2008;105:6965–9. doi:10.1073/pnas.0711084105.
179. Gile GH, Faktorova D, Castlejohn CA, Burger G, Lang BF, Farmer MA, Lukes J, Keeling PJ. Distribution and phylogeny of EFL and EF-1alpha in Euglenozoa suggest ancestral co-occurrence followed by differential loss. PLoS ONE. 2009;4:e5162. doi:10.1371/journal.pone.0005162.
180. Cocquyt E, Verbruggen H, Leliaert F, Zechman FW, Sabbe K, De Clerck O. Gain and loss of elongation factor genes in green algae. BMC Evol Biol. 2009;9:39. doi:10.1186/1471-2148-9-39.
181. Kamikawa R, Sakaguchi M, Matsumoto T, Hashimoto T, Inagaki Y. Rooting for the root of elongation factor-like protein phylogeny. Mol Phylogenet Evol. 2010;56:1082–8. doi:10.1016/j.ympev.2010.04.040.
182. Skogerson L, Engelhardt D. Dissimilarity in protein chain elongation factor requirements between yeast and rat liver ribosomes. J Biol Chem. 1977;252:1471–5.
183. Ben-Shem A, Garreau de Loubresse N, Melnikov S, Jenner L, Yusupova G, Yusupov M. The structure of the eukaryotic ribosome at 3.0 A resolution. Science. 2011; 334:1524–1529. doi:10.1126/science.1212642.
184. Van Dyke N, Pickering BF, Van Dyke MW. Stm1p alters the ribosome association of eukaryotic elongation factor 3 and affects translation elongation. Nucleic Acids Res. 2009;37:6116–25. doi:10.1093/nar/gkp645.

185. Ben-Shem A, Jenner L, Yusupova G, Yusupov M. Crystal structure of the eukaryotic ribosome. Science. 2010;330:1203–9. doi:10.1126/science.1194294.
186. Van Dyke N, Chanchorn E, Van Dyke MW. The Saccharomyces cerevisiae protein Stm1p facilitates ribosome preservation during quiescence. Biochem Biophys Res Commun. 2013;430:745–50. doi:10.1016/j.bbrc.2012.11.078.
187. Bijlsma JJ, Lie ALM, Nootenboom IC, Vandenbroucke-Grauls CM, Kusters JG. Identification of loci essential for the growth of Helicobacter pylori under acidic conditions. J Infect Dis. 2000;182:1566–9. doi:10.1086/315855.
188. Song H, Mugnier P, Das AK, Webb HM, Evans DR, Tuite MF, Hemmings BA, Barford D. The crystal structure of human eukaryotic release factor eRF1–mechanism of stop codon recognition and peptidyl-tRNA hydrolysis. Cell. 2000;100:311–21.
189. Vestergaard B, Van LB, Andersen GR, Nyborg J, Buckingham RH, Kjeldgaard M. Bacterial polypeptide release factor RF2 is structurally distinct from eukaryotic eRF1. Mol Cell. 2001;8:1375–82.
190. Shin DH, Brandsen J, Jancarik J, Yokota H, Kim R, Kim SH. Structural analyses of peptide release factor 1 from Thermotoga maritima reveal domain flexibility required for its interaction with the ribosome. J Mol Biol. 2004;341:227–39. doi:10.1016/j.jmb.2004.05.055.
191. Brenner S, Stretton AO, Kaplan S. Genetic code: the 'nonsense' triplets for chain termination and their suppression. Nature. 1965;206:994–8.
192. Nakamura Y, Ito K, Isaksson LA. Emerging understanding of translation termination. Cell. 1996;87:147–50.
193. Buckingham RH, Grentzmann G, Kisselev L. Polypeptide chain release factors. Mol Microbiol. 1997;24:449–56.
194. Jorgensen F, Adamski FM, Tate WP, Kurland CG. Release factor-dependent false stops are infrequent in Escherichia coli. J Mol Biol. 1993;230:41–50. doi:10.1006/jmbi.1993.1124.
195. Nakamura Y, Ito K, Ehrenberg M. Mimicry grasps reality in translation termination. Cell. 2000;101:349–52.
196. Kisselev LL, Buckingham RH. Translational termination comes of age. Trends Biochem Sci. 2000;25:561–6.
197. Stansfield I, Jones KM, Kushnirov VV, Dagkesamanskaya AR, Poznyakovski AI, Paushkin SV, Nierras CR, Cox BS, Ter-Avanesyan MD, Tuite MF. The products of the SUP45 (eRF1) and SUP35 genes interact to mediate translation termination in Saccharomyces cerevisiae. The EMBO J. 1995;14:4365–73.
198. Zhouravleva G, Frolova L, Le Goff X, Le Guellec R, Inge-Vechtomov S, Kisselev L, Philippe M. Termination of translation in eukaryotes is governed by two interacting polypeptide chain release factors, eRF1 and eRF3. The EMBO J. 1995;14:4065–72.
199. Alkalaeva EZ, Pisarev AV, Frolova LY, Kisselev LL, Pestova TV. In vitro reconstitution of eukaryotic translation reveals cooperativity between release factors eRF1 and eRF3. Cell. 2006;125:1125–36. doi:10.1016/j.cell.2006.04.035.
200. Pisareva VP, Pisarev AV, Hellen CU, Rodnina MV, Pestova TV. Kinetic analysis of interaction of eukaryotic release factor 3 with guanine nucleotides. J Biol Chem. 2006;281:40224–35. doi:10.1074/jbc.M607461200.
201. Hauryliuk V, Zavialov A, Kisselev L, Ehrenberg M. Class-1 release factor eRF1 promotes GTP binding by class-2 release factor eRF3. Biochimie. 2006;88:747–57. doi:10.1016/j.biochi.2006.06.001.
202. Nakamura Y, Ito K. tRNA mimicry in translation termination and beyond. Wiley Interdiscip Rev RNA. 2011;2:647–68. doi:10.1002/wrna.81.
203. Shoemaker CJ, Green R. Kinetic analysis reveals the ordered coupling of translation termination and ribosome recycling in yeast. Proc Natl Acad Sci USA. 2011;108:E1392–8. doi:10.1073/pnas.1113956108.
204. Brown A, Shao S, Murray J, Hegde RS, Ramakrishnan V. Structural basis for stop codon recognition in eukaryotes. Nature. 2015;524:493–6. doi:10.1038/nature14896.
205. Frolova LY, Tsivkovskii RY, Sivolobova GF, Oparina NY, Serpinsky OI, Blinov VM, Tatkov SI, Kisselev LL. Mutations in the highly conserved GGQ motif of class 1 polypeptide

release factors abolish ability of human eRF1 to trigger peptidyl-tRNA hydrolysis. RNA. 1999;5:1014–20.
206. Chavatte L, Seit-Nebi A, Dubovaya V, Favre A. The invariant uridine of stop codons contacts the conserved NIKSR loop of human eRF1 in the ribosome. The EMBO J. 2002;21:5302–11.
207. Bulygin KN, Khairulina YS, Kolosov PM, Ven'yaminova AG, Graifer DM, Vorobjev YN, Frolova LY, Kisselev LL, Karpova GG. Three distinct peptides from the N domain of translation termination factor eRF1 surround stop codon in the ribosome. RNA. 2010;16:1902–14. doi:10.1261/rna.2066910.
208. Shaw JJ, Green R. Two distinct components of release factor function uncovered by nucleophile partitioning analysis. Mol Cell. 2007;28:458–67. doi:10.1016/j.molcel.2007.09.007.
209. Trobro S, Aqvist J. A model for how ribosomal release factors induce peptidyl-tRNA cleavage in termination of protein synthesis. Mol Cell. 2007;27:758–66. doi:10.1016/j.molcel.2007.06.032.
210. Merkulova TI, Frolova LY, Lazar M, Camonis J, Kisselev LL. C-terminal domains of human translation termination factors eRF1 and eRF3 mediate their in vivo interaction. FEBS Lett. 1999;443:41–7.
211. Cheng Z, Saito K, Pisarev AV, Wada M, Pisareva VP, Pestova TV, Gajda M, Round A, Kong C, Lim M, Nakamura Y, Svergun DI, Ito K, Song H. Structural insights into eRF3 and stop codon recognition by eRF1. Genes Dev. 2009;23:1106–18. doi:10.1101/gad.1770109.
212. Kononenko AV, Mitkevich VA, Dubovaya VI, Kolosov PM, Makarov AA, Kisselev LL. Role of the individual domains of translation termination factor eRF1 in GTP binding to eRF3. Proteins. 2008;70:388–93. doi:10.1002/prot.21544.
213. Saito K, Kobayashi K, Wada M, Kikuno I, Takusagawa A, Mochizuki M, Uchiumi T, Ishitani R, Nureki O, Ito K. Omnipotent role of archaeal elongation factor 1 alpha (EF1alpha in translational elongation and termination, and quality control of protein synthesis. Proc Natl Acad Sci USA. 2010;107:19242–7. doi:10.1073/pnas.1009599107.
214. Ito K, Uno M, Nakamura Y. A tripeptide 'anticodon' deciphers stop codons in messenger RNA. Nature. 2000;403:680–4. doi:10.1038/35001115.
215. Klaholz BP, Pape T, Zavialov AV, Myasnikov AG, Orlova EV, Vestergaard B, Ehrenberg M, van Heel M. Structure of the Escherichia coli ribosomal termination complex with release factor 2. Nature. 2003;421:90–4. doi:10.1038/nature01225.
216. Petry S, Brodersen DE, Murphy FVt, Dunham CM, Selmer M, Tarry MJ, Kelley AC, Ramakrishnan V. Crystal structures of the ribosome in complex with release factors RF1 and RF2 bound to a cognate stop codon. Cell. 2005; 123:1255–1266. doi:10.1016/j.cell.2005.09.039.
217. Rawat UB, Zavialov AV, Sengupta J, Valle M, Grassucci RA, Linde J, Vestergaard B, Ehrenberg M, Frank J. A cryo-electron microscopic study of ribosome-bound termination factor RF2. Nature. 2003;421:87–90. doi:10.1038/nature01224.
218. Rawat U, Gao H, Zavialov A, Gursky R, Ehrenberg M, Frank J. Interactions of the release factor RF1 with the ribosome as revealed by cryo-EM. J Mol Biol. 2006;357:1144–53. doi:10.1016/j.jmb.2006.01.038.
219. Bertram G, Bell HA, Ritchie DW, Fullerton G, Stansfield I. Terminating eukaryote translation: domain 1 of release factor eRF1 functions in stop codon recognition. RNA. 2000;6:1236–47.
220. Frolova L, Seit-Nebi A, Kisselev L. Highly conserved NIKS tetrapeptide is functionally essential in eukaryotic translation termination factor eRF1. RNA. 2002;8:129–36.
221. Inagaki Y, Blouin C, Doolittle WF, Roger AJ. Convergence and constraint in eukaryotic release factor 1 (eRF1) domain 1: the evolution of stop codon specificity. Nucleic Acids Res. 2002;30:532–44.
222. Weixlbaumer A, Jin H, Neubauer C, Voorhees RM, Petry S, Kelley AC, Ramakrishnan V. Insights into translational termination from the structure of RF2 bound to the ribosome. Science. 2008;322:953–6. doi:10.1126/science.1164840.

223. Korostelev AA. Structural aspects of translation termination on the ribosome. RNA. 2011;17:1409–21. doi:10.1261/rna.2733411.
224. Jin H, Kelley AC, Ramakrishnan V. Crystal structure of the hybrid state of ribosome in complex with the guanosine triphosphatase release factor 3. Proc Natl Acad Sci USA. 2011;108:15798–803. doi:10.1073/pnas.1112185108.
225. Zhou J, Lancaster L, Trakhanov S, Noller HF. Crystal structure of release factor RF3 trapped in the GTP state on a rotated conformation of the ribosome. RNA. 2012;18:230–40. doi:10.1261/rna.031187.111.
226. Freistroffer DV, Pavlov MY, MacDougall J, Buckingham RH, Ehrenberg M. Release factor RF3 in E.coli accelerates the dissociation of release factors RF1 and RF2 from the ribosome in a GTP-dependent manner. The EMBO J. 1997;16:4126–33. doi:10.1093/emboj/16.13.4126.
227. Gao H, Zhou Z, Rawat U, Huang C, Bouakaz L, Wang C, Cheng Z, Liu Y, Zavialov A, Gursky R, Sanyal S, Ehrenberg M, Frank J, Song H. RF3 induces ribosomal conformational changes responsible for dissociation of class I release factors. Cell. 2007;129:929–41. doi:10.1016/j.cell.2007.03.050.
228. Zavialov AV, Buckingham RH, Ehrenberg M. A posttermination ribosomal complex is the guanine nucleotide exchange factor for peptide release factor RF3. Cell. 2001;107:115–24.
229. Salas-Marco J, Bedwell DM. GTP hydrolysis by eRF3 facilitates stop codon decoding during eukaryotic translation termination. Mol Cell Biol. 2004;24:7769–78. doi:10.1128/MCB.24.17.7769-7778.2004.
230. Zaher HS, Green R. Quality control by the ribosome following peptide bond formation. Nature. 2009;457:161–6. doi:10.1038/nature07582.
231. Zaher HS, Green R. A primary role for release factor 3 in quality control during translation elongation in Escherichia coli. Cell. 2011;147:396–408. doi:10.1016/j.cell.2011.08.045.
232. Kushnirov VV, Ter-Avanesyan MD, Telckov MV, Surguchov AP, Smirnov VN, Inge-Vechtomov SG. Nucleotide sequence of the SUP2 (SUP35) gene of Saccharomyces cerevisiae. Gene. 1988;66:45–54.
233. Wilson PG, Culbertson MR. SUF12 suppressor protein of yeast. A fusion protein related to the EF-1 family of elongation factors. J Mol Biol. 1988;199:559–73.
234. Grentzmann G, Brechemier-Baey D, Heurgue V, Mora L, Buckingham RH. Localization and characterization of the gene encoding release factor RF3 in Escherichia coli. Proc Natl Acad Sci USA. 1994;91:5848–52.
235. Mikuni O, Ito K, Moffat J, Matsumura K, McCaughan K, Nobukuni T, Tate W, Nakamura Y. Identification of the prfC gene, which encodes peptide-chain-release factor 3 of Escherichia coli. Proc Natl Acad Sci USA. 1994;91:5798–802.
236. Caskey CT, Beaudet AL, Scolnick EM, Rosman M. Hydrolysis of fMet-tRNA by peptidyl transferase. Proc Natl Acad Sci USA. 1971;68:3163–7.
237. Zavialov AV, Mora L, Buckingham RH, Ehrenberg M. Release of peptide promoted by the GGQ motif of class 1 release factors regulates the GTPase activity of RF3. Mol Cell. 2002;10:789–98.
238. Frolova L, Le Goff X, Rasmussen HH, Cheperegin S, Drugeon G, Kress M, Arman I, Haenni AL, Celis JE, Philippe M, et al. A highly conserved eukaryotic protein family possessing properties of polypeptide chain release factor. Nature. 1994;372:701–3. doi:10.1038/372701a0.
239. Stansfield I, Tuite MF. Polypeptide chain termination in Saccharomyces cerevisiae. Curr Genet. 1994;25:385–95.
240. Kobayashi T, Funakoshi Y, Hoshino S, Katada T. The GTP-binding release factor eRF3 as a key mediator coupling translation termination to mRNA decay. J Biol Chem. 2004;279:45693–700. doi:10.1074/jbc.M405163200.
241. Leipe DD, Wolf YI, Koonin EV, Aravind L. Classification and evolution of P-loop GTPases and related ATPases. J Mol Biol. 2002;317:41–72. doi:10.1006/jmbi.2001.5378.
242. Inagaki Y, Dacks JB, Doolittle WF, Watanabe KI, Ohama T. Evolutionary relationship between dinoflagellates bearing obligate diatom endosymbionts: insight into tertiary

endosymbiosis. Int J Syst Evol Microbiol. 2000;50(Pt 6):2075–81. doi:10.1099/00207713-50-6-2075.
243. Atkinson GC, Baldauf SL, Hauryliuk V. Evolution of nonstop, no-go and nonsense-mediated mRNA decay and their termination factor-derived components. BMC Evol Biol. 2008;8:290. doi:10.1186/1471-2148-8-290.
244. Inagaki Y, Blouin C, Susko E, Roger AJ. Assessing functional divergence in EF-1alpha and its paralogs in eukaryotes and archaebacteria. Nucleic Acids Res. 2003;31:4227–37.
245. Inagaki Y. Ford Doolittle W. Evolution of the eukaryotic translation termination system: origins of release factors. Mol Biol Evol. 2000;17:882–9.
246. Dontsova M, Frolova L, Vassilieva J, Piendl W, Kisselev L, Garber M. Translation termination factor aRF1 from the archaeon Methanococcus jannaschii is active with eukaryotic ribosomes. FEBS Lett. 2000;472:213–6.
247. Kisselev L. Polypeptide release factors in prokaryotes and eukaryotes: same function, different structure. Structure. 2002;10:8–9.
248. Ito K, Ebihara K, Nakamura Y. The stretch of C-terminal acidic amino acids of translational release factor eRF1 is a primary binding site for eRF3 of fission yeast. RNA. 1998;4:958–72.
249. Mora L, Zavialov A, Ehrenberg M, Buckingham RH. Stop codon recognition and interactions with peptide release factor RF3 of truncated and chimeric RF1 and RF2 from Escherichia coli. Mol Microbiol. 2003;50:1467–76.
250. Scolnick E, Tompkins R, Caskey T, Nirenberg M. Release factors differing in specificity for terminator codons. Proc Natl Acad Sci USA. 1968;61:768–74.
251. Mora L, Heurgue-Hamard V, Champ S, Ehrenberg M, Kisselev LL, Buckingham RH. The essential role of the invariant GGQ motif in the function and stability in vivo of bacterial release factors RF1 and RF2. Mol Microbiol. 2003;47:267–75.
252. Heurgue-Hamard V, Champ S, Engstrom A, Ehrenberg M, Buckingham RH. The hemK gene in Escherichia coli encodes the N(5)-glutamine methyltransferase that modifies peptide release factors. The EMBO J. 2002;21:769–78. doi:10.1093/emboj/21.4.769.
253. Heurgue-Hamard V, Graille M, Scrima N, Ulryck N, Champ S, van Tilbeurgh H, Buckingham RH. The zinc finger protein Ynr046w is plurifunctional and a component of the eRF1 methyltransferase in yeast. J Biol Chem. 2006;281:36140–8. doi:10.1074/jbc.M608571200.
254. Graille M, Figaro S, Kervestin S, Buckingham RH, Liger D, Heurgue-Hamard V. Methylation of class I translation termination factors: structural and functional aspects. Biochimie. 2012;94:1533–43. doi:10.1016/j.biochi.2012.01.005.
255. Hoshino S, Hosoda N, Araki Y, Kobayashi T, Uchida N, Funakoshi Y, Katada T. Novel function of the eukaryotic polypeptide-chain releasing factor 3 (eRF3/GSPT) in the mRNA degradation pathway. Biochemistry Biokhimiia. 1999;64:1367–72.
256. Hosoda N, Kobayashi T, Uchida N, Funakoshi Y, Kikuchi Y, Hoshino S, Katada T. Translation termination factor eRF3 mediates mRNA decay through the regulation of deadenylation. J Biol Chem. 2003;278:38287–91. doi:10.1074/jbc.C300300200.
257. Graille M, Chaillet M, van Tilbeurgh H. Structure of yeast Dom34: a protein related to translation termination factor Erf1 and involved in No-Go decay. J Biol Chem. 2008;283:7145–54. doi:10.1074/jbc.M708224200.
258. Moreira D, Kervestin S, Jean-Jean O, Philippe H. Evolution of eukaryotic translation elongation and termination factors: variations of evolutionary rate and genetic code deviations. Mol Biol Evol. 2002;19:189–200.
259. Drugeon G, Jean-Jean O, Frolova L, Le Goff X, Philippe M, Kisselev L, Haenni AL. Eukaryotic release factor 1 (eRF1) abolishes readthrough and competes with suppressor tRNAs at all three termination codons in messenger RNA. Nucleic Acids Res. 1997;25:2254–8.
260. Le Goff X, Philippe M, Jean-Jean O. Overexpression of human release factor 1 alone has an antisuppressor effect in human cells. Mol Cell Biol. 1997;17:3164–72.

The Unique Evolutionary Distribution of Eukaryotic Elongation Factor 3

Maria Mateyak, Arjun N. Sasikumar, Stephen Dunaway and Terri Goss Kinzy

1 Protein Synthesis Is a Multi-step Process

Protein synthesis is the process by which genetic information transcribed in messenger RNA (mRNA) is converted into specific proteins by the ribosome. As a fundamental process for life, the core mechanism of translation is conserved across all living organisms and requires soluble factors called translation factors, which catalyze a sequence of biochemical reactions at the ribosome.

Eukaryotic protein synthesis begins with the action of a series of highly regulated initiation factors (eIFs) that recognize the m^7G cap of an mRNA, bind the 40S ribosomal subunit, scan 5′ to 3′ to locate the initiation codon and position the initiator Met-$tRNA^{Met}$ at the AUG [1]. Recruitment of the 60S ribosomal subunit results in an 80S ribosome positioned at the start codon of the open reading frame (reviewed in [2]). The subsequent elongation phase of protein synthesis is a cycle of aa-tRNA delivery, peptide bond formation and translocation repeated hundreds of times during the synthesis of an average protein and facilitated by soluble elongation factor (eEF) proteins [3]. When the elongating ribosome encounters a termination codon, release factors (eRFs) specifically terminate translation and release the polypeptide from the ribosome, and ribosomal subunits are recycled to participate in a new round of translation [4].

M. Mateyak · A.N. Sasikumar · T.G. Kinzy (✉)
Department of Biochemistry and Molecular Biology, Robert Wood Johnson Medical School, Rutgers, the State University of New Jersey, 675 Hoes Lane Piscataway, NJ 08854, USA
e-mail: kinzytg@rutgers.edu

M. Mateyak
e-mail: mateyamk@rwjms.rutgers.edu

S. Dunaway
Department of Biology, Drew University, 36 Madison Ave., Madison, NJ 07940, USA

G. Hernández and R. Jagus (eds.), *Evolution of the Protein Synthesis Machinery and Its Regulation*, DOI 10.1007/978-3-319-39468-8_14

2 Translation Elongation Is a Conserved Cyclical Process

The eEFs function at the ribosome and play an important role in the accuracy of gene expression by maintaining the correct reading frame of the mRNA and assuring the specific binding of cognate aa-tRNA to the A-site [5]. *Saccharomyces cerevisiae* eEF1 consists of three subunits. eEF1A, the functional homolog of bacterial EF-Tu, binds and recruits aa-tRNAs to the A-site of the ribosome (Fig. 1; Step 1, [6]). Analogous to the bacterial factor, the initial ribosome binding step of eEF1A is thought to be codon and GTPase independent [7]. When a codon-anticodon match occurs, the ribosome stimulates eEF1A-mediated GTP hydrolysis resulting in the release of inactive GDP-bound eEF1A (Fig. 1; Step 2). Spontaneous GDP dissociation from eEF1A is slow (k_D 1.8×10^7 M^{-1} [8, 9]), requiring the eEF1Bαγ complex to stimulate GDP release [10]. The eEF1Bα subunit is essential in yeast and performs the catalytic activity in nucleotide exchange [11]. The eEF1Bγ subunit is not essential in yeast; however, strains lacking the protein are resistant to oxidative stress and have protein-processing defects [12, 13]. Following peptide bond

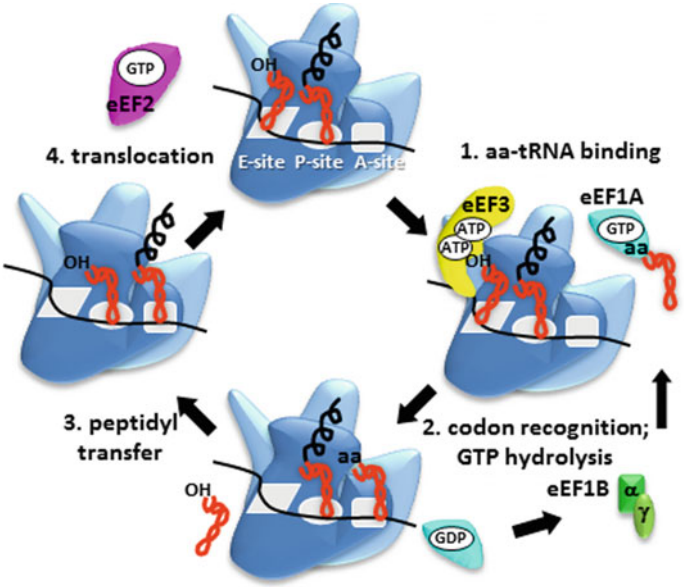


Fig. 1 Translation elongation in *S. cerevisiae*. *Step 1* eEF1A delivers aa-tRNA to the A site of the ribosome while eEF3 binds the ribosome near the E-site and facilitates the release of deacylated tRNA from the E-site. *Step 2* Accommodation of cognate tRNA activates the GTPase activity of eEF1A and releases it from the ribosome. eEF1Bαγ acts as the guanine nucleotide exchange factor for eEF1A. *Step 3* A peptide bond is formed transferring the nascent polypeptide to the A-site tRNA. *Step 4* eEF2 translocates the mRNA and places the deacylated tRNA and peptidyl tRNA at the E- and P-sites, respectively

formation (Fig. 1; Step 3), eEF2, the homolog of bacterial EF-G, catalyzes the translocation of the peptidyl-tRNA from the A-site to the P-site of the ribosome (Fig. 1; Step 4).

3 Fungal Systems Require a Third Elongation Factor

During the establishment of an in vitro fungal translation system, it was determined that a third elongation factor, the ATPase eEF3, is required for protein synthesis in budding yeast [14]. Since then, the gene encoding the protein has been identified and isolated in a range of fungal species including *S. cerevisiae* [15], *Candida albicans* [16], *Schizosaccharomyces pombe* [17] and *Cryptococcus neoformans* [18]. While bacteria contain an ATP-binding protein linked to elongation, the functional similarity to eEF3 is not clear [19]. In addition, there is no evidence for a corresponding ATPase associated with translation elongation in either mammals or plants.

3.1 Requirement for eEF3 Is Determined by the Ribosome

Initial experiments demonstrated that it is the source of the ribosomes that determines the requirement for eEF3 in vitro. *S. cerevisiae* eEF1A and eEF2 alone catalyze protein synthesis with rat liver ribosomes; however, when rat liver eEF1A and eEF2 are used with yeast ribosomes, eEF3 is essential [20]. These experiments suggest that non-fungal ribosomes have evolved to function in the absence of eEF3. Genetic studies in *S. cerevisiae* also support the pivotal role eEF3 plays in fungal translation, as deletion of the *YEF3* gene encoding eEF3 is lethal and strains harboring temperature-sensitive alleles of eEF3 display protein synthesis and translation elongation defects [21–23]. Together, these studies suggest that eEF3 plays a unique and essential role in fungal elongation.

While the exact role of eEF3 in translation elongation is unclear, previous experiments suggest that it stimulates deacylated tRNA release from the ribosome E-site [24]. In addition, eEF3 is linked to the ribosomal A-site, as it stimulates eEF1A-mediated binding of cognate aa-tRNA [25, 26] and there is a direct eEF3-eEF1A interaction [22, 23].

3.2 Unique Domains Direct the eEF3 Ribosome Interaction

Both the X-ray structure of *S. cerevisiae* eEF3 (1-980) and a cryo-EM reconstruction of the eEF3-ATP-post-translocation 80S ribosome complex have been solved [27]. The X-ray structure was solved in the apo-, ADP- and ADPNP-bound

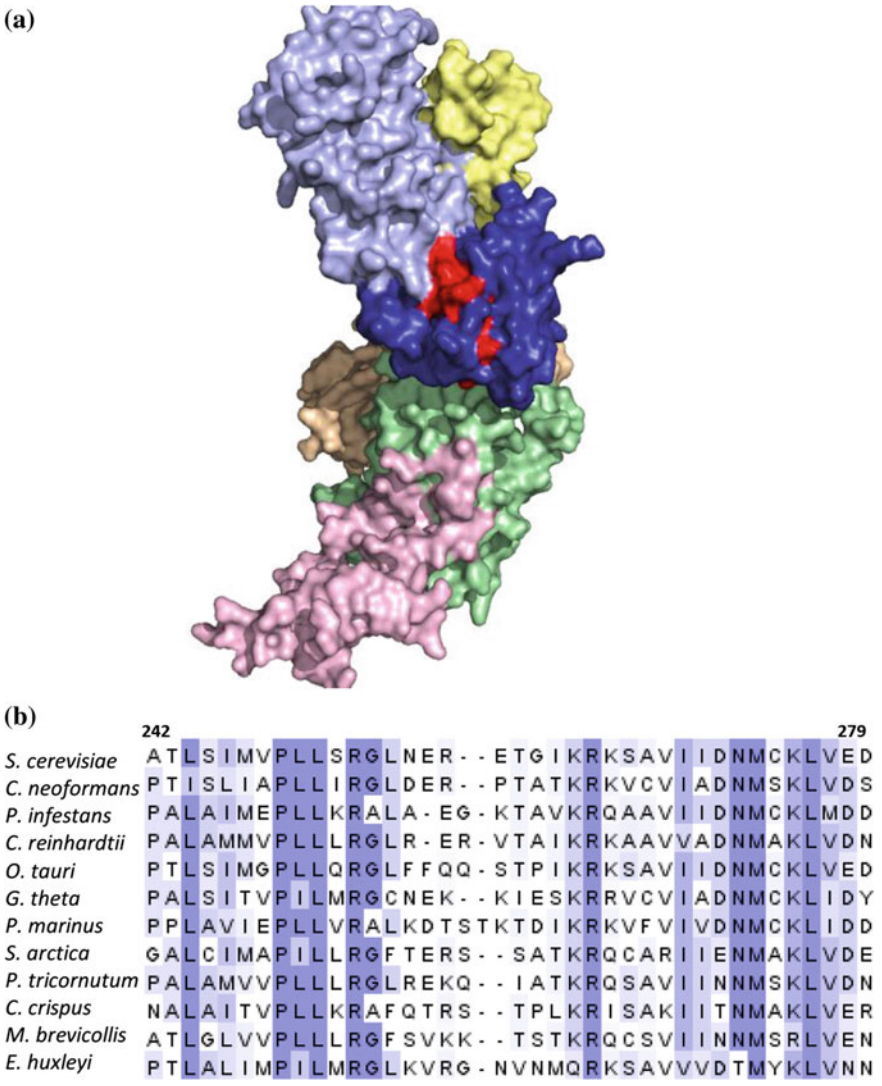


Fig. 2 Multiple sequence alignment of helical repeats shows uniquely conserved HEAT repeats. **a** Surface view of the cryo EM structure of *S. cerevisiae* eEF3 (PDB 2IX8) showing the HEAT domain (*blue* and *red*), four-helix bundle (*yellow*), ABC1 (*wheat*), ABC2 (*green*) and chromodomain (*purple*). The terminal three helical repeats of the HEAT domain are represented in *dark blue* with the conserved repeat 7 highlighted in *red*. The concave side represents the ribosome-binding interface [27]. **b** Heat domain helical repeat 7. Multiple alignment of the selected sequences was carried out using ClustalWS using the default parameters. The alignment was analyzed using Jalview 2.9 [28] to identify conserved regions. The conserved residues were color coded based on the BLOSUM62 score with the conservation color increment set to 30 and *darker blue* representing greater conservation

states [27]. All three structures showed a similar organization of five structural domains: an amino-terminal HEAT repeat domain followed by a four-helix bundle and two ATP-binding cassette (ABC)-type ATPase domains, with a chromodomain-like region inserted in ABC2. In the cryo-EM structure, eEF3 binds the ribosome near the E-site, with the chromodomain-like insertion proposed to stabilize the ribosomal L1 stalk in an open conformation allowing E-site tRNA release [27]. The major contacts with the ribosome occur via the HEAT repeats, the ABC2 domain and the chromodomain-like insertion in ABC2 (Fig. 2a). Ribosomal sites proposed to interact with eEF3 include helix 39 of the 18S rRNA, the 5S rRNA and ribosomal proteins rpS18, rpS19, rpS25, rpL5 and rpL11.

4 eEF3 Is an ABC Family ATPase

eEF3 is a member of the ABC family of proteins, the majority of which are integral membrane transporters involved in the import or export of diverse substrates across lipid bilayers [29]. Typical ABC transporters consist of two nucleotide binding domains (NBD) that bind and hydrolyze ATP. The hydrolysis of ATP results in a conformational change that is coupled to the trans-membrane domain (TMD) and allows the opening and closing of the transport channel permitting the passage of different substrates across the cell membrane [29]. In *S. cerevisiae,* there are 29 ABC proteins, 23 of which contain at least one TMD and are thus predicted to be transporters. The remaining six ABC family members, including eEF3, are soluble ATPases that lack a TMD [30]. Interestingly, the *S. cerevisiae* genome also encodes a *YEF3* paralog, *HEF3,* which is proposed to have arisen from an ancient genome duplication. BLAST analysis indicates that Hef3p is 92 % similar and 84 % identical to eEF3 at the amino acid level. While *HEF3* is not transcriptionally active in the vegetative phase of yeast growth, *HEF3* expressed from the *YEF3* promoter can support the growth of a strain deleted for the *YEF3* gene [31].

5 Identification of eEF3 Homologs in Diverse Eukaryotes

While eEF3 was originally described as a fungal specific factor, the increasing availability of whole genome sequences from a large variety of species prompted a re-evaluation of this designation. Protein BLAST searches were carried out at http://blast.ncbi.nlm.nih.gov/ using the 1044-amino-acid *S. cerevisiae* eEF3 protein (accession no. NP_013350.1) as the query sequence for a BLASTP search against the RefSeq database for Ascomycota and Basidiomycota and all non-fungal eukaryotes separately. eEF3 homologs from ten different Ascomycetes were selected for further analysis. In Basidiomycota, the second major phylum of fungi, eEF3 homologs

were identified in eight different species. Surprisingly, 18 other putative homologs were identified outside the fungal kingdom, including 7 in green algae. The similarity between the *S. cerevisiae* eEF3 sequence and its potential homologs in non-fungal species is not limited to the ATPase domains. In fact, these lower eukaryotic sequences ranged in similarity from 56 % (*E. huxleyi*) to 63 % (*P. infestans*). Thus, they are not more dissimilar to *S. cerevisiae* eEF3 than *C. neoformans* eEF3 (62 % similarity), which has already been shown to be functional as the only form of the essential protein in *S. cerevisiae* cells [18]. Together, these observations support the hypothesis that these eEF3-like proteins also retain eEF3 function.

6 Phylogenetic Tree of Putative eEF3 Homologs

An analysis of the phylogenetic relationship of putative eEF3 homologs shows clustering of species along known taxonomical lines (Fig. 3). The Ascomycetes and the Basidiomycetes partitioned into separate groups but remained joined to a common node, reflecting their shared fungal phylogeny. Of the unicellular eukaryotes where putative eEF3 homologs were identified, *Volvox carteri, Chlamydomonas reinhardtii, Ostreococcus tauri* and *Micromonas,* all members of the Chlorophyta phylum representing green algae, were part of one cluster. It is also interesting to note that the choanoflagellate *Monosiga brevicollis,* representing a group of free-living unicellular and colonial flagellate eukaryotes considered to be the closest living relatives of animals, is part of a cluster that diverged very early. The fact that eukaryotic eEF3-like proteins share a common ancestry with the fungal eEF3 strongly suggests that the emergence of eEF3 preceded the formation of the fungal kingdom of life. All of the eEF3-like orthologs were more similar to each other than to *S. cerevisiae* New1p, the soluble ATPase that is a close family member to eEF3 in *S. cerevisiae.* These observations suggest that the candidate genes identified are all likely authentic homologs.

7 Analysis of the Sequence Conservation in the of Domains eEF3-Like Proteins

To investigate the conservation of functionally important regions of eEF3, a domain-by-domain analysis of the multiple sequence alignment was carried out. A summary of these comparisons within individual phyla is presented in Table 1. Consensus sequences were derived from phyla with seven or more representative sequences, and the range of identity between the individual sequences and the consensus is displayed.

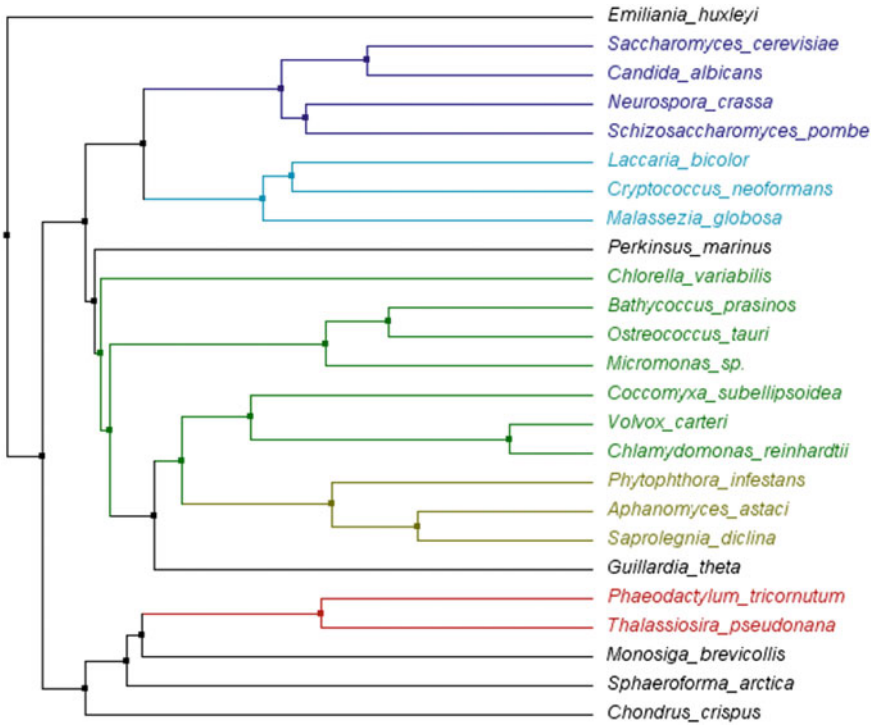


Fig. 3 Phylogenetic tree of eEF3 in eukaryotes. eEF3 homologs from the indicated species were aligned with ClustalWS, and the resulting alignment was used to construct an average distance tree using Jalview 2.9. Different taxa were color coded as follows: *dark blue*, Ascomycetes; *light blue*, Basidiomycetes; *green*, Chlorophyta; *gold*, Oomycetes; *red*, Bacillariophyta; *black*, single representatives of different taxa. Accession numbers for the sequences used to build the tree are as follows: *E. huxleyi* (XP_005765539.1), *S. cerevisiae* (NP_013350.1), *C. albicans* (XP_711404.1), *N. crassa* (XP_962438.1), *S. pombe* (NP_588285.1), *L. bicolor* (XP_001878516.1), *C. neoformans* (XP_775665.1), *M. globosa* (XP_001732572.1), *P. marinus* (XP_002783366.1), *C. variabilis* (XP_005846913.1), *B. prasinos* (XP_007514750.1), *O. tauri* (XP_003083207.1), *Micromonas_sp.* (XP_002506649.1), *C. subellipsoidea* (XP_005647408.1), *V. carteri* (XP_002951155.1), *C. reinhardtii* (XP_001692287.1), *P. infestans* (XP_002906761.1), *A. astaci* (XP_009834902.1), *S. diclina* (XP_008620178.1), *G. theta* (XP_005822594.1), *P. tricornutum* (XP_002180730.1), *T. pseudonana* (XP_002289207.1), *M. brevicollis* (XP_001747973.1), *S. artica* (XP_014157253.1) and *C. crispus* (XP_005714937.1)

7.1 Specific HEAT Repeats Show Higher Levels of Conservation

The HEAT domain contains eight helical repeats and is named after the four functionally characterized proteins containing this domain, Huntington, eEF3, 'A' subunit of PP2A and TOR1 [32]. The HEAT domain is only 55–75 % identical to

Table 1 Conservation of various domains of eEF3, expressed in terms of percent identity with respect to taxonomic consensus sequence

	Complete (1–1044)	HEAT (1–333)	4HB (334–416)	ABC1 (417–635)	ABC2 (636–759, 870–975)	Chromo (760–869)
Ascomycetes	70–83	62–75	37–68	71–84	85–95	74–93
Basidiomycetes	68–82	63–77	39–62	80–89	79–95	65–92
Green Algae	61–74	55–69	38–60	71–84	69–87	60–77

eEF3 domain sequence corresponds to that in *S. cerevisiae* [27]

the respective taxonomic consensus sequences, which is below the level of conservation of either the overall eEF3 or any of the individual domains aside from the four-helix bundle. A multiple sequence alignment of this region of the eEF3 homologs shows that while the N-terminal helical repeats, proposed to provide an extensive binding interface with the 40S subunit, show a low level of conservation, the terminal three repeats were more conserved (Fig. 2a, dark blue). While it is possible that these HEAT repeats are functionally redundant, an alternate possibility is the general conservation of proposed ribosome-binding properties in spite of sequence diversity. Unlike biochemical reactions, physical binding interactions are less influenced by specific residues, especially if they involve a large number of sites such as the case of the eEF3-ribosome interaction. Hence, it is conceivable that the evolutionary pressure to maintain a particular residue could be weak, resulting in significant sequence divergence with the passage of time.

In contrast to the remainder of the HEAT repeat, residues composing helical repeat seven are more conserved (Fig. 2b). A preliminary analysis of the position of this repeat on the cryo-EM of *S. cerevisiae* eEF3 shows that it is mostly buried with only a few residues available for intermolecular contacts (Fig. 2a, red). This suggests that helical repeat seven may play an important role in maintaining the overall fold of eEF3 through conserved intramolecular interactions and is thus less tolerant of sequence changes.

7.2 The 4-Helix Bundle of eEF3 Is Highly Divergent

An analysis of the four-helix bundle (4HB) domain shows a high degree of divergence, showing a range of only 37–68 % identity within individual phyla and much lower conservation across taxa (Table 1). From the cryo-EM data, the 4HB domain is not part of the ribosome-binding interface, and the functional relevancy of its presence in eEF3 is currently not known.

7.3 Key Motifs of ABC1 and ABC2 Domain Are Highly Conserved

The ABC1 and ABC2 domains of the eEF3-like proteins are the most highly conserved regions of the protein both within and across taxa (Table 1). The nucleotide binding domain of ABC proteins has multiple conserved motifs [29]. The Walker-A motif (GXXGXGKS/T where X is any amino acid), also known as the P-loop, binds ATP while the Walker-B motif (ΦΦΦΦD, where Φ is a hydrophobic residue) has a conserved glutamate that initiates the nucleophilic attack on ATP via a water molecule. Additional motifs important for catalysis have also been identified including the A, D, H and Q loops and the Signature domain. Not surprisingly, all of the key motifs and critical residues implicated in ATP binding and hydrolysis with the exception of the Q-loop are conserved in the ABC1 (Fig. 4) and ABC2 domains of eEF3 homologs from budding yeast to choanoflagellates. Such a high level of conservation strongly argues that the newly identified homologs in eukaryotes have been under continuous selective pressure for the maintenance of their ability to bind and hydrolyze ATP.

7.4 Chromodomain Insertion in ABC2 Is Divergent

The overall conservation of the chromodomain insertion in the ABC2 domain of eEF3 across eukaryotes is low with numerous gaps in the alignments in spite of it being flanked by more conserved ABC2 sequences on either side. As predicted for the HEAT domain, it may not be the primary sequence of this region that is

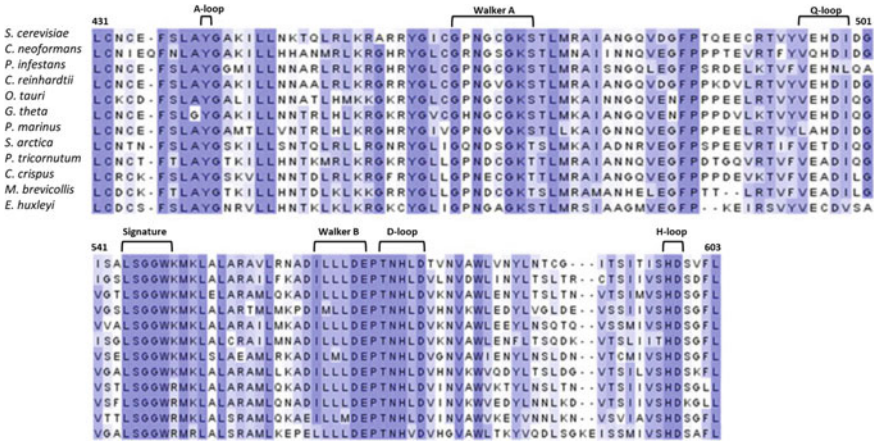


Fig. 4 Critical residues of ABC1 domain involved in ATP binding and hydrolysis are highly conserved. ABC1 domain of eEF3 from multiple sequence alignment was prepared as in Fig. 2b and is shown with sites of consensus motifs indicated

important but the tertiary structure. The chromodomain of *S. cerevisiae* eEF3p is predicted to interact with the 5S rRNA based on the homology of the tertiary structure of the fold to other nucleic acid binding proteins. Therefore, as long as the changes in the sequence maintain the structural fold, the function of the domain might be minimally affected. In contrast to the four-helix bundle domain whose overall conservation across taxa is also low, the chromodomain is well conserved within Ascomycetes, Basidiomycetes or green algae (Table 1). Mutagenesis of residues 803-808 of the chromodomain in *S. cerevisiae*, a region homologous to the DNA-binding domain of the *Sulfolobus acidocaldarius* Sac7 protein, has shown that the chromodomain is important for the maintenance of ATPase activity and eEF3 function in vivo [33].

8 Understanding the Unique Position of eEF3 in the Evolution of Protein Synthesis

Unlike other factors essential for translation elongation, such as eEF1A and eEF2, which have orthologs present in all three domains of life, eEF3 is restricted to eukaryotes. In addition, the requirement for eEF3 in eukaryotes is dependent on the species from which the ribosomes are isolated. Given this link between eEF3 and the ribosome and their known interaction, it is useful to compare the evolution of eEF3 to the phylogenetic tree created from the comparison of the eukaryotic small-subunit rRNA (Fig. 5). Interestingly, eEF3 and eEF3-like orthologs are absent in the deepest branches of the eukaryotic tree including Diplomonads, Microsporidia and Euglenozoa. The eEF3-like proteins appear in phyla that cluster together by this rRNA analysis, including Stramenopiles, Alveolates, Red Algae, Choanoflagellates and Fungi. These observations suggest that eEF3 may have evolved later in eukaryotic evolution in response to changes in the ribosome structure and/or regulation. However, since genomic sequences from unicellular eukaryotes are underrepresented in the NCBI databases, the possibility of eEF3-like sequences in the deeper branches of the eukaryotic tree cannot be excluded. In addition, no experimental evidence exists to confirm that these eEF3-like proteins retain eEF3's role in translation elongation and have not modified the core ABC domains to perform functions outside of protein synthesis.

Notably, eEF3 is absent in all animal and plant genomes sequenced to date. Given the analysis described above, these observations suggest that eEF3 was specifically lost in these lineages. Several hypotheses can be derived to explain this loss. For example, it has been shown that ribosomes from mammalian liver possess a higher level of intrinsic ATPase activity than yeast ribosomes [35–37]. This observation led to the hypothesis that the function of eEF3 at fungal ribosomes has been incorporated into the ribosome itself in higher eukaryotes. Alternatively, the ribosome in higher eukaryotes may have evolved such that the function of eEF3 is no longer required. For example, eEF3 in *S. cerevisiae* is thought to aid in the

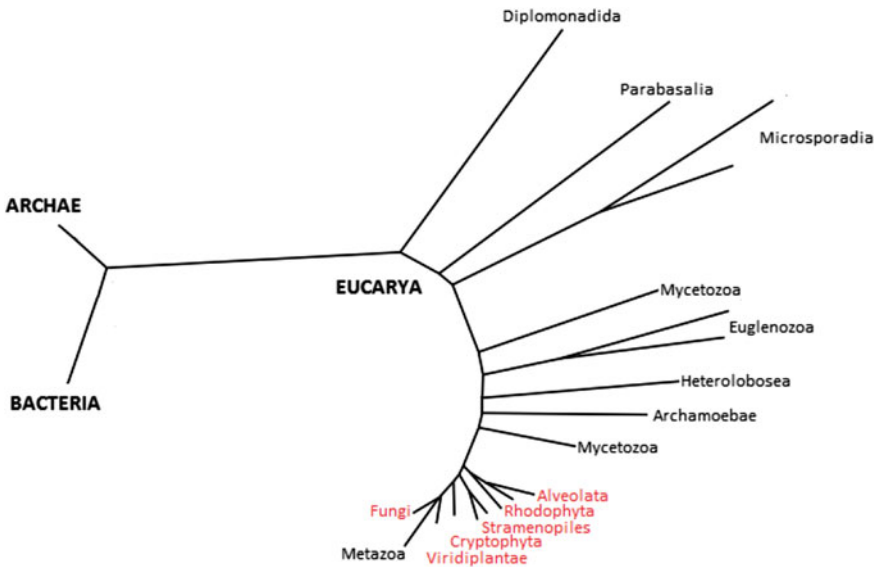


Fig. 5 Evolution of eEF3 in the tree of life. A phylogenetic tree based on SSU rRNA sequences that represents eukaryotic evolution (modified from Pace et al. [34]) to show only major eukaryotic taxonomical lineages. eEF3 bearing lineages shown in *red* indicate the emergence of eEF3 in the main trunk as a late event. Among Viridiplantae, while Chlorophytae (green algae) shows the presence of eEF3, Streptophyta, the major clade that includes land plants, lacks eEF3. Based on their position in the tree of life, the absence of eEF3 in both Streptophytes and Metazoans indicates that eEF3 was lost in these lineages independently

release of deacylated tRNA from the E-site of the ribosome. If the tRNA is bound less tightly in the E-site, eEF3 may no longer be required in higher eukaryotes. This change in the ribosome could be due to differences in either the ribosomal protein or RNA components. However, of the 78 cytoplasmic ribosomal protein families identified, all but L28e are highly conserved between *S. cerevisiae* and mammals. Differences in the ribosomal rRNA between yeast and mammals are more dramatic. For example, the human large subunit (LSU) rRNA is over 1 kb longer than the *S. cerevisiae* LSU rRNA because of the addition of numerous expansion segments that may alter the functional requirements of the ribosome.

9 Practical Implication of the Elucidation of the Function of eEF3 in Protein Synthesis

Based on its unique distribution, eEF3 has been suggested as an antifungal drug target. Fungal infections are increasingly threatening human, animal and plant health. The mortality rates for invasive infections with the three most common

species of human fungal pathogens are *C. albicans,* 20–40 %; *Aspergillus fumigatus*, 50–90 %; and *C. neoformans*, 20–70 % [38]. It is estimated that *Cryptococcal* meningitis, a fungal infection of the membranes covering the brain, affects close to 1 million people annually [39]. In addition to human health, fungal infections have consequences for plant and animal life. For example, emerging infectious diseases caused by fungi are a worldwide threat to food security with estimates of up to 125 million tons of the top five food crops being destroyed by fungus every year [40]. In animals, fungal infections caused by *Batrachochytrium dendrobatidis* and *Geomyces destructans* have caused massive declines in amphibian and bat populations, respectively [41, 42]. A better understanding of the mechanism of eEF3 function at the ribosome will pave the way to a rational approach in the design of fungal inhibitors while a study of the eEF3 homologs will provide a framework to understand, predict and plan the ecological consequences of such inhibitors on the ecosystem and the full range of organisms potentially impacted by such inhibitors.

References

1. Lorsch JR, Dever TE. Molecular view of 43 S complex formation and start site selection in eukaryotic translation initiation. J Biol Chem. 2010;285(28):21203–7.
2. Hinnebusch AG, Dever TE, Asano K. Mechanism of translation initiation in the yeast saccharomyces cerevisae. In: Sonenberg N, Hershey JWB, Mathews MB, editors. Translational control in biology and medicine. Cold Spring Harbor: Cold Spring Harbor Laboratory Press; 2007. p. 225–68.
3. Merrick WC, Nyborg J. The protein biosynthesis elongation cycle. In: Sonenberg N, Hershey JWB, Mathews MB, editors. Translational control of gene expression. Cold Spring Harbor: Cold Spring Harbor Laboratory; 2000. p. 89–126.
4. Jackson RJ, Hellen CU, Pestova TV. Termination and post-termination events in eukaryotic translation. Adv Protein Chem Struct Biol. 2012;86:45–93.
5. Valente L, Kinzy TG. Yeast as a sensor of factors affecting the accuracy of protein synthesis. Cell Mol Life Sci. 2003;60(10):2115–30.
6. Carvalho MG, Carvalho JF, Merrick WC. Biological characterization of various forms of elongation factor 1 from rabbit reticulocytes. Arch Biochem Biophys. 1984;234:603–11.
7. Rodnina MV, Pape T, Fricke R, Kuhn L, Wintermeyer W. Initial binding of the elongation factor Tu•GTP•aminoacyl-tRNA complex preceding codon recognition on the ribosome. J Biol Chem. 1996;271:646–52.
8. Pittman YR, Valente L, Jeppesen MG, Andersen GR, Patel S, Kinzy TG. Mg2 + and a key lysine modulate exchange activity of eukaryotic translation elongation factor 1B alpha. J Biol Chem. 2006;281(28):19457–68.
9. Saha SK, Chakraburtty K. Protein synthesis in yeast. Isolation of variant forms of elongation factor 1 from the yeast Saccharomyces cerevisiae. J Biol Chem. 1986;261:12599–603.
10. Slobin LI, Moller W. Purification and properties of an elongation factor functionally analogous to bacterial elongation factor Ts from embryos of Artemia salina. Eur J Biochem. 1978;84:69–77.
11. Hiraga K, Suzuki K, Tsuchiya E, Miyakawa T. Cloning and characterization of the elongation factor EF-1b homologue of Saccharomyces cerevisiae. EF-1b is essential for growth. FEBS Lett. 1993;316:165–9.

12. Olarewaju O, Ortiz PA, Chowdhury W, Chatterjee I, Kinzy TG. The translation elongation factor, eEF1B, plays a role in the oxidative stress response pathway. RNA Biol. 2004;1:12–7.
13. Esposito AM, Kinzy TG. The eukaryotic translation elongation Factor 1Bgamma has a non-guanine nucleotide exchange factor role in protein metabolism. J Biol Chem. 2010;285 (49):37995–8004.
14. Skogerson L, Wakatama E. A ribosome-dependent GTPase from yeast distinct from elongation factor 2. Proc Natl Acad Sci USA. 1976;73:73–6.
15. Dasmahapatra B, Chakraburtty K. Purification and properties of elongation factor 3 from Saccharomyces cerevisiae. J Biol Chem. 1981;256:9999–10004.
16. Colthurst DR, Santos M, Grant CM, Tuite MF. Candida albicans and three other Candida species contain an elongation factor structurally and functionally analogous to elongation factor 3. FEMS Microbiol Lett. 1991;64(1):45–9.
17. Uritani M, Miyazaki M. Characterization of the ATPase and GTPase activities of the elongation factor 3 (EF-3) from yeasts. J Biochem. 1988;103:522–30.
18. Blakely G, Hekman J, Chakraburtty K, Williamson PR. Evolutionary divergence of an elongation factor 3 from Cryptococcus neoformans. J Bact. 2001;183:2241–8.
19. Xu J, Kiel MC, Golshani A, Chosay JG, Aoki H, Ganoza MC. Molecular localization of a ribosome-dependent ATPase on Escherichia coli ribosomes. Nucleic Acids Res. 2006;34 (4):1158–65.
20. Skogerson L, Engelhardt D. Dissimilarity in protein chain elongation factor requirements between yeast and rat liver ribosomes. J Biol Chem. 1977;252(4):1471–5.
21. Qin S, Xie A, Bonato MCM, McLaughlin CS. Sequence analysis of the translational elongation factor 3 from Saccharomyces cerevisiae. J Biol Chem. 1990;265:1903–12.
22. Anand M, Chakraburtty K, Marton MJ, Hinnebusch AG, Kinzy TG. Functional interactions between yeast translation eukaryotic elongation factor (eEF) 1A and eEF3. J Biol Chem. 2003;278(9):6985–91.
23. Anand M, Balar B, Ulloque R, Gross SR, Kinzy TG. Domain and nucleotide dependence of the interaction between Saccharomyces cerevisiae translation elongation factors 3 and 1A. J Biol Chem. 2006;281(43):32318–26.
24. Triana-Alonso FJ, Chakraburtty K, Nierhaus KH. The elongation factor 3 unique in higher fungi and essential for protein biosynthesis is an E site factor. J Biol Chem. 1995;270 (35):20473–8.
25. Kamath A, Chakraburtty K. Role of yeast elongation factor 3 in the elongation cycle. J Biol Chem. 1989;264:15423–8.
26. Uritani M, Miyazaki M. Role of the yeast peptide elongation factor 3 (EF-3) at the AA-tRNA binding step. J Biochem. 1988;104:118–26.
27. Andersen CB, Becker T, Blau M, Anand M, Halic M, Balar B, et al. Structure of eEF3 and the mechanism of transfer RNA release from the E-site. Nature. 2006;443(7112):663–8.
28. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. Jalview Version 2–a multiple sequence alignment editor and analysis workbench. Bioinformatics. 2009;25(9):1189–91.
29. Hollenstein K, Dawson RJ, Locher KP. Structure and mechanism of ABC transporter proteins. Curr Opin Struct Biol. 2007;17(4):412–8.
30. Decottignies A, Goffeau A. Complete inventory of the yeast ABC proteins. Nat Genet. 1997;15(2):137–45.
31. Sarthy AV, McGonigal T, Capobianco JO, Schmidt M, Green SR, Moehle CM, et al. Identification and kinetic analysis of a functional homolog of elongation factor 3, YEF3 in Saccharomyces cerevisiae. Yeast. 1998;14:239–53.
32. Andrade MA, Bork P. HEAT repeats in the Huntington's disease protein. Nat Genet. 1995;11 (2):115–6.
33. Sasikumar AN, Kinzy TG. Mutations in the Chromodomain-like Insertion of Translation Elongation Factor 3 Compromise Protein Synthesis through Reduced ATPase Activity. J Biol Chem. 2014;289(8):4853–60.
34. Pace NR. A molecular view of microbial diversity and the biosphere. Science. 1997;276 (5313):734–40.

35. Kovalchuke O, Chakraburtty K. Comparative analysis of ribosome-associated adenosinetriphosphatase (ATPase) from pig liver and the ATPase of elongation factor 3 from Saccharomyces cerevisiae. Eur J Biochem. 1994;226(1):133–40.
36. El'skaya AV, Ovcharenko GV, Palchevskii SS, Petrushenko ZM, Triana-Alonso FJ, Nierhaus KH. Three tRNA binding sites in rabbit liver ribosomes and role of the intrinsic ATPase in 80S ribosomes from higher eukaryotes. Biochemistry. 1997;36(34):10492–7.
37. Rodnina MV, Serebryanik AI, Ovcharenko GV, El'Skaya AV. ATPase strongly bound to higher eukaryotic ribosomes. Eur J Biochem. 1994;225(1):305–10.
38. Lai CC, Tan CK, Huang YT, Shao PL, Hsueh PR. Current challenges in the management of invasive fungal infections. J Infect Chemother Official J Jpn Soc Chemother. 2008;14 (2):77–85.
39. Park BJ, Wannemuehler KA, Marston BJ, Govender N, Pappas PG, Chiller TM. Estimation of the current global burden of cryptococcal meningitis among persons living with HIV/AIDS. Aids. 2009;23(4):525–30.
40. Kupferschmidt K. Mycology: Attack of the clones. Science. 2012;337(6095):636–8.
41. Berger L, Speare R, Daszak P, Green DE, Cunningham AA, Goggin CL, et al. Chytridiomycosis causes amphibian mortality associated with population declines in the rain forests of Australia and Central America. Proc Natl Acad Sci USA. 1998;95(15):9031–6.
42. Blehert DS, Hicks AC, Behr M, Meteyer CU, Berlowski-Zier BM, Buckles EL, et al. Bat white-nose syndrome: an emerging fungal pathogen? Science. 2009;323(5911):227.

Evolution of TOR and Translation Control

Bruno D. Fonseca, Tyson E. Graber, Huy-Dung Hoang, Asier González, Alexander A. Soukas, Greco Hernández, Tommy Alain, Stephanie L. Swift, Ronit Weisman, Christian Meyer, Christophe Robaglia, Joseph Avruch and Michael N. Hall

1 Introduction

Rapamycin is a natural macrolide antibiotic originally isolated from *Streptomyces hygroscopicus*—a filamentous bacteria found in a soil sample on Easter Island [1, 2]. The quest to identify the target of rapamycin began in the late 1980s. At that time, rapamycin was known for its ability to inhibit the growth of fungi [1–3] and the proliferation of T cells [4–6]. This led my group (M. Hall) to hypothesize that the "target of rapamycin" (TOR) must have been evolutionarily conserved all the way from fungi to humans. This assumption prompted the use of the budding yeast *Saccharomyces cerevisiae*—a simple yet powerful genetically tractable organism—in the identification of the target of rapamycin [7]. We selected for mutations in *Saccharomyces cerevisiae* that confer resistance to the growth-inhibitory effects of rapamycin [7]. This selection yielded several mutants, a number of which bore

The original version of the book was revised: A spelling error in the author's name was corrected, a figure was placed correctly, and a typographical error was corrected. The erratum to this chapter is available at 10.1007/978-3-319-39468-8_23

B.D. Fonseca (✉) · T.E. Graber · H.-D. Hoang · T. Alain · S.L. Swift
Children's Hospital of Eastern Ontario Research Institute, 401 Smyth Road, Ottawa, ON K1H 8L1, Canada
e-mail: brunodfonseca@gmail.com

A. González · M.N. Hall
Biozentrum, University of Basel, 4056 Basel, Switzerland

A.A. Soukas · J. Avruch
Massachusetts General Hospital, Boston, MA, USA

G. Hernández
Division of Basic Science, National Institute of Cancer, 22 San Fernando Avenue, Tlalpan, 14080 Mexico City, Mexico

R. Weisman
Department of Molecular Microbiology and Biotechnology, Tel Aviv University, Tel Aviv, Israel

R. Weisman
Department of Natural and Life Sciences, The Open University of Israel, Raanana, Israel

G. Hernández and R. Jagus (eds.), *Evolution of the Protein Synthesis Machinery and Its Regulation*, DOI 10.1007/978-3-319-39468-8_15

gain-of-function mutations in two novel genes, which we named target of rapamycin 1 and 2 (*TOR1* and *TOR2*) [7, 8].

Over the next 3 years, four research groups raced to identify the mammalian counterpart of the yeast TOR1 and TOR2. These efforts culminated with the identification of mTOR (mammalian TOR) in 1994 [9–12]. The discovery of the two TOR proteins in budding yeast [7, 8] together with the identification of mTOR in mammals [9–12] initiated the era of TOR signaling research. The delineation of TOR in other genetically tractable model organisms, including the fission yeast *Schizosaccharomyces pombe* [13, 14], the fruit-fly *Drosophila melanogaster* [15, 16], the roundworm *Caenorhabditis elegans* [17] and the plant *Arabidopsis thaliana* [18], came about around a decade later. The meticulous study of the TOR protein in the aforementioned model organisms by dedicated investigators (for a quarter of a century) has vastly contributed to our detailed understanding of TOR and its functions today.

TOR is an important regulator of cellular and organismal growth. TOR exerts its effect on growth through the control of numerous intracellular processes. In this chapter, we devote our attention to the role of TOR in the control of protein synthesis [the reader is referred to the following reviews [19–21] for excellent coverage of other TOR outputs]. Protein synthesis (also commonly referred to as mRNA translation) is perhaps the best-understood output of TOR, which appears to be largely conserved across the eukaryotic kingdom. This chapter celebrates the contribution of each eukaryotic model organism to our present understanding of the role of TOR in the control of protein synthesis.

2 TOR Homologs in Evolution

The discovery of the TOR in budding yeast in 1991 [7] was quickly followed by the identification of its mammalian homolog in 1994 [9–12]. Four groups independently reported the identification of the mammalian homolog of TOR, mTOR [9–12]. A decade later, TOR homologs were also identified in other eukaryotes including fission yeast (*Schizosaccharomyces pombe,* TOR1/TOR2) [13, 14], plants

C. Meyer
Institut Jean-Pierre Bourgin, Unité Mixte de Recherche 1318, Institut National de la Recherche Agronomique, AgroParisTech, ERL CNRS 3559, Saclay Plant Sciences, 78026 Versailles Cedex, France

C. Robaglia
Aix-Marseille Université, Laboratoire de Génétique et Biophysique des Plantes, 13009 Marseille, France

C. Robaglia
CNRS, UMR 7265 Biologie Végétale & Microbiologie Environnementales, 13009 Marseille, France

C. Robaglia
CEA, IBEB, 13009 Marseille, France

Table 1 Identity of TORC1 components in key model organisms

	mTOR		S6K		4E-BP1		PRAS40		LARP1		RAPTOR		mLST8/GβL	
Homo sapiens	mTOR	100 %	S6K1	100 %	4E-BP1	100 %	PRAS40	100 %	LARP1	100 %	RAPTOR	100 %	mLST8	100 %
			S6K2	81 %	4E-BP2	99 %								
					4E-BP3	74 %								
Drosophila melanogaster	dTOR	54 %	dS6K	73 %	d4E-BP/THOR	49 %	dPRAS40/LOBE	38 %	dLARP	50 %	dRAPTOR	64 %	dLST8	49 %
Caenorhabditis elegans	CeTOR	35 %	CeS6K	68 %	Ce4E-BP	33 %	_		CeLARP1a	47 %	CeDAF15	45 %	CeLST8	25 %
									CeLARP1b	37 %				
Arabidopsis thaliana	AtTOR	42 %	AtS6K2	48 %	_		_		_		AtRAPTOR1	30 %	AtLST8	45 %
			AtS6K1	48 %							AtRAPTOR2	42 %		
Schizosaccharomyces pombe	TOR1	40 %	Gad8	48 %	_		_		NP_594554.1	45 %	Mip1	57 %	Pop3	45 %
	TOR2	46 %	Psk1	47 %										
Saccharomyces cerevisiae	TOR1	42 %	Sch9	46 %	Eap1	50 %	_		Sro9	45 %	Kog1	56 %	Lst8	47 %
	TOR2	43 %	Ypk3	38 %	Caf20	57 %			Slf1	39 %				

(*Arabidopsis thaliana,* AtTOR) [18], nematodes (*Caenorhabditis elegans,* CeTOR) [17] and flies (*Drosophila melanogaster,* dTOR) [15, 16], confirming our early assumption that the "target of rapamycin" was indeed widely conserved in evolution. TOR is found in every organism across the eukaryotic lineage (all the way from protozoa to metazoa) [22]. TOR is, however, conspicuously absent from prokaryotes[1] [22], suggesting that it appeared either during or soon after the bifurcation of the prokaryotic and eukaryotic lineages. Alignment of the TOR proteins in eukaryotes reveals that this protein is highly conserved at the amino acid level (Table 1), and that key structural domains of TOR are found in every eukaryotic organism. The mammalian TOR protein shares 54 % identity at the amino acid level with dTOR, 42 % with AtTOR, 35 % with CeTOR, 40–46 % with fission yeast TOR1-TOR2 and 42–43 % with budding yeast TOR1-TOR2 (Table 1). Interestingly, yeast species possess two *TOR* genes (referred to as *TOR1* and *TOR2*), while higher eukaryotes have a single *TOR* gene. The yeast TOR1 and TOR2 proteins are structurally and functionally similar but not identical [13, 25]. The two TOR proteins are found in two structurally and functionally distinct protein complexes (discussed in Sects. 4–6 of this chapter). The ubiquitous presence of TOR protein(s) across eukaryotes and the high level of conservation of TOR at the amino acid level are consistent with the fact that TOR plays a fundamental and conserved role in the control of cell growth across eukaryotes.

3 TOR Structure and Function

TOR is a serine/threonine protein kinase that belongs to the kinase family of phosphatidylinositol 3-kinase-related kinases (PIKKs), that also comprises ataxia-telangiectasia mutated (ATM), ataxia- and Rad3-related (ATR), DNA-dependent protein kinase catalytic subunit (DNA-PKcs), suppressor of morphogenesis in genitalia (SMG-1) and transformation/transcription domain-associated protein (TRRAP) [26]. Orthologous TOR proteins are highly conserved at the structural level in eukaryotic evolution, and all contain multiple domains in common: The catalytic kinase domain (homologous to that of other PIKKs [26]) is located near the C-terminus and mediates the phosphorylation of serine and threonine residues on its target proteins.[2] Two HEAT domains are present in the N-terminal region (formed by anti-parallel helix-turn-helix repeats, so-named because they were originally identified in the following proteins: Huntingtin, Elongation Factor 3, the A subunit of protein phosphatase 2A and TOR). C-terminally of the HEAT repeats lies a FAT domain (named after the PIKKs that

[1]Possibly reflecting the fact that the prevalent phospho-accepting amino acids in bacterial proteins are histidine and aspartate, and not serine, threonine or tyrosine (which are more frequently phosphorylated in eukaryotes) [23, 24].

[2]A recent study shows that mTOR (specifically mTORC2) can also catalyze the phosphorylation of tyrosine residues [27, 28].

contain this motif: FRAP, ATM and TRRAP) [29], followed by the FRB (FKBP12/rapamycin-binding) domain, which, as the name suggests, mediates the binding to the FKBP12/rapamycin complex. Finally, another FAT domain is found beyond the catalytic domain at the C-terminus of TOR and is thus termed the FATC domain (short for FAT at C-terminus).[3] Paralogous TOR1 and TOR2 found in yeasts are also structurally similar: TOR1 and TOR2 proteins in budding yeast display 67 % identity at the amino acid level, while TOR1 and TOR2 in fission yeast show 52 % identity at the amino acid level. However, TOR proteins do, in some instances, perform distinct functions. For example, in budding yeast, both TOR1 and TOR2 are involved in the control of transcription and mRNA translation programs [32], while certain cellular functions are mediated exclusively by TOR2 [33]. In budding yeast, actin polarization [34–37] and genome stability [38] are regulated exclusively by TOR2 (and not TOR1). Thus, TOR2 plays both TOR1/TOR2-shared and TOR2-specific roles in *S. cerevisiae*. Interestingly, the kinase domains of these two proteins are interchangeable [25], indicating that the functional divergence of TOR1 and TOR2 proteins in budding yeast is not the result of altered catalytic activity but rather the result of their association with different sets of proteins. The reason for the functional divergence between TOR1 and TOR2 will be addressed in further detail in Sect. 4 of this chapter. In contrast to yeasts, higher eukaryotes encode for a single TOR gene (e.g., dTOR, CeTOR, AtTOR or mTOR[4]) whose protein product integrates different protein complexes and performs distinct cellular functions. The role of the TOR protein in flies, plants, nematodes and mammals will be discussed in further detail in Sects. 5–9 of this chapter.

4 Two TOR Complexes: TORC1 and TORC2

Following the identification of TOR1 and TOR2 (in budding yeast) and mTOR (in mammals), it became apparent that TOR proteins do not function alone, but instead form large (megadalton) molecular weight protein complexes [40, 41]. It is now known that TOR proteins (in both yeast and higher eukaryotes) form two distinct complexes, in which TOR associates with different binding partners [40–42]. Studies in budding yeast [33, 40] and mammals [41, 42] described two branches of TOR signaling, each of which is specifically mediated by one of two TOR complexes: TORC1 and TORC2 in yeast (or mTORC1 and mTORC2 in mammals). In budding yeast, TORC1 is composed by either TOR1 or TOR2, the scaffolding protein Kog1 (Kontroller of Growth 1), the small GTPase-like protein Lst8 (lethal

[3]The structure of mTOR in complex with its binding proteins has now been defined at the atomic level by the following studies [30, 31].

[4]mTOR gene has been reported to encode two isoforms resultant from an alternative splicing event. The long isoform (mTOR alpha) is 289 kDa and the short isoform (mTOR beta) is 80 kDa. Most published studies refer to the long (289 kDa) mTOR alpha isoform. Further details in Ref. [39].

with Sec13 protein 8) [40] and Tco89 (89-kDa subunit of TOR complex one) [43], while TORC2 is formed by TOR2 (but not TOR1), Lst8, Bit61 (61-kDa binding partner of TOR2) [43] and three TORC2-specific proteins: Avo1, Avo2 and Avo3—so-named for their exquisite ability to adhere voraciously to TOR2 [40]. From an evolutionary standpoint, it is important to note that the majority of these TOR complex protein components are conserved from yeast to man. Briefly, in mammals, Kog1 is known as RAPTOR (regulatory-associated protein of mTOR) [41, 42], Lst8 as mammalian LST8 (mLST8) or GβL (G-protein β-subunit-like protein) [44], Avo3 as RICTOR (short for rapamycin-insensitive companion of TOR) [45, 46], and Avo1 as mSIN1 (mammalian stress-activated protein kinase-interacting protein 1) [47]. In addition to showing a unique protein composition, TORC1/mTORC1 and TORC2/mTORC2 also differ with regard to sensitivity to rapamycin: TORC1/mTORC1 is rapamycin-sensitive, whereas TORC2/mTORC2 is largely rapamycin-insensitive [40, 45, 48]. Moreover, TORC1/mTORC1 and TORC2/mTORC2 perform distinct cellular functions. TORC1/mTORC1 couples nutrient sufficiency to growth by activating anabolic processes such as protein synthesis and ribosome biogenesis and by repressing catabolic processes such as autophagy [49–52]. In budding yeast, TORC1 controls protein synthesis at multiple levels, in particular translation initiation and ribosome biogenesis (discussed in detail in Sect. 5). The best-characterized TORC1 substrates in *Saccharomyces cerevisiae* are Sch9, a serine/threonine protein kinase belonging to the AGC kinase family [53], and the Protein Phosphatase 2A (PP2A) and PP2A-like regulator Tap42 [54]. In budding yeast, TORC2 mediates the organization of the actin cytoskeleton and thereby spatial control of cell growth [55, 56]. The best-known targets of TORC2 are the AGC kinase family members Ypk1 and Ypk2 [57–59].

5 *Saccharomyces cerevisiae*

5.1 *TOR and Translation Initiation*

The first hints that TOR, specifically TORC1, stimulates protein synthesis came from experimentation with the TOR inhibitor rapamycin. Studies from budding yeast and mammalian cells demonstrated that TOR regulates both the initiation and the elongation steps of mRNA translation [32, 60–63]. Translation initiation is an important step in protein synthesis, in which the eukaryotic initiation factor (eIF)4F binds the pre-inititiation complex (PIC) that recruits the 40S subunit of the ribosome to the 5′end of an mRNA. The 40S subunit, subsequently, scans the mRNA until it finds a start codon (typically AUG), at which point a charged initiator methionyl-tRNA ($tRNA_i^{Met}$) recognizes the 40S subunit, and the 60S subunit joins to form the 80S ribosome (reviewed in [64]). In mammals, mTORC1 controls translation initiation via the eIF4E-binding proteins 4E-BPs [60]. The 4E-BPs-mediated control of translation initiation (in mammals) will be revisited in

greater detail in Sect. 10 of this chapter. For the purpose of discussing translation initiation in budding yeast, it suffices to say at this point that 4E-BPs are translational repressors that disrupt the eIF4F complex, thereby preventing cap-dependent translation initiation. mTORC1 phosphorylates 4E-BPs, causing their dissociation from eIF4E to promote translation initiation (reviewed in [65]).

In budding yeast, two functional homologs of 4E-BPs have been identified: Eap1 and Caf20 [66–68]. Disruption of *EAP1* confers partial resistance to rapamycin, suggesting that Eap1 is implicated in TORC1 signaling. Thus, budding yeast TORC1 may stimulate translation initiation via Eap1 by a mechanism similar to that described for mammalian 4E-BPs [66] (Fig. 1), although, at present, there is no evidence indicating that TORC1 affects Eap1 phosphorylation. Caf20 is phosphorylated by casein kinase, and this phosphorylation is decreased during the stationary phase [68]. However, deletion of *CAF20* does not confer resistance to rapamycin [66]. It remains to be determined whether TORC1 affects Caf20.

In budding yeast, TORC1 also promotes translation initiation by maintaining eIF4G levels via an as-yet undefined mechanism [69] (Fig. 1). Studies in yeast and mammals suggest that TORC1 controls eIF4G phosphorylation [70–72]; however, the impact of phosphorylation on eIF4G activity requires further investigation. TORC1 also targets eIF2 (reviewed in [73]). eIF2 mediates the binding of $tRNA_i^{Met}$ to the 40S ribosomal subunit in a GTP-dependent manner. eIF2 is a heterotrimer consisting of α, β and γ subunits, and its activity is negatively regulated by phosphorylation. The kinase Gcn2 phosphorylates eIF2α at Ser51, thereby inhibiting the exchange of GDP for GTP on eIF2 and ultimately blocking translation (Fig. 1). The

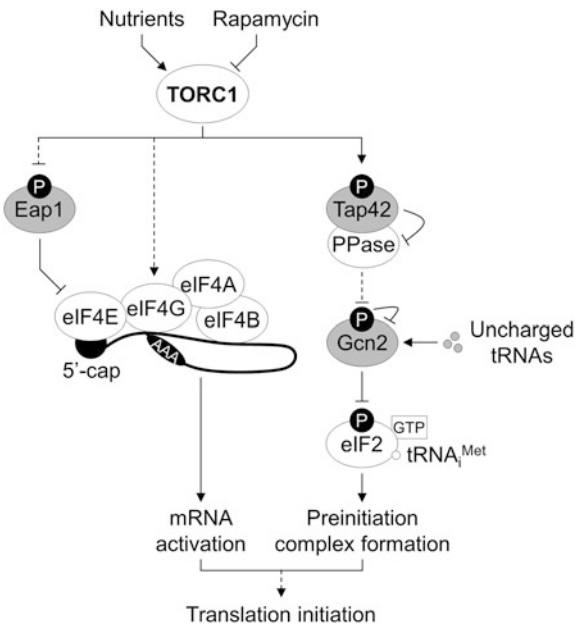


Fig. 1 Model depicting the steps of mRNA translation initiation that are regulated by TORC1 in budding yeast. Proteins shown in *white* promote mRNA translation initiation. Proteins in *gray* inhibit mRNA translation initiation. *Dashed lines* indicate indirect interactions. See main text for details

kinase activity of Gcn2 is inhibited by phosphorylation and stimulated by uncharged tRNAs that bind to Gcn2 in amino acid-starved cells. Phosphorylation of Gcn2 at Ser577 inhibits Gcn2 by reducing its tRNA-binding ability, thus promoting translation [74, 75]. The identity of the Gcn2 kinase is unknown other than it is not Sch9 [49]. It has been proposed that TORC1 increases Gcn2 phosphorylation partly by inhibiting one or more Tap42-associated phosphatases [75] (Fig. 1). Despite the conserved role of Gcn2 in translation, the involvement of TORC1 in Gcn2 regulation has only been reported in budding and fission yeast (see Sect. 6.1, last two paragraphs). A recent study showed that, upon amino acid starvation, TORC2-Ypk1 signaling inhibits the Ca^{2+}/calmodulin-dependent phosphatase calcineurin to activate Gcn2, increase eIF2α phosphorylation at Ser51 and ultimately promote autophagy [76]. It is unclear whether Ypk1-mediated eIF2α phosphorylation also affects translation.

5.2 TOR and Ribosome Biogenesis

The eukaryotic ribosome is composed of two subunits: the small 40S subunit, that contains one 18S ribosomal RNA (rRNA) and 33 ribosomal proteins (RPs), and the large 60S subunit containing three rRNAs (5S, 5.8S and 25S) and 46 RPs (for a review, see [77] and references therein). Eukaryotic ribosome assembly requires 76 different small nucleolar RNAs and more than 200 different assembly factors encoded by the Ribi (ribosome biogenesis) regulon [78]. Furthermore, ribosome synthesis involves the coordinated activity of all three RNA polymerases. RNA Pol I (Pol I) transcribes the precursor 35S rRNA, which is further processed to yield mature 18S, 5.8S and 25S rRNAs. RNA Pol II (Pol II) transcribes mRNAs (including RP and Ribi mRNAs), and RNA Pol III (Pol III) transcribes 5S rRNA and tRNAs. Inactivation of TORC1 reduces the transcription of essentially all genes involved in ribosome biogenesis because of a general inhibition of Pol I, II and III [79–83] (Fig. 2).

Several studies have shown that, in budding yeast, TORC1 stimulates Pol I recruitment to the 35S ribosomal DNA (35S rDNA) promoter by stabilizing the essential initiation factor Rrn3 [84–86]. However, additional mechanisms exist. For example, Sch9 promotes recruitment of Pol I to the 35S rDNA promoter via an Rrn3-independent, yet largely unknown mechanism [70] (Fig. 2). It has been suggested that the decrease of Pol I activity after short-term TORC1 inactivation is due to a reduction in Pol II activity and thus RP synthesis, suggesting a crosstalk between Pol I and Pol II [87]. TORC1 may support Pol I activity by indirectly promoting histone H3 acetylation at Lys56 [88]. Curiously, Li et al. reported that TOR1 translocates into the nucleus and promotes 35S rRNA synthesis through direct binding to the 35S rDNA promoter [89].

TORC1 regulates expression of RP and Ribi genes via several transcription factors (Fig. 2). The forkhead transcription factor Fhl1 is constitutively bound to most RP gene promoters, and its regulation involves two phosphoproteins: Ifh1 (a

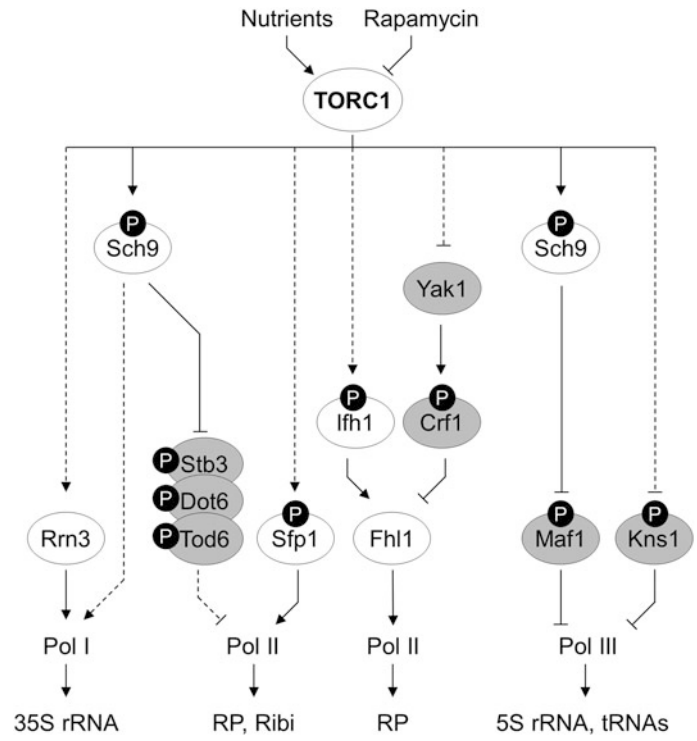


Fig. 2 Model depicting the main pathways by which TORC1 regulates the three RNA polymerases in budding yeast. Proteins shown in *white* stimulate RNA polymerase activity. Proteins shown in *gray* inhibit RNA polymerase activity. *Dashed lines* indicate indirect interactions. See main text for details

coactivator) and Crf1 (a corepressor) [90–92]. When TORC1 is active, phosphorylated Ifh1 binds Fhl1 and thereby stimulates RP gene transcription [93–95]. When TORC1 is inactive, phosphorylated Crf1 translocates into the nucleus, competes with Ifh1 for binding to Fhl1 and inhibits RP gene transcription [93]. Crf1 is phosphorylated by the TORC1- and PKA-regulated kinase Yak1. It is unclear how TORC1 affects Ifh1 phosphorylation. TORC1 also induces RP gene expression independent of Fhl1, Ifh1 and Crf1. Under favorable growth conditions, TORC1 directly phosphorylates the transcriptional activator Sfp1 to promote its binding to RP and Ribi gene promoters [78, 96, 97]. TORC1 also stimulates Pol II activity in an Sch9-dependent but Fhl1- and Sfp1-independent manner. Sch9 phosphorylates the transcriptional repressors Stb3, Dot6 and Tod6, preventing them from recruiting the RPD3L histone deacetylase to RP and Ribi gene promoters [53, 70, 78] (Fig. 2).

In mammals, mTORC1 phosphorylates and activates ribosomal S6 kinases (S6Ks), which in turn phosphorylate ribosomal protein S6 (RPS6, a component of the 40S ribosomal subunit) and ultimately promote transcription of genes required for ribosome biogenesis [98]. S6Ks may also control protein synthesis by affecting

several initiation factors (reviewed in [65, 99]). In budding yeast, Rps6 phosphorylation is differentially stimulated by TORC1 via the AGC kinase Ypk3 and by TORC2 via Ypk1/2 [100, 101]. However, neither Rps6 phosphorylation nor Ypk1-3 seems to regulate mRNA translation or ribosome biogenesis [101].

Transcription of tRNA genes relies exclusively on Pol III. TORC1 regulates tRNA transcription principally via Sch9 and the Pol III repressor Maf1 (Fig. 2). When TORC1 is active, Sch9 directly phosphorylates and inhibits Maf1. When TORC1 is inactive, dephosphorylated Maf1 translocates into the nucleus and inhibits Pol III transcription [70, 102, 103]. Sch9 may also promote Pol III activity by phosphorylating and activating Bdp1, the essential subunit of a Pol III transcription factor [104]. The observation that rapamycin reduces Pol III transcription in *sch9* mutant cells suggests that TORC1 regulates Pol III through additional, Sch9-independent mechanism(s) [103]. Such a mechanism may involve the LAMMER kinase Kns1. Upon rapamycin treatment, Kns1 is phosphorylated and accumulates in the nucleus. Nuclear Kns1 phosphorylates the Pol III subunit Rpc53 to inhibit tRNA transcription [105]. Finally, it has been suggested that TORC1 supports the activity of all three RNA polymerases by promoting binding of the prefoldin Uri/Bud27 to the Pol I, II and III subunit Rpb5 [106].

5.3 Concluding Remarks

As described in this section, TORC1 stimulates protein synthesis by promoting mRNA translation initiation and ribosome biogenesis. Interestingly, several observations suggest a feedback mechanism from the protein synthesis machinery to TORC1. Firstly, chemical inhibition of translation using cycloheximide activates TORC1 [53]. Secondly, Sfp1 negatively regulates Sch9 phosphorylation [97]. Thirdly, the leucyl-tRNA synthetase (LeuRS) has been reported to be an upstream activator of TORC1 in yeast and mammals [107, 108]. Finally, analysis of a budding yeast tRNA deletion library comprising 204 deletions of 275 tRNA genes revealed that 20 % of these mutants show growth defects under nutrient limiting conditions [109]. Follow-up studies are needed to characterize potential feedback mechanisms.

Finally, what is the role of TORC2 in protein synthesis? TORC2 associates with ribosomes in yeast and mammals [110–112]. Mammalian TORC2 is activated by its association with the ribosome to co-translationally phosphorylate and stabilize substrates [110, 113]. Experiments using budding yeast strains in which TORC2 function can be acutely and specifically inhibited by rapamycin [114] may clarify the role of (m)TORC2 in protein synthesis.

6 *Schizosaccharomyces pombe*

6.1 TOR: Cell Growth and Beyond

The TOR complexes in the fission yeast *Schizosaccharomyces pombe* harbor many similarities to its budding yeast counterpart, making *S. pombe* an excellent complementary model system to study TOR signaling in lower eukaryotes. Two TOR homologs were identified in *S. pombe*. These were named Tor1 and Tor2 following the order of their discovery [13]. However, it was later realized that Tor1 is actually the catalytic subunit of TORC2, the complex that also contains Ste20 and Sin1 (the fission yeast homologs of Avo3/RICTOR and Avo1/mSIN1). Tor2 is the catalytic subunit of TORC1, the complex that also contains Mip1, a Kog1/RAPTOR homolog [115, 116]. The *S. pombe* TOR complexes also share the Lst8/mLST8/GβL homolog, known as Wat1 or Pop3. The similarities between *S. pombe* TORC1 and TORC2 and their eukaryotic counterparts extend beyond their structural similarities and includes similar downstream and upstream effectors and regulators, as well as similarities in subcellular localization and functions. In *S. pombe*, TORC1 is activated in response to nutritional signals, in particular the availability of nitrogen and amino acids, and regulates cellular growth while inhibiting starvation responses [116–119]; TORC2 is required for growth under stress conditions, including oxidative and osmotic stresses, DNA damage conditions and nutritional stresses. The underlying molecular mechanism for the role of TORC2 in the stress response in fission yeast remains largely unknown. Recently, we (R. Weisman's group) and the Shiozaki laboratory demonstrated that TORC2 is activated in response to glucose availability [120, 121], which suggests a primitive mode of activation of TORC2 that may have developed into the insulin-dependent response of the mammalian TORC2 (mTORC2).

Disruption of TORC1 results in growth arrest, and in cells attaining physiological and morphological responses similar to those of nitrogen starved cells. Similar to mTORC1, the activation of TORC1 in response to nitrogen sufficiency leads to phosphorylation and activation of Psk1, the *S. pombe* S6K homolog, which in turn phosphorylates ribosomal protein S6 (RPS6) [122, 123]. As in mammalian cells, the phosphorylation of Rps6 is widely used as a readout for TORC1 activity in fission yeast. Yet, the relevance of Rps6 phosphorylation to the cellular functions of TORC1 is unknown. Thus, for example, replacing the serine residues that are targeted for phosphorylation with alanine residues in the two isoforms of Rps6 in *S. pombe* did not affect cellular growth [123]. The starved phenotype of *S. pombe* cells disrupted for TORC1 suggests that TORC1 mainly regulates growth, thereby regulating cell proliferation, a process that is dependent on mass accumulation. Yet without identification of the full set of TORC1 downstream effectors in fission yeast, the mechanisms by which TORC1 regulates cell growth and proliferation remain only partially understood.

While disruption of TORC1 results in growth arrest, rapamycin does not inhibit the growth of wild-type *S. pombe* cells, indicating that the core functions of TORC1 are maintained in the presence of the drug [124, 125]. Accordingly,

rapamycin does not affect the general translation in *S. pombe*, as determined by incorporation of ^{35}S-methionine into newly synthesized proteins in the presence of rapamycin [126]. However, a decrease in total protein synthesis in the presence of rapamycin occurs in tRNA modification mutant (*tit1Δ*) cells, suggesting that interference with tRNA metabolism together with TORC1 inhibition may disrupt complementary signaling to control general translation.

One of the well-known and highly conserved mechanisms to downregulate translation initiation under stress conditions is the phosphorylation of the translation-initiation factor eIF2α. In *S. pombe,* three eIF2α kinases, Gcn2, Hri1 and Hri2, contribute to phosphorylation of Ser52 of eIF2α. Of these, Gcn2 is the main kinase responsible for eIF2α phosphorylation in response to nutrient starvation, oxidative stress, DNA damage or UV light [127]. Deactivation of *S. pombe* TORC1 or treatment with rapamycin leads to phosphorylation of eIF2α in a Gcn2-dependent manner [127, 128], similar to previous findings in *S. cerevisiae* [75]. It is of interest that inactivation of TORC1 as well as rapamycin treatment should result in phosphorylation of eIF2α, since only loss of TORC1 activity results in growth arrest. The relationship between the TORC1 and Gcn2 pathways is complex. Thus, for example, TOR inactivation induces eIF2α phosphorylation only under certain conditions [123, 129], and Gcn2 becomes fully activated in response to UV-C radiation independent of TORC1 activity [127]. In response to leucine starvation, the Gcn2 pathway seems to regulate TORC1 activity rather than be affected by TORC1, indicating that Gcn2 can act both downstream and upstream of TORC1 [127]. In *S. cerevisiae,* inactivation of TOR kinases by rapamycin leads to the removal of a phosphate on Ser577 of Gcn2 via activation of phosphatases. Ser577 is not conserved in *S. pombe* or in mammalian cells. Yet the findings that TOR regulates eIF2α phosphorylation in both *S. pombe* and *S. cerevisiae*, two highly divergent yeasts, suggest that at least the wiring of the TOR and Gcn2 pathways is evolutionarily conserved, while the exact molecular mechanism differs.

6.2 Concluding Remarks

Outstanding research areas in the field include working towards a better understanding of the molecular mechanism by which TORC1 affects protein synthesis in *S. pombe*. It will also be important to understand why, in *S. pombe*, TORC1 remains largely resistant to inhibition by rapamycin with respect to protein synthesis control. Another yet unsolved question concerns the role that TORC2 may play in regulating general translational control. At present, no direct link has been suggested between TORC2 and the general control of protein synthesis.

7 Plants

In this section, we will concentrate on the features of the TOR pathway in the "green lineage" or the kingdom plantae (that we will name hereafter "plants"). Most plant TOR research has been performed in vascular plants namely in *Arabidopsis thaliana*, and in *Chlamydomonas reinhardtii*—a green unicellular alga. As much as possible, we will focus on the specific features that have allowed the tuning of the ancient unicellular TOR pathway to the "green" lifestyle such as photosynthesis, mineral nutrition and specific hormonal circuitry. For further details, the reader is invited to consult recent exhaustive reviews [130, 131].

7.1 Rapamycin and Plant TOR

The TOR pathway integrates nutrient signaling in order to process available energy towards different cellular outputs such as growth or abiotic and biotic stress adaptation. This pathway was initially described in yeast by genetic analysis of rapamycin resistance, resulting in the identification of the TOR (target of rapamycin) kinase as a central component of the pathway and of the FK506-binding-protein 12 (FKBP12) protein that, together with rapamycin and TOR, forms a ternary complex where TOR activity is inhibited [7]. Biochemical studies in vertebrate cells identified TOR as a rapamycin-binding protein, initially known as FRAP/RAFT1 [9, 10]. Rapamycin was of little help for TOR plants studies—vascular plants are resistant to growth inhibition by rapamycin. It was, in fact, the progress of genome analysis and reverse mutagenesis in the angiosperm *Arabidopsis thaliana* that allowed the characterization of the TOR kinase and associated proteins, RAPTOR and LST8 [18, 132, 133]. Interestingly, this is not the case for the entire green lineage since the unicellular green alga *Chlamydomonas reinhardtii* is susceptible to rapamycin [134]. Rapamycin susceptibility in Arabidopsis is restored by expression of yeast or mammalian FKBP12 [135–137]. Although it was originally thought that the FKBP12 protein in Arabidopsis (and other plants) lacked critical amino acids necessary for ternary complex formation, it was further shown that Arabidopsis TOR and FKBP12 proteins can interact in the presence of rapamycin and that high doses of rapamycin can inhibit plant growth and cause molecular phenotypes, such as the dephosphorylation of ribosomal S6 kinase [138]. More recently, ATP competitive inhibitors, designed to target mTOR, were found to be useful pharmacological tools and are now widely employed [139–141].

7.2 Orthology of Green TOR Pathway Components with Other Organisms

As mentioned earlier, in mammals and yeast the TOR pathway is centered around two active TOR-containing complexes, TORC1 and TORC2. A single TOR gene is present in plant genomes and encodes a protein that, in Arabidopsis, is 42 % identical to the human protein. Components of TORC1 (TOR itself, RAPTOR and LST8) are present in the genome of all groups of photosynthetic eukaryotes [142]. Direct interaction between TOR and RAPTOR and LST8 was demonstrated in *Arabidopsis thaliana* and *Chlamydomonas reinhardtii* [133, 143–145]. The situation is less clear for TORC2, since canonical components of yeast and animal TORC2 (Avo1/mSIN1 and Avo3/RICTOR) are absent from plant genomes. Clearly, biochemical studies are still lacking to decipher the composition of TOR complexes in the green lineage. Among TOR targets, although the precise role of RPS6 phosphorylation is no better understood in plants than it is in other organisms, the S6K/RPS6 axis is conserved, as is the TAP42 (46 in plants)/PP2A signaling module [138, 139, 145–147]. Orthologs of either mammalian or yeast eIF4E-binding proteins are not found in plant genomes, although proteins bearing eIF4E-binding motifs exist, but their role in translation and their possible interactions with TOR are still unknown (Ryabova, personal communication; Robaglia, unpublished). New TOR pathway targets have been identified in plants, such as E2F transcription factors [148] and Re-Initiation Supporting Protein (RISP) [139]. Concerning upstream components that may convey environmental perception to TOR, plant genomes do not contain orthologs of the Tuberous Sclerosis Complex (TSC1 and TSC2), RHEB, or class I-II PI3Ks [22, 149], but they do contain proteins homologs of the class III PI3K/VPS34 [130].

7.3 General Functions of the Green TOR Pathway

Overall, functional studies have shown that, similarly to other eukaryotes, TOR is generally required for growth control, cell cycle and global translation activity [132, 148, 150, 151], energy signaling and autophagy [136, 152]. Several plant-specific processes also appear to require a functional TOR pathway such as cell wall organization, light adaptation and starch synthesis [133, 137, 153].

7.4 Plant TOR and Plant Hormones

Plants harbor a set of small hormonal molecules that orchestrate the systemic signaling of growth control, stress and pathogen responses. Auxins, cytokinins, brassinosteroids and gibberellins represent the most prominent growth hormones,

while abscisic acid (ABA), ethylene, jasmonate and salicylic acid (SA) represent stress hormones. These hormonal pathways are linked by complex signaling circuitry. Obviously, given the basal role of TOR in the interplay between growth and stress adaptation, a key issue is to understand how its signaling pathway is wired to hormonal control. Global transcriptome analysis has shown that TOR pathway inhibition leads to activation of genes linked to stress responses through the ABA, ethylene, jasmonate and salicylic acid and as well as activation of genes linked to the repression of growth-promoting pathways controlled by auxin, cytokinin, brassinosteroid and gibberellin [133, 154]. ABA synthesis and response is reportedly repressed in conditions of TOR inactivation; Plants that overexpress TAP46 also display a higher sensitivity to ABA [155, 156], suggesting that TOR activity is required to mount the stress response through ABA synthesis. Recently, leaf chlorosis caused by the TOR inhibitor AZD8055 was found to be counteracted by mutations in the ABI4 (ABA Insensitive 4) gene [141]. ABI4 is activated by ABA and sugars to repress photosynthetic nuclear genes, such as the antenna protein that captures light and conveys photochemical energy to the core reaction center. Although the position of ABI4 relative to the TOR pathway is still unknown, its identification establishes a possible molecular link between TOR and ABA signaling.

Auxin and cytokinin activate S6K and increase S6 phosphorylation [157]. A direct connection between the growth-promoting auxin and TOR has been established by the increased TOR association with polysomes and S6K phosphorylation upon auxin addition. This promotes the translation of mRNAs containing upstream open reading frames (uORFs) such as a set of auxin response factors (ARFs); conversely, TOR inhibition blocks gravitropism, a well-known auxin-mediated response [151]. TOR also phosphorylates E2Fa transcription factors that are positively regulated by auxin and required for cell cycle entry [148]. Taken together, this suggests that auxin and TOR might be part of a positive feedback loop that enhances auxin action to promote growth. We speculate that as-yet unrecognized components involved in relaying the auxin signal to TOR await identification.

7.5 Plant TOR, Light and Nutrients

Plants are carbon autotrophs, and light is the essential energy supply that converts carbon intake into glucose through photosynthesis. Glucose activates the TOR pathway [148], and TOR inactive plants fail to grow in response to light stimulation [136]. TOR inhibition causes leaf chlorosis, represses many photosynthesis-related genes, and leads to the accumulation of lipids and reserve carbohydrates, such as starch [133, 153, 154, 158]. Direct evidence of a link between TOR and light perception comes from the phenotype of the LST8 KO mutant that displays a daylength-dependent chlorotic phenotype [133] (Fig. 3). Excess light also represses RPS6 phosphorylation and mRNA translation [159]. Besides glucose, the additional molecular actors that link light perception to the TOR pathway remain

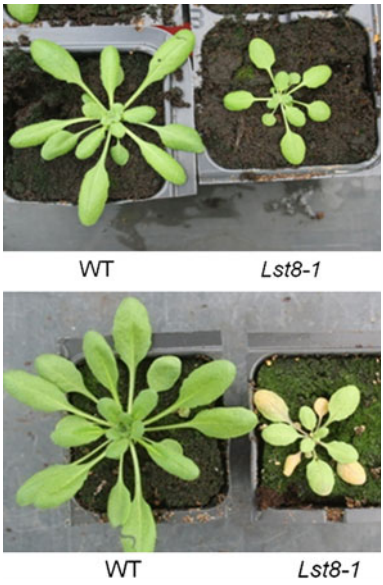


Fig. 3 Day-length adaptation: a plant-specific function of the TOR pathway. Arabidopsis *lst8-1* mutants in short days (8h, *left*) and after transfer to long days (16h, *right*). Adapted from Ref. [133]

unidentified. The cell wall, which is the primary sink for photosynthetic carbon, is also affected by TOR inactivation [137]. Soil nitrogen, as is the case of light, is an essential nutrient for plant growth. Interestingly, TOR inactivation hinders the growth-promoting effect of nitrogen. Similarly, nitrate reductase activity is enhanced by TAP46 overexpression [147].

7.6 Cellular Functions: Translation, Endomembranes, Cell Growth and Innate Immunity

TOR inactivation leads to a decrease in polysome abundance [132, 139]. However, this effect is not as strong in plants as in animals or yeast cells, and, until now, it is not known whether this is linked to the possible absence of eIF4E-binding proteins in plants. Activated TOR binds polysomes where it interacts with plant-specific protein RISP and ribosomal S6 kinase to recruit the eIF3 complex on multicistronic mRNA [139, 151]. The existence of connections, if any, between green TOR and translation initiation is presently unknown. Ribosomal RNA expression is also positively correlated to TOR expression [150]. Inhibition of TOR activity in Arabidopsis decreases RPS6 phosphorylation at Ser240, which is also negatively affected by multiple stresses such as anoxia, heat stress, light modulation and high CO_2 [160–162].

In plants and microalgae, TOR regulates autophagy [152, 163], and in Chlamydomonas, rapamycin induces the phosphorylation of the endoplasmic reticulum chaperone BiP, which is linked to ER stress [164]. TOR controls cell

growth and division in the root and shoot apical meristematic zones where stem cells generate new organs [132, 136, 140, 148]. In plants, dividing cells leaving the meristem become progressively quiescent and undergo cell expansion and polyploidization through a process termed endo-reduplication. E2FA transcription factor, mediates both cell size entry and polyploidization, depending on its location, and is a target for TOR-mediated phosphorylation [148]. Downstream of TOR, S6K is known to control the cell cycle and DNA synthesis through RBR1 (retinoblastoma-related protein 1) interaction [165]. Although TOR-independent S6K functions cannot be excluded, the TOR/S6K axis likely signals together in the developmental switch between plant cell proliferation and cell expansion during organogenesis. Careful examination of the various TOR activation patterns in the meristem and its margins is necessary to better understand these cellular processes.

Growth and immunity are known to be antagonistic [166], and TOR is involved in both pathways. Indeed, TOR inhibition has been shown to prevent plant growth while simultaneously activating the defense hormone salicylic acid-dependent pathway [133, 154]. Given its juxtaposed functions and its involvement in reserve mobilization, TOR represents an interesting target for plant pathogens. However, TOR modulation during bacterial or fungal pathogen attack has not yet been described. Several viruses, that hijack the translation machinery during infection, require an intact TOR pathway. The cauliflower mosaic virus transactivator protein (TAV) activates TOR to promote translation of the viral multicistronic mRNA [139]. Similarly, members of the potyviruses, the largest group of plant RNA viruses, also require an active TOR, since TOR inhibition leads to virus resistance and can even cure infected plants [167].

7.7 Concluding Remarks

The last 10 years of research has allowed us to draw a general picture of the green TOR pathway, but much has yet to be done to understand the molecular components involved in its different functions. Although it is probable that many outputs are similar to those of other eukaryotic organisms, the wiring is probably different as evidenced by the absence of TOR-dependent eIF4E-binding proteins in plants. Future areas of research will include mapping the connections of TOR with the plant-specific nutrition processes, particularly nitrogen assimilation and photosynthesis, as well as the interconnection of TOR with plant hormonal circuitry, innate immune responses, and other regulatory cell integrity pathways inherited through endosymbiosis, such as the stringent response. Importantly, given the role of TOR as a hub in energy signaling, growth control and reserve accumulation, it is not impossible that this new knowledge will find applications in molecular breeding of plants or microalgae for agriculture and bioenergy production.

8 *Drosophila melanogaster*

8.1 Drosophila melanogaster*: A Versatile Model Organism*

Drosophila melanogaster is a versatile model organism that enables the easy and rapid characterization of biochemical signaling pathways. *Drosophila melanogaster* has a short life cycle (compared to vertebrates) and reproduces in large numbers; it can be easily genetically manipulated (in various ways), and, from an economical standpoint, it is relatively inexpensive to maintain. Studies employing this model organism have made a vast contribution to the identification and characterization of various components of the TOR pathway. In this section, we review the contributions of *Drosophila melanogaster* research to our present understanding of the TOR pathway.

8.2 Drosophila melanogaster*: Discovery of dTOR*

Drosophila melanogaster possesses a single copy of the *dTOR* gene (FlyBase.org annotated number CG5092) encoding the dTOR protein [168, 169]. dTOR was originally identified by two research groups led by Ernst Hafen [168] and Thomas Neufeld [169], reported in two elegant papers in *Genes and Development* in 2000.

8.3 dTOR Regulates a Multitude of Cellular Functions

Earlier studies in the budding yeast *Saccharomyces cerevisiae* had shown that TOR plays a pivotal role in the control of cell growth. But, in fact, this phenotypic effect of TOR in cell growth (first observed in budding yeast) extends well beyond this unicellular model organism. The importance of TOR in the control of cell growth (a recurrent theme throughout this chapter) is conspicuous throughout eukaryotic evolution. TOR possesses the universally conserved role of integrating both nutritional as well as organism-specific endocrine cues with signaling pathways to ensure that growth is appropriate for a specific nutritional status [149, 170–173]. This is also the case in *Drosophila melanogaster*, where dTOR functions as a master regulator of cell and organismal growth [170]. In the following section, we will focus on the contribution of this model organism to our understanding of dTOR's role in the control of organismal growth [168, 169], fertility [174, 175] and survival [176, 177].

8.4 dTOR and dS6K Control Cell and Organismal Growth

The first TOR-dependent phenotype to be identified was its effect on cell and organismal growth [168, 169]. Genetic manipulation of *dTOR* revealed that the encoded protein plays a pivotal role in the control of cell growth [168, 169]. For instance, Hafen and colleagues showed that induction of tissue-specific mutant *dTOR* clones ($dTOR^{2L1}$ and $dTOR^{2L19}$) produced adult flies with reduced head size [168]. Mutant *dTOR* flies also displayed reduced growth at developmental (pupal and larval) stages [168]. Similar findings were reported by Neufeld and colleagues [169]; they observed that P-element insertions within the *dTOR* gene yielded a $dTOR^{\Delta P}$ mutant (with disrupted dTOR function) lead to defects in larval development, resulting in larva hatching with as low as 24 % of the mass of their wild-type counterparts [169]. Organismal size is the product of cell mass (with increases referred to as cell growth) and cell number (with increases referred to as cell proliferation). To determine whether the reduced size of the $dTOR^{\Delta P}$ mutant resulted from a decrease in cell size or cell number, Zhang et al. [169] analyzed the effect of disrupting dTOR function in various tissues by examining the size and number of the cells of the adult wing and the salivary gland. Both the bristles and the epithelial cells of the wing blade appeared smaller, supporting a role for dTOR in the control of cell size. Loss of dTOR also affected cell proliferation with $dTOR^{\Delta P}$ salivary glands undergoing fewer rounds of replication than wild-type cells [169]. Taken together, these data suggested that the absence of *dTOR* leads to both reduced growth and proliferative capacity. As observed in other organisms, dTOR signals downstream via two main signaling branches: the dTOR/d4E-BP and dTOR/dS6K branches (Fig. 4). Consistent with the fact that dTOR is a major controller of cell and organismal growth in the fruit fly, mutations in various upstream and downstream signaling components of the dTOR pathway phenocopy the mutant *dTOR* growth defect. For example, dS6K (a direct substrate of dTOR) is a central regulator of cell growth in *Drosophila melanogaster* [178]. P-element-induced mutagenesis of the *Drosophila melanogaster* single S6K gene (*dS6K*) gives rise to miniature flies. $dS6K^{1-1}$ homozygous females are approximately 40 % of the size of their wild-type counterparts [178]. Comparison of cells of the wing (as well as ommatidia) of dS6K mutant flies versus wild-type flies revealed that dS6K controls cell size but not cell number [178], consistent with findings in mammalian S6K1/2 double knockouts [179]. Collectively, these studies demonstrate that the dTOR/dS6K signaling branch plays an important role in the control of organismal growth, consistent with earlier findings in yeast. The ability of organisms to grow, irrespective of their uni- or multicellularity, indicates that accumulation of cell mass is a well-conserved primordial cellular function. It

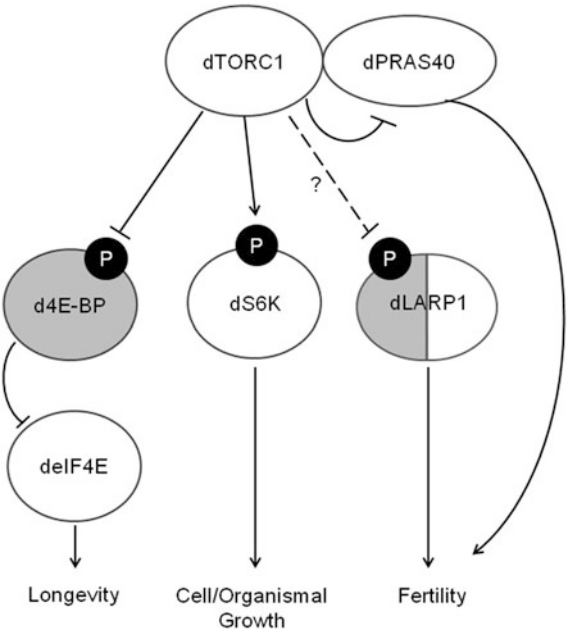


Fig. 4 Model depicting signaling pathways downstream of dTORC1 involved in the control of mRNA translation. Proteins shown in *white* activate mRNA translation. Proteins shown in *gray* inhibit mRNA translation. *Dashed lines* indicate poorly characterized pathways. See main text for details

therefore follows that the cellular circuitry involved in cell growth could also be conserved.[5] The observation that TOR's function in cell growth is conserved between yeast and flies (two phylogenetically distant organisms) suggests that this is indeed the case. Yeasts and flies differ in one major aspect; only the latter has an insulin signaling system. This system has evolved only with metazoans. Metazoan evolution was accompanied by the emergence of novel proteins that function upstream of TORC1. Notably, many of the proteins that function upstream of TORC1 are also involved in the control of cell growth. Studies in *Drosophila melanogaster* have further demonstrated that many of the upstream signaling components of dTORC1 [e.g., tuberous sclerosis complex 1 and 2 (dTSC1/dTSC2); ras-homolog enriched in the brain (dRHEB)] [180], as well as components in parallel pathways[6] (dPI3K and dPTEN) [181], play important roles in the control of cell growth. This section/chapter focuses primarily on the signaling *downstream* of

[5]One cannot rule out, however, functional adaptation of alternative signaling pathways in the control of cell growth. A shared cellular function does not necessitate a common signaling mechanism. Convergent evolution allows for different signaling mechanisms fulfilling a shared cellular role in distinct species.

[6]Metazoans have evolved several cascades that control cellular and organismal growth. One major growth pathway is the insulin/mTORC1 signaling pathway, which in mammals follows the sequence: insulin, insulin receptor, PI3K, Akt, TSC1/TSC2/RHEB and mTORC1. In *Drosophila melanogaster* the signaling mode does not follow this linear sequence. In insects, dPI3 K and dTORC1 pathways do not crosstalk—instead, they appear to function in parallel to control cellular and organismal growth in response to environmental, endocrinological and nutritional inputs [181].

dTORC1; for this reason, the analysis of these studies will not be covered here. The reader is referred to excellent reviews on the subject [171, 180].

8.5 dTOR and d4E-BP Control Lifespan

Besides regulating growth via dS6K, dTOR has been implicated in the control of various other organismal functions in the fly through other downstream substrates (Fig. 4). In this section, we will briefly discuss the role of one such dTOR target (specifically d4E-BP) in the control of the *Drosophila melanogaster* lifespan [176, 177, 182]. In mammals, mTOR (namely mTORC1) controls the phosphorylation and activity of the family of translation initiation repressors 4E-BPs (discussed in detail in Sect. 10). While the mammalian genome encodes three *4E-BP* genes known as *Eif4ebp1, Eif4ebp2* and *Eif4ebp3* [183–186] which give rise to three proteins known as 4E-BP1, 4E-BP2 and 4E-BP3, the fly genome encodes a single *4E-BP* gene, which gives rise to a single 4E-BP protein known as d4E-BP [187]. A study by Tettweiler et al. [176] in 2005 has shown that the lifespan of *Drosophila melanogaster* is controlled by d4E-BP. Specifically, the median lifespan of $d4E\text{-}BP^{null}$ male flies is reduced by 25 % compared with revertant males; the same null mutation has a similar effect in females, indicating that this phenotype is gender-independent [176]. *Drosophila melanogaster* eggs/larvae depend on protein-rich nutrition for survival. Starvation of eggs/larvae results in differential viability of $d4E\text{-}BP^{null}$ and revertant larvae: $d4E\text{-}BP^{null}$ larvae die faster than their revertant counterparts under starvation [176]. Similarly, $d4E\text{-}BP^{null}$ adult flies also exhibited reduced survival compared to revertant flies upon exposure to oxidative stress (H_2O_2) [176]. Mechanistically, the increased longevity of wild-type d4E-BP flies appears to be dependent on their interaction with deIF4E, as suggested by the observation that d4E-BP(Y54A, M59A) larvae and flies that are defective for deIF4E-binding are similarly prone to accelerated death under nutrient deprivation and oxidative stress conditions [176]. Taken together, these data indicate that the dTOR/d4E-BP/deIF4E signaling branch plays a critical role in the control of *Drosophila melanogaster* longevity in adverse nutritional and oxidative stress conditions. This is consistent with findings in mammals, where rapamycin-fed mice exhibit extended lifespans [188].

8.6 dPRAS40 and dLARP Govern Fly Fertility

In addition to governing growth and survival in metazoans, a number of studies [189, 190] provide evidence in support of a role for the dTOR pathway in the control of germline maintenance and fertility. For instance, it has been observed that mutations in *dTOR* result in pupal lethality and sterility in both male and female flies [191]. Both male and female dTOR mutant flies are sterile [191]. Mutant male

flies displayed abnormal testes while female mutant flies showed arrested development and increased cell death in reproductive tissues, indicating that dTOR is required for proper male and female sexual development. Mutations in *dTOR* have also been reported to reduce proliferation, growth and survival of germline stem cells (GSCs) and follicular stem cells (FSCs) in the ovary [192]. Notably, these effects appear to be independent of d4E-BP [192]. The fertility of *Drosophila melanogaster* is, therefore, likely governed by other effectors of dTOR. Several candidate proteins fit the bill, including: (1) dPRAS40 (*Drosophila melanogaster* proline-rich Akt substrate 40 kDa; originally identified as a component [193–198] and direct substrate [195–197] of mTORC1 in mammals) and (2) dLARP (*Drosophila melanogaster* lupus autoantigen (La) related protein; recently identified as a target of mTORC1 in mammals [199]).

dPRAS40. A recent study by Pallares-Cartes et al. [175] has shown that loss-of-function of dPRAS40 does not affect the growth or proliferation of somatic cells, but does impair ovary development. *dPRAS40* loss of function was also able to partially rescue the sterility phenotype in Chico (the equivalent of the insulin receptor in flies) mutant female flies [175]. Thus, dPRAS40 coordinates the effects of dTOR in the control of *Drosophila melanogaster* fertility.

dLARP. dLARP, the LARP1 homolog in *Drosophila melanogaster*, shares 50 % similarity with its human homolog in humans (Table 1). However, at present, there is no evidence of a genetic or physical interaction between dTOR and dLARP.[7] Interestingly, mutations in *dLARP* phenocopy the male and female sterility phenotype associated with *dTOR* mutations.[8] In male flies, mutations in the *dLARP* gene resulted in a number of meiosis defects, such as abnormal spindle pole formation, chromatid segregation and cytokinesis in male meiosis that impair spermatogenesis and reduce fertility [174]. These findings corroborate a potential role for dLARP in the control of fly sexual development downstream of dTORC1.

8.7 *Conservation of the TORC1 Pathway Components Across Insects*

Phylogenetic analyses of 12 sequenced species within the genus *Drosophila* (FlyBase.org; [201–203]), namely *melanogaster*, *grimshawi*, *mojavensis*, *virilis*, *willinstoni*, *persimilis*, *pseudoscura*, *ananassae*, *erecta*, *yakuba*, *sechellia* and *simulans*, indicate that all contain a highly conserved single-copy *TOR* ortholog (∼91 % identity with *D. melanogaster dTOR*). Outside this genus, *dTOR* shares ∼71 % identity with *TOR* from other Dipteran insects, including malaria

[7]The fact that human LARP1 and insect dLARP share a common binding partner, PABP (poly-(A) binding protein), suggests that these proteins carry functional homology across eukaryotes [174, 199, 200].

[8]It does not, however, affect fly viability.

mosquitoes (i.e., *Anopheles gambiae* and *Aedes aegypti*), sandflies (i.e., *Lutzomyia longipalpis* and *Phlebotomus papatasi*), the Hessian fly *Mayetolia destructor* and the Tsetse fly *Glossina morsitans*. All species within the order Diptera contain single-copy *TOR* genes. *dTOR* shares ~68 % identity with single-copy *TOR* genes from other orders of insects sequenced to date, including Strepsiptera (twisted-wing parasites), Coleoptera (beetles), Hymenoptera (ants, wasps and bees), Homoptera (aphids), Hemiptera (kissing bugs), Phthiraptera (body louses), Isoptera (termites) and Lepidoptera (butterflies and moths; two copies of *TOR*). Within the non-insect arthropods, *dTOR* shares ~29 % identity with, for example, the aquatic crustacean *Daphnia pulex* (water flea), the centipede *Strigamia maritima*, and the mites *Metaseiulus occidentalis* and *Ixodes scapularis* (all with single-copy genes) (FlyBase.org; [201–203]).

As alluded to earlier in this chapter, TOR exerts its cellular functions by forming two structurally and functionally distinct protein complexes termed TOR complex 1 (TORC1) and TOR complex 2 (TORC2). In *Drosophila melanogaster* (and other metazoans) the same TOR enzyme nucleates the two complexes. As mentioned in Sects. 4–6 of this chapter, the situation is different in lower eukaryotes that encode two distinct TOR proteins, "roughly" one for each complex. Phylogenetic analyses across eukaryotes have shown that the two TOR complexes, namely TORC1 (TOR, LST8 and RAPTOR) and TORC2 (TOR, LST8, RICTOR, and SIN1), originated before the last common ancestor of eukaryotes to which new proteins have been added during metazoan evolution [149, 172, 173]. The components of both TOR complexes and signaling pathways, namely RAPTOR, RICTOR, LST8 and PRAS40 (known as Lobe in insects), are conserved in *Drosophila melanogaster*. The exception would be DEPTOR and PROTOR, which appear to be missing in the fly [149, 172, 173]. Moreover, the core upstream (discussed only briefly in this book chapter) and downstream components (S6Ks, 4E-BPs, PRAS40 and LARP1) of TORC1 and (Akt, PKC and SGK) of TORC2 are also highly conserved in *Drosophila melanogaster* [149, 172, 173, 201–203]. Translation factors downstream of dTORC1/d4E-BP (such as eIF4E and eIF4G) are also highly conserved in *Drosophila melanogaster*. It is intriguing that within the genus *Drosophila* there is a striking diversity of multiple ortholog genes encoding eIF4E isoforms [204, 205]. In sharp contrast, insects outside the genus *Drosophila* contain only a single *eIF4E* gene, related to *Drosophila melanogaster* eIF4E-1 [204, 205]. *Drosophila* eIF4E-1 is regulated by the dTORC1/d4E-BP pathway [176, 206]. Since most *Drosophila* eIF4E isoforms, other than deIF4E-1, bind d4E-BP [204], they might also be regulated by the dTORC1/d4E-BP pathway, but no experimental evidence supporting this notion has thus far been reported. deIF4E-3, a testis-specific *Drosophila* isoform that is essential for spermatogenesis, is the exception to the rule in that it does not bind d4E-BP [207]. Two eIF4Gs have been reported in *Drosophila*, namely eIF4G [208] and Off-schedule [209, 210],[9] both of which are

[9]In contrast to eIF4E, both eIF4G proteins are single-copy genes within the *Drosophila* genus (FlyBase.org; [201–203]).

functional homologs of mammalian eIF4G. Whether any of these proteins are subject to the control of the dTORC1/d4E-BP is not known.

8.8 Concluding Remarks

TOR is widely expressed in all eukaryotic organisms [19, 52]. TOR is found in two physically distinct and independently regulated multi-protein complexes known as TORC1 and TORC2 (or mTORC1 and mTORC2 in mammals, see Sect. 10), as carefully delineated by studies in *Saccharomyces cerevisiae* and in mammalian systems. The TORC1 and TORC2 protein complexes are also represented in the Drosophila genus. In this section, we have reviewed the work that led to the original discovery of dTOR in flies and its pivotal importance in defining the role of dTOR in organismal growth in response to nutritional cues. Herein, we have also described some of the key signaling events downstream of dTOR (specifically dTORC1) that feed onto the control of mRNA translation. *Drosophila melanogaster* is among the most versatile model organisms for the study of cellular signaling in biology. This has been, unquestionably, the case for the elucidation of the TOR pathway in metazoans.

9 *Caenorhabditis elegans*

9.1 Discovery of CeTOR

Sequences encoding CeTOR were first identified on cosmid B0261 (on chromosome I) by C. Spycher in F. Muller's laboratory. Ann Rose had previously identified the essential genes in this region as those that were rescued by the free chromosome 1 fragment, sDp2. Ms. Spycher showed that the lethal mutations at the locus *let-363* covered by sDp2 could be rescued by microinjection of the cosmid B0261 and that several of these *let-363* mutant alleles harbored nonsense mutations in the TOR coding sequences. Subsequent DNA sequence analysis revealed that the three *let-363* mutant alleles, *h111*, *h131* and *h114*, encoded Q^{104}Amber, Q^{906}Amber and Q^{2398}Amber (CAG→TAG stop mutations) and identified a fourth allele, *h98*, as a G to A mutation that abolishes the 5′ splicing donor site exon 31, leading to a frameshift and early termination [17]. Antisense-injection experiments with RNA produced from the cDNA clone yk31h11 indicated that *Ce-tor* is an essential gene.

The *Ce-tor* mRNA is SL-1 trans-spliced, approximately 8.5 kb in length, and includes 32 exons. Expression of a nuclear-localized CeTOR::GFP fusion protein is visualized in comma-stage embryos, in all major tissues and organs and in most, if not all, cells through to adulthood. Maternal provision of *tor* RNA is indicated by positive in situ hybridization of whole-mount embryos starting at the single-cell

stage. The CeTOR polypeptide has a length of 2697 amino acids and is 35 % identical to human TOR; the most conserved region is the kinase domain, which shares approximately 72 % amino acid identity with human TOR. FAT, FATC and FRB domains are also evident; however, despite the presence of a typical FRB domain in CeTOR and the conservation of key residues in CeFKBP12 (that mediate rapamycin binding and binding of the FKBP12/rapamycin complex to TOR), TOR signaling in worms is highly resistant to rapamycin. Thus, while half-maximal rapamycin inhibition of S6 kinase (S6K1B) activity in mammalian cells is evident at 2 nM, worms incubated on plates containing 100 μM rapamycin and fed ^{35}S-labeled *E. coli* show only partial inhibition of ^{35}S incorporation into protein and modest activation of transcriptional responses strongly elicited by RNAi against components of the CeTOR pathway [211].

9.2 The Phenotype of CeTOR Deficiency

The *let-363* mutant worms also carry *dpy-5* (cuticule procollagen) and *unc-38* (nicotinic acetylcholine receptor alpha subunit) alleles, which are covered by the *sDp2* free duplication, so that the TOR phenotype is somewhat complicated by the features contributed by the linked mutations, especially the *dpy-5*. Fortunately, all of the readily definable major aspects of the *let-363* mutation are recapitulated in the F1 generation of worms fed bacteria expressing *Ce-tor* RNAi. Worms carrying the *let-363* mutant allele *h111* exhibit significantly slower development and ultimately arrest at late L3; the gonads are significantly smaller than those in *unc-38 dpy-5* control animals of the same stage, whereas the vulva is slightly more advanced; overall body length is comparable to control. The most striking changes are seen in the intestinal lining cells, which become prominent from early L3. A cohort of auto-fluorescent vesicles (Fig. 5, left) that take up the lysosomal dye Neutral Red (mixed with the *E. coli*) become numerous and progressively enlarged. These auto-fluorescent vesicles are not lysosomes, inasmuch as they are unaffected by RNAi against *lmp-1* (a homolog of LAMP and CD68 lysosomal membrane proteins); rather they are eliminated by deletion of the RAB32/38 homolog *glo-1* (Fig. 5, right) and so are lysosome-related organelles (LROs) [212–214]. In contrast to this initial accumulation and enlargement of LROs, RNAi against *Ce-tor* mRNA causes the small non-fluorescent vesicles that normally occupy a large fraction of the cytoplasm to diminish progressively in abundance. By about 4–5 days post egg laying, the LROs also start to diminish in size and number, and over the next 24–48 h the overall cytoplasmic volume of the intestinal cells is radically diminished; the intestinal lumen, normally a narrow strip, is widened dramatically (Figs. 5 and 6). Although now arrested in their development, the mutant worms continue ingestion and pharyngeal pumping for several days, so that the intestinal lumen becomes filled with clumps containing whole and partially digested bacteria. It is to be emphasized that the *let-363* mutant worms do not proceed through a dauer-like arrest, and that their arrest phenotype is distinct from dauer-arrested worms in several respects. As

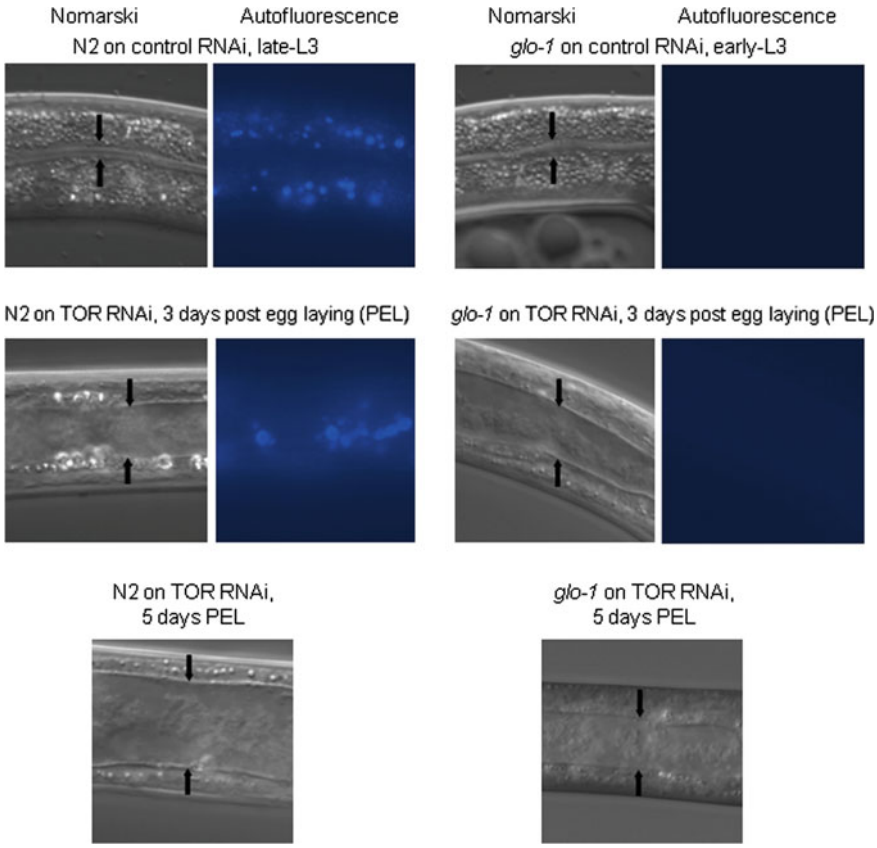


Fig. 5 Nomarski and autofluorescence micrographs of *Caenorhabditis elegans* depleted of TOR by RNAi

mentioned, there is no detectable constriction in the pharynx of *unc-38 let-363*(*h111*) *dpy-5* animals in contrast to the constricted and non-pumping pharynx of the *unc-38 dpy-5* dauers. The *unc-38 let-363*(*h111*) *dpy-5* animals never exhibit dauer-characteristic alae, prominent in *unc-38 dpy-5* dauers, whereas the latter do not have an enlarged intestinal lumen; their intestinal cells remain packed with small non-refractile intestinal vesicles and few and weakly auto-fluorescent LROs. The only characteristic shared with dauer larvae is the presence of enlarged, lipid-laden hypodermal vesicles. Starvation of wild-type worms (at the L3 stage, after the L1 dauer decision point) does cause a developmental arrest but without concomitant intestinal atrophy, and, in contrast to the lethality of *let-363*, development resumes if starved worms are refed as late as after 12 days of starvation.

A phenotype indistinguishable from that seen in the *let-363* mutant worms is reproduced in the F1 generation of normal worms by feeding bacteria producing *let-363* RNAi to P0 worms starting in early L4. These grow to normal adults; however, their progeny exhibit slowed development, late L3 arrest, gonadal degeneration,

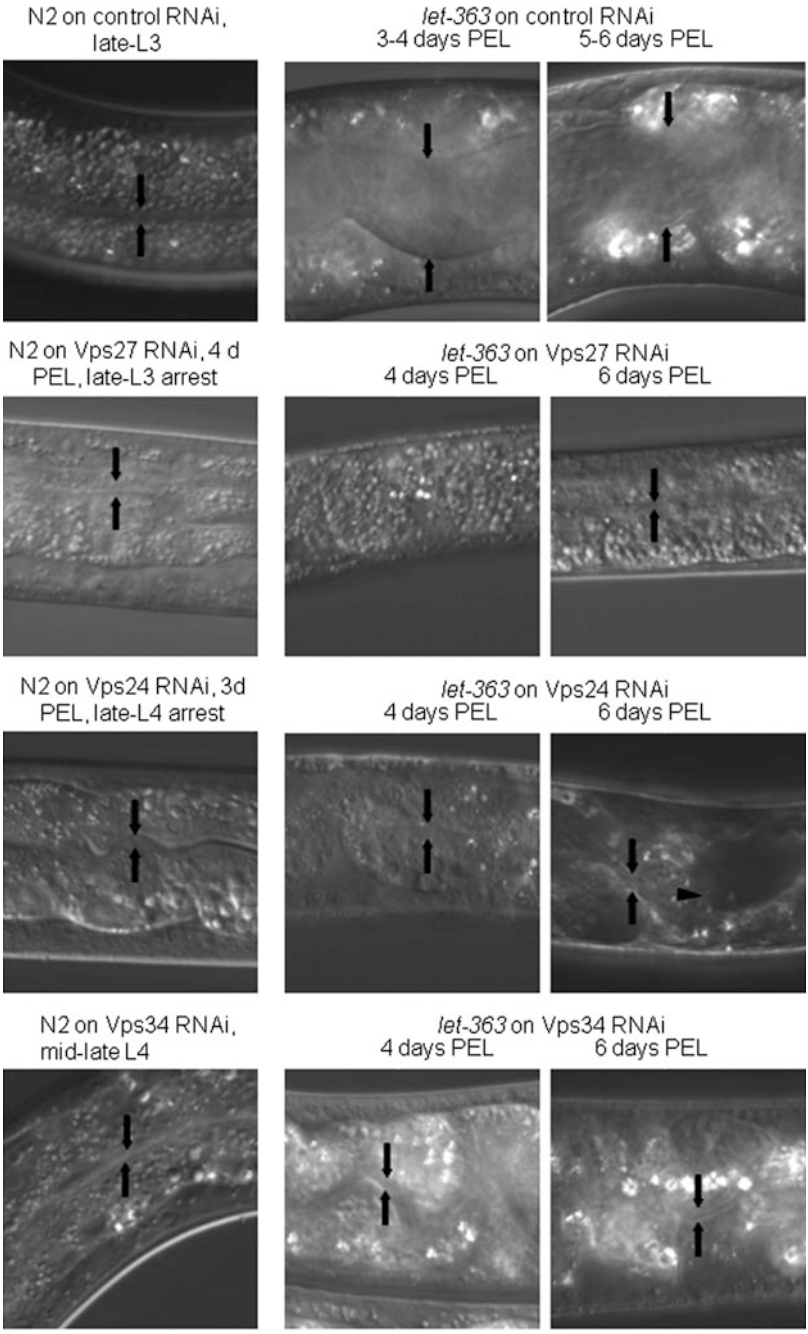


Fig. 6 Nomarski micrographs of *Caenorhabditis elegans* depleted of Vps27, Vps24 and Vps34 by RNAi

hypodermal lipid accumulation and the intestinal phenotype of the *let-363* mutant worms. The major phenotypes of *Ce-tor* deficiency are attributable to depletion of CeTOR complex 1, inasmuch as late L3 developmental arrest, gonadal degeneration, hypodermal lipid accumulation and enlarged intestinal LROs followed by the disappearance of intestinal cytoplasm are also seen with RNAi directed against *Ce-raptor/daf-15* [42, 215].

The striking intestinal phenotype caused by CeTOR or CeRAPTOR depletion, i.e., a transient increase in intestinal LROs followed by marked intestinal atrophy, is largely phenocopied after treating with RNAi directed against the worm homologs of translational initiation factors eIF4G/*ifg-1*, eIF2α/*Y37E3.10* and eIF2β/*iftb-1*; thus, a marked inhibition of overall protein synthesis is likely the underlying mechanism. The CeTORC1 target(s) mediating the regulation of overall mRNA translation is unknown; the *C. elegans* genome contains five eIF4E homologs but no homologs of the eIF4E-binding proteins. Moreover, RNAi against various TOR effectors identified in other organisms, such as the homologs of S6K1B or the Tap42/Tip41/Sit4, results in minor phenotypes [17]. The process responsible for the disappearance of the intestinal cytosolic contents when TOR is inhibited involves non-autophagic trafficking to the lysosome (Long, unpublished). While depletion of *C. elegans* homologs of several yeast Vps genes can partially (Vps34[PI-3K], Vp24 [ESCRT-II], Vps11,16,18,39[ESCRT-III/HOPS]) or completely (Vps27/Hrs [ESCRT-0]) block the intestinal cytoplasmic atrophy (Fig. 6), RNAi against the *C. elegans* homologs of the yeast autophagy genes Apg3, 7, 8, 9 and 12 is incapable of preventing the loss of intestinal cytoplasm. Although enlargement of the intestinal LRO compartment suggests their participation, RNAi against the *glo-1* Rab GTPase selectively eliminates LROs without ameliorating the loss of intestinal cytoplasm (Fig. 6, right). With regards to the proteasome, RNAi directed against many of its components (homologs of S6b, S7, S8, S10b, Rpn2, 6, 8 and 11) results in embryonic lethality in the F1 generation, but the small number of treated homozygous mutant CeTOR worms that do progress to L3 show no block of intestinal atrophy. Similarly, RNAi directed against proteasome components *that do not* cause embryonic lethality also fail to diminish the atrophy.

9.3 CeTOR and Lifespan

Developmental arrest and intestinal atrophy cannot be elicited by the introduction of *Ce-tor* RNAi into larvae, probably because maternal RNA provides sufficient TOR polypeptide to enable development. Nevertheless, several relevant phenotypes can be elicited by *Ce-tor* RNAi fed to late larval and early adult worms, such as increased autophagy, reduced protein synthesis, increased stress tolerance and, importantly, enhanced longevity. Subsequent to the initial description of *Ce-tor* deficiency, the majority of CeTOR literature has been in pursuit of the observation that TOR deficiency prolongs *C. elegans'* lifespan, as first reported by Vellai et al. [216]. This work, which employed null mutants, partial LOF mutants and RNAi, has sought to identify

the CeTOR targets/effectors that underlie this phenotype as well as the extent to which CeTOR contributes to the lifespan-prolonging effects of other interventions, especially nutrient restriction, deficient insulin-like signaling and germline inactivation. Wherever the literature presents conflicting conclusions regarding the importance of a specific gene to a response, it should be recalled that the efficacy of RNAi against a target mRNA can be variable, depending on the RNAi sequences employed as well as the feeding method used; moreover, additive effects between partial LOF mutants and RNAi, although compatible with the operation of independent pathways, are also compatible with additive inhibitions within the same pathway.

9.4 CeTOR Complex 1 and Lifespan

As in other organisms, prolongation of *C. elegans'* lifespan is primarily a result of TOR complex 1 deficiency, inasmuch as it occurs with RNAi directed against *Ce-raptor/daf-15* mRNA and in *Ce-raptor/daf-15* heterozygotes [215]. Depletion or interference with the small GTPase CeRagA [211, 217] and depletion of CeRagC [211] and CeRHEB [211, 218] also extends the lifespan. These small GTPases are known to operate in higher metazoans on the pathway that couples amino acid sufficiency to mTORC1 [219]; the RagA/B-RagC/D heterodimer binds RAPTOR in an amino acid-dependent manner and docks TORC1 on the Ragulator complex, which is located on the cytoplasmic surface of the lysosome. There TORC1 is able to interact with the v-ATPase and, most importantly, GTP-charged RHEB GTPase. The latter binds directly to the catalytic domain of TOR in Complex 1 serving as the proximate activator of TORC1 signaling. In *Drosophila* and higher metazoans, the state of RHEB GTP charging is regulated downstream of insulin/IGF receptors by the PI-3K/Akt pathway, primarily by negative regulation of the tuberous sclerosis complex (TSC), a heterotrimer of TSC1/hamartin, TSC2/tuberin and TBC1D7, which is a GTPase activator for RHEB [220]. Akt catalyzes the phosphorylation of TSC2, which displaces the TSC from the lysosomal surface, interdicting its proximity to RHEB. The tuberous sclerosis complex serves as a node through which numerous upstream inputs regulate TORC1 activity; thus, TORC1 activation requires two independent inputs: nutrients via the Rag pathway and IIS regulation of RHEB-GTP via TSC. Surprisingly, the *C. elegans* genome does not contain homologs of TSC1 or TSC2. Although unexpected, it was already known that regulation of TOR complex 1 by a TSC/RHEB module is not universal. For example, while the TSC/RHEB module is evident in fission yeast *S. pombe* [221], the budding yeast *S. cerevisiae* lacks both RHEB and a TSC [52]. Although the evidence suggests that CeRHEB/RHEB-1 is required for CeTORC1 signaling, it is not known whether CeRHEB GTP charging is subject to regulation; CeRHEB may be constitutively GTP charged such that the regulation of CeTORC1 activity occurs entirely through the amino acid-Rag pathway. There is evidence suggesting that amino acids regulate the CeTOR complex-1. For example, deletion of *bcat-1*, which catalyzes the first step in branched chain amino acid catabolism,

increases leucine levels and extends lifespans through the neuronal secretion of CeTGFβ/DAF-7, a response that requires neuronal CeTOR (shown using the *Ce-tor* RNAi in the *C. elegans* TU3311 strain, which responds to RNAi feeding preferentially in neurons) [222].

Inhibition of insulin/IGF signaling (IIS), for example through mutation or RNAi depletion of insulin-IGF receptor/*daf-2* and PI-3 kinase/*age-1,* in conditions of severe loss-of-function, causes a constitutive diversion of development from L1 into diapause, i.e., dauer development. If the loss-of-function is less severe or is introduced after the dauer decision point in L1, reduced IIS substantially prolongs lifespan [223–227] (reviewed in [228]). The developmental and prolongevity effects caused by reduced IIS are entirely reversed by deletion of the *daf-16* gene that encodes the CeFOXO/DAF-16 transcription factor protein [223]. Thus, the CeFOXO/DAF-16 protein is an indispensable regulator of lifespan regulation by IIS. Overexpression of CeFOXO/DAF-16 per se extends lifespan [229]. CeAkt phosphorylates DAF-16, thereby inhibiting its nuclear entry. CeFOXO3 is phosphorylated by CeAMPK [230] (as well as CeJnk [231] and CeMst1/2 [232]), and the extended lifespan of insulin-IGF receptor/*daf-2* mutants is partially reversed by depletion of CeAMPK/*aak-2* [233]. In mammals, IIS negatively regulates AMPK through Akt-catalyzed phosphorylation of the AMPK catalytic subunit [234, 235]. Thus, in addition to CeAkt acting directly on DAF-16, some of the negative regulation of DAF-16 by IIS may be accomplished through CeAkt phosphorylation and inhibition of CeAMPK.

What is the effect of reduced IIS on the activity of CeTORC1? The lack of a TSC in *C. elegans* eliminates the major pathway through which IIS regulates TORC1 in higher metazoans; although CeRal-GAP is suggested to function as a TSC [236], little or no evidence indicates the operation of a pathway from the insulin-IGF receptor/DAF-2 or CeAkt1/2 upstream of CeRHEB-1. Nevertheless, *insulin-IGF receptor/daf-2* gene loss-of-function is accompanied by a reduced expression of CeRAPTOR/DAF-15 [215], pointing to a convergence of the IIS and CeTOR pathways at the level of CeTOR complex 1 itself. The diminished expression of CeRAPTOR/DAF-15 in the *insulin-IGF receptor/daf-2* mutant is fully reversed by deletion of *Ce-foxo/daf-16* gene. Deletion of the *Ce-foxo/daf-16* gene also reduces the expression of RSKS-1 protein, the homolog of S6K1 in *C. elegans* [211, 215]. Consistent with this, depletion of the CeS6K1/RSKS-1 protein extends the lifespan of *C. elegans* [237, 238]. Another TSC-independent mechanism for IIS regulation of TORC1 activity is through AMPK, which in addition to its ability to inhibit TORC1 by activation of the TSC can also inhibit TORC1 by direct phosphorylation of RAPTOR in mammals [239, 240]. Thus, the disinhibition of CeAMPK engendered by reduced IIS may also reduce CeTORC1 signaling in *C. elegans.*

Based on the likelihood that TORC1 activity is diminished in long-lived *insulin-IGF receptor/daf-2* mutant worms, does reduced CeTORC1 activity contribute to the extended lifespan engendered by reduced IIS? Consistent with this possibility, CeTOR depletion does not further extend the lifespan of *insulin-IGF receptor/daf-2* mutants [216], indicating that CeTOR's major effectors of lifespan

are subsumed by *insulin-IGF receptor/daf-2* signaling. However, the prolongevity effects of *Ce-tor* mRNA depletion by RNAi are entirely insensitive to the elimination of *Ce-foxo/daf-16* [216]. Thus, if diminished CeTORC1 activity contributes to the extended lifespan caused by reduced IIS, it operates downstream of the CeFOXO/DAF-16 protein. RNAi of *Ce-tor* mRNA is capable of enhancing long-lived dauer formation in *insulin-IGF receptor/daf-2* mutants, but this observation was made at a permissive temperature where insulin-IGF receptor/DAF-2 protein is partially functional [216]. Thus, it is not clear if loss of *Ce-tor* acts on a parallel pathway or is simply enhancing the weak loss-of-function of *insulin-IGF receptor/daf-2* by acting in the same pathway. This seemingly clear-cut distinction is complicated by the finding that the prolongevity- and stress resistance-promoting effects of depleting the CeTORC1 regulatory components, CeRagA, CeRagC, CeRHEB and CeRAPTOR, are generally suppressed by deletion of *Ce-foxo/daf-16* gene as well as by deletion of the *skn-1* gene, that encodes a protein homologous to the bZip transcription NRF/NFE2 in humans [211, 215]. SKN-1 protein is a stress-activated transcription factor that is negatively regulated by IIS through CeAkt in a manner similar to the CeFOXO/DAF-16 protein [241]. *Ce-torc1* RNAi-mediated depletion does not alter the nuclear localization of SKN-1 but it does activate the expression of a cohort of SKN-1 and CeFOXO/DAF-16 target genes that overlap with, but are distinct from, those activated by reduced IIS. Unlike IIS, CeTORC1 negatively regulates nuclear localization of the CeFOXO/DAF-16f isoform and suppresses SKN-1 target gene expression by a mechanism distinct from IIS. Reciprocally SKN-1 upregulates expression of CeTORC1 pathway components. Depletion of CeRagC leads to a SKN-1-dependent upregulation of CeRAPTOR/DAF-15 and CeS6K1/RSKS-1 [211]. In contrast, metabolic consequences of TORC1 deficiency, specifically the accumulation of excess hypodermal fat droplets, are not suppressed by *Ce-foxo/daf-16* mutation [17, 215]. Although *Ce-rictor/rict-1* loss-of-function mutations also affect longevity in *C. elegans* in a complex manner [242], available information does not indicate a specific role for CeTORC2 and its major downstream effector CeSGK-1 [243] in the regulation of mRNA translation.

9.5 CeTOR and Extended Longevity Due to Dietary Restriction

Ce-rictor-deficient worms placed on plates containing an enriched diet exhibit prolonged lifespans accompanied by diminished residence time on the bacterial lawn suggestive of diminished feeding [242]. Interestingly, the differential effect of dietary source on the lifespan of *Ce-rictor* mutant worms is not observed in liquid suspension [243]. Dietary restriction (DR) is perhaps the most widely studied model of extended longevity. A considerable variety of manipulations have been employed

to achieve reduced food intake in *C. elegans* (summarized though 2009 by Greer and Brunet [244]), including growth in liquid suspensions containing different concentrations of *E. coli* [245–247] or in axenic media [246–250]; growth of worms on solid media containing different amounts of bacteria [245–248, 251], either post-reproductive worms grown on normal growth medium (NGM) or day-1 adults grown on peptone-free agar; varying the peptone concentration in NGM [252]; and intermittent feeding [218] or complete nutrient withdrawal, i.e., starvation in newly adult worms [251]. Finally, some (but not all) partial loss-of-function mutants of the nicotinic acetylcholine receptor, such as *eat-2*, have defective pharyngeal pumping and consequent lifelong reduction in food intake accompanied by an extended lifespan [246, 248, 253]. Although all of these circumstances can prolong lifespan to some extent, perhaps not surprisingly the variable forms of nutritional stress result in some variation in the transcriptional responses engendered, judging by their sensitivity to transcription factor deletion. Thus, DR-mediated longevity (induced by intermittent fasting (IF) of newly adult worms) is sensitive to *Ce-foxo/daf-16* deletion but insensitive to deletion of *skn-1* (the homolog of the human NRF/NFE2 transcription factor) or deletion of *pha-4* (a FoxA transcription factor) [218]. Similarly, the extended lifespan of post-reproductive adults (induced by feeding reduced bacterial numbers on solid media) is also sensitive to deletion of *Ce-foxo/daf-16* [230]. In contrast, *daf-16* deletion does not affect the prolongevity effect of complete fasting of adult normal [251] or sterile [254] worms; or that observed in *eat-2* (nicotinic acetylcholine receptor) mutants compared with wildtype, fed either as hatchlings [250], L4 [246–248] or early adults [245], either on solid plates [245, 253] or in liquid media with [245–248] or without (axenic) [246–248] bacteria; or that elicited by feeding newly adult worms on either solid and liquid media loaded with progressively fewer bacteria [245]. These modes of dietary restriction-induced lifespan extension are dependent upon CeNRF/CeNFE2/SKN-1 and CeFOXA/PHA-4 [245, 249]. Most studies find that dietary restriction further extends the lifespan of *insulin-IGF receptor/daf-2* mutants [247, 253], suggesting independent inputs to longevity. However, dietary restriction may further reduce IIS in these partial loss-of-function mutants. Thus, it is the insensitivity to *Ce-foxo/daf-16* deletion in the majority of dietary restriction models that argues for the predominant independence of a dietary restriction-extended lifespan from IIS.

In contrast with dietary restriction and IIS, most evidence points to a critical role for CeTOR in the prolongevity effects of dietary restriction. *Ce-tor* RNAi does not extend the lifespan of *eat-2* (nicotinic acetylcholine receptor homolog) mutants or that caused by reduced bacterial availability [237]. The expression of CeFOXA/PHA-4 protein is upregulated by dietary restriction, and depletion of *Ce-foxa/pha-4* in adult worms (i.e., bypassing its role in pharyngeal development) abolishes the extended longevity of those dietary restriction models [245], as well as that caused by depletion of *Ce-tor* or *Ce-s6k1/rsks-1* [255], but has no effect on the extended lifespan of *insulin-IGF receptor/daf-2* or mitochondrial mutants [245]. The ability of *Ce-tor* depletion to activate some PHA-4 target stress genes appears to require the GCN2 kinase [256], as had been previously observed in *S. cerevisiae* [257].

9.6 CeTOR and Extended Longevity Due to Germline Ablation

Ablation of germline stem cells (GSCs) results in infertility and in an extension of lifespan, but only if the somatic gonad remains intact [258]. This involves signaling by dafachronic acid (DA), an oxidized sterol [259] whose synthesis probably originates in the somatic gonad but involves other steroidogenic tissues, including the intestine. DA is the ligand for the LXR/FXR/VDR-like nuclear hormone receptor DAF-12 (an homolog of the vitamin D receptor, VDR) [260]. Activated CeVDR/DAF-12 promotes CeFOXO/DAF-16 nuclear localization [258] exclusively in the intestine [261, 262], where CeFOXO/DAF-16, in cooperation with several other elements (TCER-1, KRI-1, PHI-62, FTT-1/PAR-5) [263, 264], elicits a pattern of gene expression distinct from that induced by reduction of IIS [263]; lifespan extension by ablation of GSCs and reduced IIS is additive.

Other elements required for the prolongevity effect of GSCs ablation include a CeFOXA/PHA-4-dependent activation of intestinal autophagy [265], the nuclear hormone receptor NHR-80 (nuclear hormone receptor-80, a divergent ortholog of the HNF4 protein) [266], activation of CeNRF/CeNFE2/SKN-1 [267] as well as mIRs [268–270], including a neural input that is dependent on neuronal mIR-71 [269, 270]. The disinhibition of CeFOXA/PHA-4 and the activation of autophagy necessary for lifespan extension are attributable to reduced TOR signaling; GSCs ablation reduces CeTOR mRNA and protein, and RNAi targeting the *Ce-tor* mRNA does not alter the prolongevity effect of GSCs ablation [265].

In addition to the central role of CeTOR in upregulation of CeFOXA/PHA-4 and the HLH-30 (a predicted basic helix-loop-helix [bHLH] transcription factor, orthologous to the human microphthalmia-associated transcription factor [HGNC: MITF])-driven lipophagy critical to many prolongevity models (described below), GSCs ablation involves other alterations in lipid metabolism important to lifespan extension. The CeVDR/DAF-12 target, CeHNF4/NHR-80, promotes expression of a stearoyl-coenzyme A desaturase-1 homolog FAT-6, resulting in increased synthesis of monounsaturated oleic acid (OA); together, CeHNF4/NHR-80 and OA promote longevity [266].

9.7 CeTOR and Translation Control

A shared feature of *Ce-tor* deficiency, dietary restriction and reduced IIS is a reduction in protein synthesis/mRNA translation [17, 271]. As mentioned previously, RNAi-mediated depletion of the mRNAs encoding the translation initiation factors eIF4G, eIF2α and eIF2β in *C. elegans* phenocopies the developmental arrest and intestinal atrophy caused by *Ce-tor* deficiency [17], raising the question of whether reduced mRNA translation/protein synthesis per se can prolong lifespans and thus account for this response to *Ce-tor* deficiency. RNAi-induced depletion of a

variety of (but not all) ribosomal proteins mRNAs and of *Ce-s6k1/rsks-1* mRNAs extend lifespan and do so, like *Ce-tor* depletion, in a largely *Ce-foxo/daf-16*-independent manner [237, 238]. This finding is consistent with a role for reduced ribosomal biogenesis in the prolongevity effect of CeTORC1, inasmuch as TORC1 has been shown to regulate ribosomal biogenesis in mammalian cells [98]. Interestingly, inhibition of protein synthesis by depletion of eIF2β or eIF4G in *C. elegans* also prolongs lifespans. However, this response is fully prevented by deletion of *Ce-foxo/daf-16* [237] or both *Ce-foxo/daf-16* and *Ce-nfr/Ce-nfe2/skn-1* [211]. The finding that depletion of specific transcription factors can reverse the extension of lifespan induced by interference with mRNA translation indicates that these prolongevity effects are due to the expression of specific transcriptional programs and not to global protein depletion or to an overall reduction of energy consumption. Thus, depletion of different components of the translational apparatus in adult worms activates overlapping, stress-induced transcriptional responses mediated by CeFOXO/DAF-16, CeNRF/CeNFE2/SKN-1 and CeFOXA/PHA-4, which are capable, singly or in combination, of promoting lifespan extension. CeTOR, acting through both complexes 1 and 2 (not yet resolved for CeFOXA/PHA-4), suppresses the programs of all three transcription factors, whereas IIS negatively regulates CeFOXO/DAF-16 and CeNRF/CeNFE2/SKN-1. The mechanism for this selective transcriptional response to translational inhibition is not known, but presumably occurs by the introduction of a bias in mRNA translation that is driven by the specific manner in which overall mRNA translation is made deficient.

In addition to promoting mRNA translation/protein synthesis, TORC1 is also known to repress autophagy/protein breakdown. Although necessary for lifespan extension in *C. elegans*, enhanced autophagy per se is not sufficient to extend lifespan; deletion of *Ce-foxo/daf-16* from long-lived *insulin-IGF receptor/daf-2* mutant worms abolishes lifespan extension without reducing the superabundance of autophagic vesicles. A detailed description of TOR regulation of autophagy and the specific contributions of autophagy and lipophagy to lifespan extension lies beyond the scope of this chapter.

9.8 Concluding Remarks

In summary, in *C. elegans* as in other eukaryotic organisms, TOR is a master regulator of mRNA translation. CeTOR is expressed in essentially all cells, and although its intracellular program is stereotyped to a significant degree, the organismal outcome will depend on the specific cell. The worm intestine is a major site through which CeTOR regulates lifespan. The worm intestinal lining cells, in addition to mediating food digestion and nutrient absorption, share functions with the liver and white adipose tissue of higher metazoans. There, the inhibition of translation caused by *Ce-tor* depletion biases protein expression to activate/disinhibit a cohort of transcription factors that drive programs of gene expression that rewire metabolism, promote stress resistance, activate autophagy and extend lifespan [272]. CeTOR may also regulate

these transcriptional programs more directly, through pathways yet to be identified. The activation of autophagy and lysosomal biogenesis caused by *Ce-tor* depletion, in addition to providing amino acid and lipid nutrients, enhances proteostasis and generates new intra- and intercellular signals whose biology is still poorly understood. *C. elegans* provides an outstanding model in which to connect the biochemical with the organismal outcomes of TOR signaling.

10 Mammals

10.1 mTOR: Lessons of Mice and Men

mTOR is a central regulator of cellular and organismal growth in mammals. Pharmacological and genetic analyses in immortalized human cell lines and knockout mouse models have helped dissect the mechanism(s) by which mTOR (most prominently mTORC1) controls mRNA translation, and cellular and organismal growth in mammals [19]. In the following sections, we discuss findings about mTORC1 and translation control from mouse models and human cell lines studies.

10.2 Mammalian TOR: Discovery of mTOR

Shortly after the discovery of TOR in yeast [7], four research groups independently identified the mammalian TOR homolog employing a variety of techniques, including yeast-two hybrid screens [11], affinity purification of the mTOR protein in complex with rapamycin/FKBP12 (FK506 binding protein 12) [9, 10, 12] and cloning of the mTOR gene [9, 10, 12]. In the years that followed, several groups demonstrated by various means that mTOR phosphorylated two important protein families previously known to be involved in the control of mRNA translation, namely the S6Ks (ribosomal proteins S6 kinases) and the 4E-BPs (eukaryotic initiation factor 4E-binding proteins) [60, 61, 273–286]. The mechanisms by which mTOR phosphorylates and controls the activity of S6Ks and 4E-BPs in mammals will be discussed in Sects. 10.5.1 and 10.5.2.

10.3 Characterization of mTORC1 and mTORC2 in Mammals

In mammals (as is the case for all eukaryotes), mTOR nucleates two protein complexes termed mTORC1 (mTOR complex 1) and mTORC2 [19, 20]. Each complex has specific protein as well as shared protein components. mTORC1 is formed by the scaffold protein RAPTOR (regulatory associated protein of mTOR)

[41, 42], mLST8 (mammalian lethal with Sec13 protein 8; in mammals, also referred to GβL, short for G-protein β-subunit-like protein) [44], PRAS40 (proline-rich Akt substrate 40 kDa; also known as Akt1S1, Akt1 substrate 1) [193–198, 287, 288] and DEPTOR (DEP domain-containing mTOR-interacting protein) [289]. mTORC2 shares components with mTORC1, namely the mTOR kinase, mLST8/GβL and DEPTOR. mTORC2 also has unique components; these include: RICTOR (rapamycin-insensitive companion of TOR) [45], mSIN1 (mammalian stress-activated protein kinase-interacting protein 1) [290], and PRR5/PROTOR-1 and PROTOR-2 (proline-rich protein 5/protein observed with RICTOR-1 and-2) [194, 291, 292]. Notably, from an evolutionary standpoint, both mTORC1 and mTORC2 are highly structurally conserved across the entire eukaryotic kingdom—all the way from fungi to humans. For example, RAPTOR is structurally and functionally homologous to Kog1 in budding yeast [40], RICTOR is the functional equivalent of Avo3 in budding yeast [40], and mLST8/GβL is the equivalent of Lst8 in budding yeast [40]. Notably, some mammalian protein components of mTORC1 appear to be absent in lower eukaryotes: PRAS40 has only been reported in mammals [193–198, 293, 294] and insects [175] (referred to as dPRAS40 or Lobe in *Drosophila melanogaster*), while DEPTOR has only been reported in mammals and birds [289]. Database searches of the *Saccharomyces cerevisiae, Schizosaccharomyces pombe*, *Arabidopsis thaliana* and *Caenorhabditis elegans* proteomes do not yield any protein counterparts for either PRAS40 or DEPTOR, indicating that these proteins likely emerged later in eukaryotic evolution (Table 1).

10.4 Cardinal Outputs of mTOR in Translation

Both mTORC1 [295] and mTORC2 [113, 296, 297] have been linked to the control of protein synthesis. Of the two complexes, however, mTORC1 has the most prominent role in mRNA translation. mTORC1 imparts its translation control by regulating both the initiation [60] and the elongation [62] steps. Coordinated regulation of mRNA translation by mTORC1 is achieved through the direct phosphorylation of key phospho-proteins, which, in turn, exert their effects on the translational machinery and associated translation factors [99].

10.5 Downstream Targets of mTORC1 in Translation

The mTORC1 signaling pathway regulates the phosphorylation of a large number of proteins in the cell [298–301]. Pharmacological studies indicate that upwards of 1800 unique proteins are phosphorylated in an mTORC1/mTORC2-dependent manner [300]. While some of these proteins are in all likelihood *indirect* targets, other proteins are *directly* phosphorylated by mTORC1 or mTORC2. A number of mTORC1 direct targets are known to perform important roles in mRNA translation. In this section, we

will review the current literature on four *direct* targets of mTORC1 in mammals in the context of eukaryotic evolution. Specifically, we will first review the literature on S6Ks and 4E-BPs (two direct targets of mTORC1 with well-characterized roles in protein synthesis); we will then introduce PRAS40 and LARP1 (two novel direct targets of mTORC1 with growing roles in the control of protein synthesis).

10.5.1 S6Ks

Ribosomal S6 kinases (S6Ks) were the first-known direct targets of mTOR (with a defined role in mRNA translation) [273, 286, 302]. In the following section, we will review the literature pertaining to the discovery of S6Ks (section "Discovery of S6Ks"), domain architecture of S6Ks (section "Phosphorylation and Activation of S6K1"), the mechanism of phosphorylation-dependent activation of S6Ks (section "The S6K Family"), the role of mammalian S6K1 in the control of cell growth (section "S6K1 Is a Positive Regulator of Cell and Organismal Growth in Mammals") and the role of RPS6 phosphorylation in the control of translation elongation (section "S6K2 Phosphorylates RPS6 Thus Repressing Translation Elongation").

Discovery of S6Ks

S6K1 was the first of the S6Ks to be identified in the late 1980s [303, 304]. By the early 1990s, our (Joe Avruch) group and that of George Thomas had successfully cloned S6K1 [305–307].[10] Around this time, it had become evident that rapamycin decreases the phosphorylation of S6K1 [279, 285, 312–315] and its downstream substrate RPS6 [279]. In vitro experimentation (namely the inability of rapamycin in complex with FKBP12 to block S6K1 phosphorylation in vitro) indicated that rapamycin did not directly target S6K1, but rather blocked an "upstream signal" [279]. The identification of the "upstream signal" came about in 1994 with the identification of mTOR as the molecular target of the rapamycin/FKBP12 complex by four independent groups [9–12]. Definitive proof that mTOR directly phosphorylates S6K1 in vitro came later in 1995 [273] and again in 1998/1999 [286, 302]. Collectively, these studies have indelibly connected mTOR, S6K1 and RPS6 together.

Phosphorylation and Activation of S6K1

It is now appreciated that S6K1 is phosphorylated at multiple serines and threonines, including two key residues: Thr229 (amino acid numbering according to the human p70 S6K1 isoform) located within the activation loop (or T-loop) in the catalytic domain [316] and Thr389 within an evolutionary-conserved hydrophobic motif

[10]S6K2 was discovered almost a decade later by several groups [308–311].

positioned C-terminally of the kinase domain. Although the phosphorylations of Thr229 and Thr389 are both sensitive to rapamycin, they are catalyzed by different kinases. Phosphorylation of Thr229 (Thr252 in the human p85 S6K1 isoform) is catalyzed by PDK1 (3-phosphoinositide-dependent protein kinase-1) both in vitro and in vivo [316], whereas phosphorylation of Thr389 (Thr412 in human p85 S6K1 isoform) is catalyzed in vitro by mTOR [286, 302].[11] Dual phosphorylation of Thr229 and Thr389 results in a near-maximal activation of S6K1, indicating that these are the main regulatory residues in S6K1 [316]. p70 S6K1 is also phosphorylated at Ser371 within the turn motif (T-motif) (the corresponding phosphorylation site in p85 S6K1 is Ser394). Phosphorylation of Ser371 is important for S6K1 activity, since mutation of this residue to an alanine abolishes kinase activity [318]. Additionally, S6K1 is also phosphorylated at an array of proline-directed serine and threonine residues (SP or TP) located within an auto-inhibitory pseudo-substrate domain (termed SKAIPS for S6K autoinhibitory pseudo-substrate domain) located near the carboxy-terminus of S6K1 [285, 319]. These are: Ser411, Ser418, Thr421 and Thr424 in human p70 S6K1 (or Ser434, Ser441, Thr444, and Thr447 according to the human p85 S6K1 isoform amino acid numbering). The phosphorylation sites within the SKAIPS domain play a secondary role in S6K1 activation in that phosphorylation of these residues alone is necessary [280, 282, 320] but insufficient for full S6K1 activation [305, 319]. They are important nonetheless, in the sense that the ability of PDK1 to phosphorylate Thr229 in p70 S6K1 depends on the prior phosphorylation of the carboxy-terminal phosphorylation sites within the SKAIPS domain [316]. Phosphorylation of Thr229 by PDK1 also depends upon the prior phosphorylation of Thr389 within the hydrophobic motif in the linker region [316]. Therefore, phosphorylation of S6K1 follows a sequential order: the carboxy-terminal phosphorylation sites are phosphorylated first, followed by phosphorylation of Thr389 and Thr229 last. This sequential phosphorylation of S6K1 explains the sensitivity of Thr229 to rapamycin in vivo; although phosphorylation of this site is catalyzed by PDK1 both in vitro and in vivo [316], the prerequisite for prior phosphorylation of Thr389 by mTOR (namely mTORC1) renders it rapamycin sensitive in vivo. A hydrophobic pocket within the kinase domain of PDK1 (termed the PIF-binding pocket) interacts with Thr389, thus allowing for docking to S6K1 and phosphorylation of Thr229 [321–324].

The S6K Family

Ribosomal S6Ks belong to the AGC-family of serine/threonine protein kinases (that also include—among many others—the protein kinases A, G and C) [325]. Ribosomal S6Ks were named after one of their main substrates: the ribosomal protein S6 (RPS6), a component of the 40S subunit of the ribosome [326]. The

[11] PDK1 can also phosphorylate Thr389 (Thr412 according to the p85 S6K1 isoform amino acid numbering) in vitro, albeit at rather low stoichiometry indicating that PDK1 is not the predominant Thr389-kinase in vivo [317].

mammalian genome encodes two S6K genes known as *Rps6k1* and *Rps6K2*. The *Rps6k1* gene is transcribed into a single mRNA that gives rise to two S6K1 proteins termed p70 S6K1 (502 amino acids) and p85 S6K1 (525 amino acids). p70 S6K1 and p85 S6K1 protein isoforms are produced upon translation from alternative start sites and differ only in their N-termini [307], which in the case of p85 S6K1 comprises an arginine-rich region believed to be involved in subcellular localization [327–330]. The *Rps6k2* gene also encodes for two S6K2 proteins termed p54 S6K2 and p60 S6K2 [308–311]. Sequence alignment of the S6K1 and S6K2 indicates that these proteins are highly conserved at the amino acid level, sharing 70 % amino acid identity. S6K1 and S6K2 are also conserved at the structural level, and share all core structural domains, including an amino-terminal arginine-rich domain, an AGC-kinase domain, a kinase extension domain and a S6K auto-inhibitory pseudo-substrate (SKAIPS) domain located at the carboxy-terminal. Consistent with their similar structural organization, S6K1 and S6K2 exhibit a certain level of functional overlap. However, S6K1 and S6K2 are not strictly interchangeable. They do in fact display substrate selectivity. For example, while both S6K1 and S6K2 can phosphorylate RPS6, studies from *Rps6k1* and *Rps6k2* single and double knockout mouse embryo fibroblasts (MEFs) categorically demonstrate that phosphorylation of RPS6 is predominantly catalyzed by S6K2 [331].[12] The role of S6K2-mediated RPS6 phosphorylation in protein synthesis will be revisited later in section "S6K2 Phosphorylates RPS6 Thus Repressing Translation Elongation".

S6K1 Is a Positive Regulator of Cell and Organismal Growth in Mammals

The central role of the TOR pathway in the control of cell and organismal growth is a pervasive theme across evolution. Mutations in the TOR pathway in yeast, plants and *Drosophila melanogaster* result in impaired growth/proliferation phenotypes (refer to Sects. 1, 7 and 8). In Sect. 8, when introducing the subject of *Drosophila melanogaster* growth, we discussed how the dS6K branch of the dTOR pathway mediates the growth-promoting signals of the dTOR pathway [178]. In fact, as we will discuss next, the mechanism by which mTOR (specifically mTORC1) controls cell and organismal growth is conserved in mammals.

Genetic mouse models provide a fantastically pliable platform to study a myriad of mammalian phenotypes, including key outputs of TOR such as growth, fertility and viability. In 1998, Shima et al. [308] investigated the effect of homozygous disruption of the *Rps6k1* gene (which encodes for p70 S6K1 and p85 S6K1 proteins) in a genetic mouse knockout model. Disruption of *Rps6k1* did not affect the viability or fertility of mice, but did have a pronounced effect on animal growth. The body weight of homozygous $Rps6k1^{-/-}$ mice was considerably reduced compared to control $Rps6k1^{+/+}$ control littermates [308] in both male and female mice. The resultant

[12]S6K1 preferentially phosphorylates another S6K substrate: IRS-1 (insulin receptor substrate-1) [332].

reduction in global body weight (upon genetic disruption *Rps6k1* gene) was attributed to a decrease in the weight of every weighed organ (namely, brain, heart, lung, thymus, liver, spleen, kidney and testis). Embryos deleted for the *Rps6k1* gene also exhibited decreased size, indicating that maternal imprinting cannot compensate for S6K1 deficiency. Importantly, mouse embryo fibroblasts derived from *Rps6k1*$^{-/-}$ mice revealed no defects in cell proliferation, suggesting that in mammals (as is the case in flies) S6K1 controls cell size (cell mass) and not cell proliferation (cell number) [308]. The notion that S6K1 plays a role in cell size rather than cell proliferation in mammals was independently confirmed by Dowling et al. in 2010 [179]. In the process of generating the single homozygous *Rps6k1*$^{-/-}$ knockout mouse, Shima et al. [308] unexpectedly observed that disruption of *Rps6k1*$^{-/-}$ does not substantially reduce the phosphorylation of RPS6, indicating the existence of a second S6 kinase in mammals (now known as S6K2) [308–311]. Subsequent genetic analysis of the single *Rps6k2*$^{-/-}$ and double *Rps6k1*$^{-/-}$/*Rps6k2*$^{-/-}$ knockout mouse embryos revealed that S6K2 (and not S6K1) exerts the major S6 kinase activity toward ribosomal protein S6 [331]. In contrast, and importantly from a phenotypic standpoint, S6K1 (and not S6K2) is the primary positive regulator of cell and organismal growth. In summary, the role of S6K in the control of cell and organismal growth is conserved all the way from insects to mammals. One important major difference is that insects have one single *Rps6k* gene (*ds6k*), whereas the mammalian genome encodes two *S6K* genes (*Rps6k1* and *Rps6k2*), indicative of a gene duplication event late in metazoan evolution (Table 1 and Fig. 7). From a functional evolutionary perspective, between the *s6k* genes in mammals, the *Rps6k1* mammalian gene is the one that most closely mirrors the growth function of the insect *ds6k* gene.

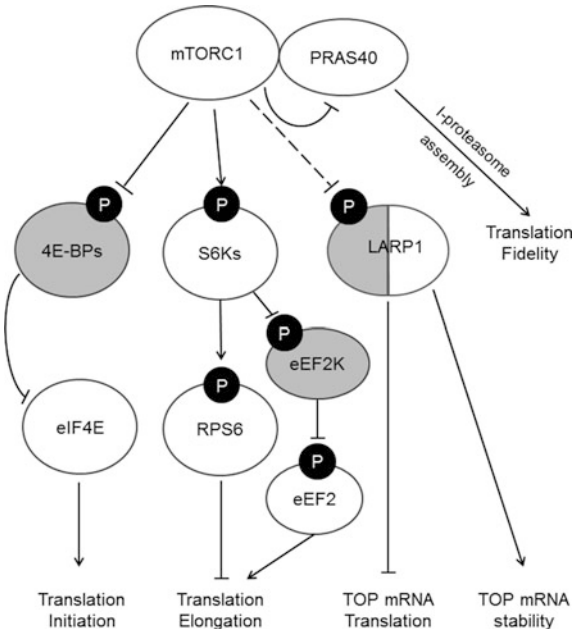


Fig. 7 Model depicting signaling pathways downstream of mTORC1 involved in the control of mRNA translation. Proteins shown in *white* activate mRNA translation. Proteins in *gray* inhibit mRNA translation. *Dashed lines* indicate poorly characterized pathways. See main text for details

S6K2 Phosphorylates RPS6 Thus Repressing Translation Elongation

What is the role of ribosomal protein S6 phosphorylation? The role of RPS6 phosphorylation in translation has been the subject of much debate over the years. RPS6 phosphorylated at multiple residues, including five well-characterized serine residues: 235, 236, 240, 244 and 247.[13] Biochemical analyses of mRNA translation initiation and elongation in mouse embryo fibroblasts (MEFs) derived from *Rps6*$^{P-/-}$ knockin mice (bearing alanine substitutions in place of serine residues at positions 235, 236, 240, 244 and 247) indicate that the RPS6 protein may play a role in the elongation step of protein synthesis [334]. Specifically, replacement of these key serine residues by non-phosphorylatable alanine residues results in increased incorporation of ^{35}S-radiolabeled cysteine and increased steady-state protein content in *Rps6*$^{P-/-}$ knockin MEFs [334]. The increase in protein synthesis rates in *Rps6*$^{P-/-}$ knockin MEFs was attributed to decreased ribosome half-transit times, suggesting accelerated elongation rates of protein synthesis in knockin MEFs [334]. This was accompanied by a concomitant decrease in cell size, presumably resulting from accelerated cell division rates. The balanced increment in cell division and decrease in cell sized explains the unchanged body weight in *Rps6*$^{P-/-}$ knockin mice reported by Ruvinsky et al. [334]. The finding that phosphorylation of RPS6 results in decreased rates of protein synthesis [334] is as important as it is surprising, in that S6Ks are typically regarded as positive regulators of protein synthesis. A crystallographic study of the 40S eukaryotic ribosome subunit in complex with the eukaryotic initiation factor 1 [326] has added another layer of complexity to the interpretation of the role of RPS6 phosphorylation in protein synthesis: structural analysis indicates that positioning of the RPS6 within the 40S subunit of the ribosome is such that the five key serine residues are not in close proximity to the mRNA tunnel [326]. Therefore, the mechanism by which RPS6 phosphorylation controls mRNA translation awaits further clarification.

RPS6 is an important target of the S6Ks (primarily S6K2), but both S6K1 and S6K2 are known to phosphorylate a large number of additional substrates, many of which also play notable roles in mRNA translation. These include, non-extensively, the eukaryotic initiation factor 4B (eIF4B) [335, 336], multiple subunits of the eukaryotic initiation 3 (eIF3) complex [151, 337, 338], the program cell death protein 4 (PDCD4) [336, 339] and the eukaryotic elongation factor 2 kinase (eEF2K) [340]. Due to space constraints, it is not feasible for us to review all the outputs of S6Ks in translation in this chapter. The reader is directed to an earlier review [99] for an in-depth analysis of this topic.

[13]Phosphorylation of RPS6 at Ser235 and Ser236 can also be catalyzed by another family of S6 kinases termed p90RSKs (short for ribosomal S6 kinases of 90 kDa). p90RSKs are not subject to regulation by the TOR pathway and are therefore not reviewed in this chapter. For additional information on p90RSKs please refer to this excellent review on the subject [333].

10.5.2 4E-BPs

4E-BPs are low-molecular-weight translation repressor proteins that function downstream of mTORC1. In vivo, the phosphorylation of 4E-BPs is partially sensitive to the mTORC1 inhibitor, rapamycin, and is fully inhibited by active site (ATP-mimic) mTOR inhibitors. In vitro mTOR can directly phosphorylate 4E-BP1 on multiple residues [274, 286]. In the following sections, we review the literature pertaining to the discovery of 4E-BPs (section "Discovery of 4E-BPs"), the mechanism of phosphorylation and inactivation of 4E-BP1 and 4E-BP2 (section "Conservation of 4E-BPs in Mammals"), and an alternative mechanism of 4E-BP2 regulation that involves asparagine deamidation (section "4E-BP1 and 4E-BP2 Regulate Adipogenesis and Insulin Resistance").

Discovery of 4E-BPs

In mammals, the 4E-BP family is comprised of three members 4E-BP1, 4E-BP2 and 4E-BP3 [183–186, 341, 342]. 4E-BP1 and 4E-BP2, the founding members of the 4E-BP family, were originally identified in 1994 [183–186, 342], while 4E-BP3, the third member of the 4E-BP family, was identified in 1998 [341].

Conservation of 4E-BPs in Mammals

4E-BP1, 4E-BP2 and 4E-BP3 are encoded by three distinct genes but show a high degree of sequence identity at the amino acid level, suggesting that they originate from a common ancestral 4E-BP that duplicated late in eukaryotic evolution (this is substantiated by the fact that three 4E-BPs can be found in mammals, whereas insects and nematodes have a single 4E-BP) (Table 1 and Figs. 4 and 7). How similar are 4E-BPs among themselves? Alignment of the amino acid sequences of the three 4E-BP proteins reveals that 4E-BP1 and 4E-BP2 are the most similar of the 4E-BP proteins, displaying 50 % identity at the amino acid level [99]. 4E-BP3 is the most divergent 4E-BP family member, exhibiting 44 and 46 % identity to 4E-BP1 and 4E-BP2 at the amino acid level, respectively [99]. 4E-BP1, 4E-BP2 and 4E-BP3 also show a high degree of regulatory homology, i.e., they are all regulated in a similar manner but with subtle variations. The activity of 4E-BPs (i.e., their association with the mRNA m^7GpppG cap-binding protein, eIF4E) is primarily regulated through multisite phosphorylation (detailed in section "Phosphorylation and Function of 4E-BP1 and 4E-BP2"). The key phospho-residues that regulate the binding of 4E-BPs to eIF4E (namely Thr37, Thr46, Ser65 and Thr70, amino acid numbering according to human 4E-BP1 protein) are conserved in both 4E-BP1 and 4E-BP2 and are subject to identical regulation. Thr37 and Thr46 residues are also conserved in mammalian 4E-BP3, but whether they are subject to regulation (phosphorylation) in this latter family member is still unclear. Some data suggest that phosphorylation of Thr23 and Thr32 (equivalent to Thr37 and Thr46 in human 4E-BP1) does not occur in human 4E-BP3 because of the

absence of an upstream regulatory motif ("RAIP" motif, named after the single amino acid code that gives rise to it) [343]. Further studies will be required to fully elucidate the regulation of 4E-BP3. By comparison, considerably more is known about the phosphorylation/regulation of 4E-BP1 and 4E-BP2. This subject will be discussed in detail in section "Phosphorylation and Function of 4E-BP1 and 4E-BP2". Phosphorylation is the best studied, but not the sole, mechanism of 4E-BP regulation. 4E-BP2, but not 4E-BP1, is subject to asparagine deamidation [344, 345]. The physiological significance of this post-translational modification to brain development will be discussed in further detail in section "Asparagine Deamidation: An Alternative Mechanism of 4E-BP2 Regulation".

Do 4E-BPs have the same function? Broadly speaking, yes. All three family members associate with the eukaryotic initiation factor 4E1 (eIF4E1, commonly referred to simply as eIF4E), but whether they act as functionally interchangeable homologs is not known. Little is known about the functional homology of 4E-BP1, 4E-BP2 and 4E-BP3, i.e., are different subset of transcripts translationally regulated by each of these proteins? To the best of our knowledge, the specificity of each of the proteins for different mRNA subclasses has not been compared head-to-head.

Phosphorylation and Function of 4E-BP1 and 4E-BP2

As the name "eukaryotic initiation factor 4E-binding proteins" suggests, 4E-BPs bind to the mRNA cap (m^7GpppG)-binding protein eIF4E and preclude the binding of eIF4E to eIF4G [183, 184, 341], which together with eIF4A form the eIF4F complex. By competing with eIF4G for binding to eIF4E, 4E-BPs effectively prevent the formation of the eIF4F complex. Formation of the eIF4F complex is critical for the initiation step of cap-dependent translation [346]. The eIF4F complex binds to the eIF3 (a multiprotein eukaryotic initiation factor). Together, eIF4F and eIF3 form the pre-initiation complex (PIC) that recruits the small (40S) subunit of the ribosome to the 5′untranslated region (5′UTR) of all nuclear-encoded cellular mRNAs. The binding of 4E-BPs to eIF4E is regulated through phosphorylation of multiple residues [346]. 4E-BP1 and 4E-BP2 undergo phosphorylation on five main residues: Thr37, Thr46, Ser65, Thr70 and Ser83 [277, 346–348]. Although these are the most commonly cited phospho-residues in the literature, additional phospho-residues exist: 4E-BP1 is also phosphorylated on Thr41 and Thr50 [349], Ser101 [350] and Ser112[14] [351]. Phosphorylation of 4E-BP1 occurs in a sequential manner [352], with Thr46 being phosphorylated first [353] followed by Thr37, Thr70 then Ser65 [352]. Initial biochemical (cap pull-down) evidence indicated that

[14]Thr37, Thr46, Thr41, Thr50, Ser65 and Thr70 (but not Ser83, Ser101 and Ser112) are conserved in between mammals and insects. This suggests that the N-terminal phosphorylation residues play a conserved regulatory role in evolution.

phosphorylation of Ser65 played a major role in the release of 4E-BP1 from eIF4E [352]. Phosphorylation of additional residues, namely of the N-terminal sites on 4E-BP1, has since been suggested to also play a role in the dissociation of 4E-BP1 from eIF4E [354]. The relative contribution of each phosphorylation event to the regulated release of 4E-BP1 and 4E-BP2 was difficult to ascertain solely by biochemical experimentation. Cap pull-down also did not clarify the underlying mechanism for phosphorylation-mediated release of 4E-BP1: does phosphorylation lead to electrostatic repulsion with the negatively charged dorsal face of eIF4E [355] or does it induce a conformational change on 4E-BP1 such that it can no longer bind eIF4E? Exact confirmation of the phosphorylation-mediated release mechanism required detailed structural and biophysical analyses. NMR and ITC studies in both 4E-BP1 [356, 357] and 4E-BP2 [358] have greatly aided our understanding of the phospho-regulatory mechanism underlying the dissociation of 4E-BPs from eIF4E. 4E-BPs are intrinsically disordered proteins [355, 359, 360] that become structured upon association with eIF4E [357, 361–364]. Phosphorylation of Thr37 and Thr46 coordinates the folding of residues Pro18-Arg62 of 4E-BP2 such that the eIF4E-binding motif (YxxxxLϕ, where x denotes any amino acid and ϕ a hydrophobic residue) becomes partly buried and thereby inaccessible to eIF4E [358]. This study [358] provides outstanding insight into how the phosphorylation of 4E-BP2 (and presumably, by association, 4E-BP1) by mTORC1 coordinates the release of 4E-BPs from eIF4E, a fundamental step in protein synthesis.

4E-BP1 and 4E-BP2 Regulate Adipogenesis and Insulin Resistance

Pioneering studies by the late John C. Lawrence and Richard Denton [185, 186, 342, 365–367] in the early 1990s remarked on an overabundance of 4E-BP1 (formerly known as PHAS-I, for phosphorylated heat- and acid-stable protein regulated by insulin [185]) mRNA and protein in adipose tissue, suggesting that 4E-BP1 played an important role in adipogenesis. Subsequent work in *Eif4ebp1* and *Eif4ebp2* double knockout mice confirmed that this is indeed the case. Combined genetic disruption of *Eif4ebp1* and *Eif4ebp2* genes simultaneously in a mixed BALB/c 129SvJ1 murine model resulted in increased adipose tissue accumulation (specifically white adipose fat) in mice fed either normal chow or high-fat diets [368]. This phenotype was accompanied by insulin resistance and altered lipid metabolism [368]. Collectively, these findings confirm the importance of both 4E-BP1 and 4E-BP2 in controlling lipid metabolism in adipose tissue in mammals.

Asparagine Deamidation: An Alternative Mechanism of 4E-BP2 Regulation

While 4E-BP1 is the dominant 4E-BP isoform in fat tissue, it is reportedly not expressed (or perhaps expressed at low levels) in the mouse brain [369]. Conversely, 4E-BP2 (formerly known as PHAS-II) is lowly expressed in fat tissue

but highly expressed in the brain [369]. This observation raised the intriguing possibility that 4E-BP2 played a prominent role in this organ. Bidinosti et al. [345] confirmed this by showing that 4E-BP2 regulates excitatory synaptic transmission in the brain. Importantly, from a biochemical perspective, this phenotype is mTORC1-regulated—but not in the mTORC1 canonical phosphorylation manner. This paper elucidates a novel molecular mechanism of 4E-BP regulation that involves a distinct post-translation modification, namely: asparagine deamidation [345]. Specifically, Bidinosti et al. show that residues Asn99 and Asn102 within an asparagine-rich region (located near the carboxy-terminus of 4E-BP2) are subject to deamidation (spontaneous asparagine conversion to asparte and iso-aspartate), resulting in augmented RAPTOR binding and reduced affinity for eIF4E [344, 345]. Deamidation takes place at a stage of brain development that is associated with low mTORC1 activity, which led the authors to postulate that its physiological purpose may be to compensate for low mTORC1 activity. The exact biochemical meaning of this post-translation modification, and the mechanism through which mTORC1 controls this effect remain undefined. Importantly, these studies [344, 345] identified a distinct mode of mTORC1-mediated regulation of 4E-BPs that does not involve phosphorylation.

10.5.3 PRAS40

PRAS40 (Proline-Rich Akt Substrate 40 kDa; also referred to as Akt1S1, for Akt1 Substrate 1) was identified in 2007 as a novel component and target of the mTORC1 pathway. PRAS40 is directly phosphorylated by mTORC1 at multiple sites. In the following sections, we review the original discovery of PRAS40 as a target for Akt (section "Discovery of PRAS40"), the identification of PRAS40 as a component and a substrate of mTORC1 (section "PRAS40 Is a Component of mTORC1") and a recent study [370] implicating PRAS40 in the assembly of the immunoproteasome (i-proteasome) in the context of dysregulated protein synthesis and improper folding of newly synthesized polypeptides (section "PRAS40 Regulates the i-Proteasome in Conditions of Reduced Translation Fidelity").

Discovery of PRAS40

PRAS40 was co-discovered independently by two research groups [293, 294] as a new target of Akt (also known as PKB, for Protein Kinase B) in 2003. Akt directly phosphorylates PRAS40 on (at least) one major residue: Thr246 [293]. In response to insulin stimulation, Akt catalyzes the phosphorylation of Thr246. Phosphorylation of this residue promotes the association with the adaptor proteins, 14-3-3s [293, 294]. Consistent with this, treatment with type 1A PI3K inhibitors decreased the interaction between PRAS40 and 14-3-3s [293, 294]. Interestingly, while testing which other signaling pathways impinged on PRAS40, Mackintosh and colleagues also observed that PRAS40's association with 14-3-3s was also

decreased by treatment with rapamycin [294], suggesting that mTORC1 phosphorylated PRAS40 thus regulating PRAS40's interaction with 14-3-3s.

PRAS40 Is a Component of mTORC1

With this in mind, we (Fonseca and co-workers) set out to investigate a potential connection between mTORC1 and PRAS40. To our excitement, immunoprecipitation of intact mTORC1 complexes from HEK293 extracts co-purified PRAS40 along with mTOR and RAPTOR (but not RICTOR, a component of the mTORC2 complex) [195], confirming a direct link between PRAS40 and the mTORC1 pathway. As previously observed for S6Ks and 4E-BPs, PRAS40 associated with RAPTOR via a TOS (TOR signaling) motif centrally located within PRAS40. As is the case with S6Ks and 4E-BPs, mutation of the TOS motif in PRAS40 markedly reduced (albeit not completely) the association of PRAS40 with RAPTOR [195, 196]. Having observed that rapamycin reduced the association between PRAS40 and 14-3-3s (which recognize and bind to phosphorylated residues) [195, 294], we set out to verify whether mTORC1 directly phosphorylated PRAS40. We [195] and others [196, 197, 288] noted that PRAS40 is phosphorylated by mTORC1 at several residues including Ser183 and Ser221 both in vitro and *in cells*. Of note, the phosphorylation of Ser183 by mTORC1 appeared particularly important, as it proved essential for the association with 14-3-3 proteins [195, 287]. At the time that our work was published, several other research groups reached similar conclusions regarding the interaction of PRAS40 with the mTORC1 pathway in mammals [193, 194, 196–198, 288]. The observation that PRAS40 associates tightly with the mTORC1 complex (PRAS40 binds more strongly to RAPTOR than S6Ks and 4E-BPs) led some groups to propose that PRAS40 may function as a core component of the mTORC1 complex [193, 198]. When PRAS40 is overexpressed at moderate-to-high levels in mammalian cells, it causes a drastic reduction in the phosphorylation of other mTORC1 substrates (namely S6Ks and 4E-BPs), suggesting that this protein might function as a substrate competitor for RAPTOR binding and phosphorylation by mTORC1 [197]. Whether PRAS40's sole cellular function is to regulate the phosphorylation of other mTORC1 targets remains an open question [371].

PRAS40 Regulates the i-Proteasome in Conditions of Reduced Translation Fidelity

A recent study by Kim's group shows that PRAS40 may, in fact, play an unique role in the control of the degradation of misfolded poly-ubiquitinated proteins downstream of mTORC1 [370]. mTORC1 hyperactivation leads to increased rates of protein synthesis, resulting in an accumulation of defective misfolded proteins that cause cellular stress. The immunoproteasome (i-proteasome) plays a key role in clearing defective ribosomal products (DRiPS) [370]. Kim and colleagues [370]

have recently elegantly shown that PRAS40 plays a prominent role in the assembly of the i-proteasome. Specifically, phosphorylation of mTORC1-mediated phosphorylation of PRAS40 at Ser183 and Ser221 is essential for the assembly of the i-proteasome. Phosphorylation of PRAS40 at these same residues is also thought to promote protein synthesis (through a poorly understood mechanism), indicating that PRAS40 phosphorylation might have a dual function in stimulating protein synthesis and preventing proteotoxicity by simultaneously stimulating i-proteasome formation, which degrades misfolded polypeptides. Consistent with this, Kazi et al. [372] have shown that PRAS40 stimulates protein synthesis in muscle cells. Further work will be required to elucidate whether PRAS40 plays a direct or indirect role in the control of protein synthesis, but the available evidence suggests that this substrate of mTORC1 is a key factor in cellular homeostasis.

10.5.4 LARP1

Lupus autoantigen (La)related protein 1 (LARP1) has recently been identified as a novel target of mTORC1 [199], and has been shown to play important roles in both TOP mRNA translation [199, 373] and stability [199, 374]. In the following sections, we review the literature that led to the discovery of this signaling component of the mTORC1 pathway and the mechanism(s) by which this important new protein controls TOP gene expression.

mTORC1 and mTORC2 Phosphorylate Numerous Proteins

S6Ks and 4E-BPs were the first direct targets of mTORC1 to be discovered, and both of these families of proteins play important roles in protein synthesis. Yet, it is becoming increasingly apparent that other direct targets of mTORC1 exist, some of which also play prominent roles in the control of protein synthesis. Two important quantitative proteome-wide phosphorylation studies from the Blenis and Sabatini laboratories [298, 300] have highlighted the vastness of the mTORC1 target protein pool in mammals. Over 1800 unique proteins were identified as *bona fide* downstream targets of mTORC1/mTORC2 [300]. Although not all of the identified targets have been shown to be directly phosphorylated by mTORC1 (some of these may lie distally downstream in the mTORC1 pathway), and not all of the phosphoproteins have roles in mRNA translation, this suggested to us that a number of these could be novel direct targets of mTORC1 with important roles in the control of protein synthesis.

LARP1 Is a New Target of mTORC1

With this in mind, in an effort to further our understanding of the role of the mTORC1 pathway in the control of protein synthesis, we (Fonseca and colleagues)

[199] searched for new proteins that specifically associate with mTORC1 and that may play a role in the control of mRNA translation downstream of this protein complex. To this end, we carried out a proteomic screen to identify novel binding proteins of RAPTOR, an mTORC1-specific component [199]. Immunoprecipitation of endogenous human RAPTOR followed by liquid chromatography-tandem mass spectometry (LC-MS/MS) analysis revealed a new direct target of mTORC1 termed LARP1 [199]. LARP1 specifically associated with RAPTOR but not RICTOR, indicating that this novel signaling component is a target of mTORC1 but not mTORC2 [199]. Having identified LARP1 as a novel target of mTORC1, we set out to further understand its function downstream of mTORC1.

Conservation of LARP1 in Eukaryotes

What is LARP1? LARP1 (or La-related protein 1) is an evolutionarily conserved RNA-binding protein that belongs to the La-related protein superfamily [375]. The LARP superfamily comprises the following: La (lupus autoantigen) protein also known as genuine La, LARP1 family, LARP4 family, LARP6 family and LARP7 family. All members of the LARP superfamily share a conserved La-motif (LAM), an ancestral motif that appeared soon after archae-eukarya radiation [375, 376]. LARPs are RNA-binding proteins with important functions in mRNA stability and translation. As such, all LARPs also comprise one or more RNA-recognition motifs (RRMs), some of which are conserved between distinct LARP families (see ref. [376] for an excellent analysis of LARP evolution). Some features are unique to certain LARP families. For example, LARP1 and LARP6 share exclusive domains, such as the domain of unknown function (DM15) and the LSA motif (for LAM- and S1-like-containing proteins) [376]. The role of the DM15 domain in mRNA binding will be discussed later in this section. The LARP1 family is conserved in eukaryotic evolution. At least one member of the LARP1 family can be found throughout the eukaryotic lineage, all the way from protista, fungi and plants to animals (Ref. [375] and Table 1). The human LARP1 family comprises three members: LARP1 (also known as LARP1a), LARP2 (also referred to as LARP1b) and LARP1c (which lacks the distinctive LSA and DM15 motifs), all of which are proximally located within the phylogenetic tree [376]. The following commonly used eukaryotic model organisms encode for at least one LARP1 family protein: *Caenorhabditis elegans* (CeLARP1) and *Drosophila melanogaster* (dLARP1) encode a single LARP1 member, *Xenopus tropicalis* encodes for two LARP1 proteins (XtLARP1a and XtLARP1b), while *Arabidopsis thaliana* encodes for three LARP1 proteins (AtLARP1a, AtLARP1b and AtLARP1c) [376]. Atypical La motif-containing proteins can also be found earlier in eukaryotic evolution: *Saccharomyces cerevisiae* encode for two proteins (Sro9 and Slf1) [377, 378], but whether these proteins are functional homologs of LARP1 (or other LARPs) is still unclear.

Function of LARP1 in Eukaryotes

What does LARP1 do? LARP1 was first cloned in *Drosophila melanogaster* (the fly homolog of LARP1 is known as dLARP) at the turn of the millennium [379]. Database analysis revealed that LARP1 is widely conserved across eukaryotes, all the way from fungi to insects, plants and animals [379, 380]. The first insights into LARP1 cellular function came from two studies in *Caenorhabditis elegans* [380] and *Drosophila melanogaster* [174] almost a decade later. LARP1 (as is the case with other LARPs) possesses a characteristic La motif (LAM), which is an ancient RNA-binding domain [376]. Consistent with this structural feature, CeLARP1 was found to locate to processing bodies (P-bodies) where it modulates the levels of specific mRNAs [380]. In *Drosophila melanogaster*, LARP1 (known as dLARP) was shown to interact directly with the poly(A) binding protein (dPABP) [174]. PABP plays important roles in both mRNA translation and stability [381], suggesting that LARP1 itself might co-regulate these cellular processes with PABP. Subsequent work in mammals and plants confirmed that LARP1 does indeed bind PABP [200] and regulate mRNA translation [200] as well as mRNA stability [374, 382].

How does LARP1 exert its effect on mRNA translation and stability? In different species, LARP1 has been reported to preferentially associate with distinct RNA sequences: in *C. elegans*, CeLARP1 was shown to bind with high affinity to poly (G), and with weaker affinity to poly(U). CeLARP1 does not bind to poly(A) or poly(C) [380]. This is in stark contrast to mammals, where LARP1 associates with an A residue, specifically to the 3′ terminus of the poly(A) tail of cellular mRNAs [374]. In mammals, LARP1 also binds to the polypyrimidine (poly(C/U)) sequences at the 5′ terminus of the cellular mRNAs [199], known as TOP (terminal oligopyrimidine) motifs.

TOP mRNA Translation

What is a TOP motif and what does it do? The eukaryotic ribosome is formed by upwards of 80 different ribosomal proteins [383]. Synthesis of stoichiometric amounts of large multi-protein complexes (such as the ribosome) requires tight translational control [384]. Coordinated synthesis of the individual protein components of the ribosome minimizes unwarranted energy expenditure. Therefore, higher eukaryotes have engendered a neat mechanism to ensure synchronized production of all the individual protein components of the ribosomes strictly when nutrients and energy become available. The mRNAs of all known ribosomal proteins together with a number of translation factors encode a TOP motif at their 5′ terminus. The TOP motif is a sequence of 4–15 pyrimidines located immediately downstream of the m^7Gppp cap of every ribosomal mRNA and some translation factors [385]. The first pyrimidine (+1 position) in the TOP motif is invariably a C residue [386]. The TOP motif displays additional characteristic features, including an even distribution of C and U residues and a high degree of conservation among

vertebrates (see [386] for an elegant review on the subject). Importantly, from a physiological standpoint, the presence of the TOP motif ensures synchronized repression of the translation of all TOP-containing mRNAs, as discussed below. Meticulous work from Meyuhas, Loreni and Amaldi research groups has, over the years, helped to pave our understanding of the mechanism(s) by which the TOP motif controls the synthesis of ribosomal proteins. Briefly, the available evidence demonstrates that the TOP motif functions as a *cis*-regulatory mRNA element [387–389] that recruits an unknown inhibitory *trans*-acting factor (protein or regulatory RNA) that represses TOP mRNA translation upon binding to the *cis*-regulatory element [390–392]. The evidence for the *trans*-acting factor being a repressor comes from a simple, yet compelling, experiment: When increasing amounts of a synthetic RNA oligonucleotide harboring the RPS16 TOP motif are added to translation-competent wheat germ extracts or rabbit reticulocyte lysate, the translation inhibition of an open reading frame reporter construct with the upstream RPS16 TOP motif is progressively relieved, providing formal evidence for the existence of a titratable repressor of TOP mRNA translation [391].

What is the identity of the *trans*-acting inhibitor of TOP mRNA translation? To date, several candidates have been put forth as potential *trans*-acting regulators[15] [295, 393–399]. However, despite intense efforts to identify the missing TOP mRNA *trans*-acting factor, definitive evidence is still lacking, and its identity remains obscure (see this review [386] for a detailed analysis on this subject).

mTOR Controls TOP mRNA Translation

How does mTOR control TOPs? Several signaling pathways have been linked to the control of TOP mRNA translation in mammals [400]. Of relevance to this chapter, the mTOR pathway plays a leading role in the control of TOP mRNA translation. Indeed, TOP mRNAs have long been considered to be under the control of mTOR; rapamycin was first shown to repress TOP mRNA translation over 20 years ago [63, 401]. Multiple studies have shown that rapamycin causes a pronounced shift of TOP mRNAs from polysomal to mRNP fractions [63, 401, 402]. Similarly, active-site mTOR inhibitors (TORIN1 and INK128) lead to a drastic shift of TOP mRNAs from polysome to mRNP fractions [199, 295, 399]. So dramatic is the shift in TOP translational control upon mTOR inhibition that TOP

[15]The following factors have been previously linked to regulation of TOP mRNA translation: ribosomal protein S6 (RPS6) and its kinases S6Ks, eukaryotic initiation factor 4E (eIF4E) and its binding partners 4E-BPs, lupus autoantigen (La) protein and the La-related protein 7, cellular nucleic acid-binding protein (CNBP) also known as zinc finger protein 9 (ZNF9), microRNAs miR-10a and miR-10b. While some of these do appear to contribute to TOP mRNA regulation on some level (e.g., TIA1/TIAR represses TOP mRNA translation in response to amino acid deprivation), definitive evidence for any these being the long-sought controller of TOP mRNA translation is lacking. Refer to the following review for an exhaustive analysis of these factors in the context of TOP mRNA regulation [386].

mRNA translation is referred to as bimodal in its nature [392], i.e., TOP mRNA translation is strictly on or off. Further to the notion that mTOR plays a fundamental role in the control of TOP mRNA translation, two important genome-scale ribosome profile studies have recently revealed that TOP mRNAs constitute the most sizeable portion of the mTOR translatome [295, 399].

How does mTOR/mTORC1 control TOP mRNA translation? It is widely accepted that mTOR plays a fundamental role in the control of TOP mRNA translation, but only recently have we begun to understand how this occurs. Which downstream substrate of mTORC1 regulates TOP mRNA translation? Both S6Ks and 4E-BPs have been previously hypothesized to regulate TOP mRNA translation, but neither of these proteins has been unanimously accepted as the elusive *trans*-acting factor. Having identified LARP1 as a novel direct target of mTORC1 [199], and in light of the recent report by Aoki et al. [374] that LARP1 stabilizes TOP mRNAs in mammals, we examined whether LARP1 could function as a repressor of TOP mRNA translation downstream of mTORC1. To investigate this possibility, we first tested whether LARP1 binds TOP mRNAs (as originally also reported by Aoki et al. [374]). Our data show that LARP1 does, indeed, interact with TOP mRNAs and does so in an mTORC1-dependent manner, i.e., pre-treatment of cells with allosteric (rapamycin) or active-site (TORIN1) mTOR inhibitors enhanced the association of LARP1 with TOP mRNAs [199]. Notably, we found LARP1 to associate directly with the TOP motif of TOP mRNAs. LARP1's association with TOP mRNAs depended on the presence of an intact TOP motif: deletion of the oligopyrimidine tract on RPS6 or RPL32 TOP mRNAs abolished their binding to LARP1 [199]. Polysome profile analysis of TOP mRNA transcripts in LARP1-depleted cells further revealed that LARP1 acts as a repressor of TOP mRNA translation, since shRNA-mediated knockdown of LARP1 led to an accumulation of TOP transcripts in heavy polysome fractions [199]. Conversely, overexpression of ectopic LARP1 repressed the translation of TOP mRNAs [199]. Moreover, we observed that LARP1 is essential for rapamycin- and TORIN1-mediated repression of TOP mRNA translation [199]. Taken together, these data led us to propose that LARP1 functions as a *trans*-acting repressor of TOP mRNA translation downstream of mTORC1.[16]

LARP1 Represses TOP mRNA Translation Downstream of mTORC1

How does LARP1 control TOP mRNA function? A recent study from the Berman laboratory has provided structural insights into how LARP1 binds and regulates TOP mRNAs, through the DM15 motif [403]. The DM15 motif represents an identifying feature of most LARP1 family members, and is highly conserved in

[16]Roux and colleagues proposed a different model in which LARP1 functions as an activator (rather than an inhibitor) of TOP mRNA translation. See Ref. [373] for an alternative perspective on this subject.

evolution, bearing 50 % similarity at the amino acid level within 90 % of LARP1 family members [376]. Crystallization of the C-terminal DM15 region of human LARP1 (residues 796–946) revealed three structurally related helix-turn-helix modules that resemble the HEAT repeats found in TOR and RAPTOR [403]. The HEAT repeats in TOR and RAPTOR play important roles in protein-protein interaction [31]. In the case of LARP1, the HEAT repeats play a role in mRNA binding, specifically to the TOP motif within ribosomal mRNAs [403]. Lahr et al. [403] showed that the TOP motif is essential for the interaction of TOP mRNAs with the DM15 protein fragment (residues 796-946), in agreement with our earlier findings that the TOP motif is required for the interaction with full-length LARP1 protein [199]. Association of TOP mRNAs with the DM15 motif is presumed to regulate the translation and stability of this class of mRNAs (at least in higher eukaryotes).

LARP1 and the Control of TOP mRNA Translation in Evolution

Is the role of LARP1 in the control of TOP mRNA functions conserved in eukaryotic evolution? The short answer is "no." TOP mRNAs are present in higher eukaryotes, namely vertebrates, insects (*Drosophila melanogaster*), plants (*Arabidopsis thaliana*) and mollusks (*Aplysia californica*), but thus far the TOP motif has not been reported in yeasts or nematodes [386]. In budding yeast, the synthesis of ribosomal proteins appears to be primarily controlled at the transcriptional level [386, 402]. Interestingly, the DM15 motif (proposed by Berman and colleagues to mediate interaction with the TOP motif) is conspicuously absent from Sro9 and Slf1 in budding yeast [376], concurrent with the finding that lower eukaryotes do not have TOP mRNAs. We speculate that LARP1 may have been adapted for alternative use in the course of eukaryotic evolution by the process of exaptation. LARP1 is a newly identified target of mTORC1, and much remains to be learned about this protein. Nonetheless, these first few studies [199, 373, 374, 403] provide compelling evidence for a link among mTORC1, LARP1 and TOP mRNA regulation in higher eukaryotes. We believe that LARP1 plays a fundamental role in TOP mRNA control. Our findings on LARP1 do not exclude the existence of other *trans*-acting TOP mRNA-regulating factors, nor do they provide formal evidence for LARP1 being *the* long sought *trans*-acting TOP mRNA repressor [391]. Additional work is required to fully understand the mechanism of TOP mRNA translation control.

So far, we have established that LARP1 plays an important role in TOP mRNA control downstream of mTORC1. Growing evidence suggests that LARP1 binds and regulates many other cellular mRNAs, including the mTOR mRNA itself [404]. Blagden and colleagues [404] have recently shown that LARP1 binds to mTOR mRNA and protects it from degradation, indicating that feedback regulatory loops from LARP1 to mTOR also exist. The positive role of LARP1 in the control of mTOR mRNA stability is entirely consistent with the role of LARP1 in stabilizing TOP mRNAs [199, 374]. Further work is also required to fully elucidate the feedback loop from LARP1 to mTOR.

10.6 Concluding Remarks

mTOR (specifically mTORC1) is an important regulator of protein synthesis. mTORC1 regulates the translation of a vast number, and perhaps the majority, of cellular mRNAs [295, 399]. Foremost among these are the TOP mRNAs [295, 399], which are remarkably sensitive to mTOR inhibition. mTORC1 has also been implicated in the translation control of other mRNAs involved in cell-cycle regulation and metabolism [179, 405, 406], though the effect of mTOR inhibitors on the translation of these mRNAs is less pronounced than their effect on TOP mRNAs.[17] 4E-BPs and S6Ks are easily the most intensely studied and, therefore, the best understood targets of mTORC1. 4E-BPs exert their effects on mRNA translation through the disassembly of the eIF4F initiation complex, while S6Ks modulate protein synthesis through the phosphorylation of multiple proteins (including initiation factors, ribosomal proteins, RNA-binding proteins and other protein kinases; reviewed in this section). It is indisputable that both 4E-BPs and S6Ks play seminal roles in translation control. There is, however, a growing appreciation for the existence of numerous other mTORC1 targets with equally important roles in protein synthesis [298, 300]. In this section, we reviewed two novel effectors (PRAS40 and LARP1) of mTORC1 that have important roles in translation control. PRAS40 regulates the assembly of the i-proteasome in conditions of reduced translation fidelity [370], while LARP1 suppresses TOP mRNA translation in conditions of nutrient deprivation [199].

11 The Exaptation of mTOR in Learning and Memory

Throughout this chapter we have explored the ample evidence that mTOR functions as a central regulator of protein synthesis that controls growth at both the organismal and cellular levels. This regulatory role is also observed in the brain, where mTOR controls the overall growth pattern of differentiating neuronal stem cells. In addition to controlling growth of neuronal stem cells, mTOR also promotes dendritic and axonal arborization, alters the balance of synaptic transmission in post-mitotic neurons and modulates the immune function of glial cells. These facets of mTOR regulation are more extensively reviewed elsewhere [407, 408]. In this section we will explore how mTORC1 exaptated from a regulator of growth in the last eukaryotic common ancestor to an enzymatic complex that orchestrates the mRNA translation that is necessary for learning and memory in the post-mitotic neuronal networks of the mammalian brain.

[17]The translation of cell-cycle-related genes has been argued to be less sensitive to acute mTOR inhibition than TOP mRNAs [295]. This subject is discussed further herein [99].

11.1 The Biochemistry of Learning and Memory

Synapses—the labile, physical connections between neurons—are key to understanding the molecular basis for the most popular theory of how we learn, store and recall information. This theory, first expounded by the Canadian psychologist Donald Hebb in 1949, tells us that persistent activity in synapsing neurons will strengthen their connections [409]. The so-called synaptic plasticity and memory (SPM) hypothesis extends the original Hebbian theory to account for a more complex landscape of synaptic plasticity found after decades of research and posits that activity-dependent changes in synaptic strength are both necessary and sufficient for information storage in Metazoan nervous systems (e.g., neurons in the human brain) [410]. Given that synaptic plasticity (the strengthening or weakening of a neuronal connection) is a cellular phenomenon, it must have a biochemical origin. Thus, if we can understand synapse biochemistry, we can begin to grasp how learning and memory are encoded in the brain. From an evolutionary perspective, there are two competing models of how neurons emerged. Phylogenetic analysis of synaptic proteomes suggests that protein complexes, including mTOR and components of the RNA translation machinery, were present in unicellular fungi in the form of protosynapse [411]. This model can also be extended to the earliest synapse (ursynapse) between cells, which is thought to have arisen before the emergence of axonal and dendritic extensions in neurons. However, in sponges and the simpler *Placozoa*, neurons and many neurotransmitters are absent [412]. The single-origin hypothesis of Grant thus dictates that these organisms lost their synapses, an unlikely possibility. Another model espoused by Moroz and Kohn is perhaps more parsimonious with convergent evolutionary theory; it proposes that neurons (and therefore synapses) evolved independently numerous times throughout evolution [412]. In either case, biochemical pathways that rely on translation control were present in synaptic ancestors, and their compartmentalization likely led to the specialization of mRNA translation that is distinct from the rest of the cell.

Multiple dendritic spines on a single neuron can receive input from a large number of afferent neurons. How then does the receiving neuron integrate such a large number of signals in quick succession? Translation control and neuronal architecture likely represent key adaptations, as the highly compartmentalized nature of dendritic spines allows not only temporal but also spatial gradients to form as a result of local protein synthesis. This allows messages present in far-reaching axons and dendrites to be translated independently of those closer to the nucleus in the neuronal soma. A unique feature of neuronal cells is the presence of large RNA transport complexes called neuronal RNA granules that are repressed at the level of translation (reviewed in [413]). This repression allows mRNA transport to and storage within dendritic spines and allows for localized translation upon an appropriate stimulus. This particular type of translation occurs during synaptic plasticity and learning [414–417]. It follows that all of the critical translation control components must be present at the synapse for synaptic mRNAs to be translated on-site. This is indeed the case, and mTORC1 as well as mTORC2 activities have

been reported in synaptic compartments [415, 418]. Indeed, stimulus-specific, localized translation has been visualized in live *Aplysia californica* (sea slug, a mollusk that has contributed greatly to our knowledge of the evolutionary origins of synaptic plasticity) and rodent hippocampal[18] neurons using fluorescent reporters [420, 421]. It is not known whether mTORC1- (or mTORC2-) dependent signaling removes this translation repression or instead simply stimulates translation following the removal of the translation repression device.

Changes in synaptic strength are necessarily mediated by biochemical pathways in both the signaling (pre-synaptic) and receiving (post-synaptic) neurons, which become activated with the appropriate stimulus (e.g., neurotransmitters secreted at a synaptic bouton that bind to receptors on a receiving dendritic spine). This process is essentially the molecular correlate of learning, and results in a molecular memory trace—a change in the local synaptic proteome that fundamentally alters the properties of that particular synapse. While changes in the synaptic proteome are not absolutely required for some forms of memory (e.g., short-term memory), it is clear that consolidation of long-term memory (LTM) in rodents is mediated by the synthesis of new proteins; as injection of protein synthesis inhibitors into hippocampal areas has long been known to prevent LTM [422]. The husband and wife team of Flexner and Flexner would go on to firmly establish this seminal finding before the end of the 1960s and are pioneers in studying memory as a biochemical phenomenon. More recent work using fear conditioning to explore LTM has shown that the biochemical traces leading to changes in synaptic plasticity are volatile. Instead of LTM being permanently stored, it is continually recalled and reconsolidated (re-encoded), and this process can be blocked by intra-hippocampal infusion of the protein synthesis inhibitor anisomycin [423, 424]. Critically, mTORC1 lies upstream of this phenomenon as rapamycin has been found to block amygdyla-, hippocampal- and gustatory cortex-dependent LTM [425–434].

11.2 The Requirement for Translation in Cellular Models of Synaptic Plasticity Across Phyla

The first evidence of long-lasting, Hebbian synaptic plasticity in the hippocampus was published by Timothy Bliss and Terje Lømo [435]. They were able to elicit long-lasting (hours) increases in the efficiency of transmission at synapses (i.e., strengthening) on granule cells within the rabbit dentate gyrus (DG, a substructure within the hippocampus) following a specific train of electrical stimuli. This phenomenon was eventually referred to as long-term potentiation (LTP) and is one of

[18]The importance of the hippocampus in consolidating new memories was described in a seminal case report by the neurosurgeon William Scoville and the psychologist Brenda Milner (a former graduate student of Donald Hebb) in 1957. A historical perspective of "Patient HM," whose hippocampi had been almost completely surgically resected has been recently published and is a highly recommended read [419].

the most extensively studied cellular correlates of LTM. At the same time, Eric Kandel was dissecting the relatively simple and easily manipulated neuronal circuitry of the mollusk *Aplysia californica* (individual neuronal processes in this sea slug can be seen by the human eye) and discovered a similar long-lasting form of synaptic plasticity that correlates with behavioral LTM in the animal (sensitization of a defensive withdrawal reflex of the animal's feeding siphon) called long-term facilitation (LTF) [436]. Perhaps not unexpectedly, protein synthesis inhibitors were later found to block LTP in rodents [437] and LTF in *Aplysia californica* [438], demonstrating that this property shares commonalities with LTM. This suggests that lower animals, such as *Aplysia californica,* retain much of the protein synthesis-dependent memory-making machinery and is consistent with a view in which translational control was a salient feature of gene expression that was co-opted in ancient synapses. It is important to note here that an opposing form of plasticity called long-term depression (LTD) has also been observed [439]. A protein synthesis-dependent form of LTD (mGluR-LTD) was discovered more recently, and we will visit this in more detail below [440].

11.3 mTOR Is an Ancient Regulator of Synaptic Plasticity

The vertebrate brain represents a complex system of neural circuits and heterogeneous neuronal niches. Synaptic plasticity itself is more complex still, with different phases present in both LTP and LTD (an early, protein synthesis-independent phase that is followed by a late, protein synthesis-dependent phase). Different neuronal sub-populations, such as those that reside in the DG versus those in the CA1 region of the hippocampus, appear to have different substrate requirements in LTP and LTD. mTORC1 is required in the late-phase of LTP in hippocampal pyramidal neurons (CA1 region), since this form of plasticity can be blocked with rapamycin treatment [441]. However, late-phase LTP in the hippocampal granule neurons of the DG (where LTP was first elicited) does not require mTORC1 as rapamycin treatment does not block the effect [442]. In this context, protein synthesis-dependent plasticity instead works through activation of the TrkB-Mnk-eIF4e pathway mediated by the neurotropic factor BDNF [443].

We speculate that these differential requirements may reflect the evolutionary origins of the DG in relation to the rest of the hippocampus. The DG evolved late, being (arguably) absent in non-mammals, while analogous regions to the CA fields and subiculum of the hippocampus are present in the pallial areas of reptiles, fish and birds [444]. Critically, LTF in *Aplysia californica* can be blocked with rapamycin [438, 445], supporting the view that mTORC1-dependent LTP was present in an ancestor that lacked a DG and that LTP in this neuronal context evolved a different molecular mechanism in vertebrates.

It stands to reason that the BDNF-Mnk-dependent plasticity observed in the DG might reflect different types and/or amounts of mRNA translation targets between DG granule cells and pyramidal cells of the hippocampus. Indeed, BDNF-TrkB-mediated

LTP can be elicited in hippocampal pyramidal cells but is not always required for protein synthesis-dependent plasticity [446]. The DG is the only site within the hippocampus capable of adult neurogenesis [447]. It is conceivable that immature granule cells that are unique to this part of the hippocampus integrate input from parallel Mnk and mTORC1 arms of translation control such that only Mnk activity is necessary for consolidation of LTP. Consistent with this, immature granule cells have been found to exhibit enhanced synaptic plasticity, with a lowered threshold to potentiation [448]. Interestingly, BDNF and TrkB receptor orthologs have also been identified in *Aplysia californica* where they reduce the threshold and increase the magnitude of plasticity, suggesting that a common ancestor may have shared a similar system to modulate plasticity [449].

11.4 mTORC1 Effectors in Different Phases of Learning and Memory

Exactly how mTORC1 induces plasticity via changes in mRNA translation is still very much an open question. It could act to cause a general increase in protein synthesis or alter the translation of specific mRNA populations. The latter argument is perhaps more likely given what we now know about how mTORC1 regulates translation in other cellular contexts (e.g., TOP mRNA translation in cancer cells; described in Sect. 10 of this chapter) and emerging data in the context of the brain. Given the complexity of the protein synthesis signaling network downstream of mTORC1, it should not be surprising that dissection of the relative mTORC1 effectors in the context of plasticity has been a daunting task. For example, S6K1 and S6K2, are both expressed in the brain and have distinct yet overlapping functions. Both *Rps6k1* and *Rps6k2* knockout mice exhibit normal translation-dependent LTP but only the *Rps6k1* knockout animals exhibited impaired translation-independent LTP [450]. Differences are even more obvious at the behavioral level, with *Rps6k1*-deficient mice expressing deficits in a variety of classical learning and memory paradigms that depend on the hippocampus such as contextual fear memory, conditioned taste aversion and spatial learning (Morris water maze). *Rps6k2*-deficient mice however displayed deficits only in more specific, temporal aspects of contextual fear memory and conditioned taste aversion while their spatial learning remained unaffected.

Several major complications restrict further dissection of the roles of S6K1/S6K2: firstly, double knockout mice are embryonic lethal, and secondly, the loss of one gene could be compensated by upregulation of the remaining one. The recent development of specific small molecule inhibitors of S6 kinase may help in this regard [451]. Using a small molecule inhibitor specific for the S6K1, Huynh et al. found that inhibition of S6K1 alone was not sufficient to reduce LTM in a reconsolidation model [452]. This model of LTM is well known to be mTORC1-dependent [428, 453]. Surprisingly, inhibition of cap-dependent translation by injecting 4EG-I, an inhibitor of eIF4G activity that blocks the formation of

the cap-binding complex, together with a small molecule S6K1 inhibitor, impaired consolidation of fear memories [452]. This suggests that the plasticity required for these memories needs translation of mRNAs, which is mediated by both eIF4F and S6K1. However, combined S6K1 and eIF4F inhibition are not sufficient to completely block memory reconsolidation, which may be in part due to other mechanisms, including the involvement of S6K2. S6K1 may be working to target the translation of mRNAs that are eIF4A-dependent as S6K1 has been shown to alter this activity through eIF4B [335], or alternatively by modulating translation elongation.

Indeed, there appears to be an important, yet under-appreciated role for elongation in synaptic plasticity, since either increasing or decreasing eEF2K, the only known kinase that regulates eEF2 activity, affects memory formation [454, 455]. In *Aplysia californica,* eEF2 acts as an eEF2K-dependent biochemical sensor that can activate distinct modes of translation regulation depending on the type of synaptic input [456]. Synaptic plasticity that works through activation of metabotropic glutamate receptors (mGluR-LTD) is a curious example of how mTORC1-dependent translation has adapted to include regulation at the elongation step. Indeed, the protein synthesis-dependent phase of mGluR-LTD has recently been shown to occur independently of translation initiation [457]. This appears to be mediated in part by FMRP, a protein whose expression is silenced in a severe monogenic form of autism called fragile X syndrome that leads to exaggerated synaptic plasticity that is no longer mTORC1-dependent [458]. FMRP acts to slow or stall elongating ribosomes on certain mRNAs involved in mediating mGluR-LTD [459]. Interestingly, there is mounting evidence that mRNAs targeted by FMRP might already be loaded with ribosomes at the synapse, where they lie in wait for the appropriate stimulus, stalled at the level of elongation [457, 460]. The ability of FMRP to repress translation of its target mRNAs is mediated by its phosphorylation [461, 462], which is rapamycin-sensitive and abolished in *Rps6k1*$^{-/-}$ hippocampal neurons [463]. FMRP phosphorylation, and thus translation repression activity, are transiently removed by protein phosphatase 2A (PP2A) activity upon stimulation of mGluRs, after which mTORC1-S6K1 signaling allows for re-phosphorylation of FMRP and a return to translation repression [464]. This attractive model of mTORC1-dependent translation, in which mGluR-dependent transient activation of PP2A releases the FMRP brake, allows for a high degree of control over how much LTD protein is produced. Moreover, since this model requires two biochemical inputs (mTORC1 and mGluR activity), it can distinguish between different synaptic inputs that could dictate whether LTP or LTD proteins are synthesized. Were similar mechanisms present in the last eukaryotic common ancestor? It is likely so, as *Aplysia californica* has an FMRP ortholog that is subject to phosphatase regulation, and it is involved in a specific protein synthesis-dependent, neuropeptide-induced LTD [465].

Other mTORC1 effectors, such as 4E-BPs, also play a role in learning and memory. In the adult brain, 4E-BP2 is the most abundant isoform, with little detectable 4E-BP1 or 4E-BP3 [369, 466]. *Eif4ebp2*$^{-/-}$ mice show impairments in late LTP, and deficits in hippocampal-dependent LTM, likely because of the enhanced

translation of mRNA(s) that inhibit this form of plasticity [466]. Conversely, mGluR-LTD is enhanced in *Eif4ebp2*$^{-/-}$ hippocampal neurons and is no longer sensitive to rapamycin, demonstrating the requirement for mTORC1/4E-BP2 signaling in this type of plasticity [467]. In *Aplysia californica,* TOR/4E-BP signaling does not appear to participate in LTF (while TOR/S6K does), as expression of a 4E-BP dominant negative mutant lacking the TOR signaling motif fails to elicit this type of plasticity [468]. 4E-BP appears to preferentially regulate translation of TOP mRNAs [295, 399], although the TOR/4E-BP may not be the main regulator of TOP mRNA translation [469]. Interestingly, TOP mRNAs are some of the more abundant localized transcripts in both *Aplysia californica* and rodent hippocampal neurons [470, 471], and at least one TOP mRNA (although not directly linked to the presence of TOP) has been shown to be translated locally at synapses during LTP [472]. Subsequent work using TOP reporters also demonstrates that mTORC1-dependent upregulation of TOP mRNA translation occurs during LTP [473]. How TOP mRNAs (which encode ribosomal proteins and translation factors) are altering synaptic plasticity remains to be elucidated. One possibility is that they modulate the translational efficiency of LTP or LTD target mRNAs by promoting ribosome specialization [474, 475]. Alternatively, they may possess extra-ribosomal functions [476].

11.5 The Emerging Role of mTORC2 in Learning and Memory

We have not yet touched on how mTORC2 contributes to synaptic plasticity. In fact, very little is known about how mTORC2 might regulate learning and memory, but we do know that is is indeed involved. mTORC2-specific inhibitors do not yet exist, and knocking out *Rictor* leads to developmental defects, confounding the study of learning and memory in these mice [477, 478]. To get around this issue, Huang et al. cleverly crossed the *Rictor*$^{-/-}$ mouse with a Cre recombinase strain driven by a *Camk2a* promoter, thus restricting the ablation of mTORC2 activity to excitatory neurons in the limbic and cortical areas of the brain after development [418]. Using this conditional knockout, they were able to convincingly show that protein synthesis dependent LTP is impaired in these mice and that this is due to a lack of mTORC2-specific Akt phosphorylation. This appeared to be mediated by activation of actin polymerization, as chemical induction of this process using jasplakinolide rescued LTP in the knockout mouse. Importantly, mice deficient in mTORC2 displayed impairments in LTM, and this mechanism appears to be evolutionarily conserved [418], since removal of *Rictor* in *Drosophila melanogaster* affected (protein synthesis-dependent) LTM.

11.6 Concluding Remarks

We have seen that mRNA translation downstream of mTORC1 is necessary for the types of synaptic plasticity that represent cellular correlates of learning and memory. These core mechanisms have likely been present since the last eukaryotic common ancestor, as they are conserved in both vertebrates and in *Aplysia californica*, although the receptors and neurotransmitters used appear to have diverged. This divergence is perhaps most striking in the dentate gyrus, a recent evolutionary adaptation in vertebrates where mTOR is not required for LTP. It is readily apparent that determining which mRNAs are being translated in these different types of plasticity and memory formation will create a more complete picture of how changes in synaptic strength are achieved as memories are formed and re-formed. Advances in techniques to assess mRNA translation and protein synthesis in synaptic compartments will hopefully achieve a clearer understanding of how mTOR can modify the local synaptic proteome to elicit long-lasting synaptic plasticity—an elegant exaptation of mTOR from its more traditional role in cellular growth and differentiation.

12 Conclusion and Future Perspectives

In the last 25 years since the original discovery of TOR in yeast, researchers have amassed a breadth of information about this fascinating central regulator of cell growth using the most varied model organisms to elucidate the intricate complexity of the TOR pathway throughout evolution. We can now appreciate that TOR plays a fundamental role in the control of many cellular functions, some of which are highly conserved throughout the evolution of the eukaryotic lineage. Protein synthesis is by far the best-understood (and certainly a very important) output of TOR. Today, we are beginning to appreciate the intricate mechanisms by which TOR regulates protein synthesis. However, it would be naïve on our part to assume that we have unraveled all (or even the majority) of the mechanisms that TOR employs to control protein synthesis. The beauty of TOR signaling in the control of protein synthesis lies, for the most part, in its layered complexity rather than our often somewhat oversimplified, reductionist models. In this chapter, we have reviewed some of the earliest findings linking TOR to protein synthesis (through 4E-BPs and S6Ks) and some of our latest research on novel direct substrates of TOR in mRNA translation, e.g., LARP1. Having witnessed a great deal of advancement in our understanding of the TOR pathway in the last quarter of a century, we now look forward to the next 25 years with great excitement about the discoveries still to come.

Acknowledgments We wish to thank all of the scientists in the TOR field for their contributions to our present understanding of the TOR pathway. We apologize to the scientists whose work was not acknowledged or discussed in further detail, in particular those who have contributed to the elucidation of the signaling pathways up- and downstream of TOR that regulate cellular functions

other than mRNA translation. The authors wish to thank the various funding agencies that have contributed to research in our laboratories over the many years of the study of TOR. The authors also wish to thank the following funding agencies that currently fund their research: A.G. and M. N.H. acknowledge support from the Louis Jeantet Foundation, the Swiss National Science Foundation and the Canton of Basel. B.D.F. and T.A. gratefully acknowledge financial support from Prostate Cancer Canada. The laboratories of C.M. and C.R. are supported by ANR grants TRANSLATOR and DECORATOR; CR is also supported by the A*MIDEX project (no. ANR-11-IDEX-0001-02).

References

1. Sehgal SN, Baker H, Vezina C. Rapamycin (AY-22,989), a new antifungal antibiotic. II. Fermentation, isolation and characterization. J Antibiot. 1975;28(10):727–32.
2. Vezina C, Kudelski A, Sehgal SN. Rapamycin (AY-22,989), a new antifungal antibiotic. I. Taxonomy of the producing streptomycete and isolation of the active principle. J Antibiot. 1975;28(10):721–6.
3. Baker H, Sidorowicz A, Sehgal SN, Vezina C. Rapamycin (AY-22,989), a new antifungal antibiotic. III. In vitro and in vivo evaluation. J Antibiot. 1978;31(6):539–45.
4. Bierer BE, Mattila PS, Standaert RF, Herzenberg LA, Burakoff SJ, Crabtree G, et al. Two distinct signal transmission pathways in T lymphocytes are inhibited by complexes formed between an immunophilin and either FK506 or rapamycin. Proc Natl Acad Sci USA. 1990;87 (23):9231–5.
5. Dumont FJ, Melino MR, Staruch MJ, Koprak SL, Fischer PA, Sigal NH. The immunosuppressive macrolides FK-506 and rapamycin act as reciprocal antagonists in murine T cells. J Immunol. 1990;144(4):1418–24.
6. Dumont FJ, Staruch MJ, Koprak SL, Melino MR, Sigal NH. Distinct mechanisms of suppression of murine T cell activation by the related macrolides FK-506 and rapamycin. J Immunol. 1990;144(1):251–8.
7. Heitman J, Movva NR, Hall MN. Targets for cell cycle arrest by the immunosuppressant rapamycin in yeast. Science. 1991;253(5022):905–9.
8. Kunz J, Henriquez R, Schneider U, Deuter-Reinhard M, Movva NR, Hall MN. Target of rapamycin in yeast, TOR2, is an essential phosphatidylinositol kinase homolog required for G1 progression. Cell. 1993;73(3):585–96.
9. Brown EJ, Albers MW, Shin TB, Ichikawa K, Keith CT, Lane WS, et al. A mammalian protein targeted by G1-arresting rapamycin-receptor complex. Nature. 1994;369(6483):756–8.
10. Sabatini DM, Erdjument-Bromage H, Lui M, Tempst P, Snyder SH. RAFT1: a mammalian protein that binds to FKBP12 in a rapamycin-dependent fashion and is homologous to yeast TORs. Cell. 1994;78(1):35–43.
11. Chiu MI, Katz H, Berlin V. RAPT1, a mammalian homolog of yeast Tor, interacts with the FKBP12/rapamycin complex. Proc Natl Acad Sci USA. 1994;91(26):12574–8.
12. Sabers CJ, Martin MM, Brunn GJ, Williams JM, Dumont FJ, Wiederrecht G, et al. Isolation of a protein target of the FKBP12-rapamycin complex in mammalian cells. J Biol Chem. 1995;270(2):815–22.
13. Weisman R, Choder M. The fission yeast TOR homolog, tor1+, is required for the response to starvation and other stresses via a conserved serine. J Biol Chem. 2001;276(10):7027–32.
14. Kawai M, Nakashima A, Ueno M, Ushimaru T, Aiba K, Doi H, et al. Fission yeast tor1 functions in response to various stresses including nitrogen starvation, high osmolarity, and high temperature. Curr Genet. 2001;39(3):166–74.
15. Oldham S, Montagne J, Radimerski T, Thomas G, Hafen E. Genetic and biochemical characterization of dTOR, the Drosophila homolog of the target of rapamycin. Genes Dev. 2000;14:2689–94.

16. Zhang H, Stallock JP, Ng JC, Reinhard C, Neufeld TP. Regulation of cellular growth by the Drosophila target of rapamycin dTOR. Genes Dev. 2000;14:2712–24.
17. Long X, Spycher C, Han ZS, Rose AM, Muller F, Avruch J. TOR deficiency in C. elegans causes developmental arrest and intestinal atrophy by inhibition of mRNA translation. Curr Biol: CB. 2002;12(17):1448–61.
18. Menand B, Desnos T, Nussaume L, Berger F, Bouchez D, Meyer C, et al. Expression and disruption of the Arabidopsis TOR (target of rapamycin) gene. Proc Natl Acad Sci USA. 2002;99(9):6422–7.
19. Laplante M, Sabatini DM. mTOR signaling in growth control and disease. Cell. 2012;149 (2):274–93.
20. Laplante M, Sabatini DM. mTOR signaling at a glance. J Cell Sci. 2009;122(Pt 20):3589–94.
21. Cornu M, Albert V, Hall MN. mTOR in aging, metabolism, and cancer. Curr Opin Genet Dev. 2013;23(1):53–62.
22. Serfontein J, Nisbet RE, Howe CJ, de Vries PJ. Evolution of the TSC1/TSC2-TOR signaling pathway. Sci Signaling. 2010;3(128):ra49.
23. West AH, Stock AM. Histidine kinases and response regulator proteins in two-component signaling systems. Trends Biochem Sci. 2001;26(6):369–76.
24. Klumpp S, Krieglstein J. Phosphorylation and dephosphorylation of histidine residues in proteins. Eur J Biochem/FEBS. 2002;269(4):1067–71.
25. Helliwell SB, Wagner P, Kunz J, Deuter-Reinhard M, Henriquez R, Hall MN. TOR1 and TOR2 are structurally and functionally similar but not identical phosphatidylinositol kinase homologues in yeast. Mol Biol Cell. 1994;5(1):105–18.
26. Lempiainen H, Halazonetis TD. Emerging common themes in regulation of PIKKs and PI3Ks. EMBO J. 2009;28(20):3067–73.
27. Wang X, Proud CG. mTORC2 is a tyrosine kinase. Cell Res. 2016;26(2):266.
28. Yin Y, Hua H, Li M, Liu S, Kong Q, Shao T, et al. mTORC2 promotes type I insulin-like growth factor receptor and insulin receptor activation through the tyrosine kinase activity of mTOR. Cell Res. 2016;26(1):46–65.
29. Bosotti R, Isacchi A, Sonnhammer EL. FAT: a novel domain in PIK-related kinases. Trends Biochem Sci. 2000;25(5):225–7.
30. Yang H, Rudge DG, Koos JD, Vaidialingam B, Yang HJ, Pavletich NP. mTOR kinase structure, mechanism and regulation. Nature. 2013;497(7448):217–23.
31. Aylett CH, Sauer E, Imseng S, Boehringer D, Hall MN, Ban N, et al. Architecture of human mTOR complex 1. Science. 2016;351(6268):48–52.
32. Barbet NC, Schneider U, Helliwell SB, Stansfield I, Tuite MF, Hall MN. TOR controls translation initiation and early G1 progression in yeast. Mol Biol Cell. 1996;7(1):25–42.
33. Zheng XF, Florentino D, Chen J, Crabtree GR, Schreiber SL. TOR kinase domains are required for two distinct functions, only one of which is inhibited by rapamycin. Cell. 1995;82(1):121–30.
34. Schmidt A, Kunz J, Hall MN. TOR2 is required for organization of the actin cytoskeleton in yeast. Proc Natl Acad Sci USA. 1996;93(24):13780–5.
35. Helliwell SB, Schmidt A, Ohya Y, Hall MN. The Rho1 effector Pkc1, but not Bni1, mediates signalling from Tor2 to the actin cytoskeleton. Curr Biol: CB. 1998;8(22):1211–4.
36. Helliwell SB, Howald I, Barbet N, Hall MN. TOR2 is part of two related signaling pathways coordinating cell growth in Saccharomyces cerevisiae. Genetics. 1998;148(1):99–112.
37. Rispal D, Eltschinger S, Stahl M, Vaga S, Bodenmiller B, Abraham Y, et al. Target of rapamycin complex 2 regulates actin polarization and endocytosis via multiple pathways. J Biol Chem. 2015;290(24):14963–78.
38. Shimada K, Filipuzzi I, Stahl M, Helliwell SB, Studer C, Hoepfner D, et al. TORC2 signaling pathway guarantees genome stability in the face of DNA strand breaks. Mol Cell. 2013;51(6):829–39.
39. Panasyuk G, Nemazanyy I, Zhyvoloup A, Filonenko V, Davies D, Robson M, et al. mTORbeta splicing isoform promotes cell proliferation and tumorigenesis. J Biol Chem. 2009;284(45):30807–14.

40. Loewith R, Jacinto E, Wullschleger S, Lorberg A, Crespo JL, Bonenfant D, et al. Two TOR complexes, only one of which is rapamycin sensitive, have distinct roles in cell growth control. Mol Cell. 2002;10(3):457–68.
41. Kim DH, Sarbassov DD, Ali SM, King JE, Latek RR, Erdjument-Bromage H, et al. mTOR interacts with raptor to form a nutrient-sensitive complex that signals to the cell growth machinery. Cell. 2002;110(2):163–75.
42. Hara K, Maruki Y, Long X, Yoshino K, Oshiro N, Hidayat S, et al. Raptor, a binding partner of target of rapamycin (TOR), mediates TOR action. Cell. 2002;110(2):177–89.
43. Reinke A, Anderson S, McCaffery JM, Yates J 3rd, Aronova S, Chu S, et al. TOR complex 1 includes a novel component, Tco89p (YPL180w), and cooperates with Ssd1p to maintain cellular integrity in Saccharomyces cerevisiae. J Biol Chem. 2004;279(15):14752–62.
44. Kim DH, Sarbassov DD, Ali SM, Latek RR, Guntur KV, Erdjument-Bromage H, et al. GbetaL, a positive regulator of the rapamycin-sensitive pathway required for the nutrient-sensitive interaction between raptor and mTOR. Mol Cell. 2003;11(4):895–904.
45. Sarbassov DD, Ali SM, Kim DH, Guertin DA, Latek RR, Erdjument-Bromage H, et al. Rictor, a novel binding partner of mTOR, defines a rapamycin-insensitive and raptor-independent pathway that regulates the cytoskeleton. Curr Biol: CB. 2004;14 (14):1296–302.
46. Jacinto E, Loewith R, Schmidt A, Lin S, Ruegg MA, Hall A, et al. Mammalian TOR complex 2 controls the actin cytoskeleton and is rapamycin insensitive. Nat Cell Biol. 2004;6 (11):1122–8.
47. Jacinto E, Facchinetti V, Liu D, Soto N, Wei S, Jung SY, et al. SIN1/MIP1 maintains rictor-mTOR complex integrity and regulates Akt phosphorylation and substrate specificity. Cell. 2006;127(1):125–37.
48. Sarbassov DD, Ali SM, Sengupta S, Sheen JH, Hsu PP, Bagley AF, et al. Prolonged rapamycin treatment inhibits mTORC2 assembly and Akt/PKB. Mol Cell. 2006;22(2):159–68.
49. Loewith R, Hall MN. Target of rapamycin (TOR) in nutrient signaling and growth control. Genetics. 2011;189(4):1177–201.
50. De Virgilio C, Loewith R. The TOR signalling network from yeast to man. Int J Biochem cell Biol. 2006;38(9):1476–81.
51. Soulard A, Cohen A, Hall MN. TOR signaling in invertebrates. Curr Opin Cell Biol. 2009;21 (6):825–36.
52. Wullschleger S, Loewith R, Hall MN. TOR signaling in growth and metabolism. Cell. 2006;124(3):471–84.
53. Urban J, Soulard A, Huber A, Lippman S, Mukhopadhyay D, Deloche O, et al. Sch9 is a major target of TORC1 in Saccharomyces cerevisiae. Mol Cell. 2007;26(5):663–74.
54. Di Como CJ, Arndt KT. Nutrients, via the Tor proteins, stimulate the association of Tap42 with type 2A phosphatases. Genes Dev. 1996;10(15):1904–16.
55. Cybulski N, Hall MN. TOR complex 2: a signaling pathway of its own. Trends Biochem Sci. 2009;34(12):620–7.
56. Eltschinger S, Loewith R. TOR complexes and the maintenance of cellular homeostasis. Trends Cell Biol. 2016;26(2):148–59.
57. Kamada Y, Fujioka Y, Suzuki NN, Inagaki F, Wullschleger S, Loewith R, et al. Tor2 directly phosphorylates the AGC kinase Ypk2 to regulate actin polarization. Mol Cell Biol. 2005;25 (16):7239–48.
58. Roelants FM, Breslow DK, Muir A, Weissman JS, Thorner J. Protein kinase Ypk1 phosphorylates regulatory proteins Orm1 and Orm2 to control sphingolipid homeostasis in Saccharomyces cerevisiae. Proc Natl Acad Sci USA. 2011;108(48):19222–7.
59. Niles BJ, Mogri H, Hill A, Vlahakis A, Powers T. Plasma membrane recruitment and activation of the AGC kinase Ypk1 is mediated by target of rapamycin complex 2 (TORC2) and its effector proteins Slm1 and Slm2. Proc Natl Acad Sci USA. 2012;109(5):1536–41.

60. Beretta L, Gingras AC, Svitkin YV, Hall MN, Sonenberg N. Rapamycin blocks the phosphorylation of 4E-BP1 and inhibits cap-dependent initiation of translation. EMBO J. 1996;15(3):658–64.
61. Azpiazu I, Saltiel AR, DePaoli-Roach AA, Lawrence JC. Regulation of both glycogen synthase and PHAS-I by insulin in rat skeletal muscle involves mitogen-activated protein kinase-independent and rapamycin-sensitive pathways. J Biol Chem. 1996;271(9):5033–9.
62. Redpath NT, Foulstone EJ, Proud CG. Regulation of translation elongation factor-2 by insulin via a rapamycin-sensitive signalling pathway. EMBO J. 1996;15(9):2291–7.
63. Jefferies HB, Reinhard C, Kozma SC, Thomas G. Rapamycin selectively represses translation of the "polypyrimidine tract" mRNA family. Proc Natl Acad Sci USA. 1994;91(10):4441–5.
64. Jackson RJ, Hellen CU, Pestova TV. The mechanism of eukaryotic translation initiation and principles of its regulation. Nat Rev Mol Cell Biol. 2010;11(2):113–27.
65. Ma XM, Blenis J. Molecular mechanisms of mTOR-mediated translational control. Nat Rev Mol Cell Biol. 2009;10(5):307–18.
66. Cosentino GP, Schmelzle T, Haghighat A, Helliwell SB, Hall MN, Sonenberg N. Eap1p, a novel eukaryotic translation initiation factor 4E-associated protein in Saccharomyces cerevisiae. Mol Cell Biol. 2000;20(13):4604–13.
67. Altmann M, Schmitz N, Berset C, Trachsel H. A novel inhibitor of cap-dependent translation initiation in yeast: p20 competes with eIF4G for binding to eIF4E. EMBO J. 1997;16 (5):1114–21.
68. Zanchin NI, McCarthy JE. Characterization of the in vivo phosphorylation sites of the mRNA.cap-binding complex proteins eukaryotic initiation factor-4E and p20 in Saccharomyces cerevisiae. J Biol Chem. 1995;270(44):26505–10.
69. Berset C, Trachsel H, Altmann M. The TOR (target of rapamycin) signal transduction pathway regulates the stability of translation initiation factor eIF4G in the yeast Saccharomyces cerevisiae. Proc Natl Acad Sci USA. 1998;95(8):4264–9.
70. Huber A, Bodenmiller B, Uotila A, Stahl M, Wanka S, Gerrits B, et al. Characterization of the rapamycin-sensitive phosphoproteome reveals that Sch9 is a central coordinator of protein synthesis. Genes Dev. 2009;23(16):1929–43.
71. Soulard A, Cremonesi A, Moes S, Schutz F, Jeno P, Hall MN. The rapamycin-sensitive phosphoproteome reveals that TOR controls protein kinase A toward some but not all substrates. Mol Biol Cell. 2010;21(19):3475–86.
72. Raught B, Gingras AC, Gygi SP, Imataka H, Morino S, Gradi A, et al. Serum-stimulated, rapamycin-sensitive phosphorylation sites in the eukaryotic translation initiation factor 4GI. EMBO J. 2000;19(3):434–44.
73. Hinnebusch AG. Translational regulation of GCN4 and the general amino acid control of yeast. Annu Rev Microbiol. 2005;59:407–50.
74. Kubota H, Obata T, Ota K, Sasaki T, Ito T. Rapamycin-induced translational derepression of GCN4 mRNA involves a novel mechanism for activation of the eIF2 alpha kinase GCN2. J Biol Chem. 2003;278(23):20457–60.
75. Cherkasova VA, Hinnebusch AG. Translational control by TOR and TAP42 through dephosphorylation of eIF2alpha kinase GCN2. Genes Dev. 2003;17(7):859–72.
76. Vlahakis A, Graef M, Nunnari J, Powers T. TOR complex 2-Ypk1 signaling is an essential positive regulator of the general amino acid control response and autophagy. Proc Natl Acad Sci USA. 2014;111(29):10586–91.
77. Woolford JL Jr, Baserga SJ. Ribosome biogenesis in the yeast Saccharomyces cerevisiae. Genetics. 2013;195(3):643–81.
78. Jorgensen P, Rupes I, Sharom JR, Schneper L, Broach JR, Tyers M. A dynamic transcriptional network communicates growth potential to ribosome synthesis and critical cell size. Genes Dev. 2004;18(20):2491–505.
79. Powers T, Walter P. Regulation of ribosome biogenesis by the rapamycin-sensitive TOR-signaling pathway in Saccharomyces cerevisiae. Mol Biol Cell. 1999;10(4):987–1000.

80. Zaragoza D, Ghavidel A, Heitman J, Schultz MC. Rapamycin induces the G0 program of transcriptional repression in yeast by interfering with the TOR signaling pathway. Mol Cell Biol. 1998;18(8):4463–70.
81. Hardwick JS, Kuruvilla FG, Tong JK, Shamji AF, Schreiber SL. Rapamycin-modulated transcription defines the subset of nutrient-sensitive signaling pathways directly controlled by the Tor proteins. Proc Natl Acad Sci USA. 1999;96(26):14866–70.
82. Shamji AF, Kuruvilla FG, Schreiber SL. Partitioning the transcriptional program induced by rapamycin among the effectors of the Tor proteins. Curr Biol: CB. 2000;10(24):1574–81.
83. Cardenas ME, Cutler NS, Lorenz MC, Di Como CJ, Heitman J. The TOR signaling cascade regulates gene expression in response to nutrients. Genes Dev. 1999;13(24):3271–9.
84. Claypool JA, French SL, Johzuka K, Eliason K, Vu L, Dodd JA, et al. Tor pathway regulates Rrn3p-dependent recruitment of yeast RNA polymerase I to the promoter but does not participate in alteration of the number of active genes. Mol Biol Cell. 2004;15(2):946–56.
85. Philippi A, Steinbauer R, Reiter A, Fath S, Leger-Silvestre I, Milkereit P, et al. TOR-dependent reduction in the expression level of Rrn3p lowers the activity of the yeast RNA Pol I machinery, but does not account for the strong inhibition of rRNA production. Nucleic Acids Res. 2010;38(16):5315–26.
86. Laferte A, Favry E, Sentenac A, Riva M, Carles C, Chedin S. The transcriptional activity of RNA polymerase I is a key determinant for the level of all ribosome components. Genes Dev. 2006;20(15):2030–40.
87. Reiter A, Steinbauer R, Philippi A, Gerber J, Tschochner H, Milkereit P, et al. Reduction in ribosomal protein synthesis is sufficient to explain major effects on ribosome production after short-term TOR inactivation in Saccharomyces cerevisiae. Mol Cell Biol. 2011;31(4):803–17.
88. Chen H, Fan M, Pfeffer LM, Laribee RN. The histone H3 lysine 56 acetylation pathway is regulated by target of rapamycin (TOR) signaling and functions directly in ribosomal RNA biogenesis. Nucleic Acids Res. 2012;40(14):6534–46.
89. Li H, Tsang CK, Watkins M, Bertram PG, Zheng XF. Nutrient regulates Tor1 nuclear localization and association with rDNA promoter. Nature. 2006;442(7106):1058–61.
90. Lee TI, Rinaldi NJ, Robert F, Odom DT, Bar-Joseph Z, Gerber GK, et al. Transcriptional regulatory networks in Saccharomyces cerevisiae. Science. 2002;298(5594):799–804.
91. Hall DB, Wade JT, Struhl K. An HMG protein, Hmo1, associates with promoters of many ribosomal protein genes and throughout the rRNA gene locus in Saccharomyces cerevisiae. Mol Cell Biol. 2006;26(9):3672–9.
92. Berger AB, Decourty L, Badis G, Nehrbass U, Jacquier A, Gadal O. Hmo1 is required for TOR-dependent regulation of ribosomal protein gene transcription. Mol Cell Biol. 2007;27(22):8015–26.
93. Martin DE, Soulard A, Hall MN. TOR regulates ribosomal protein gene expression via PKA and the Forkhead transcription factor FHL1. Cell. 2004;119(7):969–79.
94. Schawalder SB, Kabani M, Howald I, Choudhury U, Werner M, Shore D. Growth-regulated recruitment of the essential yeast ribosomal protein gene activator Ifh1. Nature. 2004;432(7020):1058–61.
95. Wade JT, Hall DB, Struhl K. The transcription factor Ifh1 is a key regulator of yeast ribosomal protein genes. Nature. 2004;432(7020):1054–8.
96. Marion RM, Regev A, Segal E, Barash Y, Koller D, Friedman N, et al. Sfp1 is a stress- and nutrient-sensitive regulator of ribosomal protein gene expression. Proc Natl Acad Sci USA. 2004;101(40):14315–22.
97. Lempiainen H, Uotila A, Urban J, Dohnal I, Ammerer G, Loewith R, et al. Sfp1 interaction with TORC1 and Mrs6 reveals feedback regulation on TOR signaling. Mol Cell. 2009;33(6):704–16.
98. Chauvin C, Koka V, Nouschi A, Mieulet V, Hoareau-Aveilla C, Dreazen A, et al. Ribosomal protein S6 kinase activity controls the ribosome biogenesis transcriptional program. Oncogene. 2014;33(4):474–83.

99. Fonseca BD, Smith EM, Yelle N, Alain T, Bushell M, Pause A. The ever-evolving role of mTOR in translation. Seminars in cell & developmental biology. 2014.
100. Gonzalez A, Shimobayashi M, Eisenberg T, Merle DA, Pendl T, Hall MN, et al. TORC1 promotes phosphorylation of ribosomal protein S6 via the AGC kinase Ypk3 in Saccharomyces cerevisiae. PLoS ONE. 2015;10(3):e0120250.
101. Yerlikaya S, Meusburger M, Kumari R, Huber A, Anrather D, Costanzo M, et al. TORC1 and TORC2 work together to regulate ribosomal protein S6 phosphorylation in Saccharomyces cerevisiae. Mol Biol Cell. 2016;27(2):397–409.
102. Upadhya R, Lee J, Willis IM. Maf1 is an essential mediator of diverse signals that repress RNA polymerase III transcription. Mol Cell. 2002;10(6):1489–94.
103. Lee J, Moir RD, Willis IM. Regulation of RNA polymerase III transcription involves SCH9-dependent and SCH9-independent branches of the target of rapamycin (TOR) pathway. J Biol Chem. 2009;284(19):12604–8.
104. Lee J, Moir RD, Willis IM. Differential phosphorylation of RNA polymerase III and the initiation factor TFIIIB in Saccharomyces cerevisiae. PLoS ONE. 2015;10(5):e0127225.
105. Lee J, Moir RD, McIntosh KB, Willis IM. TOR signaling regulates ribosome and tRNA synthesis via LAMMER/Clk and GSK-3 family kinases. Mol Cell. 2012;45(6):836–43.
106. Gstaiger M, Luke B, Hess D, Oakeley EJ, Wirbelauer C, Blondel M, et al. Control of nutrient-sensitive transcription programs by the unconventional prefoldin URI. Science. 2003;302(5648):1208–12.
107. Bonfils G, Jaquenoud M, Bontron S, Ostrowicz C, Ungermann C, De Virgilio C. Leucyl-tRNA synthetase controls TORC1 via the EGO complex. Mol Cell. 2012;46(1):105–10.
108. Han JM, Jeong SJ, Park MC, Kim G, Kwon NH, Kim HK, et al. Leucyl-tRNA synthetase is an intracellular leucine sensor for the mTORC1-signaling pathway. Cell. 2012;149(2):410–24.
109. Bloom-Ackermann Z, Navon S, Gingold H, Towers R, Pilpel Y, Dahan O. A comprehensive tRNA deletion library unravels the genetic architecture of the tRNA pool. PLoS Genet. 2014;10(1):e1004084.
110. Zinzalla V, Stracka D, Oppliger W, Hall MN. Activation of mTORC2 by association with the ribosome. Cell. 2011;144(5):757–68.
111. Betz C, Stracka D, Prescianotto-Baschong C, Frieden M, Demaurex N, Hall MN. Feature Article: mTOR complex 2-Akt signaling at mitochondria-associated endoplasmic reticulum membranes (MAM) regulates mitochondrial physiology. Proc Natl Acad Sci USA. 2013;110(31):12526–34.
112. Liu P, Gan W, Chin YR, Ogura K, Guo J, Zhang J, et al. PtdIns(3,4,5)P3-Dependent Activation of the mTORC2 Kinase Complex. Cancer Discovery. 2015;5(11):1194–209.
113. Oh WJ, Wu CC, Kim SJ, Facchinetti V, Julien LA, Finlan M, et al. mTORC2 can associate with ribosomes to promote cotranslational phosphorylation and stability of nascent Akt polypeptide. EMBO J. 2010;29(23):3939–51.
114. Gaubitz C, Oliveira TM, Prouteau M, Leitner A, Karuppasamy M, Konstantinidou G, et al. Molecular basis of the rapamycin insensitivity of target of rapamycin complex 2. Mol Cell. 2015;58(6):977–88.
115. Hayashi T, Hatanaka M, Nagao K, Nakaseko Y, Kanoh J, Kokubu A, et al. Rapamycin sensitivity of the Schizosaccharomyces pombe tor2 mutant and organization of two highly phosphorylated TOR complexes by specific and common subunits. Genes Cells: Devoted Mol Cell Mech. 2007;12(12):1357–70.
116. Matsuo T, Otsubo Y, Urano J, Tamanoi F, Yamamoto M. Loss of the TOR kinase Tor2 mimics nitrogen starvation and activates the sexual development pathway in fission yeast. Mol Cell Biol. 2007;27(8):3154–64.
117. Alvarez B, Moreno S. Fission yeast Tor2 promotes cell growth and represses cell differentiation. J Cell Sci. 2006;119(Pt 21):4475–85.
118. Urano J, Sato T, Matsuo T, Otsubo Y, Yamamoto M, Tamanoi F. Point mutations in TOR confer Rheb-independent growth in fission yeast and nutrient-independent mammalian TOR signaling in mammalian cells. Proc Natl Acad Sci USA. 2007;104(9):3514–9.

119. Weisman R, Roitburg I, Schonbrun M, Harari R, Kupiec M. Opposite effects of tor1 and tor2 on nitrogen starvation responses in fission yeast. Genetics. 2007;175(3):1153–62.
120. Cohen A, Kupiec M, Weisman R. Glucose activates TORC2-Gad8 protein via positive regulation of the cAMP/cAMP-dependent protein kinase A (PKA) pathway and negative regulation of the Pmk1 protein-mitogen-activated protein kinase pathway. J Biol Chem. 2014;289(31):21727–37.
121. Hatano T, Morigasaki S, Tatebe H, Ikeda K, Shiozaki K. Fission yeast Ryh1 GTPase activates TOR Complex 2 in response to glucose. Cell Cycle. 2015;14(6):848–56.
122. Nakashima A, Otsubo Y, Yamashita A, Sato T, Yamamoto M, Tamanoi F. Psk1, an AGC kinase family member in fission yeast, is directly phosphorylated and controlled by TORC1 and functions as S6 kinase. J Cell Sci. 2012;125(Pt 23):5840–9.
123. Nakashima A, Sato T, Tamanoi F. Fission yeast TORC1 regulates phosphorylation of ribosomal S6 proteins in response to nutrients and its activity is inhibited by rapamycin. J Cell Sci. 2010;123(Pt 5):777–86.
124. Weisman R, Choder M, Koltin Y. Rapamycin specifically interferes with the developmental response of fission yeast to starvation. J Bacteriol. 1997;179(20):6325–34.
125. Takahara T, Maeda T. TORC1 of fission yeast is rapamycin-sensitive. Genes Cells: Devoted Mol Cell Mech. 2012;17(8):698–708.
126. Lamichhane TN, Blewett NH, Crawford AK, Cherkasova VA, Iben JR, Begley TJ, et al. Lack of tRNA modification isopentenyl-A37 alters mRNA decoding and causes metabolic deficiencies in fission yeast. Mol Cell Biol. 2013;33(15):2918–29.
127. Rodland GE, Tvegard T, Boye E, Grallert B. Crosstalk between the Tor and Gcn2 pathways in response to different stresses. Cell Cycle. 2014;13(3):453–61.
128. Valbuena N, Rozalen AE, Moreno S. Fission yeast TORC1 prevents eIF2alpha phosphorylation in response to nitrogen and amino acids via Gcn2 kinase. J Cell Sci. 2012;125(Pt 24):5955–9.
129. Petersen J, Nurse P. TOR signalling regulates mitotic commitment through the stress MAP kinase pathway and the Polo and Cdc2 kinases. Nat Cell Biol. 2007;9(11):1263–72.
130. Rexin D, Meyer C, Robaglia C, Veit B. TOR signalling in plants. Biochem J. 2015;470(1):1–14.
131. Xiong Y, Sheen J. The role of target of rapamycin signaling networks in plant growth and metabolism. Plant Physiol. 2014;164(2):499–512.
132. Deprost D, Yao L, Sormani R, Moreau M, Leterreux G, Nicolai M, et al. The Arabidopsis TOR kinase links plant growth, yield, stress resistance and mRNA translation. EMBO Rep. 2007;8(9):864–70.
133. Moreau M, Azzopardi M, Clement G, Dobrenel T, Marchive C, Renne C, et al. Mutations in the Arabidopsis homolog of LST8/GbetaL, a partner of the target of Rapamycin kinase, impair plant growth, flowering, and metabolic adaptation to long days. Plant Cell. 2012;24(2):463–81.
134. Crespo JL, Diaz-Troya S, Florencio FJ. Inhibition of target of rapamycin signaling by rapamycin in the unicellular green alga Chlamydomonas reinhardtii. Plant Physiol. 2005;139(4):1736–49.
135. Sormani R, Yao L, Menand B, Ennar N, Lecampion C, Meyer C, et al. Saccharomyces cerevisiae FKBP12 binds Arabidopsis thaliana TOR and its expression in plants leads to rapamycin susceptibility. BMC Plant Biol. 2007;7:26.
136. Ren M, Venglat P, Qiu S, Feng L, Cao Y, Wang E, et al. Target of rapamycin signaling regulates metabolism, growth, and life span in Arabidopsis. Plant Cell. 2012;24(12):4850–74.
137. Leiber RM, John F, Verhertbruggen Y, Diet A, Knox JP, Ringli C. The TOR pathway modulates the structure of cell walls in Arabidopsis. Plant Cell. 2010;22(6):1898–908.
138. Xiong Y, Sheen J. Rapamycin and glucose-target of rapamycin (TOR) protein signaling in plants. J Biol Chem. 2012;287(4):2836–42.
139. Schepetilnikov M, Kobayashi K, Geldreich A, Caranta C, Robaglia C, Keller M, et al. Viral factor TAV recruits TOR/S6K1 signalling to activate reinitiation after long ORF translation. EMBO J. 2011;30(7):1343–56.

140. Montane MH, Menand B. ATP-competitive mTOR kinase inhibitors delay plant growth by triggering early differentiation of meristematic cells but no developmental patterning change. J Exp Bot. 2013;64(14):4361–74.
141. Li L, Song Y, Wang K, Dong P, Zhang X, Li F, et al. TOR-inhibitor insensitive-1 (TRIN1) regulates cotyledons greening in Arabidopsis. Front Plant Sci. 2015;6:861.
142. Shemi A, Ben-Dor S, Vardi A. Elucidating the composition and conservation of the autophagy pathway in photosynthetic eukaryotes. Autophagy. 2015;11(4):701–15.
143. Anderson GH, Veit B, Hanson MR. The Arabidopsis AtRaptor genes are essential for post-embryonic plant growth. BMC Biol. 2005;3:12.
144. Deprost D, Truong HN, Robaglia C, Meyer C. An Arabidopsis homolog of RAPTOR/KOG1 is essential for early embryo development. Biochem Biophys Res Commun. 2005;326 (4):844–50.
145. Mahfouz MM, Kim S, Delauney AJ, Verma DP. Arabidopsis TARGET OF RAPAMYCIN interacts with RAPTOR, which regulates the activity of S6 kinase in response to osmotic stress signals. Plant Cell. 2006;18(2):477–90.
146. Ahn CS, Han JA, Lee HS, Lee S, Pai HS. The PP2A regulatory subunit Tap46, a component of the TOR signaling pathway, modulates growth and metabolism in plants. Plant Cell. 2011;23(1):185–209.
147. Ahn CS, Ahn HK, Pai HS. Overexpression of the PP2A regulatory subunit Tap46 leads to enhanced plant growth through stimulation of the TOR signalling pathway. J Exp Bot. 2015;66(3):827–40.
148. Xiong Y, McCormack M, Li L, Hall Q, Xiang C, Sheen J. Glucose-TOR signalling reprograms the transcriptome and activates meristems. Nature. 2013;496(7444):181–6.
149. van Dam TJP, Zwartkruis FJT, Bos JL, Snel B. Evolution of the TOR pathway. J Mol Evol. 2011;73:209–20.
150. Ren M, Qiu S, Venglat P, Xiang D, Feng L, Selvaraj G, et al. Target of rapamycin regulates development and ribosomal RNA expression through kinase domain in Arabidopsis. Plant Physiol. 2011;155(3):1367–82.
151. Schepetilnikov M, Dimitrova M, Mancera-Martinez E, Geldreich A, Keller M, Ryabova LA. TOR and S6K1 promote translation reinitiation of uORF-containing mRNAs via phosphorylation of eIF3h. EMBO J. 2013;32(8):1087–102.
152. Liu Y, Bassham DC. TOR is a negative regulator of autophagy in Arabidopsis thaliana. PLoS ONE. 2010;5(7):e11883.
153. Caldana C, Li Y, Leisse A, Zhang Y, Bartholomaeus L, Fernie AR, et al. Systemic analysis of inducible target of rapamycin mutants reveal a general metabolic switch controlling growth in Arabidopsis thaliana. Plant J: Cell Mol Biol. 2013;73(6):897–909.
154. Dong P, Xiong F, Que Y, Wang K, Yu L, Li Z, et al. Expression profiling and functional analysis reveals that TOR is a key player in regulating photosynthesis and phytohormone signaling pathways in Arabidopsis. Front Plant Sci. 2015;6:677.
155. Hu R, Zhu Y, Shen G, Zhang H. TAP46 plays a positive role in the ABSCISIC ACID INSENSITIVE5-regulated gene expression in Arabidopsis. Plant Physiol. 2014;164(2):721–34.
156. Kravchenko A, Citerne S, Jehanno I, Bersimbaev RI, Veit B, Meyer C, et al. Mutations in the Arabidopsis Lst8 and Raptor genes encoding partners of the TOR complex, or inhibition of TOR activity decrease abscisic acid (ABA) synthesis. Biochem Biophys Res Commun. 2015;467(4):992–7.
157. Turck F, Zilbermann F, Kozma SC, Thomas G, Nagy F. Phytohormones participate in an S6 kinase signal transduction pathway in Arabidopsis. Plant Physiol. 2004;134(4):1527–35.
158. Imamura S, Kawase Y, Kobayashi I, Sone T, Era A, Miyagishima SY, et al. Target of rapamycin (TOR) plays a critical role in triacylglycerol accumulation in microalgae. Plant Mol Biol. 2015;89(3):309–18.
159. Khandal D, Samol I, Buhr F, Pollmann S, Schmidt H, Clemens S, et al. Singlet oxygen-dependent translational control in the tigrina-d.12 mutant of barley. Proc Natl Acad Sci USA. 2009;106(31):13112–7.

160. Williams AJ, Werner-Fraczek J, Chang IF, Bailey-Serres J. Regulated phosphorylation of 40S ribosomal protein S6 in root tips of maize. Plant Physiol. 2003;132(4):2086–97.
161. Turkina MV. Klang Arstrand H, Vener AV. Differential phosphorylation of ribosomal proteins in Arabidopsis thaliana plants during day and night. PLoS ONE. 2011;6(12):e29307.
162. Boex-Fontvieille E, Daventure M, Jossier M, Zivy M, Hodges M, Tcherkez G. Photosynthetic control of Arabidopsis leaf cytoplasmic translation initiation by protein phosphorylation. PLoS ONE. 2013;8(7):e70692.
163. Perez-Perez ME, Florencio FJ, Crespo JL. Inhibition of target of rapamycin signaling and stress activate autophagy in Chlamydomonas reinhardtii. Plant Physiol. 2010;152(4): 1874–88.
164. Diaz-Troya S, Perez-Perez ME, Perez-Martin M, Moes S, Jeno P, Florencio FJ, et al. Inhibition of protein synthesis by TOR inactivation revealed a conserved regulatory mechanism of the BiP chaperone in Chlamydomonas. Plant Physiol. 2011;157(2):730–41.
165. Henriques R, Magyar Z, Monardes A, Khan S, Zalejski C, Orellana J, et al. Arabidopsis S6 kinase mutants display chromosome instability and altered RBR1-E2F pathway activity. EMBO J. 2010;29(17):2979–93.
166. Reitz MU, Gifford ML, Schafer P. Hormone activities and the cell cycle machinery in immunity-triggered growth inhibition. J Exp Bot. 2015;66(8):2187–97.
167. Ouibrahim L, Rubio AG, Moretti A, Montane MH, Menand B, Meyer C, et al. Potyviruses differ in their requirement for TOR signalling. J Gen Virol. 2015;96(9):2898–903.
168. Oldham S, Montagne J, Radimerski T, Thomas G, Hafen E. Genetic and biochemical characterization of dTOR, the Drosophila homolog of the target of rapamycin. Genes Dev. 2000;14(21):2689–94.
169. Zhang H, Stallock JP, Ng JC, Reinhard C, Neufeld TP. Regulation of cellular growth by the Drosophila target of rapamycin dTOR. Genes Dev. 2000;14(21):2712–24.
170. Hafen E. Interplay between growth factor and nutrient signaling: lessons from Drosophila. In: Thomas G, Sabatini DM, Hall MN, editors. TOR Target of rapamycin. Berlin, Heidelberg, New York: Springer; 2004. p. 153–68.
171. Neufeld TP. Genetic analysis of TOR signaling in Drosophila. In: Thomas G, Sabatini DM, Hall MN, editors. TOR Target of rapamycin. Berlin, Heidelberg, New York: Springer; 2004. p. 139–52.
172. Beauchamp EM, Platanias LC. The evolution of the TOR pathway and its role in cancer. Oncogene. 2013;32:3923–32.
173. Takahara T, Maeda T. Evolutionarily conserved regulation of TOR signalling. J Biochem. 2013;154:1–10.
174. Blagden SP, Gatt MK, Archambault V, Lada K, Ichihara K, Lilley KS, et al. Drosophila Larp associates with poly(A)-binding protein and is required for male fertility and syncytial embryo development. Dev Biol. 2009;334(1):186–97.
175. Pallares-Cartes C, Cakan-Akdogan G, Teleman AA. Tissue-specific coupling between insulin/IGF and TORC1 signaling via PRAS40 in Drosophila. Dev Cell. 2012;22(1):172–82.
176. Tettweiler G, Miron M, Jenkins M, Sonenberg N, Lasko P. Starvation and oxidative stress resistance in Drosophila are mediated through the eIF4E-binding protein, d4E-BP. Genes Dev. 2005;19:1840–3.
177. Wessells R, Fitzgerald E, Piazza N, Ocorr K, Morley S, Davies C, et al. d4eBP acts downstream of both dTOR and dFoxo to modulate cardiac functional aging in Drosophila. Aging Cell. 2009;8(5):542–52.
178. Montagne J, Stewart MJ, Stocker H, Hafen E, Kozma SC, Thomas G. Drosophila S6 kinase: a regulator of cell size. Science. 1999;285(5436):2126–9.
179. Dowling RJ, Topisirovic I, Alain T, Bidinosti M, Fonseca BD, Petroulakis E, et al. mTORC1-mediated cell proliferation, but not cell growth, controlled by the 4E-BPs. Science. 2010;328(5982):1172–6.
180. Nobukini T, Thomas G. The mTOR/S6K signalling pathway: the role of the TSC1/2 tumour suppressor complex and the proto-oncogene Rheb. Novartis Foundation symposium. 2004;262:148–54; discussion 54–9, 265–8.

181. Radimerski T, Montagne J, Rintelen F, Stocker H, van der Kaay J, Downes CP, et al. dS6 K-regulated cell growth is dPKB/dPI(3)K-independent, but requires dPDK1. Nat Cell Biol. 2002;4(3):251–5.
182. Kapahi P, Zid BM, Harper T, Koslover D, Sapin V, Benzer S. Regulation of lifespan in Drosophila by modulation of genes in the TOR signaling pathway. Current Biol: CB. 2004;14(10):885–90.
183. Lin TA, Kong X, Haystead TA, Pause A, Belsham G, Sonenberg N, et al. PHAS-I as a link between mitogen-activated protein kinase and translation initiation. Science. 1994;266 (5185):653–6.
184. Pause A, Belsham GJ, Gingras AC, Donze O, Lin TA, Lawrence JC Jr, et al. Insulin-dependent stimulation of protein synthesis by phosphorylation of a regulator of 5'-cap function. Nature. 1994;371(6500):762–7.
185. Hu C, Pang S, Kong X, Velleca M, Lawrence JC Jr. Molecular cloning and tissue distribution of PHAS-I, an intracellular target for insulin and growth factors. Proc Natl Acad Sci USA. 1994;91(9):3730–4.
186. Diggle TA, Bloomberg GB, Denton RM. Further characterization of the acid-soluble phosphoprotein (SDS/PAGE apparent molecular mass of 22 kDa) in rat fat-cells by peptide sequencing and immuno-analysis: effects of insulin and isoprenaline. Biochem J. 1995;306 (Pt 1):135–9.
187. Miron M, Verdu J, Lachance PE, Birnbaum MJ, Lasko PF, Sonenberg N. The translational inhibitor 4E-BP is an effector of PI(3)K/Akt signalling and cell growth in Drosophila. Nat Cell Biol. 2001;3(6):596–601.
188. Harrison DE, Strong R, Sharp ZD, Nelson JF, Astle CM, Flurkey K, et al. Rapamycin fed late in life extends lifespan in genetically heterogeneous mice. Nature. 2009;460(7253):392–5.
189. Sun P, Quan Z, Zhang B, Wu T, Xi R. TSC1/2 tumour suppressor complex maintains Drosophila germline stem cells by preventing differentiation. Development. 2010;137 (15):2461–9.
190. Gancz D, Gilboa L. Insulin and Target of rapamycin signaling orchestrate the development of ovarian niche-stem cell units in Drosophila. Development. 2013;140(20):4145–54.
191. Zhang Y, Billington CJ Jr, Pan D, Neufeld TP. Drosophila target of rapamycin kinase functions as a multimer. Genetics. 2006;172(1):355–62.
192. LaFever L, Feoktistov A, Hsu HJ, Drummond-Barbosa D. Specific roles of Target of rapamycin in the control of stem cells and their progeny in the Drosophila ovary. Development. 2010;137(13):2117–26.
193. Vander Haar E, Lee SI, Bandhakavi S, Griffin TJ, Kim DH. Insulin signalling to mTOR mediated by the Akt/PKB substrate PRAS40. Nat Cell Biol. 2007;9(3):316–23.
194. Thedieck K, Polak P, Kim ML, Molle KD, Cohen A, Jeno P, et al. PRAS40 and PRR5-like protein are new mTOR interactors that regulate apoptosis. PLoS ONE. 2007;2(11):e1217.
195. Fonseca BD, Smith EM, Lee VH, MacKintosh C, Proud CG. PRAS40 is a target for mammalian target of rapamycin complex 1 and is required for signaling downstream of this complex. J Biol Chem. 2007;282(34):24514–24.
196. Oshiro N, Takahashi R, Yoshino K, Tanimura K, Nakashima A, Eguchi S, et al. The proline-rich Akt substrate of 40 kDa (PRAS40) is a physiological substrate of mammalian target of rapamycin complex 1. J Biol Chem. 2007;282(28):20329–39.
197. Wang L, Harris TE, Roth RA, Lawrence JC Jr. PRAS40 regulates mTORC1 kinase activity by functioning as a direct inhibitor of substrate binding. J Biol Chem. 2007;282(27):20036–44.
198. Sancak Y, Thoreen CC, Peterson TR, Lindquist RA, Kang SA, Spooner E, et al. PRAS40 is an insulin-regulated inhibitor of the mTORC1 protein kinase. Mol Cell. 2007;25(6):903–15.
199. Fonseca BD, Zakaria C, Jia JJ, Graber TE, Svitkin Y, Tahmasebi S, et al. La-related Protein 1 (LARP1) Represses Terminal Oligopyrimidine (TOP) mRNA Translation Downstream of mTOR Complex 1 (mTORC1). J Biol Chem. 2015;290(26):15996–6020.

200. Burrows C, Abd Latip N, Lam SJ, Carpenter L, Sawicka K, Tzolovsky G, et al. The RNA binding protein Larp1 regulates cell division, apoptosis and cell migration. Nucleic Acids Res. 2010;38(16):5542–53.
201. Kriventseva EV, Tegenfeldt F, Petty TJ, Waterhouse RM, Simão FA, Pozdnyakov IA, et al. OrthoDB v8: update of the hierarchical catalog of orthologs and the underlying free software. Nucleic Acids Res. 2015;43:D250–6.
202. Waterhouse RM, Zdobnov EM, Tegenfeldt F, Li J, Kriventseva EV. OrthoDB: the hierarchical catalog of eukaryotic orthologs in 2011. Nucleic Acids Res. 2011;39:D283–8.
203. Waterhouse RM, Tegenfeldt F, Li J, Zdobnov EM, Kriventseva EV. OrthoDB: a hierarchical catalog of animal, fungal and bacterial orthologs. Nucleic Acids Res. 2013;41:D358–65.
204. Hernández G, Altmann M, Sierra JM, Urlaub H, Corral RD, Schwartz P, et al. Functional analysis of seven genes encoding eight translation initiation factor 4E (eIF4E) isoforms in Drosophila. Mech Dev. 2005;122:529–43.
205. Tettweiler G, Kowanda M, Lasko P, Sonenberg N, Hernández G. The distribution of eIF4E-family members across insecta. Comp Funct Genom. 2012;2012:960420.
206. Bernal A, Kimbrell DA. Drosophila Thor participates in host immune defence and conects a translational regulator with innate immunity. Proc Natl Acad Sci U S A. 2000;97:6019–24.
207. Hernández G, Gandin V, Han H, Ferreira T, Sonenberg N, Lasko P. Translational control by Drosophila eIF4E-3 is essential for cell differentiation during spermiogenesis. Development. 2012;139:3211–20.
208. Hernández G, Castellano MM, Agudo M, Sierra JM. Isolation and characterization of the cDNA and the gene for eukaryotic translation initiation factor 4G from Drosophila melanogaster. Eur J Biochem. 1998;253:27–35.
209. Baker CC, Fuller MT. Translational control of meiotic cell cycle progression and spermatid differentiation in male germ cells by a novel eIF4G homolog. Development. 2007;134:2863–9.
210. Franklin-Dumont M, Chatterjee C, Wasserman SA, DiNardo S. A novel eIF4G homolog, Off-schedule, couples translational control of meiosis and differentiation in Drosophila spermatocytes. Development. 2007;134(15):2851–61.
211. Robida-Stubbs S, Glover-Cutter K, Lamming DW, Mizunuma M, Narasimhan SD, Neumann-Haefelin E, et al. TOR signaling and rapamycin influence longevity by regulating SKN-1/Nrf and DAF-16/FoxO. Cell Metab. 2012;15(5):713–24.
212. Hermann GJ, Schroeder LK, Hieb CA, Kershner AM, Rabbitts BM, Fonarev P, et al. Genetic analysis of lysosomal trafficking in Caenorhabditis elegans. Mol Biol Cell. 2005;16(7):3273–88.
213. Soukas AA, Carr CE, Ruvkun G. Genetic regulation of Caenorhabditis elegans lysosome related organelle function. PLoS Genet. 2013;9(10):e1003908.
214. Delahaye JL, Foster OK, Vine A, Saxton DS, Curtin TP, Somhegyi H, et al. Caenorhabditis elegans HOPS and CCZ-1 mediate trafficking to lysosome-related organelles independently of RAB-7 and SAND-1. Mol Biol Cell. 2014;25(7):1073–96.
215. Jia K, Chen D, Riddle DL. The TOR pathway interacts with the insulin signaling pathway to regulate C. elegans larval development, metabolism and life span. Development. 2004;131(16):3897–906.
216. Vellai T, Takacs-Vellai K, Zhang Y, Kovacs AL, Orosz L, Muller F. Genetics: influence of TOR kinase on lifespan in C. elegans. Nature. 2003;426(6967):620.
217. Schreiber MA, Pierce-Shimomura JT, Chan S, Parry D, McIntire SL. Manipulation of behavioral decline in Caenorhabditis elegans with the Rag GTPase raga-1. PLoS Genet. 2010;6(5):e1000972.
218. Honjoh S, Yamamoto T, Uno M, Nishida E. Signalling through RHEB-1 mediates intermittent fasting-induced longevity in C. elegans. Nature. 2009;457(7230):726–30.
219. Bar-Peled L, Sabatini DM. Regulation of mTORC1 by amino acids. Trends Cell Biol. 2014;24(7):400–6.
220. Huang J, Manning BD. The TSC1-TSC2 complex: a molecular switchboard controlling cell growth. Biochem J. 2008;412(2):179–90.

221. Aspuria PJ, Sato T, Tamanoi F. The TSC/Rheb/TOR signaling pathway in fission yeast and mammalian cells: temperature sensitive and constitutive active mutants of TOR. Cell Cycle. 2007;6(14):1692–5.
222. Mansfeld J, Urban N, Priebe S, Groth M, Frahm C, Hartmann N, et al. Branched-chain amino acid catabolism is a conserved regulator of physiological ageing. Nat Commun. 2015;6:10043.
223. Kenyon C, Chang J, Gensch E, Rudner A, Tabtiang RAC. elegans mutant that lives twice as long as wild type. Nature. 1993;366(6454):461–4.
224. Dorman JB, Albinder B, Shroyer T, Kenyon C. The age-1 and daf-2 genes function in a common pathway to control the lifespan of Caenorhabditis elegans. Genetics. 1995;141 (4):1399–406.
225. Morris JZ, Tissenbaum HA, Ruvkun G. A phosphatidylinositol-3-OH kinase family member regulating longevity and diapause in Caenorhabditis elegans. Nature. 1996;382(6591):536–9.
226. Kimura KD, Tissenbaum HA, Liu Y, Ruvkun G. daf-2, an insulin receptor-like gene that regulates longevity and diapause in Caenorhabditis elegans. Science. 1997;277(5328):942–6.
227. Dillin A, Crawford DK, Kenyon C. Timing requirements for insulin/IGF-1 signaling in C. elegans. Science. 2002;298(5594):830–4.
228. Murphy CT, Hu PJ. Insulin/insulin-like growth factor signaling in C. elegans. WormBook: the online review of C elegans. Biology. 2013:1–43.
229. Lin K, Dorman JB, Rodan A, Kenyon C. daf-16: An HNF-3/forkhead family member that can function to double the life-span of Caenorhabditis elegans. Science. 1997;278 (5341):1319–22.
230. Greer EL, Oskoui PR, Banko MR, Maniar JM, Gygi MP, Gygi SP, et al. The energy sensor AMP-activated protein kinase directly regulates the mammalian FOXO3 transcription factor. J Biol Chem. 2007;282(41):30107–19.
231. Oh SW, Mukhopadhyay A, Svrzikapa N, Jiang F, Davis RJ, Tissenbaum HA. JNK regulates lifespan in Caenorhabditis elegans by modulating nuclear translocation of forkhead transcription factor/DAF-16. Proc Natl Acad Sci USA. 2005;102(12):4494–9.
232. Lehtinen MK, Yuan Z, Boag PR, Yang Y, Villen J, Becker EB, et al. A conserved MST-FOXO signaling pathway mediates oxidative-stress responses and extends life span. Cell. 2006;125(5):987–1001.
233. Apfeld J, O'Connor G, McDonagh T, DiStefano PS, Curtis R. The AMP-activated protein kinase AAK-2 links energy levels and insulin-like signals to lifespan in C. elegans. Genes Dev. 2004;18(24):3004–9.
234. Soltys CL, Kovacic S, Dyck JR. Activation of cardiac AMP-activated protein kinase by LKB1 expression or chemical hypoxia is blunted by increased Akt activity. Am J Physiol Heart Circ Physiol. 2006;290(6):H2472–9.
235. Hawley SA, Ross FA, Gowans GJ, Tibarewal P, Leslie NR, Hardie DG. Phosphorylation by Akt within the ST loop of AMPK-alpha1 down-regulates its activation in tumour cells. Biochem J. 2014;459(2):275–87.
236. Martin TD, Chen XW, Kaplan RE, Saltiel AR, Walker CL, Reiner DJ, et al. Ral and Rheb GTPase activating proteins integrate mTOR and GTPase signaling in aging, autophagy, and tumor cell invasion. Mol Cell. 2014;53(2):209–20.
237. Hansen M, Taubert S, Crawford D, Libina N, Lee SJ, Kenyon C. Lifespan extension by conditions that inhibit translation in Caenorhabditis elegans. Aging Cell. 2007;6(1):95–110.
238. Pan KZ, Palter JE, Rogers AN, Olsen A, Chen D, Lithgow GJ, et al. Inhibition of mRNA translation extends lifespan in Caenorhabditis elegans. Aging Cell. 2007;6(1):111–9.
239. Gwinn DM, Shackelford DB, Egan DF, Mihaylova MM, Mery A, Vasquez DS, et al. AMPK phosphorylation of raptor mediates a metabolic checkpoint. Mol Cell. 2008;30(2):214–26.
240. Wolff NC, Vega-Rubin-de-Celis S, Xie XJ, Castrillon DH, Kabbani W, Brugarolas J. Cell-type-dependent regulation of mTORC1 by REDD1 and the tumor suppressors TSC1/TSC2 and LKB1 in response to hypoxia. Mol Cell Biol. 2011;31(9):1870–84.

241. Tullet JM, Hertweck M, An JH, Baker J, Hwang JY, Liu S, et al. Direct inhibition of the longevity-promoting factor SKN-1 by insulin-like signaling in C. elegans. Cell. 2008;132 (6):1025–38.
242. Soukas AA, Kane EA, Carr CE, Melo JA, Ruvkun G. Rictor/TORC2 regulates fat metabolism, feeding, growth, and life span in Caenorhabditis elegans. Genes Dev. 2009;23 (4):496–511.
243. Mizunuma M, Neumann-Haefelin E, Moroz N, Li Y, Blackwell TK. mTORC2-SGK-1 acts in two environmentally responsive pathways with opposing effects on longevity. Aging Cell. 2014;13(5):869–78.
244. Greer EL, Brunet A. Different dietary restriction regimens extend lifespan by both independent and overlapping genetic pathways in C. elegans. Aging Cell. 2009;8(2):113–27.
245. Panowski SH, Wolff S, Aguilaniu H, Durieux J, Dillin A. PHA-4/Foxa mediates diet-restriction-induced longevity of C. elegans. Nature. 2007;447(7144):550–5.
246. Houthoofd K, Braeckman BP, Lenaerts I, Brys K, De Vreese A, Van Eygen S, et al. No reduction of metabolic rate in food restricted Caenorhabditis elegans. Exp Gerontol. 2002;37 (12):1359–69.
247. Houthoofd K, Braeckman BP, Johnson TE, Vanfleteren JR. Life extension via dietary restriction is independent of the Ins/IGF-1 signalling pathway in Caenorhabditis elegans. Exp Gerontol. 2003;38(9):947–54.
248. Houthoofd K, Braeckman BP, Lenaerts I, Brys K, De Vreese A, Van Eygen S, et al. Axenic growth up-regulates mass-specific metabolic rate, stress resistance, and extends life span in Caenorhabditis elegans. Exp Gerontol. 2002;37(12):1371–8.
249. Bishop NA, Guarente L. Two neurons mediate diet-restriction-induced longevity in C. elegans. Nature. 2007;447(7144):545–9.
250. Szewczyk NJ, Udranszky IA, Kozak E, Sunga J, Kim SK, Jacobson LA, et al. Delayed development and lifespan extension as features of metabolic lifestyle alteration in C. elegans under dietary restriction. J Exp Biol. 2006;209(Pt 20):4129–39.
251. Kaeberlein TL, Smith ED, Tsuchiya M, Welton KL, Thomas JH, Fields S, et al. Lifespan extension in Caenorhabditis elegans by complete removal of food. Aging Cell. 2006;5 (6):487–94.
252. Hosono R, Nishimoto S, Kuno S. Alterations of life span in the nematode Caenorhabditis elegans under monoxenic culture conditions. Exp Gerontol. 1989;24(3):251–64.
253. Lakowski B, Hekimi S. The genetics of caloric restriction in Caenorhabditis elegans. Proc Natl Acad Sci USA. 1998;95(22):13091–6.
254. Lee GD, Wilson MA, Zhu M, Wolkow CA, de Cabo R, Ingram DK, et al. Dietary deprivation extends lifespan in Caenorhabditis elegans. Aging Cell. 2006;5(6):515–24.
255. Sheaffer KL, Updike DL, Mango SE. The Target of Rapamycin pathway antagonizes pha-4/FoxA to control development and aging. Curr Biol: CB. 2008;18(18):1355–64.
256. Rousakis A, Vlassis A, Vlanti A, Patera S, Thireos G, Syntichaki P. The general control nonderepressible-2 kinase mediates stress response and longevity induced by target of rapamycin inactivation in Caenorhabditis elegans. Aging Cell. 2013;12(5):742–51.
257. Steffen KK, MacKay VL, Kerr EO, Tsuchiya M, Hu D, Fox LA, et al. Yeast life span extension by depletion of 60s ribosomal subunits is mediated by Gcn4. Cell. 2008;133 (2):292–302.
258. Hsin H, Kenyon C. Signals from the reproductive system regulate the lifespan of C. elegans. Nature. 1999;399(6734):362–6.
259. Motola DL, Cummins CL, Rottiers V, Sharma KK, Li T, Li Y, et al. Identification of ligands for DAF-12 that govern dauer formation and reproduction in C. elegans. Cell. 2006;124 (6):1209–23.
260. Antebi A, Yeh WH, Tait D, Hedgecock EM, Riddle DL. daf-12 encodes a nuclear receptor that regulates the dauer diapause and developmental age in C. elegans. Genes Dev. 2000;14 (12):1512–27.
261. Lin K, Hsin H, Libina N, Kenyon C. Regulation of the Caenorhabditis elegans longevity protein DAF-16 by insulin/IGF-1 and germline signaling. Nat Genet. 2001;28(2):139–45.

262. Berman JR, Kenyon C. Germ-cell loss extends C. elegans life span through regulation of DAF-16 by kri-1 and lipophilic-hormone signaling. Cell. 2006;124(5):1055–68.
263. Ghazi A, Henis-Korenblit S, Kenyon C. A transcription elongation factor that links signals from the reproductive system to lifespan extension in Caenorhabditis elegans. PLoS Genet. 2009;5(9):e1000639.
264. McCormick M, Chen K, Ramaswamy P, Kenyon C. New genes that extend Caenorhabditis elegans' lifespan in response to reproductive signals. Aging Cell. 2012;11(2):192–202.
265. Lapierre LR, Gelino S, Melendez A, Hansen M. Autophagy and lipid metabolism coordinately modulate life span in germline-less C. elegans. Curr Biol: CB. 2011;21 (18):1507–14.
266. Goudeau J, Bellemin S, Toselli-Mollereau E, Shamalnasab M, Chen Y, Aguilaniu H. Fatty acid desaturation links germ cell loss to longevity through NHR-80/HNF4 in C. elegans. PLoS Biol. 2011;9(3):e1000599.
267. Steinbaugh MJ, Narasimhan SD, Robida-Stubbs S, Moronetti Mazzeo LE, Dreyfuss JM, Hourihan JM, et al. Lipid-mediated regulation of SKN-1/Nrf in response to germ cell absence. eLife. 2015;4.
268. Shen Y, Wollam J, Magner D, Karalay O, Antebi A. A steroid receptor-microRNA switch regulates life span in response to signals from the gonad. Science. 2012;338(6113):1472–6.
269. de Lencastre A, Pincus Z, Zhou K, Kato M, Lee SS, Slack FJ. MicroRNAs both promote and antagonize longevity in C. elegans. Curr Biol: CB. 2010;20(24):2159–68.
270. Boulias K, Horvitz HR. The C. elegans microRNA mir-71 acts in neurons to promote germline-mediated longevity through regulation of DAF-16/FOXO. Cell Metab. 2012;15 (4):439–50.
271. Depuydt G, Xie F, Petyuk VA, Shanmugam N, Smolders A, Dhondt I, et al. Reduced insulin/insulin-like growth factor-1 signaling and dietary restriction inhibit translation but preserve muscle mass in Caenorhabditis elegans. Mol Cell Proteomics: MCP. 2013;12 (12):3624–39.
272. Shore DE, Ruvkun G. A cytoprotective perspective on longevity regulation. Trends Cell Biol. 2013;23(9):409–20.
273. Brown EJ, Beal PA, Keith CT, Chen J, Shin TB, Schreiber SL. Control of p70 s6 kinase by kinase activity of FRAP in vivo. Nature. 1995;377(6548):441–6.
274. Brunn GJ, Hudson CC, Sekulic A, Williams JM, Hosoi H, Houghton PJ, et al. Phosphorylation of the translational repressor PHAS-I by the mammalian target of rapamycin. Science. 1997;277(5322):99–101.
275. Lin TA, Kong X, Saltiel AR, Blackshear PJ, Lawrence JC Jr. Control of PHAS-I by insulin in 3T3-L1 adipocytes. Synthesis, degradation, and phosphorylation by a rapamycin-sensitive and mitogen-activated protein kinase-independent pathway. J Biol Chem. 1995;270 (31):18531–8.
276. Graves LM, Bornfeldt KE, Argast GM, Krebs EG, Kong X, Lin TA, et al. cAMP- and rapamycin-sensitive regulation of the association of eukaryotic initiation factor 4E and the translational regulator PHAS-I in aortic smooth muscle cells. Proc Natl Acad Sci USA. 1995;92(16):7222–6.
277. Fadden P, Haystead TA, Lawrence JC Jr. Identification of phosphorylation sites in the translational regulator, PHAS-I, that are controlled by insulin and rapamycin in rat adipocytes. J Biol Chem. 1997;272(15):10240–7.
278. von Manteuffel SR, Dennis PB, Pullen N, Gingras AC, Sonenberg N, Thomas G. The insulin-induced signalling pathway leading to S6 and initiation factor 4E binding protein 1 phosphorylation bifurcates at a rapamycin-sensitive point immediately upstream of p70s6k. Mol Cell Biol. 1997;17(9):5426–36.
279. Price DJ, Grove JR, Calvo V, Avruch J, Bierer BE. Rapamycin-induced inhibition of the 70-kilodalton S6 protein kinase. Science. 1992;257(5072):973–7.
280. Weng QP, Andrabi K, Kozlowski MT, Grove JR, Avruch J. Multiple independent inputs are required for activation of the p70 S6 kinase. Mol Cell Biol. 1995;15(5):2333–40.

281. Hara K, Yonezawa K, Kozlowski MT, Sugimoto T, Andrabi K, Weng QP, et al. Regulation of eIF-4E BP1 phosphorylation by mTOR. J Biol Chem. 1997;272(42):26457–63.
282. Weng QP, Kozlowski M, Belham C, Zhang A, Comb MJ, Avruch J. Regulation of the p70 S6 kinase by phosphorylation in vivo. Analysis using site-specific anti-phosphopeptide antibodies. J Biol Chem. 1998;273(26):16621–9.
283. Hara K, Yonezawa K, Weng QP, Kozlowski MT, Belham C, Avruch J. Amino acid sufficiency and mTOR regulate p70 S6 kinase and eIF-4E BP1 through a common effector mechanism. J Biol Chem. 1998;273(23):14484–94.
284. Thomas G. p70s6k/p85s6k: mechanism of activation, effects of rapamycin and role in mitogenesis. Biochem Soc Trans. 1993;21(4):901–4.
285. Ferrari S, Pearson RB, Siegmann M, Kozma SC, Thomas G. The immunosuppressant rapamycin induces inactivation of p70s6k through dephosphorylation of a novel set of sites. J Biol Chem. 1993;268(22):16091–4.
286. Burnett PE, Barrow RK, Cohen NA, Snyder SH, Sabatini DM. RAFT1 phosphorylation of the translational regulators p70 S6 kinase and 4E-BP1. Proc Natl Acad Sci USA. 1998;95 (4):1432–7.
287. Fonseca BD, Lee VH, Proud CG. The binding of PRAS40 to 14-3-3 proteins is not required for activation of mTORC1 signalling by phorbol esters/ERK. Biochem J. 2008;411(1):141–9.
288. Wang L, Harris TE, Lawrence JC Jr. Regulation of proline-rich Akt substrate of 40 kDa (PRAS40) function by mammalian target of rapamycin complex 1 (mTORC1)-mediated phosphorylation. J Biol Chem. 2008;283(23):15619–27.
289. Peterson TR, Laplante M, Thoreen CC, Sancak Y, Kang SA, Kuehl WM, et al. DEPTOR is an mTOR inhibitor frequently overexpressed in multiple myeloma cells and required for their survival. Cell. 2009;137(5):873–86.
290. Frias MA, Thoreen CC, Jaffe JD, Schroder W, Sculley T, Carr SA, et al. mSin1 is necessary for Akt/PKB phosphorylation, and its isoforms define three distinct mTORC2s. Curr Biol: CB. 2006;16(18):1865–70.
291. Pearce LR, Huang X, Boudeau J, Pawlowski R, Wullschleger S, Deak M, et al. Identification of Protor as a novel Rictor-binding component of mTOR complex-2. Biochem J. 2007;405 (3):513–22.
292. Woo SY, Kim DH, Jun CB, Kim YM, Haar EV, Lee SI, et al. PRR5, a novel component of mTOR complex 2, regulates platelet-derived growth factor receptor beta expression and signaling. J Biol Chem. 2007;282(35):25604–12.
293. Kovacina KS, Park GY, Bae SS, Guzzetta AW, Schaefer E, Birnbaum MJ, et al. Identification of a proline-rich Akt substrate as a 14-3-3 binding partner. J Biol Chem. 2003;278(12):10189–94.
294. Harthill JE. Pozuelo Rubio M, Milne FC, MacKintosh C. Regulation of the 14-3-3-binding protein p39 by growth factors and nutrients in rat PC12 pheochromocytoma cells. Biochem J. 2002;368(Pt 2):565–72.
295. Thoreen CC, Chantranupong L, Keys HR, Wang T, Gray NS, Sabatini DM. A unifying model for mTORC1-mediated regulation of mRNA translation. Nature. 2012;485 (7396):109–13.
296. Dai N, Rapley J, Angel M, Yanik MF, Blower MD, Avruch J. mTOR phosphorylates IMP2 to promote IGF2 mRNA translation by internal ribosomal entry. Genes Dev. 2011;25 (11):1159–72.
297. Dai N, Christiansen J, Nielsen FC, Avruch J. mTOR complex 2 phosphorylates IMP1 cotranslationally to promote IGF2 production and the proliferation of mouse embryonic fibroblasts. Genes Dev. 2013;27(3):301–12.
298. Hsu PP, Kang SA, Rameseder J, Zhang Y, Ottina KA, Lim D, et al. The mTOR-regulated phosphoproteome reveals a mechanism of mTORC1-mediated inhibition of growth factor signaling. Science. 2011;332(6035):1317–22.
299. Demirkan G, Yu K, Boylan JM, Salomon AR, Gruppuso PA. Phosphoproteomic profiling of in vivo signaling in liver by the mammalian target of rapamycin complex 1 (mTORC1). PLoS ONE. 2011;6(6):e21729.

300. Yu Y, Yoon SO, Poulogiannis G, Yang Q, Ma XM, Villen J, et al. Phosphoproteomic analysis identifies Grb10 as an mTORC1 substrate that negatively regulates insulin signaling. Science. 2011;332(6035):1322–6.
301. Schultze FC, Petrova DT, Oellerich M, Armstrong VW, Asif AR. Differential proteome and phosphoproteome signatures in human T-lymphoblast cells induced by sirolimus. Cell Prolif. 2010;43(4):396–404.
302. Isotani S, Hara K, Tokunaga C, Inoue H, Avruch J, Yonezawa K. Immunopurified mammalian target of rapamycin phosphorylates and activates p70 S6 kinase alpha in vitro. J Biol Chem. 1999;274(48):34493–8.
303. Price DJ, Nemenoff RA, Avruch J. Purification of a hepatic S6 kinase from cycloheximide-treated Rats. J Biol Chem. 1989;264(23):13825–33.
304. Price DJ, Gunsalus JR, Avruch J. Insulin activates a 70-kDa S6 kinase through serine/threonine-specific phosphorylation of the enzyme polypeptide. Proc Natl Acad Sci USA. 1990;87(20):7944–8.
305. Banerjee P, Ahmad MF, Grove JR, Kozlosky C, Price DJ, Avruch J. Molecular structure of a major insulin/mitogen-activated 70-kDa S6 protein kinase. Proc Natl Acad Sci USA. 1990;87 (21):8550–4.
306. Kozma SC, Ferrari S, Bassand P, Siegmann M, Totty N, Thomas G. Cloning of the mitogen-activated S6 kinase from rat liver reveals an enzyme of the second messenger subfamily. Proc Natl Acad Sci USA. 1990;87(19):7365–9.
307. Grove JR, Banerjee P, Balasubramanyam A, Coffer PJ, Price DJ, Avruch J, et al. Cloning and expression of two human p70 S6 kinase polypeptides differing only at their amino termini. Mol Cell Biol. 1991;11(11):5541–50.
308. Shima H, Pende M, Chen Y, Fumagalli S, Thomas G, Kozma SC. Disruption of the p70 (s6 k)/p85(s6 k) gene reveals a small mouse phenotype and a new functional S6 kinase. EMBO J. 1998;17(22):6649–59.
309. Gout I, Minami T, Hara K, Tsujishita Y, Filonenko V, Waterfield MD, et al. Molecular cloning and characterization of a novel p70 S6 kinase, p70 S6 kinase beta containing a proline-rich region. J Biol Chem. 1998;273(46):30061–4.
310. Koh H, Jee K, Lee B, Kim J, Kim D, Yun YH, et al. Cloning and characterization of a nuclear S6 kinase, S6 kinase-related kinase (SRK); a novel nuclear target of Akt. Oncogene. 1999;18(36):5115–9.
311. Lee-Fruman KK, Kuo CJ, Lippincott J, Terada N, Blenis J. Characterization of S6K2, a novel kinase homologous to S6K1. Oncogene. 1999;18(36):5108–14.
312. Calvo V, Crews CM, Vik TA, Bierer BE. Interleukin 2 stimulation of p70 S6 kinase activity is inhibited by the immunosuppressant rapamycin. Proc Natl Acad Sci USA. 1992;89 (16):7571–5.
313. Terada N, Lucas JJ, Szepesi A, Franklin RA, Takase K, Gelfand EW. Rapamycin inhibits the phosphorylation of p70 S6 kinase in IL-2 and mitogen-activated human T cells. Biochem Biophys Res Commun. 1992;186(3):1315–21.
314. Kuo CJ, Chung J, Fiorentino DF, Flanagan WM, Blenis J, Crabtree GR. Rapamycin selectively inhibits interleukin-2 activation of p70 S6 kinase. Nature. 1992;358(6381):70–3.
315. Chung J, Kuo CJ, Crabtree GR, Blenis J. Rapamycin-FKBP specifically blocks growth-dependent activation of and signaling by the 70 kd S6 protein kinases. Cell. 1992;69(7):1227–36.
316. Alessi DR, Kozlowski MT, Weng QP, Morrice N, Avruch J. 3-Phosphoinositide-dependent protein kinase 1 (PDK1) phosphorylates and activates the p70 S6 kinase in vivo and in vitro. Current Biol: CB. 1998;8(2):69–81.
317. Balendran A, Currie R, Armstrong CG, Avruch J, Alessi DR. Evidence that 3-phosphoinositide-dependent protein kinase-1 mediates phosphorylation of p70 S6 kinase in vivo at Thr-412 as well as Thr-252. J Biol Chem. 1999;274(52):37400–6.
318. Moser BA, Dennis PB, Pullen N, Pearson RB, Williamson NA, Wettenhall RE, et al. Dual requirement for a newly identified phosphorylation site in p70s6k. Mol Cell Biol. 1997;17 (9):5648–55.

319. Mukhopadhyay NK, Price DJ, Kyriakis JM, Pelech S, Sanghera J, Avruch J. An array of insulin-activated, proline-directed serine/threonine protein kinases phosphorylate the p70 S6 kinase. J Biol Chem. 1992;267(5):3325–35.
320. Weng QP, Andrabi K, Klippel A, Kozlowski MT, Williams LT, Avruch J. Phosphatidylinositol 3-kinase signals activation of p70 S6 kinase in situ through site-specific p70 phosphorylation. Proc Natl Acad Sci USA. 1995;92(12):5744–8.
321. Balendran A, Biondi RM, Cheung PC, Casamayor A, Deak M, Alessi DR. A 3-phosphoinositide-dependent protein kinase-1 (PDK1) docking site is required for the phosphorylation of protein kinase Czeta (PKCzeta) and PKC-related kinase 2 by PDK1. J Biol Chem. 2000;275(27):20806–13.
322. Biondi RM, Cheung PC, Casamayor A, Deak M, Currie RA, Alessi DR. Identification of a pocket in the PDK1 kinase domain that interacts with PIF and the C-terminal residues of PKA. EMBO J. 2000;19(5):979–88.
323. Biondi RM, Kieloch A, Currie RA, Deak M, Alessi DR. The PIF-binding pocket in PDK1 is essential for activation of S6K and SGK, but not PKB. EMBO J. 2001;20(16):4380–90.
324. Biondi RM, Komander D, Thomas CC, Lizcano JM, Deak M, Alessi DR, et al. High resolution crystal structure of the human PDK1 catalytic domain defines the regulatory phosphopeptide docking site. EMBO J. 2002;21(16):4219–28.
325. Pearce LR, Komander D, Alessi DR. The nuts and bolts of AGC protein kinases. Nat Rev Mol Cell Biol. 2010;11(1):9–22.
326. Rabl J, Leibundgut M, Ataide SF, Haag A, Ban N. Crystal structure of the eukaryotic 40S ribosomal subunit in complex with initiation factor 1. Science. 2011;331(6018):730–6.
327. Coffer PJ, Woodgett JR. Differential subcellular localisation of two isoforms of p70 S6 protein kinase. Biochem Biophys Res Commun. 1994;198(2):780–6.
328. Reinhard C, Fernandez A, Lamb NJ, Thomas G. Nuclear localization of p85s6 k: functional requirement for entry into S phase. EMBO J. 1994;13(7):1557–65.
329. Rosner M, Hengstschlager M. Nucleocytoplasmic localization of p70 S6K1, but not of its isoforms p85 and p31, is regulated by TSC2/mTOR. Oncogene. 2011;30(44):4509–22.
330. Rosner M, Schipany K, Hengstschlager M. p70 S6K1 nuclear localization depends on its mTOR-mediated phosphorylation at T389, but not on its kinase activity towards S6. Amino Acids. 2012;42(6):2251–6.
331. Pende M, Um SH, Mieulet V, Sticker M, Goss VL, Mestan J, et al. S6K1(-/-)/S6K2(-/-) mice exhibit perinatal lethality and rapamycin-sensitive 5'-terminal oligopyrimidine mRNA translation and reveal a mitogen-activated protein kinase-dependent S6 kinase pathway. Mol Cell Biol. 2004;24(8):3112–24.
332. Harrington LS, Findlay GM, Gray A, Tolkacheva T, Wigfield S, Rebholz H, et al. The TSC1-2 tumor suppressor controls insulin-PI3K signaling via regulation of IRS proteins. J Cell Biology. 2004;166(2):213–23.
333. Romeo Y, Zhang X, Roux PP. Regulation and function of the RSK family of protein kinases. Biochem J. 2012;441(2):553–69.
334. Ruvinsky I, Sharon N, Lerer T, Cohen H, Stolovich-Rain M, Nir T, et al. Ribosomal protein S6 phosphorylation is a determinant of cell size and glucose homeostasis. Genes Dev. 2005;19(18):2199–211.
335. Raught B, Peiretti F, Gingras AC, Livingstone M, Shahbazian D, Mayeur GL, et al. Phosphorylation of eucaryotic translation initiation factor 4B Ser422 is modulated by S6 kinases. EMBO J. 2004;23(8):1761–9.
336. Dennis MD, Jefferson LS, Kimball SR. Role of p70S6K1-mediated phosphorylation of eIF4B and PDCD4 proteins in the regulation of protein synthesis. J Biol Chem. 2012;287 (51):42890–9.
337. Holz MK, Ballif BA, Gygi SP, Blenis J. mTOR and S6K1 mediate assembly of the translation preinitiation complex through dynamic protein interchange and ordered phosphorylation events. Cell. 2005;123(4):569–80.

338. Martineau Y, Wang X, Alain T, Petroulakis E, Shahbazian D, Fabre B, et al. Control of Paip1-eukayrotic translation initiation factor 3 interaction by amino acids through S6 kinase. Mol Cell Biol. 2014;34(6):1046–53.
339. Dorrello NV, Peschiaroli A, Guardavaccaro D, Colburn NH, Sherman NE, Pagano M. S6K1- and betaTRCP-mediated degradation of PDCD4 promotes protein translation and cell growth. Science. 2006;314(5798):467–71.
340. Wang X, Li W, Williams M, Terada N, Alessi DR, Proud CG. Regulation of elongation factor 2 kinase by p90(RSK1) and p70 S6 kinase. EMBO J. 2001;20(16):4370–9.
341. Poulin F, Gingras AC, Olsen H, Chevalier S, Sonenberg N. 4E-BP3, a new member of the eukaryotic initiation factor 4E-binding protein family. J Biol Chem. 1998;273(22):14002–7.
342. Diggle TA, Denton RM. Characterisation of a novel "22 kDa" phosphoprotein which may be important in insulin action. Biochem Soc Trans. 1995;23(2):209S.
343. Tee AR, Proud CG. Caspase cleavage of initiation factor 4E-binding protein 1 yields a dominant inhibitor of cap-dependent translation and reveals a novel regulatory motif. Mol Cell Biol. 2002;22(6):1674–83.
344. Bidinosti M, Martineau Y, Frank F, Sonenberg N. Repair of isoaspartate formation modulates the interaction of deamidated 4E-BP2 with mTORC1 in brain. J Biol Chem. 2010;285(25):19402–8.
345. Bidinosti M, Ran I, Sanchez-Carbente MR, Martineau Y, Gingras AC, Gkogkas C, et al. Postnatal deamidation of 4E-BP2 in brain enhances its association with raptor and alters kinetics of excitatory synaptic transmission. Mol Cell. 2010;37(6):797–808.
346. Gingras AC, Gygi SP, Raught B, Polakiewicz RD, Abraham RT, Hoekstra MF, et al. Regulation of 4E-BP1 phosphorylation: a novel two-step mechanism. Genes Dev. 1999;13 (11):1422–37.
347. Mothe-Satney I, Brunn GJ, McMahon LP, Capaldo CT, Abraham RT, Lawrence JC Jr. Mammalian target of rapamycin-dependent phosphorylation of PHAS-I in four (S/T)P sites detected by phospho-specific antibodies. J Biol Chem. 2000;275(43):33836–43.
348. Mothe-Satney I, Yang D, Fadden P, Haystead TA, Lawrence JC Jr. Multiple mechanisms control phosphorylation of PHAS-I in five (S/T)P sites that govern translational repression. Mol Cell Biol. 2000;20(10):3558–67.
349. Shin S, Wolgamott L, Roux PP, Yoon SO. Casein kinase 1epsilon promotes cell proliferation by regulating mRNA translation. Cancer Res. 2014;74(1):201–11.
350. Wang X, Li W, Parra JL, Beugnet A, Proud CG. The C terminus of initiation factor 4E-binding protein 1 contains multiple regulatory features that influence its function and phosphorylation. Mol Cell Biol. 2003;23(5):1546–57.
351. Yang DQ, Kastan MB. Participation of ATM in insulin signalling through phosphorylation of eIF-4E-binding protein 1. Nat Cell Biol. 2000;2(12):893–8.
352. Gingras AC, Raught B, Gygi SP, Niedzwiecka A, Miron M, Burley SK, et al. Hierarchical phosphorylation of the translation inhibitor 4E-BP1. Genes Dev. 2001;15(21):2852–64.
353. Herbert TP, Tee AR, Proud CG. The extracellular signal-regulated kinase pathway regulates the phosphorylation of 4E-BP1 at multiple sites. J Biol Chem. 2002;277(13):11591–6.
354. Livingstone M, Bidinosti M. Rapamycin-insensitive mTORC1 activity controls eIF4E:4E-BP1 binding. F1000Research. 2012;1:4.
355. Marcotrigiano J, Gingras AC, Sonenberg N, Burley SK. Cap-dependent translation initiation in eukaryotes is regulated by a molecular mimic of eIF4G. Mol Cell. 1999;3(6):707–16.
356. Tait S, Dutta K, Cowburn D, Warwicker J, Doig AJ, McCarthy JE. Local control of a disorder-order transition in 4E-BP1 underpins regulation of translation via eIF4E. Proc Natl Acad Sci USA. 2010;107(41):17627–32.
357. Tomoo K, Abiko F, Miyagawa H, Kitamura K, Ishida T. Effect of N-terminal region of eIF4E and Ser65-phosphorylation of 4E-BP1 on interaction between eIF4E and 4E-BP1 fragment peptide. J Biochem. 2006;140(2):237–46.
358. Bah A, Vernon RM, Siddiqui Z, Krzeminski M, Muhandiram R, Zhao C, et al. Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. Nature. 2015;519 (7541):106–9.

359. Fletcher CM, Wagner G. The interaction of eIF4E with 4E-BP1 is an induced fit to a completely disordered protein. Protein Sci: Publ Protein Soc. 1998;7(7):1639–42.
360. Fletcher CM, McGuire AM, Gingras AC, Li H, Matsuo H, Sonenberg N, et al. 4E binding proteins inhibit the translation factor eIF4E without folded structure. Biochemistry. 1998;37 (1):9–15.
361. Hershey PE, McWhirter SM, Gross JD, Wagner G, Alber T, Sachs AB. The Cap-binding protein eIF4E promotes folding of a functional domain of yeast translation initiation factor eIF4G1. J Biol Chem. 1999;274(30):21297–304.
362. Gross JD, Moerke NJ, von der Haar T, Lugovskoy AA, Sachs AB, McCarthy JE, et al. Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E. Cell. 2003;115(6):739–50.
363. Tomoo K, Matsushita Y, Fujisaki H, Abiko F, Shen X, Taniguchi T, et al. Structural basis for mRNA Cap-Binding regulation of eukaryotic initiation factor 4E by 4E-binding protein, studied by spectroscopic, X-ray crystal structural, and molecular dynamics simulation methods. Biochim Biophys Acta. 2005;1753(2):191–208.
364. Ptushkina M, von der Haar T, Karim MM, Hughes JM, McCarthy JE. Repressor binding to a dorsal regulatory site traps human eIF4E in a high cap-affinity state. The EMBO journal. 1999;18(14):4068–75.
365. Denton RM, Tavare JM, Borthwick A, Dickens M, Diggle TA, Edgell NJ, et al. Insulin-activated protein kinases in fat and other cells. Biochem Soc Trans. 1992;20(3):659–64.
366. Diggle TA, Denton RM. Comparison of the effects of insulin and adrenergic agonists on the phosphorylation of an acid-soluble 22 kDa protein in rat epididymal fat-pads and isolated fat-cells. Biochem J. 1992;282(Pt 3):729–36.
367. Moule SK, Edgell NJ, Welsh GI, Diggle TA, Foulstone EJ, Heesom KJ, et al. Multiple signalling pathways involved in the stimulation of fatty acid and glycogen synthesis by insulin in rat epididymal fat cells. Biochem J. 1995;311(Pt 2):595–601.
368. Le Bacquer O, Petroulakis E, Paglialunga S, Poulin F, Richard D, Cianflone K, et al. Elevated sensitivity to diet-induced obesity and insulin resistance in mice lacking 4E-BP1 and 4E-BP2. J Clin Investig. 2007;117(2):387–96.
369. Tsukiyama-Kohara K, Poulin F, Kohara M, DeMaria CT, Cheng A, Wu Z, et al. Adipose tissue reduction in mice lacking the translational inhibitor 4E-BP1. Nat Med. 2001;7 (10):1128–32.
370. Yun YS, Kim KH, Tschida B, Sachs Z, Noble-Orcutt KE, Moriarity BS, et al. mTORC1 Coordinates Protein Synthesis and Immunoproteasome Formation via PRAS40 to Prevent Accumulation of Protein Stress. Mol Cell. 2016;61(4):625–39.
371. Rapley J, Oshiro N, Ortiz-Vega S, Avruch J. The mechanism of insulin-stimulated 4E-BP protein binding to mammalian target of rapamycin (mTOR) complex 1 and its contribution to mTOR complex 1 signaling. J Biol Chem. 2011;286(44):38043–53.
372. Kazi AA, Lang CH. PRAS40 regulates protein synthesis and cell cycle in C2C12 myoblasts. Mol Med. 2010;16(9–10):359–71.
373. Tcherkezian J, Cargnello M, Romeo Y, Huttlin EL, Lavoie G, Gygi SP, et al. Proteomic analysis of cap-dependent translation identifies LARP1 as a key regulator of 5'TOP mRNA translation. Genes Dev. 2014;28(4):357–71.
374. Aoki K, Adachi S, Homoto M, Kusano H, Koike K, Natsume T. LARP1 specifically recognizes the 3' terminus of poly(A) mRNA. FEBS Lett. 2013;587(14):2173–8.
375. Deragon JM, Bousquet-Antonelli C. The role of LARP1 in translation and beyond. Wiley interdisciplinary reviews RNA. 2015.
376. Bousquet-Antonelli C, Deragon JM. A comprehensive analysis of the La-motif protein superfamily. RNA. 2009;15(5):750–64.
377. Sobel SG, Wolin SL. Two yeast La motif-containing proteins are RNA-binding proteins that associate with polyribosomes. Mol Biol Cell. 1999;10(11):3849–62.
378. Kershaw CJ, Costello JL, Castelli LM, Talavera D, Rowe W, Sims PF, et al. The yeast La related protein Slf1p is a key activator of translation during the oxidative stress response. PLoS Genet. 2015;11(1):e1004903.

379. Chauvet S, Maurel-Zaffran C, Miassod R, Jullien N, Pradel J, Aragnol D. dlarp, a new candidate Hox target in Drosophila whose orthologue in mouse is expressed at sites of epithelium/mesenchymal interactions. Dev Dyn: Official Publ Amer Assoc Anatomists. 2000;218(3):401–13.
380. Nykamp K, Lee MH, Kimble JC. elegans La-related protein, LARP-1, localizes to germline P bodies and attenuates Ras-MAPK signaling during oogenesis. RNA. 2008;14(7):1378–89.
381. Smith RW, Blee TK, Gray NK. Poly(A)-binding proteins are required for diverse biological processes in metazoans. Biochem Soc Trans. 2014;42(4):1229–37.
382. Merret R, Descombin J, Juan YT, Favory JJ, Carpentier MC, Chaparro C, et al. XRN4 and LARP1 are required for a heat-triggered mRNA decay pathway involved in plant acclimation and survival during thermal stress. Cell Rep. 2013;5(5):1279–93.
383. Korobeinikova AV, Garber MB, Gongadze GM. Ribosomal proteins: structure, function, and evolution. Biochem Biokhimiia. 2012;77(6):562–74.
384. Li GW, Burkhardt D, Gross C, Weissman JS. Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. Cell. 2014;157(3):624–35.
385. Iadevaia V, Caldarola S, Tino E, Amaldi F, Loreni F. All translation elongation factors and the e, f, and h subunits of translation initiation factor 3 are encoded by 5'-terminal oligopyrimidine (TOP) mRNAs. RNA. 2008;14(9):1730–6.
386. Meyuhas O, Kahan T. The race to decipher the top secrets of TOP mRNAs. Biochimica et biophysica acta. 2014.
387. Levy S, Avni D, Hariharan N, Perry RP, Meyuhas O. Oligopyrimidine tract at the 5' end of mammalian ribosomal protein mRNAs is required for their translational control. Proc Natl Acad Sci USA. 1991;88(8):3319–23.
388. Avni D, Shama S, Loreni F, Meyuhas O. Vertebrate mRNAs with a 5'-terminal pyrimidine tract are candidates for translational repression in quiescent cells: characterization of the translational cis-regulatory element. Mol Cell Biol. 1994;14(6):3822–33.
389. Shama S, Meyuhas O. The translational cis-regulatory element of mammalian ribosomal protein mRNAs is recognized by the plant translational apparatus. Eur J Biochem/FEBS. 1996;236(2):383–8.
390. Biberman Y, Meyuhas O. Substitution of just five nucleotides at and around the transcription start site of rat beta-actin promoter is sufficient to render the resulting transcript a subject for translational control. FEBS Lett. 1997;405(3):333–6.
391. Biberman Y, Meyuhas O. TOP mRNAs are translationally inhibited by a titratable repressor in both wheat germ extract and reticulocyte lysate. FEBS Lett. 1999;456(3):357–60.
392. Loreni F, Amaldi F. Translational control of terminal oligopyrimidine mRNAs requires a specific regulator. FEBS Lett. 1997;416(3):239–42.
393. Pellizzoni L, Cardinali B, Lin-Marq N, Mercanti D, Pierandrei-Amaldi P. A Xenopus laevis homologue of the La autoantigen binds the pyrimidine tract of the 5' UTR of ribosomal protein mRNAs in vitro: implication of a protein factor in complex formation. J Mol Biol. 1996;259(5):904–15.
394. Pellizzoni L, Lotti F, Maras B, Pierandrei-Amaldi P. Cellular nucleic acid binding protein binds a conserved region of the 5' UTR of Xenopus laevis ribosomal protein mRNAs. J Mol Biol. 1997;267(2):264–75.
395. Jefferies HB, Fumagalli S, Dennis PB, Reinhard C, Pearson RB, Thomas G. Rapamycin suppresses 5'TOP mRNA translation through inhibition of p70s6k. EMBO J. 1997;16 (12):3693–704.
396. Crosio C, Boyl PP, Loreni F, Pierandrei-Amaldi P, Amaldi F. La protein has a positive effect on the translation of TOP mRNAs in vivo. Nucleic Acids Res. 2000;28(15):2927–34.
397. Cardinali B, Carissimi C, Gravina P, Pierandrei-Amaldi P. La protein is associated with terminal oligopyrimidine mRNAs in actively translating polysomes. J Biol Chem. 2003;278 (37):35145–51.
398. Damgaard CK, Lykke-Andersen J. Translational coregulation of 5'TOP mRNAs by TIA-1 and TIAR. Genes Dev. 2011;25(19):2057–68.

399. Hsieh AC, Liu Y, Edlind MP, Ingolia NT, Janes MR, Sher A, et al. The translational landscape of mTOR signalling steers cancer initiation and metastasis. Nature. 2012;485 (7396):55–61.
400. Han K, Jaimovich A, Dey G, Ruggero D, Meyuhas O, Sonenberg N, et al. Parallel measurement of dynamic changes in translation rates in single cells. Nat Methods. 2014;11 (1):86–93.
401. Terada N, Patel HR, Takase K, Kohno K, Nairn AC, Gelfand EW. Rapamycin selectively inhibits translation of mRNAs encoding elongation factors and ribosomal proteins. Proc Natl Acad Sci USA. 1994;91(24):11477–81.
402. Tang H, Hornstein E, Stolovich M, Levy G, Livingstone M, Templeton D, et al. Amino acid-induced translation of TOP mRNAs is fully dependent on phosphatidylinositol 3-kinase-mediated signaling, is partially inhibited by rapamycin, and is independent of S6K1 and rpS6 phosphorylation. Mol Cell Biol. 2001;21(24):8671–83.
403. Lahr RM, Mack SM, Heroux A, Blagden SP, Bousquet-Antonelli C, Deragon JM, et al. The La-related protein 1-specific domain repurposes HEAT-like repeats to directly bind a 5'TOP sequence. Nucleic Acids Res. 2015;43(16):8077–88.
404. Mura M, Hopkins TG, Michael T, Abd-Latip N, Weir J, Aboagye E, et al. LARP1 post-transcriptionally regulates mTOR and contributes to cancer progression. Oncogene. 2015;34(39):5025–36.
405. Larsson O, Morita M, Topisirovic I, Alain T, Blouin MJ, Pollak M, et al. Distinct perturbation of the translatome by the antidiabetic drug metformin. Proc Natl Acad Sci USA. 2012;109(23):8977–82.
406. Morita M, Gravel SP, Chenard V, Sikstrom K, Zheng L, Alain T, et al. mTORC1 controls mitochondrial activity and biogenesis through 4E-BP-dependent translational regulation. Cell Metab. 2013;18(5):698–711.
407. Lipton JO, Sahin M. The neurology of mTOR. Neuron. 2014;84(2):275–91.
408. Takei N, Nawa H. mTOR signaling and its roles in normal and abnormal brain development. Front Mol Neurosci. 2014;7:28.
409. Hebb DO. The organization of behavior; a neuropsychological theory. New York,: Wiley; 1949. xix, 335 p.
410. Martin SJ, Grimwood PD, Morris RG. Synaptic plasticity and memory: an evaluation of the hypothesis. Annu Rev Neurosci. 2000;23:649–711.
411. Emes RD, Grant SG. Evolution of synapse complexity and diversity. Annu Rev Neurosci. 2012;35:111–31.
412. Moroz LL, Kohn AB. Independent origins of neurons and synapses: insights from ctenophores. Philos Trans R Soc Lond B Biol Sci. 2016;371(1685):20150041.
413. Sossin WS, Lacaille JC. Mechanisms of translational regulation in synaptic plasticity. Curr Opin Neurobiol. 2010;20(4):450–6.
414. Sutton MA, Schuman EM. Local translational control in dendrites and its role in long-term synaptic plasticity. J Neurobiol. 2005;64(1):116–31.
415. Costa-Mattioli M, Sossin WS, Klann E, Sonenberg N. Translational control of long-lasting synaptic plasticity and memory. Neuron. 2009;61(1):10–26.
416. Richter JD, Klann E. Making synaptic plasticity and memory last: mechanisms of translational regulation. Genes Dev. 2009;23(1):1–11.
417. Liu-Yesucevitz L, Bassell GJ, Gitler AD, Hart AC, Klann E, Richter JD, et al. Local RNA translation at the synapse and in disease. J Neurosci: official J Soc Neurosci. 2011;31 (45):16086–93.
418. Huang W, Zhu PJ, Zhang S, Zhou H, Stoica L, Galiano M, et al. mTORC2 controls actin polymerization required for consolidation of long-term memory. Nat Neurosci. 2013.
419. Dossani RH, Missios S, Nanda A. The Legacy of Henry Molaison (1926-2008) and the Impact of His Bilateral Mesial Temporal Lobe Surgery on the Study of Human Memory. World Neurosurg. 2015;84(4):1127–35.

420. Wang DO, Kim SM, Zhao Y, Hwang H, Miura SK, Sossin WS, et al. Synapse- and stimulus-specific local translation during long-term neuronal plasticity. Science. 2009;324 (5934):1536–40.
421. Tatavarty V, Ifrim MF, Levin M, Korza G, Barbarese E, Yu J, et al. Single-molecule imaging of translational output from individual RNA granules in neurons. Mol Biol Cell. 2012;23 (5):918–29.
422. Flexner JB, Flexner LB, Stellar E. Memory in mice as affected by intracerebral puromycin. Science. 1963;141(3575):57–9.
423. Debiec J, LeDoux JE, Nader K. Cellular and systems reconsolidation in the hippocampus. Neuron. 2002;36(3):527–38.
424. Nader K, Schafe GE, Le Doux JE. Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. Nature. 2000;406(6797):722–6.
425. Tischmeyer W, Schicknick H, Kraus M, Seidenbecher CI, Staak S, Scheich H, et al. Rapamycin-sensitive signalling in long-term consolidation of auditory cortex-dependent memory. Eur J Neurosci. 2003;18(4):942–50.
426. Parsons RG, Gafford GM, Helmstetter FJ. Translational control via the mammalian target of rapamycin pathway is critical for the formation and stability of long-term fear memory in amygdala neurons. J Neurosci: Official J Soc Neurosci. 2006;26(50):12977–83.
427. Bekinschtein P, Katche C, Slipczuk LN, Igaz LM, Cammarota M, Izquierdo I, et al. mTOR signaling in the hippocampus is necessary for memory formation. Neurobiol Learn Mem. 2007;87(2):303–7.
428. Blundell J, Kouser M, Powell CM. Systemic inhibition of mammalian target of rapamycin inhibits fear memory reconsolidation. Neurobiol Learn Mem. 2008;90(1):28–35.
429. Belelovsky K, Kaphzan H, Elkobi A, Rosenblum K. Biphasic activation of the mTOR pathway in the gustatory cortex is correlated with and necessary for taste learning. J Neurosci: Official J Soc Neurosci. 2009;29(23):7424–31.
430. Glover EM, Ressler KJ, Davis M. Differing effects of systemically administered rapamycin on consolidation and reconsolidation of context vs. cued fear memories. Learn Mem. 2010;17(11):577–81.
431. Gafford GM, Parsons RG, Helmstetter FJ. Consolidation and reconsolidation of contextual fear memory requires mammalian target of rapamycin-dependent translation in the dorsal hippocampus. Neuroscience. 2011;182:98–104.
432. Deli A, Schipany K, Rosner M, Hoger H, Pollak A, Li L, et al. Blocking mTORC1 activity by rapamycin leads to impairment of spatial memory retrieval but not acquisition in C57BL/6J mice. Behav Brain Res. 2012;229(2):320–4.
433. Halloran J, Hussong SA, Burbank R, Podlutskaya N, Fischer KE, Sloane LB, et al. Chronic inhibition of mammalian target of rapamycin by rapamycin modulates cognitive and non-cognitive components of behavior throughout lifespan in mice. Neuroscience. 2012;223:102–13.
434. Jobim PF, Pedroso TR, Christoff RR, Werenicz A, Maurmann N, Reolon GK, et al. Inhibition of mTOR by rapamycin in the amygdala or hippocampus impairs formation and reconsolidation of inhibitory avoidance memory. Neurobiol Learn Mem. 2012;97(1):105–12.
435. Bliss TV, Lomo T. Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. J Physiol. 1973;232 (2):331–56.
436. Pinsker HM, Hening WA, Carew TJ, Kandel ER. Long-term sensitization of a defensive withdrawal reflex in Aplysia. Science. 1973;182(4116):1039–42.
437. Stanton PK, Sarvey JM. Blockade of long-term potentiation in rat hippocampal CA1 region by inhibitors of protein synthesis. J Neurosci: Official J Soc Neurosci. 1984;4(12):3080–8.
438. Montarolo PG, Goelet P, Castellucci VF, Morgan J, Kandel ER, Schacher S. A critical period for macromolecular synthesis in long-term heterosynaptic facilitation in Aplysia. Science. 1986;234(4781):1249–54.

439. Ito M, Kano M. Long-lasting depression of parallel fiber-Purkinje cell transmission induced by conjunctive stimulation of parallel fibers and climbing fibers in the cerebellar cortex. Neurosci Lett. 1982;33(3):253–8.
440. Huber KM, Roder JC, Bear MF. Chemical induction of mGluR5- and protein synthesis–dependent long-term depression in hippocampal area CA1. J Neurophysiol. 2001;86(1):321–5.
441. Tang SJ, Reis G, Kang H, Gingras AC, Sonenberg N, Schuman EM. A rapamycin-sensitive signaling pathway contributes to long-term synaptic plasticity in the hippocampus. Proc Natl Acad Sci U S A. 2002;99(1):467–72.
442. Panja D, Dagyte G, Bidinosti M, Wibrand K, Kristiansen AM, Sonenberg N, et al. Novel translational control in Arc-dependent long term potentiation consolidation in vivo. J Biol Chem. 2009;284(46):31498–511.
443. Panja D, Kenney JW, D'Andrea L, Zalfa F, Vedeler A, Wibrand K, et al. Two-stage translational control of dentate gyrus LTP consolidation is mediated by sustained BDNF-TrkB signaling to MNK. Cell Rep. 2014;9(4):1430–45.
444. Rodriguez F, Lopez JC, Vargas JP, Broglio C, Gomez Y, Salas C. Spatial memory and hippocampal pallium through vertebrate evolution: insights from reptiles and teleost fish. Brain Res Bull. 2002;57(3–4):499–503.
445. Casadio A, Martin KC, Giustetto M, Zhu H, Chen M, Bartsch D, et al. A transient, neuron-wide form of CREB-mediated long-term facilitation can be stabilized at specific synapses by local protein synthesis. Cell. 1999;99(2):221–37.
446. Minichiello L, Calella AM, Medina DL, Bonhoeffer T, Klein R, Korte M. Mechanism of TrkB-mediated hippocampal long-term potentiation. Neuron. 2002;36(1):121–37.
447. Kempermann G. New neurons for 'survival of the fittest'. Nat Rev Neurosci. 2012;13 (10):727–36.
448. Schmidt-Hieber C, Jonas P, Bischofberger J. Enhanced synaptic plasticity in newly generated granule cells of the adult hippocampus. Nature. 2004;429(6988):184–7.
449. Kassabov SR, Choi YB, Karl KA, Vishwasrao HD, Bailey CH, Kandel ER. A single Aplysia neurotrophin mediates synaptic facilitation via differentially processed isoforms. Cell Rep. 2013;3(4):1213–27.
450. Antion MD, Merhav M, Hoeffer CA, Reis G, Kozma SC, Thomas G, et al. Removal of S6K1 and S6K2 leads to divergent alterations in learning, memory, and synaptic plasticity. Learn Mem. 2008;15(1):29–38.
451. Pearce LR, Alton GR, Richter DT, Kath JC, Lingardo L, Chapman J, et al. Characterization of PF-4708671, a novel and highly specific inhibitor of p70 ribosomal S6 kinase (S6K1). Biochem J. 2010;431(2):245–55.
452. Huynh TN, Santini E, Klann E. Requirement of Mammalian target of rapamycin complex 1 downstream effectors in cued fear memory reconsolidation and its persistence. J Neurosci: Official J Soc Neurosci. 2014;34(27):9034–9.
453. Mac Callum PE, Hebert M, Adamec RE, Blundell J. Systemic inhibition of mTOR kinase via rapamycin disrupts consolidation and reconsolidation of auditory fear memory. Neurobiol Learn Mem. 2014;112:176–85.
454. Im HI, Nakajima A, Gong B, Xiong X, Mamiya T, Gershon ES, et al. Post-training dephosphorylation of eEF-2 promotes protein synthesis for memory consolidation. PLoS ONE. 2009;4(10):e7424.
455. Gildish I, Manor D, David O, Sharma V, Williams D, Agarwala U, et al. Impaired associative taste learning and abnormal brain activation in kinase-defective eEF2K mice. Learn Mem. 2012;19(3):116–25.
456. McCamphill PK, Farah CA, Anadolu MN, Hoque S, Sossin WS. Bidirectional regulation of eEF2 phosphorylation controls synaptic plasticity by decoding neuronal activity patterns. J Neurosci: Official J Soc Neurosci. 2015;35(10):4403–17.
457. Graber TE, Hebert-Seropian S, Khoutorsky A, David A, Yewdell JW, Lacaille JC, et al. Reactivation of stalled polyribosomes in synaptic plasticity. Proc Natl Acad Sci U S A. 2013;110(40):16205–10.

458. Sharma A, Hoeffer CA, Takayasu Y, Miyawaki T, McBride SM, Klann E, et al. Dysregulation of mTOR signaling in fragile X syndrome. J Neurosci: Official J S Neurosci. 2010;30(2):694–702.
459. Darnell JC, Klann E. The translation of translational control by FMRP: therapeutic targets for FXS. Nat Neurosci. 2013;16(11):1530–6.
460. Buxbaum AR, Wu B, Singer RH. Single beta-actin mRNA detection in neurons reveals a mechanism for regulating its translatability. Science. 2014;343(6169):419–22.
461. Ceman S, O'Donnell WT, Reed M, Patton S, Pohl J, Warren ST. Phosphorylation influences the translation state of FMRP-associated polyribosomes. Hum Mol Genet. 2003;12 (24):3295–305.
462. Muddashetty RS, Nalavadi VC, Gross C, Yao X, Xing L, Laur O, et al. Reversible inhibition of PSD-95 mRNA translation by miR-125a, FMRP phosphorylation, and mGluR signaling. Mol Cell. 2011;42(5):673–88.
463. Narayanan U, Nalavadi V, Nakamoto M, Thomas G, Ceman S, Bassell GJ, et al. S6K1 phosphorylates and regulates fragile X mental retardation protein (FMRP) with the neuronal protein synthesis-dependent mammalian target of rapamycin (mTOR) signaling cascade. J Biol Chem. 2008;283(27):18478–82.
464. Niere F, Wilkerson JR, Huber KM. Evidence for a fragile X mental retardation protein-mediated translational switch in metabotropic glutamate receptor-triggered Arc translation and long-term depression. J Neurosci: Official J Soc Neurosci. 2012;32(17):5924–36.
465. Till SM, Li HL, Miniaci MC, Kandel ER, Choi YB. A presynaptic role for FMRP during protein synthesis-dependent long-term plasticity in Aplysia. Learn Mem. 2011;18(1):39–48.
466. Banko JL, Poulin F, Hou L, DeMaria CT, Sonenberg N, Klann E. The translation repressor 4E-BP2 is critical for eIF4F complex formation, synaptic plasticity, and memory in the hippocampus. J Neurosci: Official J Soc Neurosci. 2005;25(42):9581–90.
467. Banko JL, Hou L, Poulin F, Sonenberg N, Klann E. Regulation of eukaryotic initiation factor 4E by converging signaling pathways during metabotropic glutamate receptor-dependent long-term depression. J Neurosci: Official J Soc Neurosci. 2006;26(8):2167–73.
468. Weatherill DB, Dyer J, Sossin WS. Ribosomal protein S6 kinase is a critical downstream effector of the target of rapamycin complex 1 for long-term facilitation in Aplysia. J Biol Chem. 2010;285(16):12255–67.
469. Patursky-Polischuk I, Stolovich-Rain M, Hausner-Hanochi M, Kasir J, Cybulski N, Avruch J, et al. The TSC-mTOR pathway mediates translational activation of TOP mRNAs by insulin largely in a raptor- or rictor-independent manner. Mol Cell Biol. 2009;29 (3):640–9.
470. Moccia R, Chen D, Lyles V, Kapuya E, Kalachikov S, et al. An unbiased cDNA library prepared from isolated Aplysia sensory neuron processes is enriched for cytoskeletal and translational mRNAs. J Neurosci: Official J Soc Neurosci. 2003;23(28):9409–17.
471. Poon MM, Choi SH, Jamieson CA, Geschwind DH, Martin KC. Identification of process-localized mRNAs from cultured rodent hippocampal neurons. J Neurosci: Official J Soc Neurosci. 2006;26(51):13390–9.
472. Tsokas P, Grace EA, Chan P, Ma T, Sealfon SC, Iyengar R, et al. Local protein synthesis mediates a rapid increase in dendritic elongation factor 1A after induction of late long-term potentiation. J Neurosci: Official J Soc Neurosci. 2005;25(24):5833–43.
473. Gobert D, Topolnik L, Azzi M, Huang L, Badeaux F, Desgroseillers L, et al. Forskolin induction of late-LTP and up-regulation of 5' TOP mRNAs translation via mTOR, ERK, and PI3K in hippocampal pyramidal cells. J Neurochem. 2008;106(3):1160–74.
474. Slavov N, Semrau S, Airoldi E, Budnik B, van Oudenaarden A. Differential Stoichiometry among Core Ribosomal Proteins. Cell Rep. 2015;13(5):865–73.
475. Dinman JD. Pathways to Specialized Ribosomes: The Brussels Lecture. J Mol Biol. 2016.
476. McClatchy DB, Fang G, Levey AI. Elongation factor 1A family regulates the recycling of the M4 muscarinic acetylcholine receptor. Neurochem Res. 2006;31(7):975–88.

477. Siuta MA, Robertson SD, Kocalis H, Saunders C, Gresch PJ, Khatri V, et al. Dysregulation of the norepinephrine transporter sustains cortical hypodopaminergia and schizophrenia-like behaviors in neuronal rictor null mice. PLoS Biol. 2010;8(6):e1000393.
478. Carson RP, Fu C, Winzenburger P, Ess KC. Deletion of Rictor in neural progenitor cells reveals contributions of mTORC2 signaling to tuberous sclerosis complex. Hum Mol Genet. 2013;22(1):140–52.

Translational Control in Echinoderms: The Calm Before the Storm

Patrick Cormier, Héloïse Chassé, Bertrand Cosson, Odile Mulner-Lorillon and Julia Morales

1 Echinoderms and Translational Regulation, a Link with the Cell Cycle

Echinoderms, a sister group to chordates, are marine metazoans that represent the ancient phylum Echinodermata, at least 450 million years old, containing thousands of known living species present throughout the oceans [1]. Among the five main classes of modern echinoderms—Asteroida (starfish), Echinodea (sea urchin and sand dollars), Ophiuroida (brittle stars), Holothuroidea (sea cucumbers) and Crinoidea (sea lilies and feather stars)—starfish and sea urchin have been extensively used to study gametogenesis, fertilization and embryonic development. Eggs and embryos of echinoderms, which are gonochoric animals, have been used in embryological studies for more than a century, because these giant cells are easy to manipulate and to observe under optical microscopy. Furthermore, the possibility to obtain a large number of eggs from one female and the synchronism of cell divisions allow the analysis of the molecular mechanisms that govern the oocyte maturation, egg fertilization and early stages of embryonic development.

P. Cormier (✉) · H. Chassé · B. Cosson · O. Mulner-Lorillon · J. Morales
Sorbonne Universités, UPMC Univ Paris 06, UMR 8227,
Integrative Biology of Marine Models, Translation Cell Cycle and Development,
Station Biologique de Roscoff, CS 90074, 29688 Roscoff Cedex, France
e-mail: cormier@sb-roscoff.fr

P. Cormier · H. Chassé · B. Cosson · O. Mulner-Lorillon · J. Morales
CNRS, UMR 8227, Integrative Biology of Marine Models,
Station Biologique de Roscoff, CS 90074, 29688 Roscoff Cedex, France

Present Address:
B. Cosson
Université Paris Diderot, Sorbonne Paris Cité, Epigenetics and Cell Fate,
UMR7216, CNRS, 75013 Paris, France

G. Hernández and R. Jagus (eds.), *Evolution of the Protein Synthesis Machinery and Its Regulation*, DOI 10.1007/978-3-319-39468-8_16

Crucial to developmental biology, marine invertebrates have also been essential in cell cycle studies. It was in the sea urchin *Arbacia punctulata* that Tim Hunt identified the first cyclin, as he noticed the "unexpected behavior" of a protein synthesized early after fertilization that almost disappeared before cleavage and was "henceforth called cyclin" [2]. Together with studies on amphibians, data obtained in marine invertebrates, i.e., the detection in starfish oocytes of MPF, the M-phase promoting factor, and the demonstration of its universality [3], the sequencing of Clam cyclin A [4] and the purification of CDK1/cyclinB as the M-Phase promoting factor from starfish [5] were milestones in the comprehension of the mechanism underlying the control of meiotic and mitotic divisions (reviewed in [6]).

In starfish, after completion of vitellogenesis the oocytes are arrested in the prophase of first meiotic division. Meiosis reinitiation is induced by 1-methyladenine (1-MA), which acts on a yet-unidentified receptor on the oocyte surface. 1-MA initiates the transduction pathway that leads, by an activation cascade, the transformation of stored inactive pre-MPF (M-phase promoting factor) into active CDK1/cyclin B, responsible for G2/M transition. The rate of protein synthesis increases upon meiosis reinitiation, and translation is necessary for meiosis resumption and early development [7, 8].

Sea urchin eggs are metabolically quiescent cells blocked at the G1 stage of the cell cycle after completion of their meiotic divisions. Therefore, egg fertilization involves molecular mechanisms that are independent of the oocyte meiotic maturation. Fertilization of the G1 blocked haploid sea urchin egg triggers entry into S-phase and completion of the first mitotic division of the embryonic development. De novo protein synthesis is dispensable for the S-phase but is needed for the onset of M-phase and subsequent embryonic cell cycles [9, 10]. Protein synthesis is low in unfertilized eggs and is stimulated rapidly following fertilization. This rise in protein synthesis is not blocked by actinomycin D [11] and is independent of mRNA transcription and ribosome biogenesis [12, 13]. Accordingly, inhibition of new RNA transcription has no effect on the three first mitotic divisions of the sea urchin early embryos. Therefore, sea urchin eggs and early embryos provide an example of regulated gene expression at the translational level using maternal mRNAs that do not follow the traditional scheme of gene regulation governed at the transcription level [14].

The unfertilized sea urchin eggs were supposed to contain all the required components for translation [15], and the availability of the sea urchin genome further confirmed the presence of the "translational toolkit" in this model system [16, 17]. The exact mechanisms by which translation is repressed in unfertilized eggs and is increased following fertilization have been questioned for a long time, but until the 1980s there was only little evidence to explain this regulation by tuning the activity of the translation machinery [18]. Previous work reported the presence of cytoplasmic messenger ribonucleoprotein (mRNP) particles in sea urchin embryos [19] and the "unmasking" of stored mRNPs and ribosomes at fertilization [20, 21]. Because the assembly of masked mRNP complexes takes place during oogenesis, the sea urchin model is currently presented as a system that exemplifies a reversible process of mRNA repression and activation [14]. Therefore, for several

years the dogma has held that "masked" mRNAs cause protein synthesis inhibition in unfertilized eggs. The development of cell-free translation systems showed that unmasking the mRNA process is not the whole story [22]. At least two initiation factors were shown to be activated in cell-free extracts mimicking egg activation triggered by fertilization [23–26]: eIF2 [responsible for binding the initiator methionyl-transfer RNA (Met-tRNA) to the ribosome] and eIF4F (involved in the recognition of mRNA 5′ cap structures and the recruitment of the small ribosomal subunit).

Starfish oocytes and sea urchin eggs are relevant models that have helped to address the control of gene expression at the translational level in relation to cell cycle regulation in physiological conditions (Fig. 1). The mechanisms by which translation is increased following fertilization and how those controls contribute to proper development have just started to be understood in the early twenty-first century.

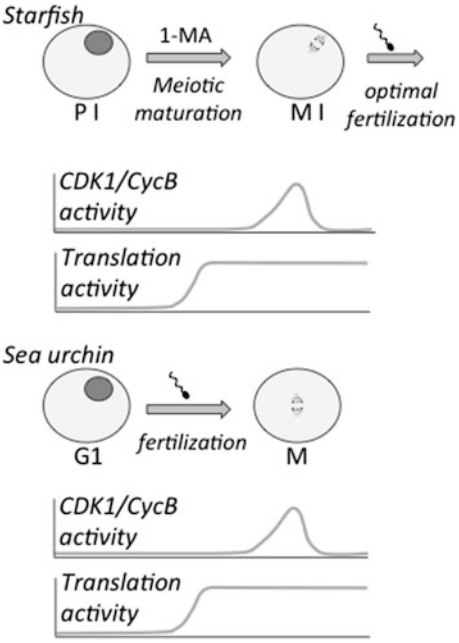


Fig. 1 Starfish oocytes and sea urchin eggs are models in which cell cycle regulation and translational control are interconnected. *Top panel* Starfish oocytes are arrested in prophase I (PI) of meiosis. Maturation is triggered by the hormone 1-methyladenine (1-MA) and proceeds through the first meiotic division. Fertilization occurs naturally between MI and MII. Translation activity increases after the onset of meiosis, preceding the CDK1/cyclin B activity responsible for cell division. *Bottom panel* Sea urchin eggs are metabolically quiescent cells blocked at the G1 stage of the cell cycle after completion of their meiotic divisions. Protein synthesis is low in unfertilized eggs and is stimulated rapidly following fertilization. Fertilization triggers the activation of CDK1/cyclin B and the entry into the first mitotic division of the embryonic development

2 Translational Regulation During Starfish Oocyte Meiotic Maturation

The twofold increase in the rate of global protein synthesis during meiosis in starfish oocyte was reported to be due in part to the activation of the eIF4F (eukaryotic initiation factor 4F) complex [27]. In mammals, eIF4F facilitates the recruitment of ribosomes to the mRNA 5′ and is described as a heterotrimeric complex [28]. eIF4F is composed of eIF4G (eukaryotic initiation factor 4G), a large scaffolding protein that interacts with eIF4E (eukaryotic initiation factor 4E), which recognizes and binds to the 5′ cap structure (m^7GpppN, where N is any nucleotide), and eIF4A (eukaryotic initiation factor 4A), an RNA-dependant ATPase and RNA helicase. eIF4G associates with eIF3 and consequently provides a physical link between the 5′ end of capped mRNA and the ribosome. eIF4G also interacts with PABP (poly (A)-binding protein) to activate translation of polyadenylated RNAs and with the MAP-kinase-interacting kinases Mnk1 and Mnk2, which phosphorylate eIF4E at serine 209 when associated with eIF4G [29]. eIF4E represents the best known substrate for Mnks, which are activated by signaling through the mitogenic MAP kinase (ERK) pathway (reviewed in [30, 31]). Although the physiological relevance of eIF4E phosphorylation is not fully elucidated, it is proposed to regulate eIF4E affinity to mRNAs [30].

The eIF4E-related proteins can be clustered in three families called eIF4E1, eIF4E2 and eIF4E3 [32]. A comparative genomics platform for the echinoderm clade is accessible (http://Echinobase.org [33]), where the genome for the *Patiria miniata* starfish is available. The three representatives of eIF4E-related genes have been found in starfishes. The first isolated starfish eIF4E1 sequence was obtained from the starfish species *Pisaster ochraceus* [34]. The starfish sequence contains the putative phosphorylation site serine 209.

In starfish oocyte, eIF4E was shown to be phosphorylated 20 min after addition of 1-MA [35]. Increased phosphorylation of eIF4E correlates with the initial protein synthesis and activity of several protein kinases such as PKC (protein kinase C), CDK1 (cyclin-dependent kinase 1) and mitogen-activated myelin basic protein kinase (MBP kinase) [35]. Using an in vitro kinase assay [34], it was shown that neither MBP kinase nor CDK1 phosphorylated the starfish eIF4E glutathione-S-transferase fusion protein, whereas starfish eIF4E was readily phosphorylated by native starfish protein kinase C-related kinase 2 (PRK2). PRK2 is a Ser/Thr kinase and Rho/Rac effector protein, required in human cells for abscission of the midbody at the end of the cell division cycle and for phosphorylation and activation of Cdc25B, the phosphatase required for activation of mitotic CDK1/cyclin B complexes at the G2/M transition [36]. In vitro phosphorylated GST-4E contains the same phosphopeptides as in vivo phosphorylated eIF4E, suggesting that PRK2 mediates the phosphorylation on eIF4E [34]. While the *P. miniata* genome contains partial sequences of MNK1, the putative role of MNKs in starfish eIF4E phosphorylation has not yet been analyzed. The respective role of PRK2 and MNK1 and the biological significance of eIF4E

phosphorylation during starfish oocyte meiotic maturation are not understood and remain important questions to be answered.

In amphibian oocyte, eIF4E was reported to be sequestered by association with a translational repressor, Maskin, which binds to CPEB (cytoplasmic polyadenylation element binding protein) and inhibits eIF4E binding to eIF4G. Dissociation of the CPEB/Maskin/eIF4E complex is essential for *cyclin B1* mRNA translational activation [37]. Studies in starfish oocytes showed that translational control of *cyclin B* mRNA is achieved through two separate but related mechanisms: polyadenylation and translational derepression [38, 39]. On the one hand, polyadenylation of pre-existing mRNAs was shown to regulate *cyclin B* translation during the reinitiation of the starfish meiosis oocyte [40, 41]. On the other hand, we recently reported that cyclin B translation is correlated with the phosphorylation and the dissociation of CPEB from eIF4E [39]. In prophase-blocked oocyte, CPEB phosphorylation is inhibited by the activity of a cytoplasmic type 1 phosphatase, which is inhibited by a nuclear type 2-inhibitor released after the disruption of the germinal vesicle induced by 1-MA [38] (Fig. 2).

Interestingly, we showed that this CPEB-correlated cyclin B translation was not affected by rapamycin, whereas the drug was found to suppress the 1-MA-induced burst of global protein synthesis occurring at the G2/M transition [39]. Rapamycin, an inhibitor of the mTOR (mechanistic target of rapamycin) pathway, is known to affect the dissociation of eIF4E from its inhibitor 4E-BP (eIF4E-binding protein).

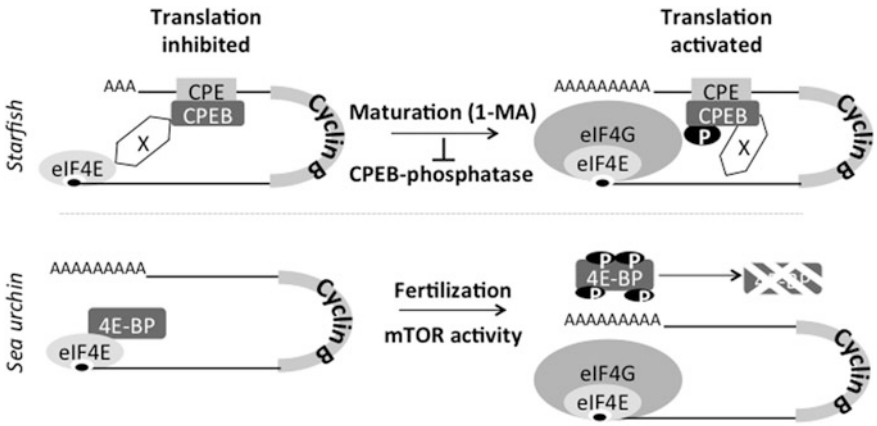


Fig. 2 Translation of *cyclin B* mRNA is controlled by different mechanisms during starfish meiotic maturation and sea urchin fertilization. In prophase-blocked starfish oocyte, the *cyclin B* mRNA is associated to CPEB and eIF4E in a translationally inactive conformation. 1-MA induces the inhibition of a CPEB phosphatase. CPEB is then phosphorylated and dissociates from eIF4E; additionally polyadenylation of *cyclin B* mRNA increases leading to its translation. In unfertilized sea urchin eggs, eIF4E is associated to its inhibitor 4E-BP, inhibiting *cyclin B* mRNA translation. Fertilization triggers the phosphorylation and the degradation of 4E-BP through the mTOR pathway, releasing eIF4E from its repressor and allowing the formation of an active initiation complex, leading to *cyclin B* mRNA translation

4E-BPs negatively regulated cap-dependent translation in mammalian cells by competing with eIF4G on the same binding site in eIF4E. Phosphorylation of 4E-BPs is regulated by mTOR and controls its association with eIF4E [42, 43]. Hypophosphorylated 4E-BPs associate with eIF4E and repress cap-dependent translation, while hyperphosphorylated forms do not [44, 45].

Therefore, during starfish meiotic maturation, cyclin B translation specifically depends on a CPEB regulation pathway, whereas global protein synthesis appears to be regulated by the mTOR pathway. Important insights were provided on this crucial translational regulation pathway by studies in sea urchin eggs where the existence of the 4E-BP ortholog in echinoderms was first revealed [46].

3 4E-BP Regulation Following Fertilization of Sea Urchin Eggs

The huge rise in the rate of translation induced by fertilization of sea urchins eggs is partly explained by an increase of the activity of the cap-binding complex eIF4F [25, 26]. It was suggested that eIF4F activity was inhibited in unfertilized eggs owing to the presence of an unidentified repressor that blocked eIF4E availability [47]. More recently, we demonstrated that 4E-BP plays a major role in eIF4E sequestration in unfertilized eggs and that eIF4E is released from its repressor in correlation with the rise of protein synthesis that occurs rapidly following egg fertilization [46, 48]. A single 4E-BP ortholog exists in sea urchin [16, 17] in agreement with the Joshi/Jagus 4E-BPs database (http://umbicc3-215.umbi.umd.edu). During metazoan evolution, 4E-BP gene duplication occurred recently in vertebrates, and consequently a single copy of the 4E-BP ortholog is found per species from radial animals to echinoderms, except in annelids that have more than one 4E-BP gene and in *Caenorhabditis elegans* where 4E-BP gene is absent [49]. Sea urchin 4E-BP shares the eIF4E-recognition motif YXXXXLΦ (where X is any amino acid and Φ is a hydrophobic residue with other eIF4E-interacting proteins such as eIF4G [50, 51]). 4E-BP competes with eIF4G for a mutually exclusive binding site on the dorsal surface of eIF4E. Multiple and hierarchical phosphorylation events on 4E-BP1, the most intensively studied of mammalian 4E-BPs, are required for its release from eIF4E [52]. Using a sea urchin cell-free translation system that mimics the increase of the rate of translation initiation in fertilized eggs, we characterized the translational repression activity of the sea urchin 4E-BP, which possesses the conserved core domain containing the phosphorylation sites (T^{37}, T^{46}, S^{65} and T^{70}, numbered according to human 4E-BP1) [53]. We showed that a variant mimicking the hyperphosphorylation of the four critical phosphorylation sites of 4E-BP is not sufficient for release from eIF4E and translation promotion. These data suggested that there are additional mechanisms to the phosphorylation at the four critical sites of 4E-BP necessary to dissociate the eIF4E/4E-BP complex. Using small-angle X-ray scattering (SAXS), we reported that sea urchin 4E-BP is

intrinsically disordered but undergoes a dramatic compaction in the bound state, forming a "fuzzy complex" with eIF4E, giving new perspectives for the understanding of eIF4E/4E-BP complex regulation. We built a new model of interaction, involving a much larger binding footprint (residues 25-88) of 4E-BP on eIF4E [54]. Interestingly, our low-resolution model derived from SAXS data on sea urchin 4E-BP resembles the published dynamic model of human 4E-BP2 interaction with eIF4E [55]. 4E-BPs bind to eIF4E through a bipartite interface that consists of the well-known eIF4E-binding domain YXXXXLΦ connected by a linker of 15-30 residues to a "non-canonical" eIF4E-binding motif (NC 4E-BM) [56]. The fuzzy nature results from the NC 4E-BM motif that is highly dynamic in the complex. Between orthologs across species, the non-canonical motifs do not share a conserved sequence, but they all seem to bind to a conserved lateral domain of eIF4E that is not required for eIF4G binding.

We showed in the sea urchin *Sphaerechinus granularis* that 4E-BP degradation triggered by fertilization allows eIF4E association with eIF4G and consequent protein synthesis increase (Fig. 3) [48, 57]. The control of 4E-BPs stability was reported to play a regulatory role to control eIF4E availability [48, 58–60]. Down- and upregulation of the 4E-BP protein level was first observed during embryonic development of sea urchin [61]. Two important cellular stresses, hypoxia and bleomycin prolonged checkpoint mobilization, triggered 4E-BP protein overexpression in developing sea urchin embryos [62]. Chromium (III) induced DNA-damage provoked a time- and dose-dependent increase in the level of 4E-BP protein in sea urchin embryos [63]. We provided experimental evidence supporting that 4E-BP degradation, and consequently eIF4E release from its repressor, is under the control of Ca2+ and pH-dependent events [64]. 4E-BP degradation induced by fertilization of sea urchin eggs is affected by rapamycin treatment, suggesting an important role of mTOR-mediated regulation of translation initiation at the early embryonic development [48]. Modelization of the translational regulation following fertilization in sea urchin showed a requirement for a strong stimulation of the 4E-BP-degradation mechanism [65]. Therefore, sea urchin egg fertilization represents a powerful model to analyze the regulation of the 4E-BP level and the dynamic of protein synthesis that may regulate stem cell homeostasis and determine developmental changes [66].

mTOR is a serine/threonine protein kinase that associates with several proteins to form two distinct complexes named mTOR complex 1 (TORC1) and 2 (TORC2), which are evolutionarily conserved from yeast to mammals. Both complexes have been involved in the regulation of cell growth in response to various stimuli, TORC1 controlling the cell mass while TORC2 is implicated in the control of cell surface area [67]. TORC1 phosphorylates p70 ribosomal S6 protein kinase, 4E-BP1 and eEF2K (eukaryotic elongation factor 2 kinase), which in turn promote protein synthesis (reviewed in [68]).

In mammalian cells, rapamycin inhibits TOR kinase activity by association with FKBP12 (12-kDa FK506 binding protein), and together they affect TOR enzymatic

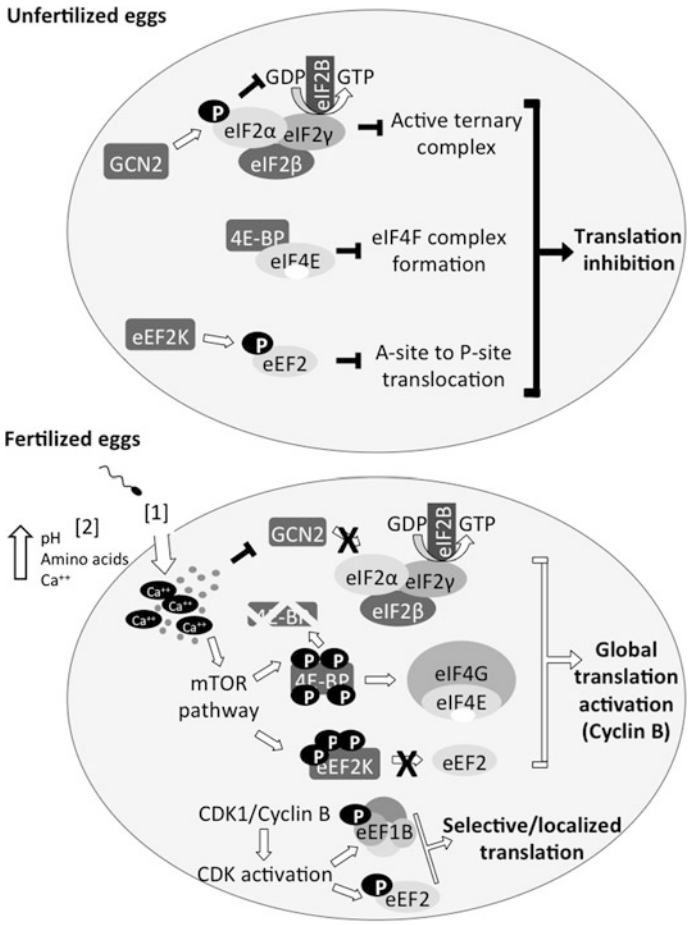


Fig. 3 Translational control at fertilization involves the regulation of several translation factors. In unfertilized eggs, several initiation and elongation factors are targeted to maintain a low translation activity. The ternary complex activity is inhibited by the phosphorylation of eIF2α by GCN2. The eIF4F complex formation is inhibited by the interaction of eIF4E with its inhibitor 4E-BP. The A-site to P-site translocation of the elongation step is inhibited by the phosphorylation of eEF2 by eEF2 kinase. After sperm entry [1], early events after fertilization include a calcium and a pH increase and entry of amino acids [2]. GCN2 is inhibited by the amino acid increase, and eIF2α is dephosphorylated and therefore active for the initiation step. The activated mTOR pathway triggers 4E-BP phosphorylation and degradation, allowing eIF4E to engage in the initiation complex with eIF4G. mTOR also activates eEF2 and the elongation step by phosphorylating and inhibiting eEF2 kinase. Activation of the initiation and elongation steps of translation contributes to the protein synthesis increase and mRNA translation after fertilization, in particular to cyclin B synthesis leading to CDK1/cyclin B complex formation. Furthermore, a pool of eEF2 and eEF1B are phosphorylated following the activation of CDK1/cyclin B, probably inducing localized or selective translation. See text for details

activities by binding next to the kinase region of mTOR but not competing with ATP binding [69]. Using PP242, a novel and specific ATP-competitive inhibitor of mTOR kinase [70], we recently confirmed that 4E-BP is degraded by an mTOR-sensitive pathway triggered by sea urchin fertilization [71]. A few reports [59, 60] had been published related to the ubiquitination of 4E-BPs and degradation by the proteasome. In mammalian cells, ^{57}Lys was the potential ubiquitination site in 4E-BP1. Only hypophosphorylated 4E-BP1 could be degraded, suggesting that 4E-BP degradation proceeds only when 4E-BP is inaccessible to mTOR kinase activity. Since mTOR inhibitors alter 4E-BP phosphorylation and degradation following fertilization, it is unlikely that this mechanism of 4E-BP degradation corresponds to the one observed in mammalian cells. Moreover, neither ^{57}Lys nor ^{69}Lys and ^{105}Lys present in human 4E-BP1 are conserved in sea urchins [17, 53]. Thus it is likely that mTOR-mediated regulation of 4E-BP degradation after fertilization relies on another molecular mechanism that remains to be elucidated.

Using two sea urchin species (*Paracentrotus lividus* and *Sphaerechinus granularis*) separated by 20 million years of evolutionary time, we showed that the involvement of the mTOR signaling pathway in 4E-BP degradation, *cyclin B* mRNA recruitment into polysomes and cyclin B protein accumulation following fertilization is conserved [71]. Altogether, these data bring evidence of the control of cyclin B translation via mTOR signaling and highlight the fine-tuning orchestration of mitotic cyclin translation required for CDK1 activation following fertilization in sea urchins (Fig. 2).

Interestingly, the amount of *cyclin B* mRNA present in active polysomes appears to be insensitive to PP242 treatment, suggesting that an alternative mTOR-independent signaling pathway acts in parallel to control *cyclin B* mRNA translation triggered by fertilization [71]. Some reports have put forward 3′ UTR-binding proteins in the control of *cyclin B* mRNA translation in different species (reviewed in [72]). Whether a similar mechanism could be involved in the alternative pathway for *cyclin B* translation remains to be uncovered in sea urchin.

Sea urchin contains genes involved in cytoplasmic polyadenylation such as Symplekin, CPSF and CPEB (reviewed in [17]). Polyadenylation of mRNAs increases following egg fertilization [73]. However, cordycepin, an inhibitor of RNA adenylation, does not affect sea urchin development prior to hatching [74], suggesting that CPEB-mediated polyadenylation is not required for the first mitotic divisions in sea urchin. Interestingly, in starfish eggs, cyclin A and B are translated after fertilization in a mechanism that appears not to be regulated by elongation of poly(A) tail length [75], in contrast to meiotic maturation as described in the previous paragraph [39]. Whether similar mechanisms implicated in *cyclin* mRNA translation are conserved between sea urchin and starfish following fertilization remains to be determined.

4 eIF2α Regulation in Protein Synthesis Control in Sea Urchin

Another regulatory mode at the level of eIF2 (eukaryotic translation initiation factor 2) plays a key role in the regulation of mRNA translation initiation. eIF2 is a complex of three subunits (α, β and γ). Within a ternary complex with GTP and initiator methionyl-tRNA, it mediates the delivery of initiator methionyl-tRNA to the ribosomes. The GTP bound to eIF2 is hydrolyzed during the initiation step of translation. The GDP to GTP exchange is catalyzed by eIF2B, a guanine-nucleotide exchange factor. In response to stresses, the alpha subunit of eIF2 is phosphorylated at a conserved serine (Ser-51 in mammals) and sequesters the guanine-nucleotide exchange complex eIF2B, thus inhibiting translation initiation and general protein synthesis while inducing the translation of specific mRNAs [76].

Phosphorylation of eIF2α is performed by four known serine/threonine protein kinases that share a related kinase domain but respond to different stimuli through specific regulatory domains: the general control non-derepressible 2 (GCN2), the double-stranded RNA protein kinase (PKR), the PKR-like endoplasmic reticulum kinase (PERK) and the heme-regulated inhibitor kinase (HRI) [77]. Only three eIF2α kinases were described in the sea urchin genome, GCN2, PERK and HRI [17]. It is currently accepted that translational control by eIF2α phosphorylation is a conserved adaptation to cell stress that has existed since the onset of eukaryotes [49].

In sea urchin, it had been suggested that eIF2 activity could play a role in the protein synthesis regulation occurring at fertilization since the addition of eIF2B to sea urchin egg extracts stimulated the protein synthesis rate [24, 78]. We showed that the alpha subunit of eIF2 is regulated after fertilization [79]. In the sea urchin unfertilized egg the eIF2α subunit is phosphorylated, as revealed by an antibody directed against the phospho-Ser51-containing domain. Fertilization triggers the dephosphorylation of eIF2α, while the amount of the protein is not modified. By using a phosphomimetic mutant of eIF2α, we showed that dephosphorylation of eIF2α is induced by fertilization and is necessary for the cell cycle division. Therefore, our data indicate that eIF2α should contribute to the regulation of protein synthesis required for the first mitotic division in sea urchin embryos [79] (Fig. 3).

Several lines of evidence suggest that GCN2, the most ancestral eIF2α kinase, is likely to be involved in the regulation of eIF2α occurring in sea urchin fertilization. GCN2 is expressed in early development (http://Echinobase.org) [33, 80], and phosphopeptides corresponding exclusively to GCN2 were detected in unfertilized eggs [81]. Phosphorylation of threonine residues in the activation loop of the GCN2 kinase domain is required for its activation [82]. The sea urchin GCN2 possesses two potential autophosphorylation sites in the conserved kinase domain. Using an antibody recognizing a phospho-peptide of the mouse activated kinase, we showed that GCN2 is phosphorylated in unfertilized eggs and fertilization triggers the dephosphorylation of the kinase [79]. These data suggest that GCN2 is active in unfertilized eggs and is inactivated following fertilization.

GCN2 is known to regulate protein synthesis in response to amino acid starvation. Several studies have shown that changes in amino acid uptake occur at fertilization [83, 84], and recently functional amino acid transporters were cloned in sea urchin [85]. GCN2 kinase possesses a HisRS domain homologous to Histidyl-tRNA synthetases, which can be activated by the binding of uncharged tRNAs. Increasing the pool of charged tRNAs would lead to inactivation of GCN2 [77]. We hypothesize that the increase in amino acid uptake at fertilization may lead to an increase in the charged tRNAs pool. Interestingly, by using alcoholic derivatives of amino acids, which are competitive inhibitors of tRNA synthetases and inhibit amino acid charging of tRNA and aminoacylation of tRNA [86], we could induce the phosphorylation of eIF2α and the inhibition of a protein synthesis increase in sea urchin fertilized eggs [79].

GCN2 has recently been shown to act also as a cell cycle regulator: UV radiation and DNA damaging agent MMS activate GCN2 in yeast and in human [87–90]. GCN2 regulates a G1/S checkpoint in yeast by acting upon the pre-replicative complex [90]. Interestingly, exposure of sea urchin embryos to MMS induces the phosphorylation of eIF2α and the inhibition of protein synthesis [91]. A report showed recently that the checkpoint kinase pathway is involved in the G1 arrest in sea urchin eggs [92]. Moreover, cell cycle arrested mouse oocytes contain active GCN2 kinase [93]. Natural cell cycle arrests often use existing checkpoint control occurring in dividing cells: an appealing hypothesis is that the GCN2 kinase, in addition to its role in maintaining a low protein synthesis activity in reproductive cells, may be implicated in the maintenance of cell cycle arrest in oocytes prior to fertilization.

Interestingly, regulation of eIF2α phosphorylation has been shown in mouse [93] and *Xenopus* [94] meiotic maturation and in drosophila egg activation [95] in addition to our data in sea urchin fertilization. Therefore, the GCN2 kinase and its downstream target eIF2α would not only be involved in the stress response, but would also be implicated in protein synthesis, which is physiologically regulated in unfertilized eggs and following fertilization.

5 Elongation in Sea Urchin

Besides the evident regulatory role of the initiation step, a number of data support a role of polypeptide elongation in regulating important physiological processes such as early development, neural function, cell stress, cell growth and proliferation, and cancer development [96–99].

Regulating the elongation step is at first essential to ensure the coordination with changes in the initiation rate in order to guarantee translational accuracy and fidelity. Furthermore, inhibiting elongation, which results in freezing mRNAs loaded on polysomes, the so-called stalled ribosomes, is a powerful means to ensure that the cell is prepared for a fast response to a stimulus or for a rapid reversion from a transient stimulation. Stalled ribosomes also have a protective function for

the mRNAs by limiting access to the nucleases. Besides its role in the regulation of the overall protein synthesis rate, it has been proposed that transitory inhibition of elongation could be a means of specifically regulating the level of short-lived proteins such as the pro-apoptotic proteins [100] and of favoring the translation of mRNAs with weak translational efficiency as the mRNAs bearing TOP sequences [101–103].

The process of peptide chain elongation involves the selection of an aminoacyl-tRNA according to the codon sequence of the mRNA and the peptide bond formation between the growing peptide and the incoming amino acid [104]. The elongation process is mediated by elongation factors. The first step depends on eEF1A, a G-protein, and eEF1B, a guanine nucleotide exchange complex, and the translocation step relies on a second G-protein, eEF2.

In sea urchin embryos, the fertilization-induced increase in protein synthesis was reported to depend on the simultaneous increase of translational initiation and elongation rates of maternal mRNAs a long time ago [105, 106]. It was further shown that, in addition, the increase in the elongation rate displayed a significant slowing down at the time of CDK1 activation in M-phase [107]. Altogether, the sea urchin embryo model sounded suitable to study the regulation of protein synthesis at the level of the elongation step.

The sea urchin genome revealed the existence of one single gene for eEF1A and the different major subunits of eEF1B. The canonical nucleotide exchange protein eEF1Bα, its associated structural protein, eEF1Bγ, present in all eukaryotes from fungi to human, and the second nucleotide exchange leucine zipper containing protein, eEF1Bδ, specific for the metazoan complexes, are present in the sea urchin genome with more than 75 % identity with their human homologs [108]. ValyltRNA synthetase (VRS), the unique tRNA synthetase recovered in association with eEF1B complex in most vertebrates, is present under two isoforms in the sea urchin genome [17]. Remarkably, further analysis showed that none of them contained the eEF1Bδ-binding domain, responsible for anchoring VRS to the eEF1B complex in vertebrates. Since eEF1Bδ is already present in the sea urchin complex, these data indicate that the association of eEF1Bδ and VRS in eEF1B complex arises sequentially during metazoan evolution. In vertebrates, the function of this specific association with a unique tRNA synthetase in vertebrates is not yet elucidated; it was reported to potentially regulate the valine-rich protein translation in reticulocyte lysate experiments [109] and may also be related to the potential regulatory role of the tRNAs in the elongation process reviewed in [99]. Sea urchin eEF1B thus represents a powerful opportunity to elucidate this function by comparison with VRS-containing complexes.

The functional reason for an increasing complexity of eEF1B composition through evolution is still a matter of research (reviewed in [98, 110]). Obviously, the presence of two proteins, eEF1Bα and eEF1Bδ, both sharing the nucleotide exchange function in the same complex, remains puzzling. Biochemical analyses in sea urchin embryos further complexify the situation. Indeed, sea urchin eEF1B complex was shown to contain two eEF1Bδ isoforms [111]. Both eEF1Bδ isoforms differ only by a 26AA insert present in the N-terminus of one of the proteins,

resulting from the alternative splicing of the same eEF1Bδ gene. Interestingly, *in silico* genomic analysis revealed that an insert, sharing a comparable sequence and the same localization (ahead of the leucine zipper motif), was found in all available metazoan transcripts from cnidarian to human [98]. In sea urchin embryos, it was further demonstrated that these two isoforms gave rise to two populations of eEF1B complex, one containing only the insert-free eEF1Bδ isoform, the other containing the insert-free and the insert-containing isoform with a 1:1 ratio [108]. This structural feature, whose conservation through evolution supports a functional relevance, must now be taken into account in the search for the specific role of eEF1Bδ in the eEF1B complex.

The various subunits of eEF1B are phosphorylated by a number of serine/threonine kinases (reviewed in [98]). The protein eEF1Bγ was indeed identified as the first physiological substrate for CDK1/cyclin B, the universal cell cycle regulator [112]. eEF1δ was also showed to be a CDK1/cyclin B substrate in *Xenopus* eggs [113] as well as in human cells [114]. Searches in the data bank showed that almost all deuterostome eEF1Bγ sequences and metazoan eEF1Bδ sequences display CDK phosphorylation sites. Strikingly, the consensus site was not found in sea urchin EF1Bγ but is present in the sequence of EF1Bα. Further investigation in the published sequences showed that in each case where the CDK1 consensus site is absent from the EF1Bγ sequence it is found in the EF1Bα sequence [98].

Mitotic phosphorylation of eEF1B therefore appears universally conserved on one or the other of its subunits. Mitosis is commonly thought to be associated with reduced protein translation. In sea urchin embryos, such inhibition was observed as M-phase associated pauses in the global fertilization-induced elongation increase [107]. In the human cell, CDK1-induced eEF1Bδ phosphorylation results in hindering the tRNA delivery to ribosomes [114]. In sea urchin, concomitantly to the elongation pauses, eEF1B complex showed intracellular localization changes during mitosis [111]. In human fetal brain cells, eEF1Bδ was found to anchor the eEF1B complex to the endoplasmic reticulum by binding to kinectin, potentially favoring the synthesis of membrane proteins. On the other hand, eEF1Bα appears involved in eEF1A activities on the actin/tubulin cytoskeleton (reviewed in [110]). The implication of eEF1B phosphorylation(s) in those localizations has not yet been reported. Nevertheless, we hypothesize that mitotic eEF1B phosphorylation would regulate elongation by transiently decreasing translation in the context of the huge global protein synthesis increase involved in early development whereas favoring localized or selective proteins required for cell division (Fig. 3).

The second step of peptide chain elongation has been the subject of more attention to translational regulation. This step is catalyzed by eEF2, a monomeric protein that is inactivated by phosphorylation. The phosphorylation/activation state of eEF2 relies on the activity of EF2 kinase, an unusual calcium-calmodulin kinase, whose only substrate is eEF2. eEF2 kinase activity is regulated by multiple phosphorylation sites, which either activated (PKA, AMPK) or inhibited (mTOR/S6 kinase, ERK/RSK) the kinase (reviewed in [115]). eEF2 K is therefore

the integrating target of several signaling pathways that lead to protein synthesis regulation through the phosphorylation of its unique substrate eEF2.

Immunological studies in sea urchin eggs demonstrated for the first time that eEF2 is present in the cell under two isoforms differing by 2-kDa MW [116]. These two isoforms most probably issued from the alternative translation at an internal initiation codon existing in the transcript of the unique eEF2 gene present in the sea urchin genome [17]. The longer isoform is ten times more abundant than the shorter one. Both isoforms are phosphorylated in unfertilized eggs on the threonine residue (threonine 57), a substrate for eEF2 kinase (Fig. 3). This immunological determination has recently been supported by a phosphoproteomic analysis in sea urchin eggs [81].

Both isoforms are phosphorylated in unfertilized eggs in accordance with the low rate of protein synthesis in these cells. Fertilization induces the rapid dephosphorylation of both isoforms correlated with the global increase in protein synthesis. At first sight, this result seems surprising since fertilization induces calcium signaling pathway activation, which would rather result in activating eEF2 kinase. The most probable explanation is that mTOR pathway activation induced by fertilization (see above) triggers a dominant negative inhibition of eEF2K through direct or indirect phosphorylation at multiple sites (review in [115]). Whatever the mechanism, these data demonstrate that, in addition to the activation of the initiation step, a stimulation of elongation by eEF2 dephosphorylation contributes to the activation of global protein synthesis induced at fertilization in sea urchin embryos.

Interestingly, the longer eEF2 isoform, the more abundant one, remains dephosphorylated during the first cell cycles, probably ensuring the sustained increase in global protein synthesis, which goes on throughout early development. On the other hand, the shorter isoform, which accounts for the less abundant eEF2 fraction, is progressively re-phosphorylated on the same eEF2 kinase-specific site (threonine 57), with a cell-cycle dependent behavior. It was further shown that this re-phosphorylation is required for proper cell division. A recent report indicated that human eEF2 contains a CDK phosphorylation site (serine 595, conserved in the sea urchin sequence) and that phosphorylation at this site is required for efficient phosphorylation of the eEF2 kinase-inhibiting site [117]. We suggest that in sea urchin a specific and/or localized pool of eEF2, a target for CDK, would be re-phosphorylated at the M-phase stage and involved in the already-mentioned elongation pauses [107], possibly on selective mitotic protein(s) translation, as proposed in [116].

Altogether, analysis of sea urchin elongation factors revealed increasing composition complexity through eukaryotic organism evolution as well as significant conservation of a number of domains and phosphorylation sites, highlighting a functional relevance. This thus opens important new roads in the study of translational control at the level of elongation. Due to its occurrence more than 30 times over the initiation process for a mean size protein, it was long considered that elongation regulation is particularly advantageous for a cell. It is rapid, easily reversible and energy saving. Deciphering the mechanism of elongation regulation is therefore an important challenge. Sea urchin analyses appeared a promising way in this domain, particularly at the level of a localized or selective translation control.

6 Conclusions and Perspectives

The last few years have witnessed a surge of new information on the regulation of translational actors involved in the physiological inhibition of protein synthesis in echinoderm oocytes and unfertilized eggs. It is now clear that there are multiple ways in which overall rates of mRNA translation can be controlled in response to fertilization of sea urchin eggs (summarized in Fig. 3). At least three translation factors are activated in vivo in response to egg fertilization, leading to an increase of protein synthesis. Combining biochemical and cellular analyses of the translational actors activated by fertilization will continue to provide valuable information about the physiological importance of the translational machinery involved in the translation of the proteins required for the mitotic division and the early development of the embryos. Studies in echinoderms allowed for the comparison of translation regulation occurring at meiotic maturation and fertilization and have shed light on pathways controlling cyclin B translation that differ according to the stages of the cell cycle. Further work is needed to tackle the temporal and molecular switch between polyadenylation regulation and translational derepression.

mTOR-mediated regulation of 4E-BP degradation represents a new means to regulate eIF4E availability, and the molecular mechanism of 4E-BP degradation in sea urchin egg remains to be elucidated. mTOR signaling plays a major role in the fine-tuning orchestration of *cyclin B* mRNA translation following fertilization in sea urchin. Translatome analysis, by carrying out polysome profiling coupled with high-throughput sequencing technologies, should allow the identification of the subset of mRNAs present in unfertilized eggs that are actively translated and that are under the control of the mTOR signaling pathway activation triggered by fertilization. The identification of the sets of protein that are translated following fertilization should help establish translational regulatory networks that control the early events of the embryonic development.

Combining mathematical and biological approaches, and taking advantage of the quasi-spherical shape and constant volume of the sea urchin egg, the analysis of the spatio-temporal dynamics of the mTOR pathway and its impact on protein synthesis should provide important information on the dynamics of the system.

Acknowledgments We apologize to those whose work was not cited or discussed here because of the broad scope of this review and space limitation. Research in the authors' laboratories is supported by "La Ligue contre le Cancer (coordination du Grand Ouest (comités Finistère, Côtes d'Armor, Deux-Sèvres, Morbihan))," the "Région Bretagne" and the "Conseil Général du Finistère." B. Cosson is supported by Université Sorbonne Paris Cité (USPC) by the Research Project 2014 grant and 2015 International Fellowship. H. Chassé is a Ph.D. fellow supported by the Région Bretagne.

References

1. Ettensohn CA, Wessel GM, Wray GA. The invertebrate deuterostomes: an introduction to their phylogeny, reproduction, development, and genomics. Methods Cell Biol. 2004; 74:1–13.
2. Evans T, Rosenthal ET, Youngblom J, Distel D, Hunt T. Cyclin: a protein specified by maternal mRNA in sea urchin eggs that is destroyed at each cleavage division. Cell. 1983;33:389–96.
3. Kishimoto T, Kuriyama R, Kondo H, Kanatani H. Generality of the action of various maturation-promoting factors. Exp Cell Res. 1982;137:121–6.
4. Swenson KI, Farrell KM, Ruderman JV. The clam embryo protein cyclin A induces entry into M phase and the resumption of meiosis in Xenopus oocytes. Cell. 1986;47:861–70.
5. Labbe JC, Capony JP, Caput D, Cavadore JC, Derancourt J, Kaghad M, Lelias JM, Picard A, Doree M. MPF from starfish oocytes at first meiotic metaphase is a heterodimer containing one molecule of cdc2 and one molecule of cyclin B. EMBO J. 1989;8:3053–8.
6. Hunt T. Nobel lecture. Protein synthesis, proteolysis, and cell cycle transitions. Biosci Rep. 2002;22:465–86.
7. Kishimoto T. A primer on meiotic resumption in starfish oocytes: the proposed signaling pathway triggered by maturation-inducing hormone. Mol Reprod Dev. 2011;78:704–7. doi:10.1002/mrd.21343.
8. Ochi H, Aoto S, Tachibana K, Hara M, Chiba K. Block of CDK1-dependent polyadenosine elongation of Cyclin B mRNA in metaphase-i-arrested starfish oocytes is released by intracellular pH elevation upon spawning. Mol Reprod Dev. 2016;83:79–87. doi:10.1002/mrd.22599.
9. Wagenaar EB. The timing of synthesis of proteins required for mitosis in the cell cycle of the sea urchin embryo. Exp Cell Res. 1983;144:393–403.
10. Dube F. Effect of reduced protein synthesis on the cell cycle in sea urchin embryos. J Cell Physiol. 1988;137:545–52. doi:10.1002/jcp.1041370321.
11. Gross PR, Malkin LI, Moyer WA. Templates for the first proteins of embryonic development. Proc Natl Acad Sci USA. 1964;51:407–14.
12. Epel D. Protein synthesis in sea urchin eggs: a "late" response to fertilization. Proc Natl Acad Sci USA. 1967;57:899–906.
13. Brandhorst BP. Two-dimensional gel patterns of protein synthesis before and after fertilization of sea urchin eggs. Dev Biol. 1976;52:310–7.
14. Mathews MB, Sonenberg N, Hershey JW. Origins and principles of translational control. In: Sonenberg N, Hershey JW, Mathews MB, editors. Translational control of gene expression. Cold Spring Harbor: Cold Spring Harbor Laboratory Press; 2000. p. 1–31.
15. Davidson EH, Hough-Evans BR, Britten RJ. Molecular biology of the sea urchin embryo. Science. 1982;217:17–26.
16. Sea Urchin Genome Sequencing C, Sodergren E, Weinstock GM, Davidson EH, Cameron RA, Gibbs RA, Angerer RC, Angerer LM, Arnone MI, Burgess DR, Burke RD, Coffman JA, Dean M, Elphick MR, Ettensohn CA, Foltz KR, Hamdoun A, Hynes RO, Klein WH, Marzluff W, McClay DR, Morris RL, Mushegian A, Rast JP, Smith LC, Thorndyke MC, Vacquier VD, Wessel GM, Wray G, Zhang L, Elsik CG, Ermolaeva O, Hlavina W, Hofmann G, Kitts P, Landrum MJ, Mackey AJ, Maglott D, Panopoulou G, Poustka AJ, Pruitt K, Sapojnikov V, Song X, Souvorov A, Solovyev V, Wei Z, Whittaker CA, Worley K, Durbin KJ, Shen Y, Fedrigo O, Garfield D, Haygood R, Primus A, Satija R, Severson T, Gonzalez-Garay ML, Jackson AR, Milosavljevic A, Tong M, Killian CE, Livingston BT, Wilt FH, Adams N, Belle R, Carbonneau S, Cheung R, Cormier P, Cosson B, Croce J, Fernandez-Guerra A, Geneviere AM, Goel M, Kelkar H, Morales J, Mulner-Lorillon O, Robertson AJ, Goldstone JV, Cole B, Epel D, Gold B, Hahn ME, Howard-Ashby M, Scally M, Stegeman JJ, Allgood EL, Cool J, Judkins KM, McCafferty SS, Musante AM, Obar RA, Rawson AP, Rossetti BJ, Gibbons IR, Hoffman MP,

Leone A, Istrail S, Materna SC, Samanta MP, Stolc V, Tongprasit W, Tu Q, Bergeron KF, Brandhorst BP, Whittle J, Berney K, Bottjer DJ, Calestani C, Peterson K, Chow E, Yuan QA, Elhaik E, Graur D, Reese JT, Bosdet I, Heesun S, Marra MA, Schein J, Anderson MK, Brockton V, Buckley KM, Cohen AH, Fugmann SD, Hibino T, Loza-Coll M, Majeske AJ, Messier C, Nair SV, Pancer Z, Terwilliger DP, Agca C, Arboleda E, Chen N, Churcher AM, Hallbook F, Humphrey GW, Idris MM, Kiyama T, Liang S, Mellott D, Mu X, Murray G, Olinski RP, Raible F, Rowe M, Taylor JS, Tessmar-Raible K, Wang D, Wilson KH, Yaguchi S, Gaasterland T, Galindo BE, Gunaratne HJ, Juliano C, Kinukawa M, Moy GW, Neill AT, Nomura M, Raisch M, Reade A, Roux MM, Song JL, Su YH, Townley IK, Voronina E, Wong JL, Amore G, Branno M, Brown ER, Cavalieri V, Duboc V, Duloquin L, Flytzanis C, Gache C, Lapraz F, Lepage T, Locascio A, Martinez P, Matassi G, Matranga V, Range R, Rizzo F, Rottinger E, Beane W, Bradham C, Byrum C, Glenn T, Hussain S, Manning G, Miranda E, Thomason R, Walton K, Wikramanayke A, Wu SY, Xu R, Brown CT, Chen L, Gray RF, Lee PY, Nam J, Oliveri P, Smith J, Muzny D, Bell S, Chacko J, Cree A, Curry S, Davis C, Dinh H, Dugan-Rocha S, Fowler J, Gill R, Hamilton C, Hernandez J, Hines S, Hume J, Jackson L, Jolivet A, Kovar C, Lee S, Lewis L, Miner G, Morgan M, Nazareth LV, Okwuonu G, Parker D, Pu LL, Thorn R, Wright R. The genome of the sea urchin Strongylocentrotus purpuratus. Science. 2006; 314:941–952. doi:10.1126/science.1133609.

17. Morales J, Mulner-Lorillon O, Cosson B, Morin E, Belle R, Bradham CA, Beane WS, Cormier P. Translational control genes in the sea urchin genome. Dev Biol. 2006;300:293–307. doi:10.1016/j.ydbio.2006.07.036.
18. Humphreys T. Efficiency of translation of messenger-RNA before and after fertilization in sea urchins. Dev Biol. 1969;20:435–58.
19. Spirin AS, Nemer M. Messenger RNA in early sea-urchin embryos: cytoplasmic particles. Science. 1965;150:214–7.
20. Kaumeyer JF, Jenkins NA, Raff RA. Messenger ribonucleoprotein particles in unfertilized sea urchin eggs. Dev Biol. 1978;63:266–78.
21. Jenkins NA, Kaumeyer JF, Young EM, Raff RA. A test for masked message: the template activity of messenger ribonucleoprotein particles isolated from sea urchine eggs. Dev Biol. 1978;63:279–98.
22. Clemens M. Translational control. Developments in development. Nature. 1987;330:699–700. doi:10.1038/330699a0.
23. Huang WI, Hansen LJ, Merrick WC, Jagus R. Inhibitor of eukaryotic initiation factor 4F activity in unfertilized sea urchin eggs. Proc Natl Acad Sci U S A. 1987;84:6359–63.
24. Colin AM, Brown BD, Dholakia JN, Woodley CL, Wahba AJ, Hille MB. Evidence for simultaneous derepression of messenger RNA and the guanine nucleotide exchange factor in fertilized sea urchin eggs. Dev Biol. 1987;123:354–63.
25. Lopo AC, Lashbrook CC, Hershey JW. Characterization of translation systems in vitro from three developmental stages of Strongylocentrotus purpuratus. Biochem J. 1989;258:553–61.
26. Jagus R, Huang WI, Hansen LJ, Wilson MA. Changes in rates of protein synthesis and eukaryotic initiation factor-4 inhibitory activity in cell-free translation systems of sea urchin eggs and early cleavage stage embryos. J Biol Chem. 1992;267:15530–6.
27. Xu Z, Hille MB. Cell-free translation systems prepared from starfish oocytes faithfully reflect in vivo activity; mRNA and initiation factors stimulate supernatants from immature oocytes. Cell Regul. 1990;1:1057–67.
28. Merrick WC. eIF4F: a retrospective. J Biol Chem. 2015;290:24091–9. doi:10.1074/jbc.R115.675280.
29. Gingras AC, Raught B, Sonenberg N. eIF4 initiation factors: effectors of mRNA recruitment to ribosomes and regulators of translation. Annu Rev Biochem. 1999;68:913–63. doi:10.1146/annurev.biochem.68.1.913.
30. Proud CG. Mnks, eIF4E phosphorylation and cancer. Biochim Biophys Acta. 2015;1849:766–73. doi:10.1016/j.bbagrm.2014.10.003.

31. Pyronnet S. Phosphorylation of the cap-binding protein eIF4E by the MAPK-activated protein kinase Mnk1. Biochem Pharmacol. 2000;60:1237–43.
32. Joshi B, Cameron A, Jagus R. Characterization of mammalian eIF4E-family members. Eur J Biochem. 2004;271:2189–203. doi:10.1111/j.1432-1033.2004.04149.x.
33. Cameron RA, Samanta M, Yuan A, He D, Davidson E. SpBase: the sea urchin genome database and web site. Nucleic Acids Res. 2009; 37:D750–754. doi:gkn887 [pii]/nar/gkn887.
34. Lee SJ, Stapleton G, Greene JH, Hille MB. Protein kinase C-related kinase 2 phosphorylates the protein synthesis initiation factor eIF4E in starfish oocytes. Dev Biol. 2000;228:166–80. doi:10.1006/dbio.2000.9943.
35. Xu Z, Dholakia JN, Hille MB. Maturation hormone induced an increase in the translational activity of starfish oocytes coincident with the phosphorylation of the mRNA cap binding protein, eIF-4E, and the activation of several kinases. Dev Genet. 1993;14:424–39. doi:10.1002/dvg.1020140604.
36. Schmidt A, Durgan J, Magalhaes A, Hall A. Rho GTPases regulate PRK2/PKN2 to control entry into mitosis and exit from cytokinesis. EMBO J. 2007;26:1624–36. doi:10.1038/sj.emboj.7601637.
37. Cao Q, Richter JD. Dissolution of the maskin-eIF4E complex by cytoplasmic polyadenylation and poly(A)-binding protein controls cyclin B1 mRNA translation and oocyte maturation. The EMBO journal. 2002;21:3852–62. doi:10.1093/emboj/cdf353.
38. Lapasset L, Pradet-Balade B, Lozano JC, Peaucellier G, Picard A. Nuclear envelope breakdown may deliver an inhibitor of protein phosphatase 1 which triggers cyclin B translation in starfish oocytes. Dev Biol. 2005;285:200–10. doi:10.1016/j.ydbio.2005.06.016.
39. Lapasset L, Pradet-Balade B, Verge V, Lozano JC, Oulhen N, Cormier P, Peaucellier G. Cyclin B synthesis and rapamycin-sensitive regulation of protein synthesis during starfish oocyte meiotic divisions. Mol Reprod Dev. 2008;75:1617–26. doi:10.1002/mrd.20905.
40. Rosenthal ET, Brandhorst BP, Ruderman JV. Translationally mediated changes in patterns of protein synthesis during maturation of starfish oocytes. Dev Biol. 1982;91:215–20.
41. Standart N, Minshull J, Pines J, Hunt T. Cyclin synthesis, modification and destruction during meiotic maturation of the starfish oocyte. Dev Biol. 1987;124:248–58.
42. Bah A, Vernon RM, Siddiqui Z, Krzeminski M, Muhandiram R, Zhao C, Sonenberg N, Kay LE, Forman-Kay JD. Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. Nature. 2014;. doi:10.1038/nature13999.
43. Lin TA, Kong X, Haystead TA, Pause A, Belsham G, Sonenberg N, Lawrence JC Jr. PHAS-I as a link between mitogen-activated protein kinase and translation initiation. Science. 1994;266:653–6.
44. Fonseca BD, Smith EM, Yelle N, Alain T, Bushell M, Pause A. The ever-evolving role of mTOR in translation. Semin Cell Dev Biol. 2014;36:102–12. doi:10.1016/j.semcdb.2014.09.014.
45. Laplante M, Sabatini DM. mTOR signaling in growth control and disease. Cell. 2012;149:274–93. doi:10.1016/j.cell.2012.03.017.
46. Cormier P, Pyronnet S, Morales J, Mulner-Lorillon O, Sonenberg N, Belle R. eIF4E association with 4E-BP decreases rapidly following fertilization in sea urchin. Dev Biol. 2001;232:275–83. doi:10.1006/dbio.2001.0206.
47. Jagus R, Huang W, Hiremath LS, Stern BD, Rhoads RE. Mechanism of action of developmentally regulated sea urchin inhibitor of eIF-4. Dev Genet. 1993;14:412–23. doi:10.1002/dvg.1020140603.
48. Salaun P, Pyronnet S, Morales J, Mulner-Lorillon O, Belle R, Sonenberg N, Cormier P. eIF4E/4E-BP dissociation and 4E-BP degradation in the first mitotic division of the sea urchin embryo. Dev Biol. 2003;255:428–39.
49. Hernandez G, Altmann M, Lasko P. Origins and evolution of the mechanisms regulating translation initiation in eukaryotes. Trends Biochem Sci. 2010;35:63–73. doi:10.1016/j.tibs.2009.10.009.
50. Marcotrigiano J, Gingras AC, Sonenberg N, Burley SK. Cap-dependent translation initiation in eukaryotes is regulated by a molecular mimic of eIF4G. Mol Cell. 1999;3:707–16.

51. Mader S, Lee H, Pause A, Sonenberg N. The translation initiation factor eIF-4E binds to a common motif shared by the translation factor eIF-4 gamma and the translational repressors 4E-binding proteins. Mol Cell Biol. 1995;15:4990–7.
52. Gingras AC, Raught B, Gygi SP, Niedzwiecka A, Miron M, Burley SK, Polakiewicz RD, Wyslouch-Cieszynska A, Aebersold R, Sonenberg N. Hierarchical phosphorylation of the translation inhibitor 4E-BP1. Genes Dev. 2001;15:2852–64.
53. Oulhen N, Boulben S, Bidinosti M, Morales J, Cormier P, Cosson B. A variant mimicking hyperphosphorylated 4E-BP inhibits protein synthesis in a sea urchin cell-free, cap-dependent translation system. PLoS ONE. 2009;4:e5070. doi:10.1371/journal.pone.0005070.
54. Gosselin P, Oulhen N, Jam M, Ronzca J, Cormier P, Czjzek M, Cosson B. The translational repressor 4E-BP called to order by eIF4E: new structural insights by SAXS. Nucleic Acids Res. 2011;39:3496–503. doi:10.1093/nar/gkq1306.
55. Lukhele S, Bah A, Lin H, Sonenberg N, Forman-Kay JD. Interaction of the eukaryotic initiation factor 4E with 4E-BP2 at a dynamic bipartite interface. Structure. 2013;21:2186–96. doi:10.1016/j.str.2013.08.030.
56. Peter D, Igreja C, Weber R, Wohlbold L, Weiler C, Ebertsch L, Weichenrieder O, Izaurralde E. Molecular architecture of 4E-BP translational inhibitors bound to eIF4E. Mol Cell. 2015;57:1074–87. doi:10.1016/j.molcel.2015.01.017.
57. Oulhen N, Salaun P, Cosson B, Cormier P, Morales J. After fertilization of sea urchin eggs, eIF4G is post-translationally modified and associated with the cap-binding protein eIF4E. J Cell Sci. 2007;120:425–34. doi:10.1242/jcs.03339.
58. Lin TA, Kong X, Saltiel AR, Blackshear PJ, Lawrence JC Jr. Control of PHAS-I by insulin in 3T3-L1 adipocytes. Synthesis, degradation, and phosphorylation by a rapamycin-sensitive and mitogen-activated protein kinase-independent pathway. J Biol Chem. 1995;270:18531–8.
59. Elia A, Constantinou C, Clemens MJ. Effects of protein phosphorylation on ubiquitination and stability of the translational inhibitor protein 4E-BP1. Oncogene. 2008;27:811–22. doi:10.1038/sj.onc.1210678.
60. Yanagiya A, Suyama E, Adachi H, Svitkin YV, Aza-Blanc P, Imataka H, Mikami S, Martineau Y, Ronai ZA, Sonenberg N. Translational homeostasis via the mRNA cap-binding protein, eIF4E. Mol Cell. 2012;46:847–58. doi:10.1016/j.molcel.2012.04.004.
61. Salaun P, Boulben S, Mulner-Lorillon O, Belle R, Sonenberg N, Morales J, Cormier P. Embryonic-stage-dependent changes in the level of eIF4E-binding proteins during early development of sea urchin embryos. J Cell Sci. 2005;118:1385–94. doi:10.1242/jcs.01716.
62. Le Bouffant R, Cormier P, Mulner-Lorillon O, Belle R. Hypoxia and DNA-damaging agent bleomycin both increase the cellular level of the protein 4E-BP. J Cell Biochem. 2006;99:126–32. doi:10.1002/jcb.20856.
63. Le Bouffant R, Mulner-Lorillon O, Morales J, Cormier P, Belle R. Chromium(III) triggers the DNA-damaged checkpoint of the cell cycle and induces a functional increase of 4E-BP. Chem Res Toxicol. 2008;21:542–9. doi:10.1021/tx700244h.
64. Oulhen N, Mulner-Lorillon O, Cormier P. eIF4E-binding proteins are differentially modified after ammonia versus intracellular calcium activation of sea urchin unfertilized eggs. Mol Reprod Dev. 2010;77:83–91. doi:10.1002/mrd.21110.
65. Laurent S, Richard A, Mulner-Lorillon O, Morales J, Flament D, Glippa V, Bourdon J, Gosselin P, Siegel A, Cormier P, Belle R. Modelization of the regulation of protein synthesis following fertilization in sea urchin shows requirement of two processes: a destabilization of eIF4E:4E-BP complex and a great stimulation of the 4E-BP-degradation mechanism, both rapamycin-sensitive. Front Genet. 2014;5:117. doi:10.3389/fgene.2014.00117.
66. Buszczak M, Signer RAJ, Morrison SJ. Cellular differences in protein synthesis regulate tissue homeostasis. Cell. 2014;2014(159):242–51. doi:10.1016/J.Cell.09.016.
67. Eltschinger S, Loewith R. TOR complexes and the maintenance of cellular homeostasis. Trends Cell Biol. 2016;26:148–59. doi:10.1016/j.tcb.2015.10.003.

68. Ma XM, Blenis J. Molecular mechanisms of mTOR-mediated translational control. Nat Rev Mol Cell Biol. 2009;10:307–18. doi:10.1038/nrm2672.
69. Guertin DA, Sabatini DM. The pharmacology of mTOR inhibition. Science Signaling. 2009; 2:pe24. doi:10.1126/scisignal.267pe24.
70. Feldman ME, Apsel B, Uotila A, Loewith R, Knight ZA, Ruggero D, Shokat KM. Active-site inhibitors of mTOR target rapamycin-resistant outputs of mTORC1 and mTORC2. PLoS Biol. 2009;7:e38. doi:10.1371/journal.pbio.1000038.
71. Chassé H, Mulner-Lorillon O, Boulben S, Glippa V, Morales J, Cormier P. Cyclin B translation depends on mTOR activity after fertilization in sea urchin embryos. Plos One. 2016; doi:10.1371/journal.pone.0150318. doi:10.1371/journal.pone.0150318.
72. Kronja I, Orr-Weaver TL. Translational regulation of the cell cycle: when, where, how and why? Philos Trans R Soc Lond B Biol Sci. 2011;366:3638–52. doi:10.1098/rstb.2011.0084.
73. Wilt FH. The dynamics of maternal poly(A)-containing mRNA in fertilized sea urchin eggs. Cell. 1977;11:673–81.
74. Spieth J, Whiteley AH. Effect of 3′-deoxyadenosine (cordycepin) on the early development of the sand dollar. Dendraster excentricus. Dev Biol. 1980;79:95–106.
75. Hara M, Mori M, Wada T, Tachibana K, Kishimoto T. Start of the embryonic cell cycle is dually locked in unfertilized starfish eggs. Development. 2009;136:1687–96. doi:10.1242/dev.035261.
76. Jackson RJ, Hellen CU, Pestova TV. The mechanism of eukaryotic translation initiation and principles of its regulation. Nat Rev Mol Cell Biol. 2010;11:113–27. doi:10.1038/nrm2838.
77. Dever TE, Dar AC, Sicheri F. The eIF2α Kinases. In: Mathews MB, Sonenberg N, Hershey JWB (eds) Translational control in biology and medicine cold spring harbor. Cold Spring Harbor Laboratory Press; 2007. pp. 319–344.
78. Akkaraju GR, Hansen LJ, Jagus R. Increase in eukaryotic initiation factor 2B activity following fertilization reflects changes in redox potential. J Biol Chem. 1991;266:24451–9.
79. Costache V, Bilotto S, Laguerre L, Belle R, Cosson B, Cormier P, Morales J. Dephosphorylation of eIF2alpha is essential for protein synthesis increase and cell cycle progression after sea urchin fertilization. Dev Biol. 2012;365:303–9. doi:10.1016/j.ydbio.2012.03.002.
80. Tu Q, Cameron RA, Davidson EH. Quantitative developmental transcriptomes of the sea urchin Strongylocentrotus purpuratus. Dev Biol. 2014;385:160–7. doi:10.1016/j.ydbio.2013.11.019.
81. Guo H, Garcia-Vedrenne AE, Isserlin R, Lugowski A, Morada A, Sun A, Miao Y, Kuzmanov U, Wan C, Ma H, Foltz K, Emili A. Phosphoproteomic network analysis in the sea urchin Strongylocentrotus purpuratus reveals new candidates in egg activation. Proteomics. 2015;15:4080–95. doi:10.1002/pmic.201500159.
82. Romano PR, Garcia-Barrio MT, Zhang X, Wang Q, Taylor DR, Zhang F, Herring C, Mathews MB, Qin J, Hinnebusch AG. Autophosphorylation in the activation loop is required for full kinase activity in vivo of human and yeast eukaryotic initiation factor 2alpha kinases PKR and GCN2. Mol Cell Biol. 1998;18:2282–97.
83. Epel D. Activation of an Na + -dependent amino acid transport system upon fertilization of sea urchin eggs. Exp Cell Res. 1972;72:74–89.
84. Timourian H, Watchmaker G. Protein synthesis in sea urchin eggs. II. Changes in amino aicd uptake and incorporation at fertilization. Dev Biol. 1970;23:478–91.
85. Meyer E, Manahan DT. Nutrient uptake by marine invertebrates: cloning and functional analysis of amino acid transporter genes in developing sea urchins (Strongylocentrotus purpuratus). Biol Bull. 2009; 217:6–24. doi:217/1/6 [pii].
86. Iiboshi Y, Papst PJ, Kawasome H, Hosoi H, Abraham RT, Houghton PJ, Terada N. Amino acid-dependent control of p70(s6 k). Involvement of tRNA aminoacylation in the regulation. J Biol Chem. 1999;274:1092–9.
87. Deng J, Harding HP, Raught B, Gingras AC, Berlanga JJ, Scheuner D, Kaufman RJ, Ron D, Sonenberg N. Activation of GCN2 in UV-irradiated cells inhibits translation. Curr Biol. 2002; 12:1279–1286. doi:S0960982202010370 [pii].

88. Krohn M, Skjolberg HC, Soltani H, Grallert B, Boye E. The G1-S checkpoint in fission yeast is not a general DNA damage checkpoint. J Cell Sci. 2008; 121:4047–4054. doi:jcs.035428 [pii] 10.1242/jcs.035428.
89. Menacho-Marquez M, Perez-Valle J, Arino J, Gadea J, Murguia JR. Gcn2p regulates a G1/S cell cycle checkpoint in response to DNA damage. Cell Cycle. 2007; 6:2302–2305. doi:4668 [pii].
90. Tvegard T, Soltani H, Skjolberg HC, Krohn M, Nilssen EA, Kearsey SE, Grallert B, Boye E. A novel checkpoint mechanism regulating the G1/S transition. Genes Dev. 2007; 21: 649–654. doi:21/6/649 [pii] 10.1101/gad.421807.
91. Le Bouffant R, Boulben S, Cormier P, Mulner-Lorillon O, Belle R, Morales J. Inhibition of translation and modification of translation factors during apoptosis induced by the DNA-damaging agent MMS in sea urchin embryos. Exp Cell Res. 2008;314:961–8. doi:10.1016/j.yexcr.2007.12.014.
92. Aze A, Fayet C, Lapasset L, Geneviere AM. Replication origins are already licensed in G1 arrested unfertilized sea urchin eggs. Dev Biol. 2010;340:557–70. doi:10.1016/j.ydbio.2010.02.009.
93. Alves VS, Motta FL, Roffe M, Delamano A, Pesquero JB, Castilho BA. GCN2 activation and eIF2alpha phosphorylation in the maturation of mouse oocytes. Biochem Biophys Res Commun. 2009;378:41–4.
94. Lee S, Truesdell SS, Bukhari SI, Lee JH, LeTonqueze O, Vasudevan S. Upregulation of eIF5B controls cell-cycle arrest and specific developmental stages. Proc Natl Acad Sci U S A. 2014;111:E4315–22. doi:10.1073/pnas.1320477111.
95. Krauchunas AR, Horner VL, Wolfner MF. Protein phosphorylation changes reveal new candidates in the regulation of egg activation and early embryogenesis in D. melanogaster. Dev Biol. 2012;370:125–34. doi:10.1016/j.ydbio.2012.07.024.
96. Traugh JA. Insulin, phorbol ester and serum regulate the elongation phase of protein synthesis. Prog Mol Subcell Biol. 2001;26:33–48.
97. Browne GJ, Proud CG. Regulation of peptide-chain elongation in mammalian cells. Eur J Biochem. 2002;269:5360–8.
98. Le Sourd F, Boulben S, Le Bouffant R, Cormier P, Morales J, Belle R, Mulner-Lorillon O. eEF1B: At the dawn of the 21st century. Biochim Biophys Acta. 2006;1759:13–31. doi:10.1016/j.bbaexp.2006.02.003.
99. Richter JD, Coller J. Pausing on polyribosomes: Make way for elongation in translational control. Cell. 2015;163:292–300. doi:10.1016/j.cell.2015.09.041.
100. White-Gilbertson S, Rubinchik S, Voelkel-Johnson C. Transformation, translation and TRAIL: an unexpected intersection. Cytokine Growth Factor Rev. 2008;19:167–72. doi:10.1016/j.cytogfr.2008.01.007.
101. Gismondi A, Caldarola S, Lisi G, Juli G, Chellini L, Iadevaia V, Proud CG, Loreni F. Ribosomal stress activates eEF2 K-eEF2 pathway causing translation elongation inhibition and recruitment of terminal oligopyrimidine (TOP) mRNAs on polysomes. Nucleic Acids Res. 2014;42:12668–80. doi:10.1093/nar/gku996.
102. Ryazanov AG, Rudkin BB, Spirin AS. Regulation of protein synthesis at the elongation stage. New insights into the control of gene expression in eukaryotes. FEBS Lett. 1991;285:170–5.
103. Walden WE, Thach RE. Translational control of gene expression in a normal fibroblast. Characterization of a subclass of mRNAs with unusual kinetic properties. Biochemistry. 1986;25:2033–41.
104. Merrick WC, Nyborg J. The protein synthesis elongation cycle. In: Sonenberg N, Hershey J, Mathews M, editors. Translational control of gene expression. New York: CSHL Press; 2000. p. 89–125.
105. Brandis JW, Raff RA. Elevation of protein synthesis is a complex response to fertilisation. Nature. 1979;278:467–9.
106. Hille MB, Albers AA. Efficiency of protein synthesis after fertilisation of sea urchin eggs. Nature. 1979;278:469–71.

107. Monnier A, Morales J, Cormier P, Boulben S, Belle R, Mulner-Lorillon O. Protein translation during early cell divisions of sea urchin embryos regulated at the level of polypeptide chain elongation and highly sensitive to natural polyamines. Zygote. 2001;9:229–36.
108. Le Sourd F, Cormier P, Bach S, Boulben S, Belle R, Mulner-Lorillon O. Cellular coexistence of two high molecular subsets of eEF1B complex. FEBS Lett. 2006;580:2755–60. doi:10.1016/j.febslet.2006.04.038.
109. Monnier A, Belle R, Morales J, Cormier P, Boulben S, Mulner-Lorillon O. Evidence for regulation of protein synthesis at the elongation step by CDK1/cyclin B phosphorylation. Nucleic Acids Res. 2001;29:1453–7.
110. Sasikumar AN, Perez WB, Kinzy TG. The many roles of the eukaryotic elongation factor 1 complex. Wiley Interdiscip Rev RNA. 2012;3:543–55. doi:10.1002/wrna.1118.
111. Boulben S, Monnier A, Le Breton M, Morales J, Cormier P, Belle R, Mulner-Lorillon O. Sea urchin elongation factor 1delta (EF1delta) and evidence for cell cycle-directed localization changes of a sub-fraction of the protein at M phase. Cell Mol Life Sci. 2003;60:2178–88. doi:10.1007/s00018-003-3201-x.
112. Mulner-Lorillon O, Poulhe R, Cormier P, Labbe JC, Doree M, Belle R. Purification of a p47 phosphoprotein from Xenopus laevis oocytes and identification as an in vivo and in vitro p34cdc2 substrate. FEBS Lett. 1989;251:219–24.
113. Mulner-Lorillon O, Minella O, Cormier P, Capony JP, Cavadore JC, Morales J, Poulhe R, Belle R. Elongation factor EF-1 delta, a new target for maturation-promoting factor in Xenopus oocytes. J Biol Chem. 1994;269:20201–7.
114. Sivan G, Aviner R, Elroy-Stein O. Mitotic modulation of translation elongation factor 1 leads to hindered tRNA delivery to ribosomes. J Biol Chem. 2011;286:27927–35. doi:10.1074/jbc.M111.255810.
115. Kenney JW, Moore CE, Wang X, Proud CG. Eukaryotic elongation factor 2 kinase, an unusual enzyme with multiple roles. Adv Biol Regul. 2014;55:15–27. doi:10.1016/j.jbior.2014.04.003.
116. Belle R, Pluchon PF, Cormier P, Mulner-Lorillon O. Identification of a new isoform of eEF2 whose phosphorylation is required for completion of cell division in sea urchin embryos. Dev Biol. 2011;350:476–83. doi:10.1016/j.ydbio.2010.12.015.
117. Hizli AA, Chi Y, Swanger J, Carter JH, Liao Y, Welcker M, Ryazanov AG, Clurman BE. Phosphorylation of eukaryotic elongation factor 2 (eEF2) by cyclin A-cyclin-dependent kinase 2 regulates its inhibition by eEF2 kinase. Mol Cell Biol. 2013;33:596–604. doi:10.1128/MCB.01270-12.

Unique and Conserved Features of the Protein Synthesis Apparatus in Parasitic Trypanosomatid (*Trypanosoma* and *Leishmania*) Species

Osvaldo P. de Melo Neto, Christian R.S. Reis, Danielle M.N. Moura, Eden R. Freire and Mark Carrington

1 Trypanosomatids: Taxonomy, Life Cycle, Genome Organization and Post-transcriptional Control of Gene Expression

Trypanosomatids are protozoans characterized by a single flagellum and distinct cellular morphologies. The best-characterized trypanosomatids are *Leishmania* and *Trypanosoma* species, mainly those that are pathogenic in humans and other mammals. The pathogenic species tend to have complex life cycles with many developmental forms that occur successively during the physical transit between different tissues of the invertebrate vector and mammalian host. These are responsible for different diseases of worldwide impact, targeting millions of people mainly in low-income countries and are classified as neglected diseases by the World Health Organization. Chagas disease and human African trypanosomiasis are caused respectively by two *Trypanosoma* species, *T. cruzi* and *T. brucei*, while different *Leishmania* species are responsible for the various forms of Leishmaniasis.

Trypanosomatids are classified within the order Kinetoplastida (phylum Euglenozoa, supergroup Excavata), which includes both free-living and pathogenic organisms. The Kinetoplastida are characterized by a specialized single mitochondrion containing an internal array of concatenated mitochondrial DNA molecules forming a structure, visible with the light microscope, called the kinetoplast [1]. The Kinetoplastida probably represent one of the earliest diverged groups of

O.P. de Melo Neto (✉) · C.R.S. Reis · D.M.N. Moura · E.R. Freire
Centro de Pesquisas Aggeu Magalhães, Fundação Oswaldo Cruz, Av. Moraes Rego S/N, Campus UFPE, Recife, PE 50670-420, Brazil
e-mail: opmn@cpqam.fiocruz.br

M. Carrington
Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB21QW, UK

G. Hernández and R. Jagus (eds.), *Evolution of the Protein Synthesis Machinery and Its Regulation*, DOI 10.1007/978-3-319-39468-8_17

eukaryotes and consequently have had ~1 billion years of separate evolution. Within the Kinetoplastida, the family Trypanosomatidae (the trypanosomatids) contains species that are all symbiotic/parasitic and probably co-evolved with invertebrates. The majority of species retain a single host, but some have evolved to infect a second host. The range of the secondary host species is amazing; some infect mammals, others infect vascular plants and others everything in between. The evolution of the ability to infect a secondary host must have occurred many times [2, 3].

The environment encountered by the trypanosomatid cell can change dramatically, especially when transferring between hosts, and its adaptive response includes changes in the cell morphology, cell surface and catabolic metabolism; all of course downstream of regulated changes in gene expression. For example, *Leishmania* sp. are intracellular within their mammalian hosts, where they multiply as amastigotes, rounded cells with no protruding flagellum, within the phagolysosomes of macrophages, at conditions of acidic pH at 37 °C [4, 5]. When transferred to a hematophagous sandfly vector, they first differentiate into extracellular, elongated, flagellated forms (promastigotes and others), which proliferate at neutral pH and ambient temperature in the insect gut. These then migrate to the salivary glands where they differentiate once again to metacyclic forms, endowed with a very long flagellum and pre-adapted for survival within the mammalian host [6]. A new cycle is started when the infected insect once again feeds upon a mammal.

The ability to survive through their complex life cycles requires that the trypanosomatids precisely regulate their gene expression in order to adapt to the different extracellular environments. This regulation has to be superimposed on a gene expression mechanism that evolved while the world was a simpler place. A decade ago, the availability of the first three trypanosomatid genomes, the TriTryps [7–10], followed by subsequent genomes from major lineages [11], led to a major boost in the study of genome structure and gene expression [12–14]. The most unexpected aspects of the genome structures were the arrangement of protein coding genes into long tandem arrays and the conservation of synteny within the arrays despite divergence in chromosome number and size. This conservation of gene order is linked to the mechanism of RNA polymerase II transcription and mRNA maturation; tens of protein-coding genes are encoded in tandem arrays and co-transcribed from occasional transcription start sites. No defined promoters for RNA polymerase II-dependent protein coding genes have been characterized so far, but changes in the pattern of histone modifications have been seen to be associated with these regions and might be required for transcription to start/stop [15–17].

Transcription in trypanosomatids is thought to be constitutive as there is little or no evidence for selective use of RNA pol II [18], although the overall rate of transcription initiation is regulated. Whether RNA pol II transcribes different tandem gene arrays at the same rate remains to be tested. There is co-transcriptional processing of the pre-mRNAs to monocistronic mRNAs, through *trans*-splicing of a capped ~39 nucleotide mini-exon donated from the Spliced Leader (SL) RNA to the 5′ end of each mRNA and a linked cleavage and polyadenylation of the upstream mRNA. Both processes are mediated by the same set of sequence motifs

found within the precursor mRNA [19–21]. Prior to *trans*-splicing, the SL RNA is modified with a 'cap 4' structure at its 5′ end consisting of the typical inverted 7-methyl-GTP commonly followed by methylations at various positions on the first four nucleotides of the SL sequence (AACU) [22, 23]. All mature trypanosomatid mRNAs then are characterized by the SL sequence plus cap 4 at their 5′ ends with the typical eukaryotic poly-A tails at their 3′ ends.

The constitutive nature of transcription in trypanosomatids, lack of defined promoters and the long polycistronic transcription units indicate an absence of most known mechanisms for selective transcriptional control of gene expression. It follows that most processes associated with the regulation of gene expression are mediated at the post-transcriptional level. Many of these may act at the level of mRNA processing, transport, storage, translation and half-life, while others likely act at the post-translational level, through the regulation of protein modification and degradation, in most cases by yet undefined mechanisms [24–28]. Translation is a key step in quantitative regulation of gene expression in trypanosomatids, and a major role is expected to be played during its initiation stage. Several unique features of the trypanosomatid translation initiation apparatus have been identified and will be discussed in detail below. This review also provides an update on what is known regarding other major elements of the protein synthesis machinery in trypanosomatids, including other translation factors, ribosomal subunits and RNA-binding proteins, always comparing data from the best studied species belonging to the genera *Leishmania* and *Trypanosoma*. Here, when discussing individual trypanosomatid polypeptides, all will be named in capital letters following the proposed nomenclature for trypanosomatid proteins [29].

1.1 Eukaryotic Initiation Factors (eIFs) and Translation Initiation

The picture that emerges from searches for genes encoding translation initiation factors (eIFs) within the available trypanosomatid genomes is the conservation in regard to the presence of a complete set of initiation factors. Apart from the automatic annotation made with the original sequencing of the first trypanosomatid genomes, however, little has progressed regarding the characterization of many of their eIFs and the evaluation as to the degree they are conserved in function when compared to other eukaryotes. Nevertheless, substantial progress has been made concerning the characterization of individual polypeptides/complexes. Here, in order to systematically evaluate what is known regarding their translation initiation factors, this review will first discuss those factors that consist of single polypeptides (eIF1, eIF1A, eIF4B, eIF5, eIF5B and eIF6), followed by less well-studied complexes of multiple subunits (eIF2, eIF2B, eIF3) and a more detailed description of the multiple eIF4F complexes, and their subunits, which have been characterized recently in these organisms.

1.1.1 Single Polypeptide Translation Initiation Factors

Most single polypeptide eIFs—eIF1, eIF1A, eIF5, eIF5B and eIF6—have readily identifiable trypanosomatid orthologs, annotated as such within the various genome sequences available at the TriTrypDB database. The single exception is eIF4B, which is more divergent in sequence, as exemplified by a comparison of human, yeast and plant eIF4Bs [30]. Trypanosomatids lack clearly identifiable eIF4B orthologs, and it will not be discussed further. None of the above remaining factors have been specifically targeted for more detailed studies in trypanosomatids, although an eIF6 sequence was early on described from *Trypanosoma cruzi* after being found encoded within a genomic DNA fragment containing a short interspersed repetitive element (SIRE) [31]. Several of these factors, however, were found to co-purify with other translation initiation complexes studied in these organisms, therefore providing clues as to their functional properties. *Leishmania* EIF1, for instance, strongly associated with the eIF3 complex in a co-precipitation assay using polyclonal antibodies directed against its EIF3E subunit. In this experiment, no other eIF, or even the EIF3J subunit of eIF3, co-precipitated with the 11-subunit eIF3 complex, highlighting the strict conditions in which the purification was carried out [32]. The strong interaction between the *Leishmania* EIF1 and the eIF3 complex was independently observed using a slightly different approach where a streptavidin-binding-peptide (SBP)-tagged EIF3E was also used to precipitate eIF3 using streptavidin-Sepharose beads. In this assay the EIF1A and EIF5 orthologs also co-precipitated with the eIF3 complex, as well as the orthologs for the three eIF2 subunits. This pattern of co-precipitation, also seen using a tagged *Leishmania* EIF3A subunit, was proposed to indicate the presence in the trypanosomatids of the multi-factor complex (MFC), formed by EIF1, EIF1A, EIF5 and both eIF3 and eIF2 complexes, independently of the ribosome [33]. *Leishmania* EIF5 also co-precipitated with two eIF4E homologs (EIF4E1 and EIF4E4), when these two proteins were tagged with the SBP, in an assay that also brought down several eIF3 and all eIF2 subunits, but no EIF1 or EIF1A [34]. EIF6 has also been seen to co-precipitate with eIF3 [33] and EIF5B with EIF4E1/EIF4E4 [34] in the assays described above, but no further functional data have been described associated with them. In an extensive analysis of the parasite's phosphoproteins [35], the *T. brucei* EIF6 ortholog was found to be phosphorylated at a single serine residue near to its C-terminus; however, the serine is not conserved in *L. major* EIF6, and the phosphorylation is distinct from the previously reported eIF6 phosphorylation from other organisms [36]. Overall the analysis of these different eIFs in trypanosomatids highlights their overall conservation within the eukaryotes but emphasizes the need for a better investigation in order to define conserved and divergent aspects of their functions.

1.1.2 eIF2, eIF2B and eIF2α Kinases

Orthologs of all three eIF2 subunits are found in the trypanosomatid genome sequences, but so far only the EIF2α subunit has been investigated. The two other eIF2 subunits (EIF2β and EIF2γ) have yet to be characterized, although both, and EIF2α, co-precipitated with tagged eIF3 and eIF4F subunits [33, 34], as described elsewhere in this review. Orthologs to all five eIF2B subunits are also present in the genome sequences although they have not been characterized either. The overall conservation is much higher for the three regulatory subunits (EIF2Bα, EIF2Bβ and EIF2Bδ) when compared with the catalytic ones (EIF2Bγ and EIF2Bε) [de Melo Neto et al., unpublished observation], which is consistent with only the regulatory subunits having defined archaean homologs [37, 38]. The pattern seen then from both eIF2 and eIF2B subunits is consistent with a conserved role for these factors throughout the eukaryotes, although unique features have been found for those polypeptides studied in more detail. One noteworthy feature has to do with the regulation of eIF2 and eIF2B function by novel phosphorylation events. Not only EIF2α, but also EIF2γ and two regulatory subunits of eIF2B (EIF2Bβ and EIF2Bδ) have been identified as phosphoproteins [35], and, in an independent analysis, EIF2Bβ has also been found to be targeted by tyrosine phosphorylation [39].

A trypanosomatid EIF2α was first identified within the first published genome sequences from *T. brucei* in a study that had as its main focus the characterization of its EIF2α kinases [40]. The *T. brucei* EIF2α gene was found to codify for an extended protein containing many features conserved in eukaryotic eIF2α sequences but including an N-terminal extension only present in other trypanosomatids. This extension is roughly 110 amino acids long and shows little variation between the *Trypanosoma* and *Leishmania* sequences. When ectopically expressed in *Saccharomyces cerevisiae* cells, the full-length *T. brucei* sequence could not complement a lack of the endogenous yeast protein; however, a truncated *T. brucei* EIF2α, missing its N-terminal extension, was functional. Despite an overall conservation in the predicted loop that includes the serine 51 (S51) residue targeted for phosphorylation by eIF2α kinases in other eukaryotes, this residue is replaced by a threonine in different trypanosomatid sequences (T169 in *T. brucei* EIF2α). This T169 residue, nevertheless, was efficiently phosphorylated by a *T. brucei* EIF2α kinase although it did not seem to be recognized by other eukaryotic eIF2α kinases [40].

Three potential eIF2α kinases (TbEIF2K1 to TbEIF2K3) have been identified in trypanosomatids, originally from *T. brucei*. TbEIF2K2 and its orthologs have been the focus of several functional studies in both *Leishmania* and *Trypanosoma* species and will be discussed in more detail. In *T. brucei*, it is a transmembrane glycoprotein located in or near the flagellar pocket that has been shown to specifically phosphorylate yeast eIF2α at S51 and also phosphorylates the *T. brucei* EIF2α at T169. This kinase is also targeted by phosphorylation, with its phosphorylation pattern changing during the parasite's life cycle, and it has been suggested that it could be involved in sensing protein or nutrient transport [40]. Its ortholog in *T. cruzi* has been recently characterized (TcK2) and seen to localize to

the endosomal compartment where it binds heme and is inhibited by it [41]. In *Leishmania*, the TbEIF2K2 ortholog (~44 % identical) has been identified as a PERK kinase homolog, a glycosylated transmembrane protein that localizes to the endoplasmic reticulum, phosphorylates *Leishmania* EIF2α at T166 (equivalent to the *T. brucei*'s T169) and also undergoes autophosphorylation. It has been shown that this PERK kinase homolog is differentially targeted by post-translational modifications during *Leishmania*'s life cycle and that these modifications are also seen after treatment with an agent that induces stress to the endoplasmic reticulum [42]. These modifications are presumed to be phosphorylation, which would likely be associated with the kinase activation, EIF2α phosphorylation and translation inhibition. In *T. cruzi*, a nutritional stress has also been shown to induce EIF2α phosphorylation [43], highlighting the central role that the EIF2α kinases and EIF2α phosphorylation may have in the regulation of translation in response to different cellular stresses. However, this may not always be so, since in *T. brucei* the phosphorylation of the EIF2α T169 residue was not found to be associated with a decrease in polysomes seen in cells stressed through heat shock [44]. A second EIF2α kinase, TbEIF2K3 or PK3, has only recently been studied in more detail and also seen to partially localize to the endoplasmic reticulum, but, upon stress induction and phosphorylation, translocates to the nucleus where it seems to phosphorylate factors associated with transcription events [45].

Several different studies have indicated a requirement for a tight control of protein synthesis associated with EIF2α phosphorylation and linked to the differentiation processes seen during different trypanosomatid life cycles. In *Leishmania* an increase in EIF2α phosphorylation has been seen to correlate with a decrease in overall protein synthesis, which occurs during the differentiation to amastigote forms (which live in their mammalian hosts) [46, 47]. Furthermore, the overexpression of a dominant version of the PERK kinase, which inhibits EIF2α phosphorylation, delays differentiation [42]. In *T. cruzi*, blocking the attenuation of protein synthesis by the overexpression of a mutant form of EIF2α, which cannot be phosphorylated, abolished the differentiation of the parasite's epimastigote form (which grows in the insect vector) into infective metacyclic forms [43]. More recently, the study of the *T. cruzi* TcK2 kinase has indicated a critical role for this protein during the differentiation process, linking it to the metabolism of heme. In its absence, activation of TcK2 leads to EIF2α phosphorylation, translation arrest and differentiation [41].

1.1.3 eIF3

The eIF3 complex in trypanosomatids has only recently been the focus of more detailed studies, although seven eIF3 subunits (A, C, E, G, I, K and L) were initially detected in *Leishmania* in complexes that were pulled down with an eIF4E homolog, EIF4E3, but not with its partner, EIF4G4 [48]. Recently, 12 *Leishmania* eIF3 subunits (A through L, with the exception of the M subunit) were identified using an in-depth bioinformatics study with subsequent biochemical validation.

This study used the *Leishmania* EIF3E as a target to immunoprecipitate the whole complex followed by mass spectrometry analysis. The biochemical data indicate a lack of association of the EIF3J subunit to the complex, and an important finding was the strong interaction between EIF1 and the whole eIF3 complex [32]. An independent study combining bioinformatics with affinity purification and mass spectrometry identified the same subunits and additionally demonstrated that the EIF3J subunit may be part of the complex in *Leishmania* [33].

Comparative functional studies of the eIF3 complex in trypanosomatids have just been started, and, as expected, all identified eIF3 subunits were found in polysome fractions in *T. brucei* [49]. In *L. amazonensis*, pull-down of a tagged eIF4G homolog, EIF4G3, expressed in transgenic cells, resulted in efficient capture of eIF3 subunits, implicating this eIF4G homolog as a possible link to recruit eIF3 and the 43S pre-initiation complex to the mRNAs in these protozoans, in a manner reminiscent to what is seen in others eukaryotes. The same work identified a direct interaction between eIF3 and an eIF4E homolog, EIF4E1, which in turn does not to bind any eIF4G or eIF4G-like protein. Through yeast two-hybrid experiments, this interaction was mapped between the C-terminus of the EIF3A subunit and EIF4E1 [33]. The biological function of this interaction between EIF3A/EIF4E1 is not yet clear, but it could serve as a novel mode of eIF3 recruitment to the mRNA, presumably directed to selected populations of mRNAs, and which would bypass the need for any eIF4G function.

Phosphorylation sites for multiple eIF3 subunits have also been identified in *T. brucei*, with single sites reported for EIF3B, EIF3D, EIF3E, EIF3I and EIF3K, while four sites were found targeting EIF3C [35]. When compared to known phosphorylation events targeting the mammalian eIF3 subunits [50], most were not conserved, with the exception of those targeting EIF3C. Three of the four *T. brucei* EIF3C sites mapped to its N-terminal end, remarkably conserved between homologs from different eukaryotes and required for its interaction with eIF5 [32]. The equivalent segment from both human and plant eIF3c is also targeted by multiple phosphorylation events [50, 51], and in plants these have seen to be mediated by the CK2 kinase, which also targets several other initiation factors that are part of the MFC complex. It has been proposed that the CK2 kinase, through the phosphorylation of the plant eIF3c subunit and other translation factors, may have a role in regulating translation initiation in plants [51], and the conservation in target sites among trypanosomatids, plants and mammals might indicate a mechanism conserved through most eukaryotic lineages.

1.1.4 eIF4F

In trypanosomatids the eIF4F subunits have been the focus of several recent papers that have begun to unravel their role during translation initiation and have highlighted a novel pattern for eIF4F-like complexes. Multiple homologs for the eIF4A, eIF4E and eIF4G subunits have then been described, conserved in different *Leishmania* and *Trypanosoma* species, which associate into multiple eIF4F-like

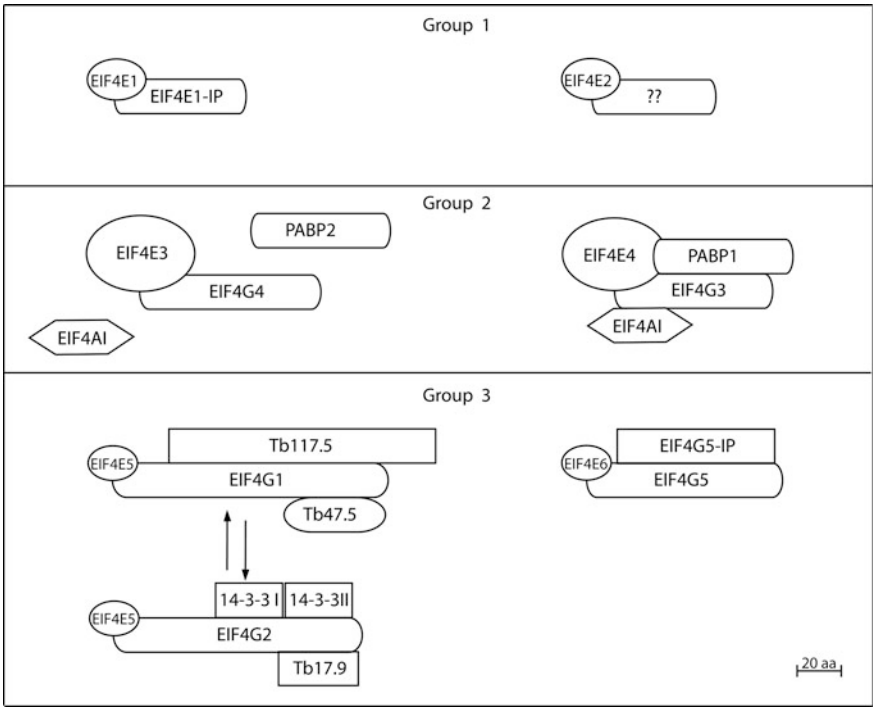


Fig. 1 Comparative scheme of the trypanosomatid eIF4E homologs and their respective eIF4F-like complexes, mostly as defined from *T. brucei*. Group 1, formed by eIF4E homologs that do not interact with eIF4G partners: EIF4E1/EIF4E1-IP [34] and EIF4E2, with so far unknown partner(s). Group 2, formed by eIF4E homologs that form complexes involved in translation: EIF4E4/EIF4G3/EIF4AI/PABP1 and EIF4E3/EIF4G4 (which may or not also interact with EIF4AI and PABP2) [34, 48, 63]. Group 3, formed by eIF4E homologs that interact with eIF4G partners and also with putative cap-generating proteins: EIF4E5/EIF4G1/Tb117.5/Tb47.5 and/or EIF4E5/EIF4G2/Tb17.9/14-3-3 I/14-3-3 II [56]; EIF4E6/ EIF4G5/ EIF4G5-IP [75]

complexes with yet unresolved roles in translation [23, 52–56]. Here we review what is known about these factors, the complexes in which they participate, protein partners and possible biological roles. Figure 1 summarizes then the data available regarding these subunits that are discussed in more detail below.

eIF4A

The first translation initiation factor identified in trypanosomatids was an eIF4A homolog, identified after the screening of a genomic expression library from *Leishmania braziliensis* for immunologically relevant proteins [57]. Subsequently, the *L. major* ortholog was cloned and characterized as part of an effort to investigate its role in inducing protective immunity against leishmaniasis [58]. With the focus on studying translation initiation and the eIF4F complex in *Leishmania* and other

trypanosomatids, sequences encoding two eIF4A homologs were first identified through Blast searches using the human eIF4A as query against the *L. major* genomic sequences. These were first named EIF4A1 and EIF4A2, and conserved orthologs were found from other trypanosomatid sequences. When used to evaluate their evolutionary relationship with other eIF4A sequences in a Phylogenetic tree based on their sequence alignment, the trypanosomatid eIF4As were found to cluster together in branches distinct from the human eIF4A homologs. Nevertheless, EIF4A1 was >100-fold more abundant than EIF4A2 and was the only one able to bind to a *Leishmania* eIF4G homolog (EIF4G3), although it did not bind to human eIF4G [52]. Further functional assays were carried out in yeast where the *Leishmania* EIF4A1 gene was not able to complement the deletion of the endogenous eIF4A genes, and the overexpressed protein, despite being able to interact with yeast eIF4G, inhibited cell growth [59]. Using *T. brucei* as a model, the two eIF4A sequences were further characterized, with its EIF4A1 ortholog seen to localize to the cytoplasm while EIF4A2 localized strictly within the nucleus. Metabolically labeling with ^{35}S-methionine after RNA interference induction and growth curve analysis of cells expressing dominant negative versions of the two proteins were then carried out to investigate their function more directly. The data derived from these experiments conclusively defined EIF4A1 as the eIF4A homolog involved in translation initiation. Further sequence analysis, with the focus on identifying individual amino acid residues that could discriminate between eIF4AI and eIF4AIII orthologs, clearly grouped the second trypanosomatid eIF4A homolog with EIF4AIII orthologs from major eukaryotic lineages and pinpointed several residues that could have functional relevance for both proteins [60]. The identification of an eIF4AIII ortholog in trypanosomatids was an unexpected event considering its known role in mRNA splicing as part of the exon-junction complex [61]. Nevertheless, subsequently, further core components of this complex, such as the conserved Mago protein, were found in trypanosomatids, although no direct interaction was seen between a dimer of the Mago and Y14-like subunits and the EIF4AIII ortholog [62].

eIF4E

The initial studies carried out with the purpose of characterizing the trypanosomatid eIF4E homologs were based on homology searches using mammalian or yeast eIF4E sequences as queries against *Leishmania* genomic sequences. These resulted in the identification of four homologs that were subsequently targeted for functional studies [52, 53]. More recently, two further eIF4E homologs, more divergent in sequence, were identified in *T. brucei*, with orthologs in other trypanosomatid species [56]. To date then, six trypanosomatid eIF4E homologs (EIF4E1 to E6) have been identified, conserved in both *Leishmania* and *Trypanosoma* species and that can be paired in three distinct groups, based on structural and molecular properties and on associated protein partners [52, 56, 63]: Group 1 consists of EIF4E1 and EIF4E2; group 2 is formed by EIF4E3 and EIF4E4; group 3, more

recently identified, is formed by EIF4E5 and EIF4E6. None of these trypanosomatid eIF4E groups can be confidently assigned to any of the better known classes of eIF4E homologs found in multicellular organisms [64], suggesting independent duplication events.

Group 1 This group is formed by proteins with sizes similar to human and yeast eIF4Es, which interact with the cap structure but not with eIF4G homologs. EIF4E1, the first eIF4E homolog identified in *L. major* (LeishIF4E-1) [65], is a 24-kDa protein (26 kDa for *T. brucei*) with 23 % identity to the human eIF4E-1 homolog and containing seven of the eight conserved tryptophan residues typical of eIF4E family members. These include the three residues involved in cap binding (W56, W102 and W166 positions in human eIF4E-1), with the tryptophan involved in eIF4G interaction (equivalent to W73 in human eIF4E) being replaced by a phenylalanine. EIF4E2 was described from *Leishmania* as a 31.5-kDa protein (28.3 in *T. brucei*) possessing all conserved tryptophan residues typical of the eIF4E protein family [52] and sharing 29 % identity with the human eIF4E-1. EIF4E1 and EIF4E2 were identified as cytoplasmic proteins in *Leishmania* [53], but in *T. brucei*, when overexpressed as fluorescent fusions, both behaved as nucleo-cytoplasmic polypeptides [63]. Neither of these two eIF4E homologs are abundant polypeptides, and they are present in levels far below those estimated for their group 2 counterparts (EIF4E3 and EIF4E4; group 3 eIF4Es have not been quantified so far), in both *Leishmania* and *T. brucei* [52, 63].

EIF4E1 from both *Leishmania* and *T. brucei* is able to bind to m^7GTP-Sepharose, and the same applies for the EIF4E2 from *T. brucei*. The EIF4E2 from *Leishmania* differs from its *T. brucei* ortholog in that it does not bind to m^7GTP-Sepharose beads [52, 53, 63, 65, 66]. The difference between the EIF4E2 orthologs may be caused by an insertion between W113–W130 in the *Leishmania* protein that could hypothetically change the protein folding, making it unable to bind m^7GTP-Sepharose [55]. Additional analysis using cap analogs in vitro have shown that the *Leishmania* EIF4E2 binds preferentially to the methylated cap 4 structure, while EIF4E1 binds well to both m^7GTP cap and the cap 4 [53].

Polysome distribution analysis reported for *Leishmania* EIF4E2 placed it in both monosome and polysome fractions and suggested an association with translating mRNAs. In contrast, the data for EIF4E1 were contradictory since they were first reported as present exclusively on the top fraction of a sucrose gradient, which is not associated with ribosomes, but subsequently was also seen to be associated with polysomes [53, 67]. Most of the early work with the *Leishmania* eIF4Es was carried out with the insect promastigote form, since it is easily cultured. The cultured vertebrate stage, the intracellular amastigote, is not easily available and doesn't necessarily reflect the pattern of gene expression present in the true intracellular parasite [68]. Nevertheless, experiments using axenic amastigote cells from *L. amazonensis* found that EIF4E1 was the single eIF4E with increased expression in amastigotes, and it was proposed that EIF4E1 would be a functional eIF4E in amastigotes [34]. This possibility was reinforced by the recent identification of a direct interaction between EIF4E1 and the EIF3A subunit of eIF3, an interaction that would eliminate the need for an eIF4G intermediate [33]. In addition, a

Leishmania EIF4E1-binding protein was identified, later called Leish4E-IP (here 4E1-IP), that has no homology to eIF4G or eIF4E-BP but contains the short conserved eIF4E-binding motif. This protein has been seen to preferentially bind to EIF4E1 only during its promastigote stage, leading to a model where the release of EIF4E1 by 4E1-IP during amastigote differentiation would lead to changes in mRNA translation mediated by the freed EIF4E1 [34]. The evidence from *T. brucei*, however, does not support this model. In *T. brucei*, neither EIF4E1 nor EIF4E2 is essential for cell survival in the insect procyclic forms although EIF4E1 is essential for the human bloodstream forms. Depletion of both together in procyclics led to a rapid cell death without any apparent effect on general translation. The fact that neither is able to bind to any of the described eIF4G homologs also makes them unlikely candidates for *bona fide* eIF4Es acting during translation initiation [63]. More recently, no association between EIF4E1 or EIF4E2 and polysomes could be detected in *T. brucei* in a proteomic analysis of the polysome constituents from both life stages of its life cycle [49]. In fact, in a search for translation regulators from *T. brucei*, using a tethering assay where selected factors were bound to the 3′ UTR of reporter mRNAs, both EIF4E1 and the 4E-IP ortholog were seen to cause translational repression of the mRNA to which they were bound [69, 70].

Group 2 This group consists of EIF4E3 (29 % identity between *L. major* EIF4E3 and human eIF4E-1) and EIF4E4 (31 % identity with the human eIF4E). Both proteins are characterized by N-terminal extensions not seen in other eIF4Es from eukaryotes outside the kinetoplastids. In *L. major*, EIF4E3 is a 38-kDa protein (48 kDa in *T. brucei*) while EIF4E4 is slightly larger in size, 48 kDa (46 kDa in *T. brucei*). The group 2 eIF4Es are characterized by substitutions in key residues involved in cap binding: the human W56 is replaced by a phenylalanine or tyrosine in both EIF4E3 and EIF4E4 and the almost universally conserved WED motif, which includes the W102 residue, is replaced by WEH in EIF4E3 (also seen in the trypanosomatids' EIF4E2 homologs). Several other tryptophan residues generally conserved in eIF4E homologs (three for EIF4E3 and two for EIF4E4) are also modified, although mostly replaced by aromatic residues. Both proteins are cytoplasmic and the most abundant trypanosomatid eIF4E homologs. EIF4E4 binds efficiently to m^7GTP-Sepharose beads and to the soluble m^7GTP cap or cap 4, while EIF4E3, despite being able to bind to m^7GTP in solution, binds poorly to the m^7GTP beads or to soluble cap 4 [52, 53, 55, 63].

EIF4E3 and EIF4E4 are constitutively expressed in the insect and vertebrate developmental forms of *T. brucei* [63], but they have been found to be downregulated in the amastigote, vertebrate stage of *L. amazonensis* [34]. Both proteins undergo post-translational modifications, specifically multiple phosphorylation events, during cell growth of *T. brucei* and *L. amazonensis*. EIF4E4 phosphorylation in both organisms was found to be associated with exponentially growing cells while phosphorylation of EIF4E3, in *L. amazonensis* only was associated with stationary phase cells [71]. For *L. infantum* EIF4E4, which had its phosphorylation characterized in more detail, it was found to be constitutively expressed in the two stages of the parasite cell cycle, and phosphorylation was seen to be typical of

exponential growth in both life stages. These phosphorylation events did not require binding to eIF4G homologs, ruling out a Mnk-like pattern of phosphorylation, and the target sites were mapped to several serine/proline or threonine/proline motifs localized within its N-terminal extension, in a pattern reminiscent of phosphorylation by CDK or MAP kinases [72]. In *T. brucei*, a high throughput analysis of its phosphoproteins identified both EIF4E3 and EIF4E4 as targets for multiple phosphorylation events directed at serine/threonine residues within their N-terminal halves [35]. Both proteins displayed a phosphorylation pattern reminiscent of *Leishmania* EIF4E4 and in marked contrast to what is observed for *Leishmania* EIF4E3.

Sucrose density gradient analysis in *Leishmania* showed that while both group 2 proteins co-sedimented with monosomes and polysomes, most of the EIF4E4 was present at the top of the gradients, away from ribosomes [53]. Later experiments found that EIF4E3 is present only in nuclease-resistant 80S particles and enters stress granules during starvation [48]. A recent analysis in *T. brucei*, however, has shown that both EIF4E3 and EIF4E4 are present in polysomes and most of the polysomal mRNAs are bound to EIF4E4, strongly implying a role for EIF4E4 in translation [49]. RNAi experiments showed that EIF4E3 is required for proliferation of both mammalian bloodstream and insect procyclic forms of *T. brucei*. This contrasts with EIF4E4, which is required in only the bloodstream form. Nevertheless, a double knockdown of EIF4E1 and EIF4E4 does lead to a major inhibition of translation, for reasons not yet understood [63]. Failed attempts to create double EIF4E4 knockout and complementation studies in *L. infantum* indicate that this protein may be essential for cell survival in insect stage promastigotes [72].

EIF4E3 and EIF4E4 have the ability to interact with eIF4G homologs forming potential eIF4F complexes: EIF4E3 binds to EIF4G4 and EIF4E4 to EIF4G3, and both complexes can interact with EIF4AI [34, 48, 63]. The conservation in sequence within the putative eIF4G-binding regions implicates the DVECFW motif in *T. brucei* EIF4E3 (DVESFW in *L. major*) as being required for the interaction with EIF4G4, while *T. brucei* EIF4E4 most likely binds to EIF4G3 through the similar ISSFW motif (ILTFW in *L. major*) [23, 55, 73]. EIF4E4 can also interact directly with PABP1 through its N-terminal extension, a unique eIF4E/PABP interaction that so far has only been described in trypanosomatids [34]. Recently this interaction was further characterized, and it was found to require three small regions, or boxes, mapped to the N-terminal extension of EIF4E4. Those boxes are based on the consensus L/MN/DXXAXXY/FXP (where X can be any amino acid) and are found conserved not only in different EIF4E4 homologs but also in the EIF4E3 sequences (within their N-terminus) from several trypanosomatid species. Mutational analysis revealed that amino acid changes disrupting individual boxes were not enough to prevent the EIF4E4/PABP1 binding, but this interaction was abolished when all three boxes were mutated. The functional relevance of this interaction was evaluated by assaying whether ectopic overexpression of EIF4E4, wild type and mutants, could complement the absence of the endogenous protein. Both wild-type EIF4E4 and a mutant impaired on its ability to bind to EIF4G3

could compensate for the knockout of the endogenous alleles, although the cells grown with the mutant protein could not differentiate into amastigotes. A mutant EIF4E4 impaired on its binding to PABP1, however, could not compensate for the loss of the endogenous protein, highlighting the importance of the novel EIF4E4/PABP1 interaction and confirming that it is more critical for EIF4E4 function than its interaction with EIF4G3 [72].

Most of the evidence generated so far implicates EIF4E4 directly as the most likely eIF4E homolog to be required for translation in trypanosomatids. Nevertheless, the possibility that EIF4E3 is also involved in translation in some manner cannot be ruled out, especially considering the impact on translation of its knockdown in *T. brucei* [63]. This was reinforced by the data from tethering experiments where both EIF4E3 and EIF4E4 were able to stimulate the translation of a reporter mRNA when tethered to its 3′ UTR [69]. The possibility remains that the two different eIF4F-like complexes, based on EIF4E3 and EIF4E4, could be required for the translation of distinct mRNA subsets, with the EIF4E4-based complex being responsible for the translation of the bulk of the mRNAs. In *Leishmania*, the presence of EIF4E3 in stress granules led to the proposal that it would be involved in stress granule formation only [48], but mammalian eIF4E is also present in stress granules despite being active in translation [74]. Perhaps the dissociation of EIF4E3 from EIF4G4 under stress conditions [48] is an indication of specific translation repression, but additional data are required to clarify the real role of the EIF4E3/EIF4G4 complex and the degree to which it is conserved between different trypanosomatid species.

Group 3 This group is based on the two recently identified eIF4E homologs, EIF4E5 and EIF4E6, which are both very small, fewer than 200 amino acids, and more closely related to each other than to the other trypanosomatid eIF4Es. When compared with group 1 and group 2 eIF4Es, these are the most divergent, with pair-wise alignments between EIF4E5 or EIF4E6 with the human eIF4E-1 producing identity values equal to or below 25 %. Nevertheless, both contain several critical residues conserved in eIF4Es from different organisms and required for their function. So far, these two proteins have only been properly studied in *T. brucei* [56, 75].

T. brucei EIF4E5 is a cytoplasmic, 22-kDa protein, having the W56 and W166 residues required for cap binding and with W102 replaced by a tyrosine. It binds to soluble cap and cap 4 structures with affinities similar to those reported for *Leishmania* EIF4E4. So far, this is the only *T. brucei* eIF4E homolog with two putative eIF4G partners, EIF4G1 and EIF4G2, although it seems that, at least in the procyclic stage, EIF4G2 is the preferred binding partner. Mass spectrometry analysis confirmed that EIF4E5 forms two independent complexes. When bound to EIF4G1, it also associates with two hypothetical proteins of 117.5 kDa (named Tb117.5) and 47.5 kDa (named Tb47.5), plus one of the two *T. brucei* homologs of the conserved phosphoserine/phosphothreonine-binding protein 14-3-3 (14-3-3 II). Tb117.5 has some intriguing characteristics since it harbors two cap-generating domains, with possible guanylyltransferase and methyltransferase activities, while Tb47.5 has two RNA-binding domains. The complex based on EIF4E5/EIF4G1

only interacts with its Tb117.5 partner in the absence of 14-3-3 II, suggesting that the latter might act as a regulator of this interaction. A second EIF4E5-based complex, dependent on the interaction with EIFG2, is associated with yet another hypothetical protein of 17.9 kDa (Tb17.9) and both 14 3-3 homologs (14 3-3 I and 14-3-3 II). Knockdown of EIF4E5 indicated that the protein is not essential for cell survival for procyclic *T. brucei* in culture but caused a motility phenotype since the cells did not remain in suspension in liquid culture and growth patterns on agarose plates were altered [56].

EIF4E6 is also a small cytoplasmic protein, 21 kDa in size, having a conserved W166 but with both W56 and W102 being replaced by phenylalanine. It also can bind to cap analogs in vitro, but its affinity for soluble m^7GTP cap or cap 4 is lower than those observed for EIF4E4 and EIF4E5 and more similar to that observed for EIF4E1. This eIF4E homolog forms a complex with EIF4G5 and a 70.3-kDa hypothetical protein that interacts directly with EIF4G5, named TbG5-IP. Surprisingly, TbG5-IP also possesses two domains usually found in nuclear cap-generating proteins, a nucleoside triphosphate hydrolase and a guanylyltransferase domain. Interestingly, knocking down EIF4E6 leads to a flagellum detachment phenotype and alters growth on agarose plates, but it does not influence overall translation rates, interfere with cell growth or reduce the targeted cell's ability to remain in suspension [75].

eIF4G

Five distinct trypanosomatid eIF4G homologs were originally identified through searches using the mammalian HEAT1-MIF4G domain as queries against the *L. major* genome sequences (named EIF4G1 to EIF4G5). All five identified polypeptides shared the central conserved HEAT1-MIF4G domain and on reciprocal BLAST searches against mammalian, plant and yeast sequences yielded eIF4G homologs as the most similar proteins. Orthologs to all five proteins were easily identifiable in other trypanosomatid species, but apart from EIF4G3 and EIF4G4, which are clearly related, the homology between these proteins is restricted to the HEAT1-MIF4G domain [52]. In both *T. brucei* and *Leishmania*, EIF4G3 and EIF4G4 have subsequently been shown to specifically bind to two different eIF4E homologs, EIF4E3 and EIF4E4, forming distinct eIF4F-like complexes that have both been implicated in protein synthesis and have subsequently been studied in more detail [34, 48, 63, 67, 73]. As described above, EIF4G1, EIF4G2 and EIF4G5 have only recently been shown to form novel eIF4F complexes in *T. brucei* that have not been directly linked to the translation initiation process [56, 75]. All three proteins were nevertheless seen to stimulate translation of reporter mRNAs when tethered to their 3'UTR [69, 70], although none were found in *T. brucei* polysomal fractions [49]. In the phosphoproteomic analysis of *T. brucei* proteins, both EIF4G1 and EIF4G2 were seen to be phosphorylated at multiple serine (EIF4G1) or serine/threonine (EIF4G2) residues. EIF4G1 was also found to be phosphorylated at a unique tyrosine residue while a single serine

phosphorylation site was identified for EIF4G5 [35]. Further work will be required then in order to understand what, if any, roles these proteins and their partners play during initiation of protein synthesis in trypanosomatids and how they might be associated with translation control mechanisms. Figure 1 summarizes the data available for the trypanosomatid eIF4G homologs within the described eIF4F-like complexes.

Apart from the common HEAT1-MIF4G domain, EIF4G3 and EIF4G4 also share a very short, similar N-terminus and homologous C-termini, including divergent HEAT2-MA3 and HEAT3-W2 domains. The structure of both proteins thus indicates an ancient origin of the eIF4G tripartite structure, which is apparently missing from plant and yeast eIF4G homologs [73]. Both EIF4G3 and EIF4G4 are localized predominantly to the cytoplasm, and both are moderately expressed, with EIF4G3 being roughly three- to tenfold more abundant than EIF4G4. When compared with the previous estimates for the abundance of their corresponding eIF4E partners, EIF4G3 is found in levels similar to EIF4E4, while EIF4G4 is present at levels at least tenfold lower than EIF4E3, although the reasons for the difference in abundance of the latter protein pair is not understood [63, 73]. As seen for their eIF4E partners, EIF4G3 and EIF4G4 are also targeted for phosphorylation in both *L. amazonensis* and *T. brucei*, although no differences in phosphorylation were observed between logarithmic growth and stationary phase [71]. In the high-throughput search for *T. brucei* phosphoproteins, EIF4G4 was indeed found to be targeted by six different phosphorylation events directed at serine residues concentrated at its N- and C-terminal halves. In contrast, EIF4G3 was not identified as a phosphoprotein in this assay [35].

In sucrose gradients in *Leishmania*, EIF4G3 co-migrates with EIF4E4 and EIF4AI in polysome-containing fractions [67], but EIF4G4 does not [48]. *Leishmania* EIF4G3 also binds to the initiation complex eIF3 through a direct interaction observed in vitro between the fully assembled eIF3 complex and recombinant EIF4G3 [33]. EIF4G4, however, did not pull down any eIF3 subunits in co-precipitation assays from native extracts, contrasting with its partner EIF4E3, which efficiently pulled down seven eIF3 subunits [48]. In *T. brucei*, knockdown of EIF4G3 strongly reduced cell growth and global translation very soon after RNAi induction, but, in contrast, knockdown of EIF4G4 does not impact significantly on protein synthesis although it does result in cell death over a longer time scale. The conclusion from these experiments is that the two proteins are essential for viability, but only EIF4G3 is a major participant in the initiation of translation. Nevertheless, the observation that knockdown of EIF4G4 induces changes in morphology prior to cell death may indicate a selective role in the translation of specific mRNAs [73]. Indeed, also in *T. brucei*, the tethering assay confirmed that both EIF4G3 and EIF4G4 and their eIF4E partners are able to stimulate the translation of mRNAs to which they are tethered [69, 70], and the two protein pairs are found in polysomes [49].

The similarities in sequence and structure observed between EIF4G3 and EIF4G4 and also between their eIF4E-binding partners, EIF4E3 and EIF4E4, are indicative of gene duplication events preceding the split of the *Trypanosoma* and *Leishmania* lineages. A single original complex would then evolve into the two

distinct complexes seen today, with acquisition of new functions for at least one of these [73]. As discussed above, the evidence from both *Leishmania* and *T. brucei* definitely implicates the EIF4G3/EIF4E4 complex in translation initiation, and the data from *T. brucei* are also in agreement for a similar role for the second complex EIF4G4/EIF4E3 [33, 49, 67, 70, 73]. Nevertheless, the data from *Leishmania* are not consistent with a clear role for EIF4G4/EIF4E3 in translation, and the data from the two subunits are not entirely in agreement with both functioning as a pair, since EIF4E3 is found in polysomes and pulls down eIF3 subunits but not its partner EIF4G4 [48]. Likewise, in *T. brucei* RNAi-mediated depletion clearly implicates EIF4E3 in translation but not its partner EIF4G4 [63, 73]. A clear difference regarding the two eIF4F complexes has to do with the requirements for eIF4A binding. EIF4G3 interacts strongly with EIF4AI in vitro, requiring only its HEAT1-MIF4G domain, and overexpression of an EIF4G3 mutant that is impaired on its interaction with EIF4AI inhibits cell growth, an indication of a dominant negative phenotype. In contrast, the binding between EIF4G4 and EIF4AI seems much less efficient and requires the full length EIF4G4, and overexpression of an EIF4G4 mutant that does not bind EIF4AI does not impact on cell growth [73]. Both EIF4G3 and EIF4E4 have also been seen to interact directly with one of the *Leishmania* PABP homologs, PABP1, and these interactions seem to be specific and required for proper function of all three proteins during translation initiation [34, 72, 73, 76]. So far, an interaction between the EIF4G4/EIF4E3 complex and PABP homologs has not been definitively proven, and, in vitro at least, no clear interaction between EIF4G4 and any of the *Leishmania* PABP homologs has been detected [73]. *Leishmania* PABP2 specifically co-precipitated with EIF4G4 in assays using whole parasite extract, but it was not found in similar co-precipitation assays using EIF4E3 as bait [48], so the possibility exists of a yet undemonstrated specific interaction between PABP2 and the second eIF4F-like complex based on EIF4G4/EIF4E3.

One aspect of the study involving the two complexes based on EIF4G3/EIF4E4 and EIF4G4/EIF4E3 deals with the motifs in the two eIF4Gs responsible for these interactions and that reside in the short N-terminal segments from both proteins. It has been proposed that an eIF4E-binding motif is present in the N-terminus of EIF4G3 between positions 20 and 26 (YPGFSLD) in *Leishmania*. Mutations in the tyrosine, leucine and phenylalanine residues within this motif completely abrogate binding to EIF4E4 [67]. A subsequent work has suggested that there is not a consensus motif in *Leishmania* EIF4G4 for binding to EIF4E3, but instead this interaction is mainly based on the secondary structure of the EIF4G4 binding peptide, and preventing it by replacing the L26 residue with a proline would eliminate the interaction [48]. However, more recently, it has been proposed that the EIF4G3 and EIF4G4 proteins do share a common eIF4E-binding motif (F/MXXXXIL/R—*Leishmania* consensus) that should include residues in equivalent positions in both eIF4Gs (I8 and R9 in EIF4G3/I25 and L26 in EIF4G4). In agreement with this hypothesis, overexpression of *T. brucei* EIF4G3 and EIF4G4 variants having mutations targeting the equivalent residues effectively abolished the interaction with their eIF4E partners. Furthermore, overexpression of these EIF4G3 and EIF4G4 mutants induced minor but reproducible reductions in the cell

proliferation rate [73]. Surprisingly, as previously stated, loss of the direct EIF4E4/EIF4G3 interaction does not affect the survival of *Leishmania* promastigotes, possibly a consequence of the compensatory interactions between EIF4E4/PABP1 and PABP1/EIF4G3, which under some circumstances might bypass the need for the direct EIF4E4/EIF4G3 interaction [72].

1.2 Other Translation Factors

This topic covers the current knowledge of individual proteins participating in other steps of the translation process in trypanosomatids. Classical elongation, termination and release factors, which have been studied in trypanosomatids, are discussed as well as eIF5A, which has been studied with some detail. Poly-A binding proteins (PABPs) are also discussed, since they have multiple roles not only during translation initiation but also during other stages of mRNA translation, processing and degradation.

1.2.1 Elongation Factors

EF-1 and EF-2

Phylogenetic analysis of the distribution of EF-1A and its paralog EF-Like (EFL) in the Euglenozoa, and in other eukaryotes, showed that trypanosomatids have only a single gene encoding EF-1A [77, 78]. This factor is essential for *T. brucei* cell viability since its silencing through RNA interference led to an almost instant cessation of growth, eventually causing cell death [79]. EF-1A depletion also caused a lethal phenotype in a high-throughput experiment that evaluated loss of fitness following RNAi in procyclic and bloodstream forms and also during differentiation [80], consistent with an essential function during protein synthesis.

Comparison of structural models shows that EF-1A orthologs from *Leishmania*, yeast and mammals resemble each other closely. Despite its conservation in function and general structure, the EF-1A in trypanosomatids has several distinct differences when compared to its mammalian counterpart. A major difference is the deletion of 12 amino acids from the *Leishmania* protein that is associated with a hairpin loop found in mammalian EF-1A [81]. Absence of this hairpin loop has been found to expose part of the main body of EF-1A. The function of this particular region is not well defined, but when it was ‘blocked’ by a specific antibody, which did not bind mammalian EF-1A because of the presence of the hairpin loop, protein synthesis was inhibited in vitro in a *Leishmania* cell-free translation system. A similar inhibitory effect was observed by a peptide based on the exposed region of the *Leishmania* protein [82]. Indeed this region has been evaluated as a target for small molecules that bind specifically to the *Leishmania* EF-1A and that inhibit specifically the *Leishmania* translation [83].

Similarly to mammalian and plant EF-1As, which are modified by two ethanolamine phosphoglycerol (EPG) moieties attached to conserved glutamic acid residues, *T. brucei* EF-1A is also modified accordingly at a single EPG modification site, residue E362. Although a second potential EPG modification site, E289, equivalent to the one found in mammals and plants, is found in the *T. brucei* protein, it does not seem to be targeted by this modification. Despite the fact that the EPG modification is exclusive to EF-1A, and found in many species, it is not necessary for cell growth in *T. brucei*, and its function remains unknown [84, 85].

Despite their importance in translation, recent data from a number of laboratories have supported the idea of secondary functions and biological significance for the EF1 subunits [86]. Notably, many studies have linked the EF-1A to immunological responses during pathogenesis in *Leishmania* and it should also be noted that two previously unrecognized putative immunoreceptors have been revealed through its sequence analysis. In a similar context, EF-1B subunits in trypanosomatids were identified as part of a trypanothione S-transferase (TST) complex, with the active site localized to the γ subunit, with the authors suggesting that it may have a role in translational response to oxidative and xenobiotic stress [87, 88].

In *T. cruzi*, EF-2 is encoded by two identical gene copies, and the protein has 60 % identity to *S. cerevisiae* EF-2. It is highly conserved in all trypanosomatids, sharing over 96 % homology with EF-2 of *Leishmania* species and over 98 % with other trypanosomes. Comparative analysis of secondary structures shows an overall conserved architecture, displaying the four canonical alpha helices in the GTPase domain as well as a motif involved in nucleotide binding. However, all *Trypanosoma* EF-2 proteins seem to lack regulation by phosphorylation, since they do not have the targeted threonine residue that mediates its binding to the ribosome [89].

Both EF-1 subunits (1A and 1B) and EF-2 were detected in pull-down assays targeting the cap-binding protein EIF4E4, as well as the polypeptides EIF4G3 and eIF3, all currently suggested as functional translation initiation factors that act during the initial steps of protein synthesis in trypanosomatids [33, 34]. As expected, they were also found in a proteomic analysis of the polysome contents of *T. brucei* [49].

eIF5A

In trypanosomatids, eIF5A homologs have been described from *Trypanosoma cruzi, Leishmania donovani* and *T. brucei* [90–92]. As expected, the *Leishmania* EIF5A was cytoplasmic and displayed high sequence identity with *T. cruzi* (76 %) and *T. brucei* (77 %) orthologs and less identity to the human ortholog (45 % identity) [91]. *Trypanosoma* EIF5A was also found to localize to the cytoplasm and to be modified by hypusination as in other eukaryotes [90, 92]. Indeed, the two enzymes that catalyze hypusination, deoxyhypusine synthase and deoxyhypusine hydroxylase, have been found and characterized from *Leishmania* [93, 94]. Other post-translational modifications targeting eIF5A in trypanosomatids were found through mass spectrometry analysis. *T. cruzi* EIF5A was phosphorylated at a single

conserved serine residue, S2, but a novel tyrosine phosphorylation event, at Y21, has also been identified. Three methylations events were also found, mapped to residues E73, E99 and E116 [90]. Phosphorylation at S2 was also identified for *T. brucei* EIF5A [35, 95].

In exponentially growing *T. cruzi* cells, EIF5A was phosphorylated and partially found to be associated with polysomes. In stationary phase cells, EIF5A was dephosphorylated, and a greater fraction was present in polysomes. Overexpression of a phosphomimetic mutant S2D of EIF5A increased cell proliferation and protein synthesis but seems to be toxic to stationary phase cells. Overall EIF5A seems to cycle between phosphorylated and dephosphorylated forms, which might be required to regulate translation in response to growth conditions [90]. Complementary work has shown that EIF5A is essential for cell growth in *T. brucei*; RNAi-mediated knockdown causes growth arrest and also leads to morphological abnormalities including cell rounding and detached flagella. The RNAi-induced growth defect was complemented by expression of wild-type human eIF5A, showing that the trypanosome and human proteins are functionally homologous, but not by a K52 mutant (K50 in human eIF5A) that blocks modification by deoxyhypusine, confirming this modification to be an essential requirement for cell viability. The expression levels of two representative polyprolyl proteins involved in the actin cytoskeleton were also analyzed after *T. brucei* EIF5A knockdown, and both were shown to be reduced [92]. The selective reduction of the polyprolyl-containing protein suggests a preferential requirement for EIF5A for the translation of proteins containing consecutive proline tracts, as demonstrated recently in other eukaryotes [96].

1.2.2 Termination/Release Factors

Trypanosomatid orthologs of eRF1 and eRF3 are present and annotated as putative proteins within the various genome sequences available at the TriTrypDb database [10]. Due to the high degree of conservation of sequence in release factors, an analysis of the *T. brucei* ERF1 and ERF3 sequences has been included in phylogenetic works that helped improve the relationships and the understanding of the translation process between eukaryotic groups [97–99]. As for some other factors with major roles in translation, RNAi knockdown of ERF1 and ERF3 resulted in cessation of proliferation following RNAi in both procyclic and bloodstream forms and also in an experiment that evaluated loss of fitness post-RNAi during differentiation of *T. brucei* [80].

1.2.3 Poly-A-Binding Proteins (PABPs)

The first trypanosomatid PABP homolog, from *T. cruzi*, was identified through early studies in the 1990s, which started with the biochemical characterization of the native protein after its purification through affinity chromatography using

poly-A Sepharose. This was followed by the screening of DNA libraries with polyclonal serum directed against the purified protein, which led to the identification of a single PABP homolog encoded by two gene copies [100]. The *T. brucei* ortholog was subsequently identified [101, 102] followed by the identification of a more divergent *Leishmania* PABP homolog [103]. With the completion of the genome sequences, the differences in PABP genes were clarified, with three distinct PABP homologs identified first in *L. major*, and subsequently in other *Leishmania* species, and two in *T. brucei* and *T. cruzi*, orthologous to *Leishmania* PABP1 and PABP2. In *Leishmania*, the three PABPs are simultaneously expressed as abundant proteins at least during the promastigote insect stage of its life cycle [76].

PABP1, the first *Leishmania* PABP homolog identified, is a phosphoprotein that binds specifically to poly-A and localizes to the cytoplasm. It interacts in vivo and in vitro with EIF4G3, the *Leishmania* eIF4G homolog most directly implicated in translation initiation. The second PABP homolog, PABP2, is the ortholog to the one originally identified in *T. cruzi* and *T. brucei*. Its specificity in binding to poly-A seems to be reduced, and it contains several polymorphisms in residues almost universally conserved in PABP sequences and previously implicated in poly-A recognition. PABP3, absent from *T. brucei* and *T. cruzi*, co-precipitates with PABP2 in a mRNA independent manner, and it seems likely that both bind to the same population of mRNAs and may interact with each other [76].

There is evidence that the PABP orthologs have discrete functions: PABP1 binds to a distinct set of mRNAs and also differs from PABP2 and PABP3 in trafficking between the cytoplasm and nucleus, since both PABP2 and PABP3, but not PABP1, migrate to the nucleus upon inhibition of transcription by actinomycin D [76]. A subsequent phylogenetic analysis of PABP sequences from different trypanosomatids and related kinetoplastids revealed that the lack of PABP3 from the *Trypanosoma* genus is a secondary loss that occurred after its divergence from the remaining trypanosomatid lineages [104], in agreement with PABP2 and PABP3 in *Leishmania* having related and perhaps redundant functions.

In *T. brucei*, the two PABP homologs were seen to stimulate translation of a reporter mRNA when tethered to their 3′ UTR, compatible with both proteins having relevant roles during translation initiation [69]. Nevertheless, both proteins are individually essential for viability and seem to have non-redundant roles [76]. In *T. cruzi*, both PABP homologs were shown to migrate to cytoplasmic mRNP granules [105], an observation also confirmed in *T. brucei* where the two proteins were found to localize to different sets of granules in response to different stress conditions or inhibition of transcription/*trans*-splicing, implying again a differential association with distinct mRNA populations [104, 106]. So far, however, not much has been described regarding specific mRNA targets differentially bound by these two PABP homologs, with the exception of an early report from *Crithidia fasciculata*, a related trypanosomatid, where PABP2 was found in a complex that binds to mRNAs with cell cycle-dependent regulation [107].

The evidence regarding the association of the different PABP homologs with the eIF4F complexes in *Leishmania* is still controversial and needs to be resolved. As discussed above, a clear interaction between PABP1 and EIF4E4 has been

identified through different means [34, 72], and PABP1, but not PABP2 or PABP3, was seen to co-precipitate with native EIF4G3 [76]. Other co-precipitation assays targeting EIF4E4 brought down both PABP1 and PABP2, but in these assays EIF4E3 and also EIF4G4 were found in the precipitated fractions [34]. Subsequently, PABP2 was found to co-precipitate with EIF4G4 but not with EIF4E3 [48]. The data available so far then favor an exclusive association of PABP1 with the EIF4E4/EIF4G3 complex; an association that seems to be conserved in *T. brucei*, since co-precipitation experiments have also seen that PABP1 co-migrates with the complex formed by EIF4E4/EIF4G3 [104, 106]. In the same study, PABP2 co-migrated with EIF4E1 so it is still not clear whether a similar association exists between PABP2 and the second eIF4F complex, formed by EIF4E3/EIF4G4. No PABP homologs co-precipitated with the recently described eIF4F-like complexes based on EIF4E5 and EIF4E6, which indeed do seem to be functionally distinct [56, 75].

A number of RNA-binding proteins from both *T. brucei* and *T. cruzi* have been seen to associate with either of the PABP homologs, and their characterizations have led to relevant data that can be useful for the understanding of the functional distinctions between these proteins. In *T. cruzi*, PABP2 has been found to associate specifically with proteins named as UBPs, which bind to U-rich sequences found within the 3′ UTRs of trypanosomatid mRNAs and are involved in mRNA destabilization events [108]. In *T. brucei*, both PABPs were also found to co-precipitate with four distinct Alba-domain proteins, two of which are associated with polysomes and might be required for the stage-specific translation of selected mRNAs [109]. Also in *T. brucei*, the two PABP homologs have been seen to interact with the *T. brucei* ortholog of PBP, a non-essential protein in yeast that interacts with PABP and co-sediments with polysomes. In *T. brucei* the PBP ortholog seems to bridge an interaction between the PABPs and ZC3H11, a zinc finger RNA-binding protein that binds selectively to AU-rich elements in mRNAs and stimulates translation in tethering assays. A model has been proposed, according to which these interactions would allow the selective translation of mRNAs bound by ZC3H11 [110].

1.3 Ribosomes and Ribosomal Proteins

A remarkable characteristic of the trypanosomatid ribosomes, first reported in the 1980s, is the processing events targeting the ribosomal rRNAs. These lead to the generation of a very long 18S rRNA, the presence of only one form of 5.8S rRNA and the 25/28S rRNA being fragmented into six mature transcripts [111–114]. More recently, with the solving of the trypanosomatid's ribosome structure, and the characterization of its protein content, much more is understood regarding similarities and differences in ribosome structure and function in comparison with other eukaryotes, as described below.

1.3.1 Ribosome Structure

In 2005, the structure of the *Trypanosoma cruzi* 80S ribosome was resolved by cryo-electron microscopy (cryo-EM) and the general structures of the 40S and 60S subunits display evolutionary conserved features seen in other eukaryotes. The major characteristics of a typical ribosome and the phylogenetically conserved eukaryotic rRNA core structure are also maintained. However, in comparison to other species there are some distinctive differences in both small and large subunits. Indeed, the density map of the *T. cruzi* 80S ribosome includes unusual structural components, mostly related to large expanded segments in the rRNA molecules [115]. Subsequently, the structure of the *T. brucei* ribosome was solved, also through cryo-EM but with a higher resolution. As noted for *T. cruzi*, the main characteristic of the ribosome is the unusual size of several expanded segments (ES) in the rRNAs. Some of these segments, found in both ribosomal subunits, seem to be several times larger than in other ribosomes of known structure; with the 60S subunit also including an additional domain, called the kinetoplastid-specific domain (KSD) [116]. The functions of the large expanded segments of rRNA remain unknown, although the arrangement and localization of the largest expansions within the 18S rRNA (ES6/ES7) near the mRNA channel suggest that it may be important for translation control. Indeed, part of the ES6/ES7 in the 40S subunit makes up a structure first reported in *T. cruzi* and which was named 'turret'. This structure was considered the longest helical structure ever observed in a ribosome, and it only exists in trypanosomatids, with high conservation in sequence and size. The extension of a lower part of the turret also forms a bridge with the 60S subunit, being responsible for a unique type of connection between the small and large subunits. Within the 60S subunit, a comparison of its structure with other eukaryotes revealed that the trypanosomatid subunit does not have typical eukaryotic features, such as a planar surface near the exit site of the polypeptide; instead, it has a shape more similar to the one seen in bacteria [115, 116]. Based on the differences cited above, mainly the turret structure, and the cap 4 5′ end of trypanosomatid mRNAs, a role in translation initiation for the turret structure has been proposed to provide an explanation for the low affinity of the trypanosomatid eIF4Es to the cap 4, when compared to the affinity between mouse eIF4E and the monomethylated cap [115].

1.3.2 Ribosomal Proteins

Sequence analysis of ribosomal proteins showed that almost all yeast ribosomal proteins have counterparts in *T. cruzi* and *T. brucei*, with the exception of L41 and S31 [115, 116]. In regard to individual ribosomal proteins, the P complex present in the stalk of the ribosome, involved in the translocation step of protein synthesis, was described quite early on from *T. cruzi*. Four components were identified prior to the availability of the genome sequence (P0, P1, P2α and P2β). P0 was found to have a divergent C terminus, while P1, P2α and P2β have more typical C-terminal

ends [117]. Later, a novel putative P complex component, named P1β, was identified within the *T. cruzi* genome. Further sequence analysis revealed that *T. brucei* and *Leishmania* also contain a similar set of five P proteins, suggesting that the composition of the different P protein subtypes is a feature found in their common ancestor. Yeast two-hybrid and surface plasmon resonance analysis revealed a unique protein-protein interaction pattern for this complex, when compared to other eukaryotes, as P0 interacts with each of the four P1/P2 proteins [89, 118]. More recently, the P proteins have been investigated as potential drug targets, with the identification of a single chain recombinant antibody directed against the C-terminal end of the *T. cruzi* P2β, which specifically inhibits the translation of various trypanosomatid species [119].

In an expanded data-mining search for trypanosomatid ribosomal proteins, using *S. cerevisiae* sequences as queries, several putative ribosomal proteins were identified within the *T. cruzi* genome database. However, the average amino acid identity was very low (approximately 50 %), contrasting with the high degree of conservation of the 80S ribosome during evolution. *T. cruzi* ribosomal proteins are, on average, longer than their *S. cerevisiae* counterparts, with the extra regions usually at their N- or C-terminal ends. Thirty-two genes encoding proteins with significant sequence identity to polypeptides found within the *S. cerevisiae* 40S subunit were found, where 29 of those were also identified after mass spectrometry of the *T. cruzi* ribosome. For the 60S subunit, genes encoding putative orthologs for all 48 yeast proteins were found, with the mass spectrometry confirming the presence of most proteins within the native ribosome, with the exceptions of L1, L35, L39, L40, P1 and P2 [120].

One component of the large ribosomal subunit characterized in more detail is the L5 protein, tightly associated with the 5S rRNA to protect and stabilize it from degradation by nucleases. In trypanosomes, two novel 5S rRNA-binding proteins were found as well as L5, named P34 and P37 [121]. While it is reported that 90 % of mammalian 5S rRNA is bound by L5, only about 25 % of the trypanosomatid 5S rRNA is associated to its L5 ortholog. The amino acid sequence of the trypanosomatid L5 differs from the eukaryotic consensus at potentially significant positions; however, L5 still is essential to cell viability in *T. brucei*. Considering their characteristics, the trypanosome-specific P34 and P37 probably compensates for the weaker binding of L5 to the 5S rRNA, and it has been suggested that the trimolecular complex is necessary for adequate 5S rRNA stability in trypanosomatids [122]. The scaffold protein RACK1, which is a constituent of all eukaryotic ribosomes, was also identified in its conserved binding site within the small subunit of *T. brucei* ribosomes. For some unresolved reason, RACK1 was missing from the *T. cruzi* ribosome cryo-EM analysis, but data mining indicates the presence of one RACK1 ortholog within the *T. cruzi* genome database [115, 116].

1.4 tRNAs and tRNAs Synthetases

The total number of tRNA genes in the Tritryps was early on found to be low when compared to most eukaryotes, which usually have hundreds of tRNA genes [123]. Subsequent bioinformatic analysis confirmed 83 tRNA genes annotated in the *L. major* genome and indicated 66 genes in *T. brucei*. The distribution of these genes does not seem to be random as they are normally organized into clusters that are mostly confined to a subset of chromosomes [124]. Sixteen anticodons were not found in the tRNA genes of trypanosomatids, although their corresponding codons are present in many mRNAs of these organisms [125]. It is well known that some relaxed base pairing at the third position of the codon occurs (wobble), and this seems to be the case in trypanosomatids since there are 46 isoacceptors types, which are able to read the 61 canonical codons [124].

Genes encoding tRNA-Sec, which reads UGA as a selenocysteine codon in some specific mRNAs, were also identified in all three TriTryp species, but with variations in the number of copies [124]. The presence of selenoproteins is reported in trypanosomatids [126–128], although they do not seem to be essential for parasite viability [129, 130]. Unlike other tRNAs that are transcribed by RNA pol III, tRNA-Sec of *L. major* and *T. brucei* are reported to be transcribed only by RNA pol II or by RNA pol II and III [131, 132].

Few aminoacyl-tRNA synthetases (aaRSs) have been individually characterized in trypanosomatids. Twenty-four nuclear genes have been identified in the genome of *T. brucei*, all highly conserved in *T. cruzi,* which encode enzymes to aminoacylate all 20 amino acids [133]. Besides the 24 conserved genes, *Leishmania spp*. have an extra copy for the AsnRS [134]. Unlike most eukaryotes that usually encode two genes for each of the 20 standard aminoacyl synthetases, trypanosomatids encode only one copy of each gene, except for AspRS, TrpRS and LysRS, which have two copies each, and PheRS, which is formed by two subunits (α and β). Amino acid sequence comparison shows a good identity with other eukaryotic aminoacyl-tRNA synthetases, and the enzymes identified play roles in either cytoplasmic or mitochondrial translation, with some of them working in both compartments. The dual localization is the result of many different processes such as alternative *trans*-splicing, alternative start codons or protein modification and/or translocation to the mitochondria. All gene products are essential for cell growth in *T. brucei* procyclic and bloodstream forms. As in mammals and yeast, several *T. brucei* aminoacyl-tRNA synthetases and other related proteins seem to be associated in a multiprotein complex. This complex, called MARS (Multiple Aminoacyl-tRNA Synthetase complex), is responsible for enhancing the efficiency and accuracy of tRNA aminoacylation and improving general translation, by avoiding charged tRNAs diffusion and facilitating tRNA reutilization [133, 135].

1.5 Translation in Trypanosomatid Mitochondria

The trypanosomatid mitochondrial genome encodes similar polypeptides to other eukaryotes as well as two rRNAs. As in other organisms, translation then is required for synthesis of several components of the electron transport chain, but the process by which some of these components are synthesized is distinct from the mammalian system. First, transcripts are post-transcriptionally modified through RNA editing, an extensive and precise insertion and deletions of uridines, usually within the coding regions, correcting frameshifts and producing translatable mRNAs [136, 137]. Second, major differences in the mitochondrial translation apparatus have also been identified, which are described below.

1.5.1 Mitochondrial tRNAs and tRNA Synthetases

In contrast to yeast and mammals, the mitochondrial genome of trypanosomes are devoid of tRNAs genes, and so all the tRNAs needed for translation in the organelle are imported in small amounts from the cytosol through an unique import machinery [138, 139]. This is dependent on the binding of the tRNAs by EF1A in the cytosol [140], and both tRNA and protein imports share elements in common [141]. Upon reaching the mitochondria, most of these tRNAs are aminoacylated by the same synthetases also found in the cytosol. One exception is the $tRNA^{ASP}$, which is targeted by distinct synthetases in the two compartments [142].

As in all eukaryotes, the mitochondrial translation in trypanosomatids is of the bacterial type although it only uses eukaryotic-type tRNAs. This has resulted in some adaptations to integrate the imported eukaryotic-type tRNAs [143]. For instance, in mitochondria the stop codon UGA has been reassigned to tryptophan. Trypanosomatids have adapted to this through a mitochondria-specific RNA-editing event that converts the CCA anticodon of the imported $tRNA^{Trp}$ to UCA and by evolving a highly divergent eukaryotic-type TrpRS that is specific for mitochondria [144, 145]. Generation of the formylated initiator methionine is required for mitochondrial initiation of translation, but in trypanosomatids the single initiator $tRNA^{Met}$ remains in the cytosol and is required solely for cytosolic translation. Within the mitochondria only the elongator $tRNA^{Met}$ is found, but a fraction of it is targeted by the formylase activity to generate the formylated $tRNA^{Met}$, which is then used during the mitochondrial translation. This unique activity is carried out by a polypeptide homologous to other prokaryotic and mitochondrial enzymes of similar function [146, 147].

1.5.2 Mitochondrial Translation Factors

The orthologs of mitochondrial translation factors (EF-Tu, EF-G1, EF-Ts and RF1) are readily detected in trypanosomatid genomes, and the depletion of each of these

four factors results in growth impairment of the procyclic, insect forms. Despite having much reduced mitochondria and not depending on oxidative phosphorylation, the mammalian bloodstream forms of *T. brucei* also require the functional mitochondrial translation for cell survival, since the depletion of EF-Tu through RNAi in bloodstream cells also led to growth arrest and death [145, 148]. EF-Tu of trypanosomatids is very similar to other species, although it includes a C-terminal extension of approximately 30 amino acids. Sequence comparison has showed that the cytosolic EF-Tu ortholog, EF-1A, also has a similar C-terminal extension enriched in charged amino acids. RNAi depletion assays have confirmed that the trypanosomatid-specific C-terminal extension of EF-Tu is essential, and complementation with a chimeric EF-Tu harboring the corresponding segment of EF-1A is able to restore cell growth. Thus, it has been suggested that this specific motif may be linked to unique features of mitochondrial translation, as, for example, an adaptation that allows the protein to fit in the peculiar organellar ribosomes of these parasites [148].

1.5.3 Mitochondrial Ribosomes

The mitochondrial ribosome of the non-pathogenic *Leishmania tarentolae* is a minimal ribosomal rRNA-containing structure. It was isolated as a 50S protein-rich complex, and its two subunits were characterized as 28-30S small and 40S large subunits, and an unusual ≈45S SSU complex, representing a dimer of the small subunits and additional proteins. The cryo-EM of this ribosome shows most of the main characteristics of a typical ribosome, but with a size smaller than its eubacterial or mammalian counterparts. Comparing the morphology of each subunit, they resemble eubacterial subunits, containing several homologs of bacterial ribosomal proteins. Despite these similarities, more than 60 % of the eubacterial rRNA segments are absent while other differences in the density map of the small subunit are the result of specific proteins and extensions of the *Leishmania* mitochondrial ribosome. In addition, the subunits are held together by only nine inter-subunit bridges, fewer than those found in bacterial (13) or mammalian mitochondrial (15) ribosomes. The overall size and the maintenance of minimal key architectural components reflect a conserved functioning of the ribosomes. However, distinct topological differences in the *Leishmania* mitochondrial ribosome suggest that the process of protein synthesis is mechanistically different [149].

The remarkably small size of the mitochondrial ribosomes in trypanosomatids is due to the very short rRNAs found in their small and large subunits (9S and 12S, respectively). The main rRNA core regions are conserved, but other known domains are completely absent. A total of 133 proteins were identified in association with purified ribosomal subunits in the *T. brucei* mitochondrial ribosomes, and, when compared, the orthologs of all 49 proteins that were found in 45S SSU-related ribosomal complex in *L. tarentolae* were identified [143]. The specific function of the 45S SSU-related complex is not fully understood although it has been suggested that it could be involved in the translational control of specific

mRNAs. More recently, this complex has been shown in *T. brucei* to be indispensable for normal cellular growth and translation of mitochondrial mRNAs, such as those encoding cytochrome *c* oxidase and cytochrome *bc1*, in procyclic insect forms. The same complex, however, seems to be absent or downregulated in the bloodstream mammalian forms, in which the respiratory chain is not required [150].

It is possible that the higher number of ribosomal proteins identified in the mitochondrial ribosomes of trypanosomatids works in a way to compensate for the shorter rRNAs. Only 30 % of the ribosome-associated proteins in the *T. brucei* mitochondrial ribosome have orthologs in bacteria or in other eukaryotes, while 70 % are trypanosomatid-specific or may have diverged beyond confident recognition when compared to proteins from other species. To date, 56 proteins with high confidence assignment to purified mitochondrial ribosome of *T. brucei* have no recognizable homology outside the Kinetoplastida while several other proteins have no evident homology to other ribosomal proteins despite having motifs suggestive of associated functions, for example, PPR, GTP binding, GTPase and methyltransferase motifs [143, 151].

1.6 RNA Binding Proteins

The lack of transcriptional regulation and the reliance of trypanosomatids on post-transcriptional mechanisms to control gene expression, mainly at the mRNA level, imply the existence of elaborated mechanisms associated with regulation of mRNA translation, possibly associated with a diversified set of RNA-binding proteins (RBPs). It is believed that the association of specific RNA-binding proteins with 3′ UTRs of different classes of regulated transcripts may coordinately regulate the stability and translation of groups of mRNAs in response to environmental signals. This process likely constitutes a major mechanism for regulation of gene expression in these organisms, mediated by *cis*-acting signals through the binding of mRNA-specific or generic RNA-binding proteins [25, 28, 152]. Several recent reviews have discussed the different classes of RNA-binding proteins found in trypanosomatids, their diversity and conservation in comparison to other eukaryotes. These include RNA helicases of the DEAD-box subfamily, the RRM containing proteins, the CCCH family of zinc-finger proteins, PUF and ALBA domain-containing proteins [25, 153–156]. This review will briefly discuss what is known regarding these proteins with the focus on those that have been shown to play a role in mRNA metabolism and/or translation.

1.6.1 RNA Helicases

The first reports of DEAD-box RNA helicases in trypanosomatids were in 1995 and included a description of a 64-kDa polypeptide (HEL64) from *T. brucei* homologous to a yeast nuclear RNA helicase [157] and the first *Leishmania* eIF4A

homolog [57]. Subsequently, a great number of polypeptides were identified and annotated as putative DEAD/DEAH helicases within the TriTryp genomes. A systematic assessment of the full range of trypanosomatids' helicases, however, revealed that the total number of true RNA helicases of the DEAD-Box subfamily does not significantly exceed the number described for other eukaryotes of similar life cycle. Within the TriTryps 27 to 30 different DEAD-Box helicases were identified [155, 158], a number comparable to the 25 helicases reported from *S. cerevisae* [159]. Indeed, only four of the trypanosomatid DEAD-Box RNA helicases were identified as components of the *T. brucei* polysomes, and this include orthologs to two DED1-related proteins originally described from *Leishmania* [49], as detailed below.

The two DED1 homologs (DED1-1 and DED1-2) that have been identified in trypanosomatids have evolved from a gene duplication event that preceded the origin of the different trypanosomatid lineages. The two *Leishmania* proteins are able to complement a yeast mutant lacking the endogenous proteins Ded1p. Both co-precipitated with EIF4E1 and EIF4E4 but only DED1-2 with EIF4G3. In both instances, however, these interactions were RNA dependent and were also not detected through yeast two-hybrid assay. A stage-specific expression pattern was observed in *L. amazonensis*, with DED1-2 being preferentially expressed in the insect promastigote stage, while DED1-1 was more abundant in the amastigote mammalian stage. RNAi assays using the conserved *T. brucei* orthologs in procyclic cells lead to a very mild reduction in proliferation rate after DED1-1 depletion, while depletion of DED1-2 induces a proliferation arrest. These experiments suggest redundant functions for the two proteins since simultaneous depletion of both homologs lead to a synthetic phenotype with a rapid growth arrest and inhibition of translation [160]. *Leishmania* DED1-1 was also investigated in an independent study that identified it in a search for proteins that bound to the large ribosomal subunit rRNA and was found to have a protective role in preventing anti-sense ribosomal rRNA fragmentation [161].

Yet another member of the DEAD-Box family of RNA helicases that has roles implicated in the regulation of gene expression, and possibly translation, in trypanosomatids, is Dhh1. It was first identified in a search for eIF4A homologs in *Leishmania* where it was found to be the third nearest match to the human eIF4A within the *L. major* genomic sequences, although clearly distinct from the EIF4AI and EIF4AIII sequences [52]. Subsequently, the Dhh1 orthologs were characterized from *T. cruzi*, where they were found to be associated with P-bodies and other granules and with translationally repressed mRNAs [162, 163], and also from *T. brucei*. In fact, *T. brucei* Dhh1 was found to be essential for cell viability and the overexpression of either wild type or an ATPase-inactive form of Dhh1 led to a decrease in polysomes, increase in P-bodies and growth arrest. A selective effect on the expression of developmentally regulated mRNAs was observed, implicating this enzyme in critical roles during translation regulation in trypanosomatids [164]. Interestingly, Dhh1 was more recently found not to interact with the *T. brucei* ortholog of the known translational repressor SCD6. *T. brucei* SCD6 shares the same general domain structure seen in orthologs from different eukaryotic lineages

and localizes to P-bodies and other granules, and its depletion leads to an increase in global translation rates, demonstrating similar roles to the ones described in other organisms. Lack of an association with Dhh1, seen in all other organisms studied so far, suggests a clear difference in the mode of action of the two proteins during translation control in trypanosomatids [165].

1.6.2 RRM Domain Proteins

RNA recognition motif (RRM) proteins are involved in several processes that are conserved between trypanosomatid species. There are around 75 genes encoding RBPs with one or more RRM in trypanosomatid genomes [154, 166]. Some are involved in differentiation control, like RBP10, a protein involved in regulation of several mRNAs differentially expressed in the *T. brucei*'s mammalian form [167]. Another example is RBP6, which binds to AU-rich elements at 3′ UTR of mRNAs and is involved in differentiation/progression between different life stages [168, 169]. RBP42 binds mainly within the coding region of mRNAs that encode proteins involved in cellular energy metabolism [170]. An RRM protein homolog to hnRNP F/H controls *trans*-splicing efficiency and mRNA stability by binding to a purine-rich motif, AAGAA, found in some bloodstream form mRNAs [171]. UBP1 and UBP2, also RRM-containing proteins from *T. cruzi*, bind to a 43-nt U-rich mRNA element and stabilize the mRNAs from selective life stages [172, 173]. In *T. brucei* UBP1 and UBP2 are essential for normal growth and have different targets from *T. cruzi*, binding to a group of mRNAs from the F-box protein family, some of them transmembrane proteins. Overexpression of *T. brucei* UBP2 inhibited cell growth and upregulated the levels of several mRNAs [174, 175]. Yet another pair of RRM proteins, which has been better characterized in trypanosomatids, is the PTB homologs named PTB1/PTB2 (DRBD3/DRBD4). These are mainly nuclear proteins, but can localize to the cytoplasm and have been implicated in different events such as *trans*- and *cis*-splicing, mRNA transport, regulation of mRNA stability and translation [176–179].

1.6.3 CCCH Family Proteins

Several genes encoding proteins with one to five zinc-finger motifs (CCCH) were identified in the trypanosomatid genomes. Most are conserved and probably evolved before speciation events, and some may act in regulatory roles of mRNA metabolism [25, 153]. Relevant examples are ZFP1, ZFP2 and ZFP3, small CCCH-type proteins. In *T. brucei*, ZFP1 is transiently enriched during differentiation to tsetse fly midgut procyclic forms, and while ZFP1 RNAi had no phenotype in the mammalian bloodstream form, ZFP1 null bloodstream cell lines were unable to differentiate into procyclics. ZFP2 RNAi affects the ability of cells to differentiate to procyclic form, and ectopic expression of ZFP2 induced differentiation with elongated cell phenotypes due to polar extension of the cytoskeleton and cell cycle

arrest in G1 [180, 181]. ZFP2 from *T. cruzi* was found associated with mRNAs upregulated in the metacyclic trypomastigote forms [182]. Ectopic expression of ZFP3 in *T. brucei* also induced a differentiation phenotype similar to ZFP2 [183], and this protein is associated to a subset of mRNAs enriched in transmissible forms of the parasite [184]. Cycling sequence binding proteins (CSBP) A and B are CCCH proteins identified as binding to mRNAs containing the sequence CAUAGAAG, expressed during the S-phase, but their specific function remains unclear [185, 186].

1.6.4 PUF Domain Proteins

Eleven PUF proteins were identified in Trypanosomatids [187, 188], but only a few have been characterized. PUF2 exhibits a low level of in vivo RNA binding, is not associated with polysomes and may have a repressive role [189]. PUF6 regulates target mRNA levels by association with degradation complexes [190]. PUF7, PUF8 and PUF10 are nucleolar proteins involved in rRNA metabolism and do not seem to be involved in translation [191, 192]. PUF9 controls the expression of proteins involved in replicative processes by binding and stabilizing a small number of mRNAs that increase in G1 phase [193].

1.6.5 ALBA Proteins

In *Trypanosoma brucei* there are four ALBA (*acetylation lowers binding affinity*) proteins found in homo- and heterodimers in the cytoplasm [109]. ALBA1 and 2 interact with 3′ UTR elements of GPEET procyclin. ALBA3 and 4 are involved in cell growth and linked to differentiation. They are expressed in all stages of the *T. brucei* life cycle, except the transition from procyclic to epimastigote form. Complexes of ALBA1/2 and ALBA3/4 were found as components of stress granules, and they also seem to be associated with polysomes and can be co-purified with the translation initiation machinery, which indicates they may be involved in translation control [109, 194]. In *Leishmania infantum* there are two ALBA proteins, LiAlba1 and LiAlba3, which form a complex interacting with other RNA-binding proteins, ribosomal subunits and translation factors and may have a role in translational repression. They display differential subcellular localization since in the insect promastigote stage they localize predominantly in the cytoplasm but move to the nucleolus and the flagellum upon amastigote differentiation [195, 196].

1.7 Concluding Remarks

Throughout this review, novel features of translation in trypanosomatids have been described that highlight differences to other, better characterized eukaryotes. For instance, the canonical eIF4F complex EIF4G3/EIF4E4/EIF4AI binds PABP1 directly through EIF4E4 and appears to be regulated by phosphorylation at novel sites; in animals the interaction is between eIF4G and PABP. There is a second eIF4F complex also likely to be involved in translation, based on EIF4G4/EIF4E3, and it also may make similar interactions with PABP homologs. The evidence from trypanosomatids then reinforces the strong link between PABP and eIF4F function during eukaryotic translation initiation, and further characterization might help clarify the role of PABP in the process. Two further eIF4F complexes, recently identified and based on EIF4E5 and EIF4E6, are associated with proteins homologous to nucleotide-modifying enzymes associated with cap formation in the nucleus, but with so far undefined functions. Could these enzymes be involved in a novel cytoplasmic mRNA capping or recapping, or maybe translation repression or selective translation of small subsets of mRNAs? What are the mechanisms involved? It is also possible that the multiple eIF4F-like complexes identified in trypanosomatids constitute multiple hubs integrating signals from different sources and leading to the selective translation of specific mRNAs, reminiscent of nematodes [197]. Regarding ribosome function, the distinctive features seen in both cytoplasmic and mitochondrial ribosomes in trypanosomatids expand the range of ribosome diversity in eukaryotes considerably. They also raise questions regarding to what extent these ribosomes use conserved mechanisms in order to achieve proper translation or whether they depend on novel features/processes in order to accomplish their functions successfully. The great number of RNA-binding proteins and other mRNA associated proteins in trypanosomatids [70], most of which are novel and have unknown function, also highlights the diversity associated with mRNA metabolism and translation. Further approaches will be required in different fields in order to solve these and other questions raised by the study of translation in trypanosomatids, but undoubtedly the answers generated shall further highlight the exquisite and complex model for translation found in such extraordinary organisms.

Acknowledgments Work in trypanosomatid protein synthesis in OPDMN's laboratory is currently funded with grants provided by the Brazilian funding agencies CNPq (480899/2013-4) and CAPES (23038.007656/2011-92). The authors would like to thank the Science Without Borders Program from the Brazilian government, maintained by CNPq and Capes, for funding and the 'Special Visiting Researcher' fellowship awarded to MC (401888/2013-4). ERF is a recipient of a 'Young Talents' award, also from the Science Without Borders Program (401282/2014-7).

References

1. Stuart K, Brun R, Croft S, Fairlamb A, Gurtler RE, McKerrow J, et al. Kinetoplastids: related protozoan pathogens, different diseases. J Clin Invest. 2008;118(4):1301–10.
2. Simpson AG, Stevens JR, Lukes J. The evolution and diversity of kinetoplastid flagellates. Trends Parasitol. 2006;22(4):168–74.
3. Lukes J, Skalicky T, Tyc J, Votypka J, Yurchenko V. Evolution of parasitism in kinetoplastid flagellates. Mol Biochem Parasitol. 2014;195(2):115–22.
4. Cecilio P, Perez-Cabezas B, Santarem N, Maciel J, Rodrigues V, Cordeiro da SA. Deception and manipulation: the arms of leishmania, a successful parasite. Front Immunol. 2014;5:480.
5. Lievin-Le Moal V, Loiseau PM. *Leishmania* hijacking of the macrophage intracellular compartments. FEBS J. 2016;283(4):598–607.
6. Dostalova A, Volf P. *Leishmania* development in sand flies: parasite-vector interactions overview. Parasit Vectors. 2012;5:276.
7. Ivens AC, Peacock CS, Worthey EA, Murphy L, Aggarwal G, Berriman M, et al. The genome of the kinetoplastid parasite, *Leishmania major*. Science. 2005;309(5733):436–42.
8. Berriman M, Ghedin E, Hertz-Fowler C, Blandin G, Renauld H, Bartholomeu DC, et al. The genome of the African trypanosome *Trypanosoma brucei*. Science. 2005;309(5733):416–22.
9. El-Sayed NM, Myler PJ, Bartholomeu DC, Nilsson D, Aggarwal G, Tran AN, et al. The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. Science. 2005;309(5733):409–15.
10. Aslett M, Aurrecoechea C, Berriman M, Brestelli J, Brunk BP, Carrington M, et al. TriTrypDB: a functional genomic resource for the Trypanosomatidae. Nucleic Acids Res. 2010;38(Database issue):D457–D462.
11. Jackson AP. Genome evolution in trypanosomatid parasites. Parasitology. 2015;142(Suppl 1):S40–56.
12. Choi J, El-Sayed NM. Functional genomics of trypanosomatids. Parasite Immunol. 2012;34 (2–3):72–9.
13. Teixeira SM, de Paiva RM, Kangussu-Marcolino MM, Darocha WD. Trypanosomatid comparative genomics: Contributions to the study of parasite biology and different parasitic diseases. Genet Mol Biol. 2012;35(1):1–17.
14. Bartholomeu DC, de Paiva RM, Mendes TA, Darocha WD, Teixeira SM. Unveiling the intracellular survival gene kit of trypanosomatid parasites. PLoS Pathog. 2014;10(12): e1004399.
15. Daniels JP, Gull K, Wickstead B. Cell biology of the trypanosome genome. Microbiol Mol Biol Rev. 2010;74(4):552–69.
16. Martinez-Calvillo S, Vizuet-de-Rueda JC, Florencio-Martinez LE, Manning-Cela RG, Figueroa-Angulo EE. Gene expression in trypanosomatid parasites. J Biomed Biotechnol. 2010;2010:525241.
17. Alsford S, duBois K, Horn D, Field MC. Epigenetic mechanisms, nuclear architecture and the control of gene expression in trypanosomes. Expert Rev Mol Med. 2012;14:e13.
18. Muhich ML, Boothroyd JC. Polycistronic transcripts in trypanosomes and their accumulation during heat shock: evidence for a precursor role in mRNA synthesis. Mol Cell Biol. 1988;8 (9):3837–46.
19. Campbell DA, Thomas S, Sturm NR. Transcription in kinetoplastid protozoa: why be normal? Microbes Infect. 2003;5(13):1231–40.
20. Liang XH, Haritan A, Uliel S, Michaeli S. *trans* and *cis* splicing in trypanosomatids: mechanism, factors, and regulation. Eukaryot Cell. 2003;2(5):830–40.
21. Gunzl A. The pre-mRNA splicing machinery of trypanosomes: complex or simplified? Eukaryot Cell. 2010;9(8):1159–70.
22. Bangs JD, Crain PF, Hashizume T, McCloskey JA, Boothroyd JC. Mass spectrometry of mRNA cap 4 from trypanosomatids reveals two novel nucleosides. J Biol Chem. 1992;267 (14):9805–15.

23. Zinoviev A, Shapira M. Evolutionary conservation and diversification of the translation initiation apparatus in trypanosomatids. Comp Funct Genomics. 2012;2012:813718.
24. Ouellette M, Papadopoulou B. Coordinated gene expression by post-transcriptional regulons in African trypanosomes. J Biol. 2009;8(11):100.
25. De Gaudenzi JG, Noe G, Campo VA, Frasch AC, Cassola A. Gene expression regulation in trypanosomatids. Essays Biochem. 2011;51:31–46.
26. Schwede A, Kramer S, Carrington M. How do trypanosomes change gene expression in response to the environment? Protoplasma. 2012;249(2):223–38.
27. Kramer S. Developmental regulation of gene expression in the absence of transcriptional control: the case of kinetoplastids. Mol Biochem Parasitol. 2012;181(2):61–72.
28. Clayton CE. Networks of gene expression regulation in *Trypanosoma brucei*. Mol Biochem Parasitol. 2014;195(2):96–106.
29. Clayton C, Adams M, Almeida R, Baltz T, Barrett M, Bastien P, et al. Genetic nomenclature for *Trypanosoma* and *Leishmania*. Mol Biochem Parasitol. 1998;97(1–2):221–4.
30. Hernandez G, Vazquez-Pianzola P. Functional diversity of the eukaryotic translation initiation factors belonging to eIF4 families. Mech Dev. 2005;122(7–8):865–76.
31. Lorenzi HA, Vazquez MP, Levin MJ. The genes for a DEAH RNA helicase, a NifU like protein and the translation factor eIF6 constitute the SZ5 locus of *Trypanosoma cruzi*. Mol Biochem Parasitol. 2000;111(1):207–11.
32. Rezende AM, Assis LA, Nunes EC, da Costa Lima TD, Marchini FK, Freire ER, et al. The translation initiation complex eIF3 in trypanosomatids and other pathogenic excavates–identification of conserved and divergent features based on orthologue analysis. BMC Genom. 2014;15:1175.
33. Meleppattu S, Kamus-Elimeleh D, Zinoviev A, Cohen-Mor S, Orr I, Shapira M. The eIF3 complex of *Leishmania*-subunit composition and mode of recruitment to different cap-binding complexes. Nucleic Acids Res. 2015;43(13):6222–35.
34. Zinoviev A, Leger M, Wagner G, Shapira M. A novel 4E-interacting protein in *Leishmania* is involved in stage-specific translation pathways. Nucleic Acids Res. 2011;39(19):8404–15.
35. Urbaniak MD, Martin DM, Ferguson MA. Global quantitative SILAC phosphoproteomics reveals differential phosphorylation is widespread between the procyclic and bloodstream form lifecycle stages of *Trypanosoma brucei*. J Proteome Res. 2013 3;12(5):2233–44.
36. Guo J, Jin Z, Yang X, Li JF, Chen JG. Eukaryotic initiation factor 6, an evolutionarily conserved regulator of ribosome biogenesis and protein translation. Plant Signal Behav. 2011;6(5):766–71.
37. Kakuta Y, Tahara M, Maetani S, Yao M, Tanaka I, Kimura M. Crystal structure of the regulatory subunit of archaeal initiation factor 2B (aIF2B) from hyperthermophilic archaeon *Pyrococcus horikoshii* OT3: a proposed structure of the regulatory subcomplex of eukaryotic IF2B. Biochem Biophys Res Commun. 2004;319(3):725–32.
38. Dev K, Santangelo TJ, Rothenburg S, Neculai D, Dey M, Sicheri F, et al. Archaeal aIF2B interacts with eukaryotic translation initiation factors eIF2*alpha* and eIF2B*alpha*: Implications for aIF2B function and eIF2B regulation. J Mol Biol. 2009;392(3):701–22.
39. Nett IR, Davidson L, Lamont D, Ferguson MA. Identification and specific localization of tyrosine-phosphorylated proteins in *Trypanosoma brucei*. Eukaryot Cell. 2009;8(4):617–26.
40. Moraes MC, Jesus TC, Hashimoto NN, Dey M, Schwartz KJ, Alves VS, et al. Novel membrane-bound eIF2alpha kinase in the flagellar pocket of *Trypanosoma brucei*. Eukaryot Cell. 2007;6(11):1979–91.
41. da Silva Augusto L, Moretti NS, Ramos TC, de Jesus TC, Zhang M, Castilho BA, et al. A membrane-bound eIF2 *alpha* kinase located in endosomes is regulated by heme and controls differentiation and ROS levels in *Trypanosoma cruzi*. PLoS Pathog. 2015;11(2): e1004618.
42. Chow C, Cloutier S, Dumas C, Chou MN, Papadopoulou B. Promastigote to amastigote differentiation of *Leishmania* is markedly delayed in the absence of PERK eIF2*alpha* kinase-dependent eIF2*alpha* phosphorylation. Cell Microbiol. 2011;13(7):1059–77.

43. Tonelli RR, da Silva Augusto L, Castilho BA, Schenkman S. Protein synthesis attenuation by phosphorylation of eIF2*alpha* is required for the differentiation of *Trypanosoma cruzi* into infective forms. PLoS One. 2011;6(11):e27904.
44. Kramer S, Queiroz R, Ellis L, Webb H, Hoheisel JD, Clayton C, et al. Heat shock causes a decrease in polysomes and the appearance of stress granules in trypanosomes independently of eIF2(*alpha*) phosphorylation at Thr169. J Cell Sci. 2008;121(Pt 18):3002–14.
45. Hope R, Ben-Mayor E, Friedman N, Voloshin K, Biswas D, Matas D, et al. Phosphorylation of the TATA-binding protein activates the spliced leader silencing pathway in *Trypanosoma brucei*. Sci Signal. 2014;7(341):ra85.
46. Lahav T, Sivam D, Volpin H, Ronen M, Tsigankov P, Green A, et al. Multiple levels of gene regulation mediate differentiation of the intracellular pathogen *Leishmania*. FASEB J. 2011;25(2):515–25.
47. Cloutier S, Laverdiere M, Chou MN, Boilard N, Chow C, Papadopoulou B. Translational control through eIF2alpha phosphorylation during the *Leishmania* differentiation process. PLoS ONE. 2012;7(5):e35085.
48. Zinoviev A, Manor S, Shapira M. Nutritional stress affects an atypical cap-binding protein in *Leishmania*. RNA Biol. 2012;9(12).
49. Klein C, Terrao M, Inchaustegui GD, Clayton C. Polysomes of *Trypanosoma brucei*: Association with Initiation Factors and RNA-Binding Proteins. PLoS ONE. 2015;10(8): e0135973.
50. Damoc E, Fraser CS, Zhou M, Videler H, Mayeur GL, Hershey JW, et al. Structural characterization of the human eukaryotic initiation factor 3 protein complex by mass spectrometry. Mol Cell Proteomics. 2007;6(7):1135–46.
51. Dennis MD, Person MD, Browning KS. Phosphorylation of plant translation initiation factors by CK2 enhances the in vitro interaction of multifactor complex components. J Biol Chem. 2009;284(31):20615–28.
52. Dhalia R, Reis CR, Freire ER, Rocha PO, Katz R, Muniz JR, et al. Translation initiation in *Leishmania major*: characterisation of multiple eIF4F subunit homologues. Mol Biochem Parasitol. 2005;140(1):23–41.
53. Yoffe Y, Zuberek J, Lerer A, Lewdorowicz M, Stepinski J, Altmann M, et al. Binding specificities and potential roles of isoforms of eukaryotic initiation factor 4E in *Leishmania*. Eukaryot Cell. 2006;5(12):1969–79.
54. Clayton C, Shapira M. Post-transcriptional regulation of gene expression in trypanosomes and leishmanias. Mol Biochem Parasitol. 2007;156(2):93–101.
55. Jagus R, Bachvaroff TR, Joshi B, Place AR. Diversity of eukaryotic translational initiation factor eIF4E in protists. Comp Funct Genomics. 2012;2012:134839.
56. Freire ER, Vashisht AA, Malvezzi AM, Zuberek J, Langousis G, Saada EA, et al. eIF4F-like complexes formed by cap-binding homolog TbEIF4E5 with TbEIF4G1 or TbEIF4G2 are implicated in post-transcriptional regulation in *Trypanosoma brucei*. RNA. 2014;20 (8):1272–86.
57. Skeiky YA, Guderian JA, Benson DR, Bacelar O, Carvalho EM, Kubin M, et al. A recombinant *Leishmania* antigen that stimulates human peripheral blood mononuclear cells to express a Th1-type cytokine profile and to produce interleukin 12. J Exp Med. 1995;181(4):1527–37.
58. Skeiky YA, Kennedy M, Kaufman D, Borges MM, Guderian JA, Scholler JK, et al. LeIF: a recombinant *Leishmania* protein that induces an IL-12-mediated Th1 cytokine profile. J Immunol. 1998;161(11):6171–9.
59. Barhoumi M, Tanner NK, Banroques J, Linder P, Guizani I. *Leishmania infantum* LeIF protein is an ATP-dependent RNA helicase and an eIF4A-like factor that inhibits translation in yeast. FEBS J. 2006;273(22):5086–100.
60. Dhalia R, Marinsek N, Reis CR, Katz R, Muniz JR, Standart N, et al. The two eIF4A helicases in *Trypanosoma brucei* are functionally distinct. Nucleic Acids Res. 2006;34 (9):2495–507.

61. Le HH, Andersen GR. Structural insights into the exon junction complex. Curr Opin Struct Biol. 2008;18(1):112–9.
62. Bercovich N, Levin MJ, Clayton C, Vazquez MP. Identification of core components of the exon junction complex in trypanosomes. Mol Biochem Parasitol. 2009;166(2):190–3.
63. Freire ER, Dhalia R, Moura DM, da Costa Lima TD, Lima RP, Reis CR, et al. The four trypanosomatid eIF4E homologues fall into two separate groups, with distinct features in primary sequence and biological properties. Mol Biochem Parasitol. 2011;176(1):25–36.
64. Joshi B, Lee K, Maeder DL, Jagus R. Phylogenetic analysis of eIF4E-family members. BMC Evol Biol. 2005;5:48.
65. Yoffe Y, Zuberek J, Lewdorowicz M, Zeira Z, Keasar C, Orr-Dahan I, et al. Cap-binding activity of an eIF4E homolog from *Leishmania*. RNA. 2004;10(11):1764–75.
66. Lewdorowicz M, Yoffe Y, Zuberek J, Jemielity J, Stepinski J, Kierzek R, et al. Chemical synthesis and binding activity of the trypanosomatid cap-4 structure. RNA. 2004;10 (9):1469–78.
67. Yoffe Y, Leger M, Zinoviev A, Zuberek J, Darzynkiewicz E, Wagner G, et al. Evolutionary changes in the *Leishmania* eIF4F complex involve variations in the eIF4E-eIF4G interactions. Nucleic Acids Res. 2009;37(10):3243–53.
68. Gupta N, Goyal N, Rastogi AK. *In vitro* cultivation and characterization of axenic amastigotes of *Leishmania*. Trends Parasitol. 2001;17(3):150–3.
69. Erben ED, Fadda A, Lueong S, Hoheisel JD, Clayton C. A genome-wide tethering screen reveals novel potential post-transcriptional regulators in *Trypanosoma brucei*. PLoS Pathog. 2014;10(6):e1004178.
70. Lueong S, Merce C, Fischer B, Hoheisel JD, Erben ED. Gene expression regulatory networks in *Trypanosoma brucei*: Insights into the role of the mRNA-binding proteome. Mol Microbiol. 2016.
71. Pereira MM, Malvezzi AM, Nascimento LM, da Costa Lima TD, Alves VS, Palma ML, et al. The eIF4E subunits of two distinct trypanosomatid eIF4F complexes are subjected to differential post-translational modifications associated to distinct growth phases in culture. Mol Biochem Parasitol. 2013;190(2):82–6.
72. de Melo Neto OP, da Costa Lima TD, Xavier CC, Nascimento LM, Romao TP, Assis LA, et al. The unique *Leishmania* EIF4E4 N-terminus is a target for multiple phosphorylation events and participates in critical interactions required for translation initiation. RNA Biol. 2015;12(11):1209–21.
73. Moura DM, Reis CR, Xavier CC, da Costa Lima TD, Lima RP, Carrington M, et al. Two related trypanosomatid eIF4G homologues have functional differences compatible with distinct roles during translation initiation. RNA Biol. 2015;12(3):305–19.
74. Anderson P, Kedersha N, Ivanov P. Stress granules, P-bodies and cancer. Biochim Biophys Acta. 2015;1849(7):861–70.
75. Freire ER, Malvezzi AM, Vashisht AA, Zuberek J, Saada EA, Langousis G, et al. *Trypanosoma brucei* Translation Initiation Factor Homolog EIF4E6 Forms a Tripartite Cytosolic Complex with EIF4G5 and a Capping Enzyme Homolog. Eukaryot Cell. 2014;13 (7):896–908.
76. da Costa Lima TD, Moura DM, Reis CR, Vasconcelos JR, Ellis L, Carrington M, et al. Functional characterization of three leishmania poly(a) binding protein homologues with distinct binding properties to RNA and protein partners. Eukaryot Cell. 2010;9(10):1484–94.
77. Gile GH, Faktorova D, Castlejohn CA, Burger G, Lang BF, Farmer MA, et al. Distribution and phylogeny of EFL and EF-1*alpha* in Euglenozoa suggest ancestral co-occurrence followed by differential loss. PLoS ONE. 2009;4(4):e5162.
78. Kamikawa R, Brown MW, Nishimura Y, Sako Y, Heiss AA, Yubuki N, et al. Parallel re-modeling of EF-1*alpha* function: divergent EF-1alpha genes co-occur with EFL genes in diverse distantly related eukaryotes. BMC Evol Biol. 2013;13:131.
79. Szabova J, Ruzicka P, Verner Z, Hampl V, Lukes J. Experimental examination of EFL and MATX eukaryotic horizontal gene transfers: coexistence of mutually exclusive transcripts predates functional rescue. Mol Biol Evol. 2011;28(8):2371–8.

80. Alsford S, Turner DJ, Obado SO, Sanchez-Flores A, Glover L, Berriman M, et al. High-throughput phenotyping using parallel sequencing of RNA interference targets in the African trypanosome. Genome Res. 2011;21(6):915–24.
81. Nandan D, Cherkasov A, Sabouti R, Yi T, Reiner NE. Molecular cloning, biochemical and structural analysis of elongation factor-1 *alpha* from *Leishmania donovani*: comparison with the mammalian homologue. Biochem Biophys Res Commun. 2003;302(4):646–52.
82. Lopez M, Cherkasov A, Nandan D. Molecular architecture of leishmania EF-1*alpha* reveals a novel site that may modulate protein translation: a possible target for drug development. Biochem Biophys Res Commun. 2007;356(4):886–92.
83. Nandan D, Lopez M, Ban F, Huang M, Li Y, Reiner NE, et al. Indel-based targeting of essential proteins in human pathogens that have close host orthologue(s): discovery of selective inhibitors for *Leishmania donovani* elongation factor-1alpha. Proteins. 2007;67 (1):53–64.
84. Greganova E, Butikofer P. Ethanolamine phosphoglycerol attachment to eEF1A is not essential for normal growth of *Trypanosoma brucei*. Sci Rep. 2012;2:254.
85. Eltschinger S, Greganova E, Heller M, Butikofer P, Altmann M. Eukaryotic translation elongation factor 1A (eEF1A) domain I from *S. cerevisiae* is required but not sufficient for inter-species complementation. PLoS ONE. 2012;7(7):e42338.
86. Sasikumar AN, Perez WB, Kinzy TG. The many roles of the eukaryotic elongation factor 1 complex. Wiley Interdiscip Rev RNA. 2012;3(4):543–55.
87. Vickers TJ, Wyllie S, Fairlamb AH. *Leishmania major* elongation factor 1B complex has trypanothione S-transferase and peroxidase activity. J Biol Chem. 2004;279(47):49003–9.
88. Vickers TJ, Fairlamb AH. Trypanothione S-transferase activity in a trypanosomatid ribosomal elongation factor 1B. J Biol Chem. 2004;279(26):27246–56.
89. Smulski CR, Longhi SA, Ayub MJ, Edreira MM, Simonetti L, Gomez KA, et al. Interaction map of the *Trypanosoma cruzi* ribosomal P protein complex (stalk) and the elongation factor 2. J Mol Recognit. 2011;24(2):359–70.
90. Chung J, Rocha AA, Tonelli RR, Castilho BA, Schenkman S. Eukaryotic initiation factor 5A dephosphorylation is required for translational arrest in stationary phase cells. Biochem J. 2013;451(2):257–67.
91. Singh S, Raju K, Jatekar D, Dinesh N, Paul MS, Sobhia ME. *Leishmania donovani* eukaryotic initiation factor 5A: molecular characterization, localization and homology modelling studies. Microb Pathog. 2014;73:37–46.
92. Nguyen S, Leija C, Kinch L, Regmi S, Li Q, Grishin NV, et al. Deoxyhypusine Modification of eukaryotic translation initiation factor 5A (eIF5A) is essential for *Trypanosoma brucei* growth and for expression of polyprolyl-containing proteins. J Biol Chem. 2015;290 (32):19987–98.
93. Chawla B, Jhingran A, Singh S, Tyagi N, Park MH, Srinivasan N, et al. Identification and characterization of a novel deoxyhypusine synthase in *Leishmania donovani*. J Biol Chem. 2010;285(1):453–63.
94. Chawla B, Kumar RR, Tyagi N, Subramanian G, Srinivasan N, Park MH, et al. A unique modification of the eukaryotic initiation factor 5A shows the presence of the complete hypusine pathway in *Leishmania donovani*. PLoS ONE. 2012;7(3):e33138.
95. Nett IR, Martin DM, Miranda-Saavedra D, Lamont D, Barber JD, Mehlert A, et al. The phosphoproteome of bloodstream form *Trypanosoma brucei*, causative agent of African sleeping sickness. Mol Cell Proteomics. 2009;8(7):1527–38.
96. Dever TE, Gutierrez E, Shin BS. The hypusine-containing translation factor eIF5A. Crit Rev Biochem Mol Biol. 2014;49(5):413–25.
97. Inagaki Y, Ford DW. Evolution of the eukaryotic translation termination system: origins of release factors. Mol Biol Evol. 2000;17(6):882–9.
98. Inagaki Y, Blouin C, Doolittle WF, Roger AJ. Convergence and constraint in eukaryotic release factor 1 (eRF1) domain 1: the evolution of stop codon specificity. Nucleic Acids Res. 2002;30(2):532–44.

99. Inagaki Y, Blouin C, Susko E, Roger AJ. Assessing functional divergence in EF-1*alpha* and its paralogs in eukaryotes and archaebacteria. Nucleic Acids Res. 2003;31(14):4227–37.
100. Batista JA, Teixeira SM, Donelson JE, Kirchhoff LV, de Sa CM. Characterization of a *Trypanosoma cruzi* poly(A)-binding protein and its genes. Mol Biochem Parasitol. 1994;67 (2):301–12.
101. Pitula J, Ruyechan WT, Williams N. *Trypanosoma brucei*: identification and purification of a poly(A)-binding protein. Exp Parasitol. 1998;88(2):157–60.
102. Hotchkiss TL, Nerantzakis GE, Dills SC, Shang L, Read LK. *Trypanosoma brucei* poly(A) binding protein I cDNA cloning, expression, and binding to 5 untranslated region sequence elements. Mol Biochem Parasitol. 1999;98(1):117–29.
103. Bates EJ, Knuepfer E, Smith DF. Poly(A)-binding protein I of *Leishmania*: functional analysis and localisation in trypanosomatid parasites. Nucleic Acids Res. 2000;28(5):1211–20.
104. Kramer S, Bannerman-Chukualim B, Ellis L, Boulden EA, Kelly S, Field MC, et al. Differential localization of the two *T. brucei* poly(A) binding proteins to the nucleus and RNP granules suggests binding to distinct mRNA pools. PLoS ONE. 2013;8(1):e54004.
105. Cassola A, De Gaudenzi JG, Frasch AC. Recruitment of mRNAs to cytoplasmic ribonucleoprotein granules in trypanosomes. Mol Microbiol. 2007;65(3):655–70.
106. Kramer S, Marnef A, Standart N, Carrington M. Inhibition of mRNA maturation in trypanosomes causes the formation of novel foci at the nuclear periphery containing cytoplasmic regulators of mRNA fate. J Cell Sci. 2012;125(Pt 12):2896–909.
107. Mittra B, Ray DS. Presence of a poly(A) binding protein and two proteins with cell cycle-dependent phosphorylation in *Crithidia fasciculata* mRNA cycling sequence binding protein II. Eukaryot Cell. 2004;3(5):1185–97.
108. D'Orso I, Frasch AC. TcUBP-1, an mRNA destabilizing factor from trypanosomes, homodimerizes and interacts with novel AU-rich element- and Poly(A)-binding proteins forming a ribonucleoprotein complex. J Biol Chem. 2002;277(52):50520–8.
109. Mani J, Guttinger A, Schimanski B, Heller M, costa-Serrano A, Pescher P, et al. Alba-domain proteins of *Trypanosoma brucei* are cytoplasmic RNA-binding proteins that interact with the translation machinery. PLoS One. 2011;6(7):e22463.
110. Singh A, Minia I, Droll D, Fadda A, Clayton C, Erben E. Trypanosome MKT1 and the RNA-binding protein ZC3H11: interactions and potential roles in post-transcriptional regulatory networks. Nucleic Acids Res. 2014;42(7):4652–68.
111. White TC, Rudenko G, Borst P. Three small RNAs within the 10 kb trypanosome rRNA transcription unit are analogous to domain VII of other eukaryotic 28S rRNAs. Nucleic Acids Res. 1986;14(23):9471–89.
112. Campbell DA, Kubo K, Clark CG, Boothroyd JC. Precise identification of cleavage sites involved in the unusual processing of trypanosome ribosomal RNA. J Mol Biol. 1987;196 (1):113–24.
113. Hartshorne T, Toyofuku W. Two 5'-ETS regions implicated in interactions with U3 snoRNA are required for small subunit rRNA maturation in *Trypanosoma brucei*. Nucleic Acids Res. 1999;27(16):3300–9.
114. Sakyiama J, Zimmer SL, Ciganda M, Williams N, Read LK. Ribosome biogenesis requires a highly diverged XRN family 5'-> 3' exoribonuclease for rRNA processing in *Trypanosoma brucei*. RNA. 2013;19(10):1419–31.
115. Gao H, Ayub MJ, Levin MJ, Frank J. The structure of the 80S ribosome from *Trypanosoma cruzi* reveals unique rRNA components. Proc Natl Acad Sci USA. 2005;102(29):10206–11.
116. Hashem Y, des GA, Fu J, Buss SN, Jossinet F, Jobe A, et al. High-resolution cryo-electron microscopy structure of the *Trypanosoma brucei* ribosome. Nature. 2013;494(7437):385–9.
117. Levin MJ, Vazquez M, Kaplan D, Schijman AG. The *Trypanosoma cruzi* ribosomal P protein family: classification and antigenicity. Parasitol Today. 1993;9(10):381–4.
118. Ayub MJ, Smulski CR, Nyambega B, Bercovich N, Masiga D, Vazquez MP, et al. Protein-protein interaction map of the *Trypanosoma cruzi* ribosomal P protein complex. Gene. 2005;357(2):129–36.

119. Ayub MJ, Nyambega B, Simonetti L, Duffy T, Longhi SA, Gomez KA, et al. Selective blockade of trypanosomatid protein synthesis by a recombinant antibody anti-*Trypanosoma cruzi* P2*beta* protein. PLoS ONE. 2012;7(5):e36233.
120. Ayub MJ, Atwood J, Nuccio A, Tarleton R, Levin MJ. Proteomic analysis of the *Trypanosoma cruzi* ribosomal proteins. Biochem Biophys Res Commun. 2009;382(1):30–4.
121. Ciganda M, Williams N. Characterization of a novel association between two trypanosome-specific proteins and 5S rRNA. PLoS ONE. 2012;7(1):e30029.
122. Umaer K, Ciganda M, Williams N. Ribosome biogenesis in african trypanosomes requires conserved and trypanosome-specific factors. Eukaryot Cell. 2014;13(6):727–37.
123. Tan TH, Pach R, Crausaz A, Ivens A, Schneider A. tRNAs in *Trypanosoma brucei*: genomic organization, expression, and mitochondrial import. Mol Cell Biol. 2002;22(11):3707–17.
124. Padilla-Mejia NE, Florencio-Martinez LE, Figueroa-Angulo EE, Manning-Cela RG, Hernandez-Rivas R, Myler PJ, et al. Gene organization and sequence analyses of transfer RNA genes in Trypanosomatid parasites. BMC Genom. 2009;10:232.
125. Horn D. Codon usage suggests that translational selection has a major impact on protein expression in trypanosomatids. BMC Genom. 2008;9:2.
126. Cassago A, Rodrigues EM, Prieto EL, Gaston KW, Alfonzo JD, Iribar MP, et al. Identification of *Leishmania* selenoproteins and SECIS element. Mol Biochem Parasitol. 2006;149(2):128–34.
127. Lobanov AV, Gromer S, Salinas G, Gladyshev VN. Selenium metabolism in *Trypanosoma*: characterization of selenoproteomes and identification of a Kinetoplastida-specific selenoprotein. Nucleic Acids Res. 2006;34(14):4012–24.
128. Geslain R, Aeby E, Guitart T, Jones TE, de Castro MM, Charriere F, et al. *Trypanosoma* seryl-tRNA synthetase is a metazoan-like enzyme with high affinity for tRNASec. J Biol Chem. 2006;281(50):38217–25.
129. Aeby E, Seidel V, Schneider A. The selenoproteome is dispensable in bloodstream forms of *Trypanosoma brucei*. Mol Biochem Parasitol. 2009;168(2):191–3.
130. Aeby E, Palioura S, Pusnik M, Marazzi J, Lieberman A, Ullu E, et al. The canonical pathway for selenocysteine insertion is dispensable in Trypanosomes. Proc Natl Acad Sci USA. 2009;106(13):5088–92.
131. Aeby E, Ullu E, Yepiskoposyan H, Schimanski B, Roditi I, Muhlemann O, et al. tRNASec is transcribed by RNA polymerase II in *Trypanosoma brucei* but not in humans. Nucleic Acids Res. 2010;38(17):5833–43.
132. Padilla-Mejia NE, Florencio-Martinez LE, Moreno-Campos R, Vizuet-de-Rueda JC, Cevallos AM, Hernandez-Rivas R, et al. The selenocysteine tRNA gene in *Leishmania major* is transcribed by both RNA polymerase II and RNA polymerase III. Eukaryot Cell. 2015;14(3):216–27.
133. Cestari I, Kalidas S, Monnerat S, Anupama A, Phillips MA, Stuart K. A multiple aminoacyl-tRNA synthetase complex that enhances tRNA-aminoacylation in African trypanosomes. Mol Cell Biol. 2013;33(24):4872–88.
134. Gowri VS, Ghosh I, Sharma A, Madhubala R. Unusual domain architecture of aminoacyl tRNA synthetases and their paralogs from *Leishmania major*. BMC Genom. 2012;13:621.
135. Kalidas S, Cestari I, Monnerat S, Li Q, Regmi S, Hasle N, et al. Genetic validation of aminoacyl-tRNA synthetases as drug targets in *Trypanosoma brucei*. Eukaryot Cell. 2014;13 (4):504–16.
136. Simpson L, Thiemann OH, Savill NJ, Alfonzo JD, Maslov DA. Evolution of RNA editing in trypanosome mitochondria. Proc Natl Acad Sci U S A. 2000;97(13):6986–93.
137. Simpson L, Sbicego S, Aphasizhev R. Uridine insertion/deletion RNA editing in trypanosome mitochondria: a complex business. RNA. 2003;9(3):265–76.
138. Mukherjee S, Basu S, Home P, Dhar G, Adhya S. Necessary and sufficient factors for the import of transfer RNA into the kinetoplast mitochondrion. EMBO Rep. 2007;8(6):589–95.
139. Schneider A. Mitochondrial tRNA import and its consequences for mitochondrial translation. Annu Rev Biochem. 2011;80:1033–53.

140. Bouzaidi-Tiali N, Aeby E, Charriere F, Pusnik M, Schneider A. Elongation factor 1a mediates the specificity of mitochondrial tRNA import in *T. brucei*. EMBO J. 2007;26 (20):4302–12.
141. Seidman D, Johnson D, Gerbasi V, Golden D, Orlando R, Hajduk S. Mitochondrial membrane complex that contains proteins necessary for tRNA import in *Trypanosoma brucei*. J Biol Chem. 2012;287(12):8892–903.
142. Charriere F, O'Donoghue P, Helgadottir S, Marechal-Drouard L, Cristodero M, Horn EK, et al. Dual targeting of a tRNAAsp requires two different aspartyl-tRNA synthetases in *Trypanosoma brucei*. J Biol Chem. 2009;284(24):16210–7.
143. Niemann M, Schneider A, Cristodero M. Mitochondrial translation in trypanosomatids: a novel target for chemotherapy? Trends Parasitol. 2011;27(10):429–33.
144. Charriere F, Helgadottir S, Horn EK, Soll D, Schneider A. Dual targeting of a single tRNA (Trp) requires two different tryptophanyl-tRNA synthetases in *Trypanosoma brucei*. Proc Natl Acad Sci USA. 2006;103(18):6847–52.
145. Cristodero M, Seebeck T, Schneider A. Mitochondrial translation is essential in bloodstream forms of *Trypanosoma brucei*. Mol Microbiol. 2010;78(3):757–69.
146. Tan TH, Bochud-Allemann N, Horn EK, Schneider A. Eukaryotic-type elongator tRNAMet of *Trypanosoma brucei* becomes formylated after import into mitochondria. Proc Natl Acad Sci USA. 2002;99(3):1152–7.
147. Charriere F, Tan TH, Schneider A. Mitochondrial initiation factor 2 of *Trypanosoma brucei* binds imported formylated elongator-type tRNA(Met). J Biol Chem. 2005;280(16):15659–65.
148. Cristodero M, Mani J, Oeljeklaus S, Aeberhard L, Hashimi H, Ramrath DJ, et al. Mitochondrial translation factors of *Trypanosoma brucei*: elongation factor-Tu has a unique subdomain that is essential for its function. Mol Microbiol. 2013;90(4):744–55.
149. Sharma MR, Booth TM, Simpson L, Maslov DA, Agrawal RK. Structure of a mitochondrial ribosome with minimal RNA. Proc Natl Acad Sci USA. 2009;106(24):9637–42.
150. Ridlon L, Skodova I, Pan S, Lukes J, Maslov DA. The importance of the 45 S ribosomal small subunit-related complex for mitochondrial translation in *Trypanosoma brucei*. J Biol Chem. 2013;288(46):32963–78.
151. Zikova A, Panigrahi AK, Dalley RA, Acestor N, Anupama A, Ogata Y, et al. *Trypanosoma brucei* mitochondrial ribosomes: affinity purification and component identification by mass spectrometry. Mol Cell Proteomics. 2008;7(7):1286–96.
152. Fernandez-Moya SM, Estevez AM. Posttranscriptional control and the role of RNA-binding proteins in gene regulation in trypanosomatid protozoan parasites. Wiley Interdiscip Rev RNA. 2010;1(1):34–46.
153. Kramer S, Kimblin NC, Carrington M. Genome-wide *in silico* screen for CCCH-type zinc finger proteins of *Trypanosoma brucei*, *Trypanosoma cruzi* and *Leishmania major*. BMC Genom. 2010;11:283.
154. Kramer S, Carrington M. *Trans*-acting proteins regulating mRNA maturation, stability and translation in trypanosomatids. Trends Parasitol. 2011;27(1):23–30.
155. Gargantini PR, Lujan HD, Pereira CA. *In silico* analysis of trypanosomatids' helicases. FEMS Microbiol Lett. 2012;335(2):123–9.
156. Clayton C. The regulation of trypanosome gene expression by RNA-binding proteins. PLoS Pathog. 2013;9(11):e1003680.
157. Missel A, Norskau G, Shu HH, Goringer HU. A putative RNA helicase of the DEAD box family from *Trypanosoma brucei*. Mol Biochem Parasitol. 1995;75(1):123–6.
158. Marchat LA, rzola-Rodriguez SI, Hernandez-de la CO, Lopez-Rosas I, Lopez-Camarillo C. DEAD/DExH-Box RNA helicases in selected human parasites. Korean J Parasitol. 2015;53 (5):583–95.
159. Linder P. Dead-box proteins: a family affair–active and passive players in RNP-remodeling. Nucleic Acids Res. 2006;34(15):4168–80.

160. Zinoviev A, Akum Y, Yahav T, Shapira M. Gene duplication in trypanosomatids—Two DED1 paralogs are functionally redundant and differentially expressed during the life cycle. Mol Biochem Parasitol. 2012;185(2):127–36.
161. Padmanabhan PK, Samant M, Cloutier S, Simard MJ, Papadopoulou B. Apoptosis-like programmed cell death induces antisense ribosomal RNA (rRNA) fragmentation and rRNA degradation in Leishmania. Cell Death Differ. 2012;19(12):1972–82.
162. Holetz FB, Correa A, Avila AR, Nakamura CV, Krieger MA, Goldenberg S. Evidence of P-body-like structures in *Trypanosoma cruzi*. Biochem Biophys Res Commun. 2007;356 (4):1062–7.
163. Holetz FB, Alves LR, Probst CM, Dallagiovanna B, Marchini FK, Manque P, et al. Protein and mRNA content of TcDHH1-containing mRNPs in *Trypanosoma cruzi*. FEBS J. 2010;277(16):3415–26.
164. Kramer S, Queiroz R, Ellis L, Hoheisel JD, Clayton C, Carrington M. The RNA helicase DHH1 is central to the correct expression of many developmentally regulated mRNAs in trypanosomes. J Cell Sci. 2010;123(Pt 5):699–711.
165. Cristodero M, Schimanski B, Heller M, Roditi I. Functional characterization of the trypanosome translational repressor SCD6. Biochem J. 2014;457(1):57–67.
166. De Gaudenzi J, Frasch AC, Clayton C. RNA-binding domain proteins in Kinetoplastids: a comparative analysis. Eukaryot Cell. 2005;4(12):2106–14.
167. Wurst M, Seliger B, Jha BA, Klein C, Queiroz R, Clayton C. Expression of the RNA recognition motif protein RBP10 promotes a bloodstream-form transcript pattern in *Trypanosoma brucei*. Mol Microbiol. 2012;83(5):1048–63.
168. Kolev NG, Ramey-Butler K, Cross GA, Ullu E, Tschudi C. Developmental progression to infectivity in *Trypanosoma brucei* triggered by an RNA-binding protein. Science. 2012;338 (6112):1352–3.
169. Najafabadi HS, Lu Z, MacPherson C, Mehta V, Adoue V, Pastinen T, et al. Global identification of conserved post-transcriptional regulatory programs in trypanosomatids. Nucleic Acids Res. 2013;41(18):8591–600.
170. Das A, Morales R, Banday M, Garcia S, Hao L, Cross GA, et al. The essential polysome-associated RNA-binding protein RBP42 targets mRNAs involved in *Trypanosoma brucei* energy metabolism. RNA. 2012;18(11):1968–83.
171. Gupta SK, Kosti I, Plaut G, Pivko A, Tkacz ID, Cohen-Chalamish S, et al. The hnRNP F/H homologue of *Trypanosoma brucei* is differentially expressed in the two life cycle stages of the parasite and regulates splicing and mRNA stability. Nucleic Acids Res. 2013;41(13):6577–94.
172. De Gaudenzi JG, D'Orso I, Frasch AC. RNA recognition motif-type RNA-binding proteins in *Trypanosoma cruzi* form a family involved in the interaction with specific transcripts in vivo. J Biol Chem. 2003;278(21):18884–94.
173. Li ZH, De Gaudenzi JG, Alvarez VE, Mendiondo N, Wang H, Kissinger JC, et al. A 43-nucleotide U-rich element in 3'-untranslated region of large number of *Trypanosoma cruzi* transcripts is important for mRNA abundance in intracellular amastigotes. J Biol Chem. 2012;287(23):19058–69.
174. Hartmann C, Benz C, Brems S, Ellis L, Luu VD, Stewart M, et al. Small trypanosome RNA-binding proteins TbUBP1 and TbUBP2 influence expression of F-box protein mRNAs in bloodstream trypanosomes. Eukaryot Cell. 2007;6(11):1964–78.
175. Hartmann C, Clayton C. Regulation of a transmembrane protein gene family by the small RNA-binding proteins TbUBP1 and TbUBP2. Mol Biochem Parasitol. 2008;157(1):112–5.
176. Estevez AM. The RNA-binding protein TbDRBD3 regulates the stability of a specific subset of mRNAs in trypanosomes. Nucleic Acids Res. 2008;36(14):4573–86.
177. Stern MZ, Gupta SK, Salmon-Divon M, Haham T, Barda O, Levi S, et al. Multiple roles for polypyrimidine tract binding (PTB) proteins in trypanosome RNA metabolism. RNA. 2009;15(4):648–65.
178. Fernandez-Moya SM, Garcia-Perez A, Kramer S, Carrington M, Estevez AM. Alterations in DRBD3 ribonucleoprotein complexes in response to stress in *Trypanosoma brucei*. PLoS ONE. 2012;7(11):e48870.

179. Das A, Bellofatto V, Rosenfeld J, Carrington M, Romero-Zaliz R, del VC, et al. High throughput sequencing analysis of *Trypanosoma brucei* DRBD3/PTB1-bound mRNAs. Mol Biochem Parasitol. 2015;199(1–2):1–4.
180. Hendriks EF, Robinson DR, Hinkins M, Matthews KR. A novel CCCH protein which modulates differentiation of *Trypanosoma brucei* to its procyclic form. EMBO J. 2001;20 (23):6700–11.
181. Hendriks EF, Matthews KR. Disruption of the developmental programme of *Trypanosoma brucei* by genetic ablation of TbZFP1, a differentiation-enriched CCCH protein. Mol Microbiol. 2005;57(3):706–16.
182. Morking PA, Rampazzo RC, Walrad P, Probst CM, Soares MJ, Gradia DF, et al. The zinc finger protein TcZFP2 binds target mRNAs enriched during *Trypanosoma cruzi* metacyclogenesis. Mem Inst Oswaldo Cruz. 2012;107(6):790–9.
183. Paterou A, Walrad P, Craddy P, Fenn K, Matthews K. Identification and stage-specific association with the translational apparatus of TbZFP3, a CCCH protein that promotes trypanosome life-cycle development. J Biol Chem. 2006;281(51):39002–13.
184. Walrad PB, Capewell P, Fenn K, Matthews KR. The post-transcriptional trans-acting regulator, TbZFP3, co-ordinates transmission-stage enriched mRNAs in *Trypanosoma brucei*. Nucleic Acids Res. 2012;40(7):2869–83.
185. Pasion SG, Hines JC, Ou X, Mahmood R, Ray DS. Sequences within the 5' untranslated region regulate the levels of a kinetoplast DNA topoisomerase mRNA during the cell cycle. Mol Cell Biol. 1996;16(12):6724–35.
186. Brown LM, Ray DS. Cell cycle regulation of RPA1 transcript levels in the trypanosomatid *Crithidia fasciculata*. Nucleic Acids Res. 1997;25(16):3281–9.
187. Caro F, Bercovich N, Atorrasagasti C, Levin MJ, Vazquez MP. *Trypanosoma cruzi*: analysis of the complete PUF RNA-binding protein family. Exp Parasitol. 2006;113(2):112–24.
188. Luu VD, Brems S, Hoheisel JD, Burchmore R, Guilbride DL, Clayton C. Functional analysis of *Trypanosoma brucei* PUF1. Mol Biochem Parasitol. 2006;150(2):340–9.
189. Jha BA, Fadda A, Merce C, Mugo E, Droll D, Clayton C. Depletion of the Trypanosome Pumilio domain protein PUF2 or of some other essential proteins causes transcriptome changes related to coding region length. Eukaryot Cell. 2014;13(5):664–74.
190. Dallagiovanna B, Correa A, Probst CM, Holetz F, Smircich P, de Aguiar AM, et al. Functional genomic characterization of mRNAs associated with TcPUF6, a pumilio-like protein from *Trypanosoma cruzi*. J Biol Chem. 2008;283(13):8266–73.
191. Droll D, Archer S, Fenn K, Delhi P, Matthews K, Clayton C. The trypanosome Pumilio-domain protein PUF7 associates with a nuclear cyclophilin and is involved in ribosomal RNA maturation. FEBS Lett. 2010;584(6):1156–62.
192. Schumann Burkard G, Kaser S, de Araujo PR, Schimanski B, Naguleswaran A, Knusel S, et al. Nucleolar proteins regulate stage-specific gene expression and ribosomal RNA maturation in *Trypanosoma brucei*. Mol Microbiol. 2013;88(4):827–40.
193. Archer SK, Luu VD, de Queiroz RA, Brems S, Clayton C. *Trypanosoma brucei* PUF9 regulates mRNAs for proteins involved in replicative processes over the cell cycle. PLoS Pathog. 2009;5(8):e1000565.
194. Subota I, Rotureau B, Blisnick T, Ngwabyt S, Durand-Dubief M, Engstler M, et al. ALBA proteins are stage regulated during trypanosome development in the tsetse fly and participate in differentiation. Mol Biol Cell. 2011;22(22):4205–19.
195. Dupe A, Dumas C, Papadopoulou B. An Alba-domain protein contributes to the stage-regulated stability of amastin transcripts in *Leishmania*. Mol Microbiol. 2014;91(3):548–61.
196. Dupe A, Dumas C, Papadopoulou B. Differential subcellular localization of *Leishmania* Alba-Domain proteins throughout the parasite development. PLoS ONE. 2015;10(9): e0137243.
197. Friday AJ, Keiper BD. Positive mRNA translational control in germ cells by initiation factor selectivity. Biomed Res Int. 2015;2015:327963.

Evolutionary Aspects of Translation Regulation During Abiotic Stress and Development in Plants

René Toribio, Alfonso Muñoz, Ana B. Castro-Sanz, Alejandro Ferrando, Marta Berrocal-Lobo and M. Mar Castellano

1 Introduction

Plant growth requires a high spatial and temporal regulation of gene expression. Transcription has been considered the major player in this process. However, growing evidence points out that post-transcriptional regulation, especially translational control, also plays an important role in plant development and in response to environmental cues.

In the last decade, different studies have shown that translation factors and translational regulation are quite conserved in eukaryotes. Nevertheless, some diversity has been observed across different eukaryotes [1–3], including plants, where some of the mechanisms described in other eukaryotes are also operative but key players are missing in their genomes [4]. Despite this conservation, the knowledge of the mechanisms that regulate translation is very poor in plants.

In this chapter we will focus on the mechanisms that regulate selective mRNA translation, paying special attention to their possible conservation in the plant kingdom.

R. Toribio, A. Muñoz and A. B. Castro-Sanz have contributed equally to this work.

R. Toribio · A. Muñoz · A.B. Castro-Sanz · M. Berrocal-Lobo · M.M. Castellano (✉)
Centro de Biotecnología y Genómica de Plantas. Universidad Politécnica de Madrid (UPM) - Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA), Campus de Montegancedo (UPM), Pozuelo de Alarcón 28223, Madrid, Spain
e-mail: castellano.mar@inia.es

R. Toribio
e-mail: toribio.francisco@inia.es

A. Ferrando
Instituto de Biología Molecular y Celular de Plantas, CSIC-Universidad Politécnica de Valencia, Valencia, Spain

G. Hernández and R. Jagus (eds.), *Evolution of the Protein Synthesis Machinery and Its Regulation*, DOI 10.1007/978-3-319-39468-8_18

2 Tight Control at the Translation Initiation Phase Under Specific Developmental Cues and Stress Conditions

Under different conditions, eukaryotic cells finely regulate mRNA translation. In some cases, such as under stress conditions, a general regulation of translation affects a large part of the bulk of mRNAs. In other cases, translational regulation is subtler and only affects the translation of specific mRNAs. The mechanisms involved in both types of regulation have been characterized in mammals and yeast, where it is known to be mostly directed at the translation initiation step through the action of eIF2α and eIF4E [5–10]. In contrast, the mechanisms of translation regulation in plants are mainly unknown, and the processes by which some mRNAs may be selectively translated still remain to be elucidated [4, 11, 12].

3 Unique Characteristics of the eIF4E Complex in Plants

Throughout evolution, plants have conserved most translation factors involved in initiation. However, a clear difference exclusive of the higher plant translational machinery is the presence of an extra eIF4E isoform, termed eIFiso4E, which only shares around 50 % of homology with eIF4E [13]. Moreover, plants also have an additional eIF4G isoform, termed eIFiso4G, which usually lacks a large part of the N-terminal that is present in the eIF4G protein [14]. eIFiso4G seems to have appeared in basal plant lineages before eIFiso4E during evolution, which emerged around the period when flowering plants evolved [14]. Both eIF4E and eIFiso4E interact selectively with eIF4G and eIFiso4G to form eIF4F and eIFiso4F complexes, respectively. However, in the absence of their respective binding partners, they can form mixed complexes that are also functional in vitro [15]. Both eIF4F isoforms seem to have specialized translational activities that could impinge on the selective translation of different mRNAs. Indeed, different mRNAs have been shown to be initiated preferentially by the eIF4F or the eIFiso4F isoforms [16]. In this sense, the eIF4F complex has been shown to be more efficient than the eIFiso4F complex in translating highly structured mRNAs in wheat germ extracts [17].

4 Regulation of eIF4E Activity: 4E-BPs

The regulation of eIF4E under stress conditions is one of the best-studied mechanisms in other eukaryotes. One of the most interesting aspects of eIF4E regulation in mammals and *Drosophila* is its inhibition by eIF4E–binding proteins (4E–BPs). In these organisms, hypo-phosphorylated 4E–BPs interact with eIF4E through the 4E-binding domain (YXXXXLØ) that is also used by the eIF4G for its binding to eIF4E. For this reason, binding of the 4E-BPs to the eIF4E prevents the effective

formation of the eIF4F complex and the subsequent recruitment of the ribosome to the mRNA cap structure [18–21]. This generates a global inhibition of translation that can only be bypassed by those mRNAs that could be initiated in a cap-independent mechanism [22–24]. Interestingly, this regulation is coupled to the physiological state of the cell through the activity of the TOR kinase [25]. Thus, under TOR phosphorylation mammalian 4E-BPs dissociate from eIF4E, allowing the formation of the eIF4F complex and cap-dependent translation [22, 23].

The widespread distribution of the 4E-BP orthologs across eukaryotes, the conservation of the 4E-binding motif and phosphorylation sites of the 4E-BPs of distantly related species and the conservation of TOR in all eukaryotes have prompted speculation that a single ancestral 4E-BP emerged in early eukaryotes to shut down cap-dependent translation in response to nutritional and environmental stresses [2].

Despite lacking 4E-BP ortholog genes, plants also tightly regulate eIF4E. Indeed, plant translational apparatus is known to support cap–independent translation under stress conditions. First, plant viruses use a cap–independent translation strategy to translate their mRNAs in the host cells [26]. More relevantly, in maize, different cellular mRNAs, the alcohol dehydrogenase *ADH1* and the heat shock protein *HSP101,* are translated in a cap–independent manner in oxygen-deprived roots [27] and during heat stress [28]. Quite recently, the IRES-dependent translation of WUSCHEL has been proposed to control the stem cell homeostasis of *Arabidopsis* in response to environmental hazards [29]. Finally, TOR regulates protein synthesis in plants as it does in mammals and yeast [30–34]. Despite the parallelisms with other eukaryotes, no homolog of the 4E–BPs has been found in plants to date, suggesting that the ancestral 4E-BP that arose in the last common ancestor of extant eukaryotes (LCAEE) was lost early in the evolution of this kingdom [2]. Supporting this notion, it has been proposed that the existence of the 4E-BPs is unlikely based on the fact that the plant eIF4E and eIFiso4E form tighter complexes (on the order of nanomolar and subnanomolar) with their respective eIF4G subunits than their mammalian counterparts [13].

5 Regulation of eIF4E Activity: Other 4E-Binding Partners

Besides 4E-BPs, during the last decade other eIF4E-binding proteins have been described in mouse, *Drosophila* and *Xenopus* [24, 35, 36]. These other eIF4E-interacting proteins are not related in sequence or structure to the mammalian 4E-BPs, and for this reason it has been speculated that they have evolved independently [2].

While some of these proteins interact with the eIF4E through the canonical 4E-binding site, some others possess a similar structure instead that allow binding to eIF4E. Interestingly, some of them play additional roles to translation, suggesting that the binding to eIF4E evolved by co-opting pre-existing molecules from various

cellular and viral processes [2]. Another main difference between canonical 4E-BPs and the other eIF4E-interacting proteins is that they are usually recruited by specific proteins that only associate to and repress translation of a small subset of target mRNAs [24].

One of the best examples of this eIF4E-driven specific translational repression is the one carried out by MASKIN during the germline development in mice and *Xenopus* [37]. Mouse and *Xenopus* oocytes contain silent or "masked" mRNAs that are translated only when the cell re-enters the meiotic divisions or after fertilization. Some of these mRNAs contain a cytoplasmic polyadenylation element (CPE) in their 3′UTR that is recognized by the CPEB protein. CPEB binding promotes the further association of the eIF4E binding partner Maskin and the subsequent loading of the eIF4E. The interaction of CPEB, Maskin and eIF4E inhibits the formation of the eIF4F complex over the mRNAs targeted by CPEB and represses their translation until maturation [38, 39]. Alternatively to Maskin, the eIF4E binding partner neuroguidin controls the neural tube closure and neural crest migration in *Xenopus laevis* embryos, repressing the translation of CPE-containing mRNAs [40].

Another example of this translational inhibition is exerted through the activity of the eIF4E-binding partner Cup, which controls the proper spatio-temporal translation of *nanos* and *oskar* mRNAs during the embryogenesis and oogenesis in *Drosophila*. These specific mRNAs are bound by Smaug (Smg) and Bruno (Brn) through the Smg recognition elements (SRE) and Brn recognition elements (BRE) present in their respective sequences. As in the previous case, binding of Smg and Brn to these mRNAs allows the recruitment of the translational repressive Cup-eIF4E complex until they reach their correct localization [41, 42]. Moreover, a different eIF4E-binding partner, CYFIP1, inhibits the translation of mRNAs that are associated to FMRP, controlling the translation of different RNAs involved in synaptic plasticity and maturation [43]. Finally, an alternative control of specific mRNA translation by the selective recruitment of d4E-HP to certain mRNAs such as *caudal* and *hunchback* has also been described in *Drosophila.* d4E-HP is an eIF4E homologous protein that binds to the cap structure but does not contain the eIF4G domain [44]. Binding of d4E-HP inhibits the translation of the bound RNA as it also precludes the formation of the eIF4F complex. In this case, the selective recruitment of d4E-HP is directed by RNA-binding proteins, such as Bicoid (Bcd), Nanos (Nos) and Pumilio (Pum). Some of these proteins, such as Bcd, interact directly with d4E-HP [45]; others, such as Pum, form a complex with Brat that also interacts with d4E-HP [46], causing the translation inhibition of the mRNAs bound to them.

In plants, some of the RNA-binding proteins that in other eukaryotes target specific mRNAs for translational repression by their association to 4E-binding partners are also conserved. This is the case of the BRUNO-like proteins that, as in the mammals *Drosophila* and *Xenopus*, are also present in different plant species, such as *Oryza sativa* and *Arabidopsis* [47]. In Arabidopsis, the two BRN proteins bind to a BRE sequence present in the 3′UTR of *SOC1*. This binding regulates *SOC1* transcript and protein abundances, controlling in this way the flowering time in plants [47].

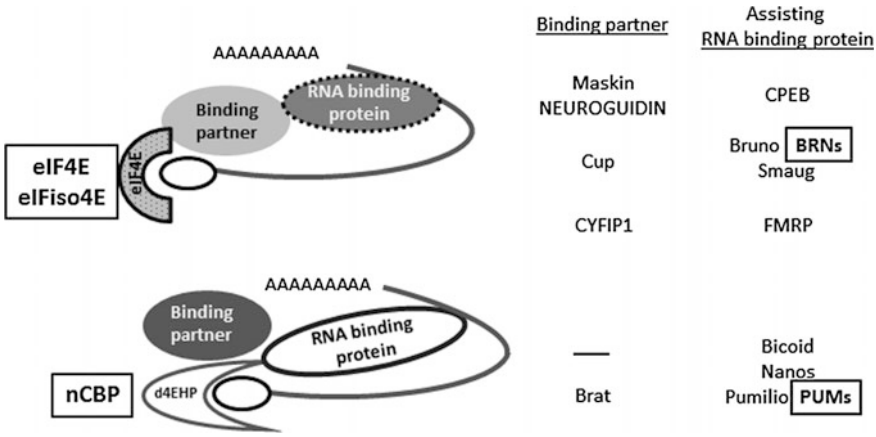


Fig. 1 Mechanisms for translation inhibition of specific mRNAs through eIF4E and d4EHP regulation. eIF4E and 4E-HP binding partners and their interacting RNA binding proteins are shown on the *left*. In the case of d4EHP, RNA binding proteins could interact directly with d4E-HP. Alternatively, this interaction could be assisted by Brat. The homologs of these proteins in the plant kingdom, if any, are boxed

In addition, the RNA-binding protein Pumilio has also been conserved throughout the evolution in mammals, fungi, protozoa and plants. Indeed, at least six PUM homologs to the *Drosphila* Pumilio have been identified in *Arabidopsis* [48]. These PUMs bind to a consensus sequence that is closely related to the NRE sequence recognized in *Drosophila*, suggesting that AtPUMs may also bind their target RNAs through evolutionary-conserved mechanisms. These sequences are present in the 3′UTRs of transcripts involved in stem cell maintenance in plants, reinforcing the idea that the role of PUM in the maintenance of the stem cell niche seems to be conserved throughout evolution [48].

Despite the existence of some RNA-binding proteins involved in the selective translational repression, in plants, no homologous proteins to the eIF-4E binding partners described in other eukaryotes (Maskin, Cup, CYFP1, etc.) have been identified. Nevertheless, the possible existence of some analogs is still an open question. Indeed, a simple search in the *Arabidopsis* database retrieved more than 6900 proteins that contain one or more canonical eIF4E binding domains (YXXXXLØ) and that, therefore, are susceptible to binding the eIF4E and inhibiting canonical initiation of translation. This number may be even larger if we take into account that the canonical domain could have variations in their 3′ end and that some structures such as the reversed L-shaped could also promote the eIF4E binding [43].

Finally, as cited above, Pumilio binds to d4E-HP to inhibit translational repression of certain mRNAs, and NCBP, a homologous protein to the d4E-HP, has also been previously identified in plants. This protein with unknown function is able to bind the cap structure but, unlike its animal counterparts, has shown a modest

ability to stimulate translation in the presence of eIFiso4G in wheat germ extracts [49]. Despite the presence of PUMs and NCBP in plants, whether PUMs are able to control the translation of their target genes by binding nCBPs still remains unsolved.

The presence of a large number of proteins susceptible to binding the eIF4E and eIFiso4E factors along with the conservation through evolution of some specific RNA binding proteins that modulate the translation of specific mRNAs through their association to eIF4E and d4E-HP (Fig. 1) opens a hot and exciting new field of research in plants that awaits further experimental evidence.

6 Regulation of Translation Initiation by eIF2α Phosphorylation

In mammals and yeast, one of the main mechanisms of translation inhibition in response to stress is the regulation by phosphorylation of the α subunit of the eIF2 translation factor. eIF2α phosphorylation by eIF2α kinases promotes the inhibition of translation by hindering the formation of the eIF2/GTP/$tRNA_i^{Met}$ternary complex [23]. All known eIF2α kinases share related kinase domains and phosphorylate highly conserved sites among the eIF2α orthologs. However, the number and the activation requirements of the different eIF2α kinases by stress conditions are different among eukaryotes. In vertebrates four different eIF2α kinases, namely GCN2, PERK, PKR and HRI, have been described. These eIF2α kinases are activated respectively by nutrient limitation [50], protein misfolding in the endoplasmic reticulum (ER) [51], virus infection [52] and heme group availability [53]. In other organisms the number of eIF2α kinases varies: HRI and GCN2 have been characterized in *Schizosaccharomyces pombe,* PERK and GCN2 are found in *D. melanogaster* and *Caenorhabditis elegans*, and GCN2 is the sole eIF2α kinase present in *Saccharomyces cerevisiae* [54]. The observation that GCN2, alone among eIF2α kinases, is ubiquitously distributed across eukaryotes supports the idea that GCN2 is the ancestral eIF2α kinase. This kinase seems to have emerged in the LCAEE after multicellularity diversified in several taxa into different linages to respond to diverse stress conditions [2].

A strong inhibition of protein synthesis by eIF2α phosphorylation under different stress conditions has also been reported in plants, demonstrating that this mechanism of regulation of translation is conserved in these organisms [55]. Genome-wide searches for the presence of eIF2α kinases suggest that higher plants only contain a GCN2-like eIF2α kinase [55], and, in agreement, so far only GCN2 has been characterized in plants [55, 56]. *Arabidopsis* GCN2 complements its yeast counterpart, suggesting that some aspects of the yeast amino acid control mechanism may also be conserved in the plant kingdom. Indeed, AtGCN2 is activated under amino acid deprivation promoted by herbicide treatment [56]. In addition, plant GCN2 is also activated by purine deprivation, cadmium, UV, cold shock,

wounding and in response to different hormones involved in defense response [55, 56].

Interestingly, inhibition of translation mediated by eIF2α phosphorylation is less dramatic in plants as compared to the one described in mammals. This observation is in line with a report demonstrating closer values of dissociation constants for wheat eIF2 with GDP/GTP than the values displayed in rabbit reticulocytes [57]. This opens the possibility that eIF2B recycling (the main effector of the inhibition of translation through the eIF2α phosphorylation) may be less critical in plants than it is in mammals, and it may explain the maintenance of a more efficient translation in plants as compared to animals when eIF2α is phosphorylated [57].

7 The Increasing Importance of SGs in the Translational Control

For years, research in translational control has focused on the regulation of the translation factors. However, with the identification of stress granules (SGs) and processing bodies (PBs), a new mechanism for translation regulation under stress conditions has emerged that seems to be conserved through evolution.

As stated before, different kinds of stresses cause a translational inhibition that promotes the general dissociation of the mRNAs from the actively translating ribosomes (polysomes). In such conditions, the mRNAs no longer engaged in translation are recruited into cytoplasmic ribonucleoprotein (RNP) complexes called SGs. In these complexes, the mRNAs are not available to the translation machinery but are safeguarded until they are exported to P-bodies for degradation or are released to resume translation when the conditions are favorable. SGs in mammals include, among others, eIF4E, eIF4G, eIF4A, eIF4B, eIF3, eIF2, the 40S ribosomal subunit and poly(A) binding protein (PABP). SGs also contain RNA binding proteins TIA1 and TIAR, which possess self–aggregation-mediating domains necessary for SGs formation [58].

Plant SGs do not seem to differ importantly in protein composition from SGs of the rest of eukaryotes. Indeed, the characterization in plants of UBP1 proteins, a family with the highest amino acid sequence similarity to TIA1, has shed light on the mechanism for the reversible aggregation of translationally repressed mRNAs to SGs under conditions of hypoxia in *Arabidopsis*. More importantly, this mechanism partially explains the rapid oscillation of mRNA translational activity in response to a transient energy stress [59].

In the absence of mechanisms that clearly explain the drastic inhibition of translation, particularly under certain abiotic stress conditions, the possibility that SGs could play a prevalent role in the regulation of translation in plants is being considered.

8 Translational Regulation at the Elongation Step: The Case of eIF5A

Elongation of translation is strongly affected under stress conditions as the mRNAs being translated disassemble from polysomes and are recruited to SGs to prevent their degradation. One atypical translation elongation factor involved in the formation of SGs after stress-mediated polysome disassembly is eIF5A [60], initially described as IF-M2Bα and later renamed as eIF4D and finally cataloged as eIF5A following successive uniform nomenclature. Early biochemical characterization of eIF5A as a protein factor loosely bound to the ribosome and capable of stimulating in vitro formation of methionyl-puromycin suggested that eIF5A is not a canonical initiation factor [61]. In fact, it took until 2009 to decipher eIF5A's role within the ribosome as a translation elongation factor in yeast [62]. Its original name as an initiation factor remains controversial, in spite of recent suggestions to change its name to eEF5 [63], which seems more appropriate to recapitulate its activity within the ribosome.

The eIF5A activation by posttranslational modification is a unique feature of this elongation factor, which converts a highly conserved lysine into the unusual amino acid hypusine [64]. The biological function of amino acid hypusine on the activity of the translation elongation factor eIF5A remained elusive for decades since its initial description in 1971 [65] and the elucidation of its biosynthesis from spermidine a few years later. Two enzymes act sequentially on eIF5A to carry out the hypusination process, namely deoxyhypusine synthase (DHS), which forms the intermediate residue named deoxyhypusine, and deoxyhypusine hydroxylase (DOHH), which catalyzes the subsequent hydroxylation of deoxyhypusine to generate the hypusine residue [66]. The requirement of spermidine for eIF5A hypusination is essential for eukaryotic cell viability and development [67–71]. Archaea contains both eIF5A (aIF5A) and DHS homologs, although no DOHH enzyme has been identified [72]. Although no homolog to either eIF5A or the hypusination enzymes has been found in prokaryotic genomes, a structural similar translation elongation factor, EFP, has been shown to suffer similar posttranslational modifications by completely different mechanisms in bacteria [73, 74]. Moreover, EFP posttranslational modifications are critical for its role during translation elongation [75]. Structural studies have located EFP between the P- and the E-sites of the 70S ribosome, thus supporting its role as an elongation factor [76]. Hydroxyl radical mapping of eIF5A on the yeast 80S ribosome indicated that eIF5A binds alongside the P-site, occupying a similar ribosomal location to the prokaryotic EFP [77].

A remarkable breakthrough on the function of both EFP and eIF5A within the ribosome has been the recent elucidation of their role to facilitate translation of mRNAs encoding proline rich-repeat proteins, which otherwise cause ribosome stalling [77–79]. These findings have led to the identification of underlying functions for spermidine, for instance, its role in yeast mating to promote translation of the polyproline-rich protein formin needed for the remodeling of the actin

cytoskeleton [80]. The high degree of structural and functional similarity between eIF5A and EFP points to an essential function on translation conserved throughout evolution, raising the hypothesis that these proteins may have co-evolved as sequence-specific elongation factors together with functionally specialized polyproline-rich proteins. To validate this hypothesis, recent studies have searched for the evolutionary distribution of genes encoding polyproline-containing proteins whose translation is expected to depend on eIF5A/EFP [81]. The authors found that polyproline motifs are poorly represented within prokaryotic proteomes, whereas their presence increases with the complexity of the eukaryotic organism, thus pointing to an evolutionary association between specific translation elongation factors and proline-rich repeat proteins. This functional association has been recently confirmed in mice comparing the polyproline-rich protein network and the phenotypic consequences after the genetic inactivation of the eIF5A hypusination pathway [68]. In spite of the evidence of the involvement of EFP and eIF5A in the synthesis of polyproline-rich proteins, recent genome-wide studies with bacterial *efp* mutants using ribosome footprint profiling [82] have uncovered that not every mRNA encoding proline-rich repeat proteins can be cataloged as a *bona fide* EFP target [83, 84]. These studies have revealed that specific mRNA features such as the strength and location of the pause and the translational efficiency may all contribute to the final protein output. Similar studies are expected to precisely clarify the role of eIF5A in the ribosome to help define the unequivocal portfolio of canonical mRNA clients. Ideally the molecular data based on technologies such as ribosome profiling and the like should be complemented with functional studies based on genetic or pharmacological approaches to inactivate the pathway as has been done in mice [68].

The available information on eIF5A in higher plants is scarce and limited to either overexpression or antisense genetic approaches in *Arabidopsis*. These experiments have involved eIF5A in both cell-death-related developmental processes, as well as in the response to challenging growth conditions [85–88], for which a biochemical description of the pathway has been recently reported [89]. Eventually, we need to elaborate a list of mRNA targets for eIF5A to elucidate eIF5A function in plants. Finally, eIF5A has also been involved in other mRNA-related functions out of the ribosome whose details are not yet clarified. One example is its involvement in nucleus-cytoplasm mRNA transport [90]. Another example of eIF5A moonlighting function in *Arabidopsis* has assigned a role to the eIF5A2 isoform in cytokinin signaling as a partner of the cytokinin receptor machinery that controls root vasculature [91].

9 Conclusions

Although most translation factors are conserved in plants, our current knowledge about translation regulation in this kingdom is still scarce. As mentioned, plants have unique eIFiso4F complexes. Thus, the existence of other plant-specific features of the translation machinary could not be ruled out. In addition, we are still far

away from deciphering the general processes of translation regulation in plants. On one hand, it is of most relevant interest to investigate whether or not eIF4E is regulated in plants by similar mechanisms to other eukaryotes, namely by the action of eIF4E-interacting proteins/partners and/or phosphorylation. On the other hand, although eIF2α regulation by phosphorylation is also conserved in plants, it seems unlikely that it could account for the drastic inhibition of translation under different stress conditions. The involvement of eIF5A in plant cell-death processes demands a detailed description of its *bona fide* mRNA targets that might have evolved specific functions. Moreover, its expected role in plant SG assembly under environmental challenges remains to be investigated.

References

1. Hernandez G, Vazquez-Pianzola P. Functional diversity of the eukaryotic translation initiation factors belonging to eIF4 families. Mech Dev. 2005;122:865–76. doi:10.1016/j.mod.2005.04.002.
2. Hernandez G, Altmann M, Lasko P. Origins and evolution of the mechanisms regulating translation initiation in eukaryotes. Trends Biochem Sci. 2010;35:63–73. doi:10.1016/j.tibs.2009.10.009.
3. Hernandez G, Proud CG, Preiss T, Parsyan A. On the diversification of the translation apparatus across eukaryotes. Comp Funct Genomics. 2012;2012:256848. doi:10.1155/2012/256848.
4. Browning KS, Bailey-Serres J. Mechanism of cytoplasmic mRNA translation. The Arabidopsis book/Am Soc Plant Biol. 2015;13:e0176. doi:10.1199/tab.0176.
5. Sonenberg N, Hinnebusch AG. Regulation of translation initiation in eukaryotes: mechanisms and biological targets. Cell. 2009;136:731–45. doi:10.1016/j.cell.2009.01.042.
6. Jackson RJ, Hellen CU, Pestova TV. The mechanism of eukaryotic translation initiation and principles of its regulation. Nat Rev Mol Cell Biol. 2010;11:113–27. doi:10.1038/nrm2838.
7. Aitken CE, Lorsch JR. A mechanistic overview of translation initiation in eukaryotes. Nat Struct Mol Biol. 2012;19:568–76. doi:10.1038/nsmb.2303.
8. Hershey JW, Sonenberg N, Mathews MB. Principles of translational control: an overview. Cold Spring Harbor perspectives in biology. 2012; 4. doi:10.1101/cshperspect.a011528.
9. Lomakin IB, Steitz TA. The initiation of mammalian protein synthesis and mRNA scanning mechanism. Nature. 2013;500:307–11. doi:10.1038/nature12355.
10. Mead EJ, Masterton RJ, von der Haar T, Tuite MF, Smales CM. Control and regulation of mRNA translation. Biochem Soc Trans. 2014;42:151–4. doi:10.1042/BST20130259.
11. Muñoz A, Castellano MM. Regulation of translation initiation under abiotic stress conditions in plants: is it a conserved or not so conserved process among eukaryotes? Comp Funct Genomics. 2012;2012:8. doi:10.1155/2012/406357.
12. Echevarria-Zomeno S, Yanguez E, Fernandez-Bautista N, Castro-Sanz AB, Ferrando A, Castellano MM. Regulation of Translation Initiation under Biotic and Abiotic Stresses. Int J Mol Sci. 14:4670–4683. doi:10.3390/ijms14034670.
13. Mayberry LK, Allen ML, Nitka KR, Campbell L, Murphy PA, Browning KS. Plant cap binding complexes eukaryotic initiation factors eIF4F and eIFiso4F: molecular specificity of subunit binding. J Biol Chem. 2011;. doi:10.1074/jbc.M111.280099.
14. Patrick RM, Browning KS. The eIF4F and eIFiso4F Complexes of plants: an evolutionary perspective. Comp Funct Genomics. 2012;2012:287814. doi:10.1155/2012/287814.
15. Mayberry LK, Allen ML, Nitka KR, Campbell L, Murphy PA, Browning KS. Plant cap-binding complexes eukaryotic initiation factors eIF4F and eIFISO4F: molecular

specificity of subunit binding. J Biol Chem. 2011;286:42566–74. doi:10.1074/jbc.M111.280099.
16. Mayberry LK, Allen ML, Dennis MD, Browning KS. Evidence for variation in the optimal translation initiation complex: plant eIF4B, eIF4F, and eIF(iso)4F differentially promote translation of mRNAs. Plant Physiol. 2009;150:1844–54. doi:10.1104/pp.109.138438.
17. Gallie DR, Browning KS. eIF4G functionally differs from eIFiso4G in promoting internal initiation, cap-independent translation, and translation of structured mRNAs. J Biol Chem. 2001;276:36951–60. doi:10.1074/jbc.M103869200.
18. Mader S, Lee H, Pause A, Sonenberg N. The translation initiation factor eIF-4E binds to a common motif shared by the translation factor eIF-4 gamma and the translational repressors 4E-binding proteins. Mol Cell Biol. 1995;15:4990–7.
19. Marcotrigiano J, Gingras AC, Sonenberg N, Burley SK. Cap-dependent translation initiation in eukaryotes is regulated by a molecular mimic of eIF4G. Mol Cell. 1999;3:707–16.
20. Altmann M, Schmitz N, Berset C, Trachsel H. A novel inhibitor of cap-dependent translation initiation in yeast: p20 competes with eIF4G for binding to eIF4E. The EMBO journal. 1997;16:1114–21. doi:10.1093/emboj/16.5.1114.
21. Matsuo H, Li H, McGuire AM, Fletcher CM, Gingras AC, Sonenberg N, Wagner G. Structure of translation factor eIF4E bound to m7GDP and interaction with 4E-binding protein. Nature structural biology. 1997;4:717–24.
22. Siddiqui N, Sonenberg N. Signalling to eIF4E in cancer. Biochem Soc Trans. 2015;43:763–72. doi:10.1042/BST20150126.
23. Clemens MJ. Translational regulation in cell stress and apoptosis. Roles of the eIF4E binding proteins. J Cell Mol Med. 2001; 5:221–239. doi:005.003.01
24. Rhoads RE. eIF4E: new family members, new binding partners, new roles. J Biol Chem. 2009;284:16711–5. doi:10.1074/jbc.R900002200.
25. Carrera AC. TOR signaling in mammals. J Cell Sci. 2004;117:4615–6. doi:10.1242/jcs.01311.
26. Kneller EL, Rakotondrafara AM, Miller WA. Cap-independent translation of plant viral RNAs. Virus Res. 2006;119:63–75. doi:10.1016/j.virusres.2005.10.010.
27. Mardanova ES, Zamchuk LA, Skulachev MV, Ravin NV. The 5′ untranslated region of the maize alcohol dehydrogenase gene contains an internal ribosome entry site. Gene. 2008;420:11–6. doi:10.1016/j.gene.2008.04.008.
28. Dinkova TD, Zepeda H, Martinez-Salas E, Martinez LM, Nieto-Sotelo J, de Jimenez ES. Cap-independent translation of maize Hsp101. Plant J. 2005;41:722–31. doi:10.1111/j.1365-313X.2005.02333.x.
29. Cui Y, Rao S, Chang B, Wang X, Zhang K, Hou X, Zhu X, Wu H, Tian Z, Zhao Z, Yang C, Huang T. AtLa1 protein initiates IRES-dependent translation of WUSCHEL mRNA and regulates the stem cell homeostasis of Arabidopsis in response to environmental hazards. Plant, Cell Environ. 2015;38:2098–114. doi:10.1111/pce.12535.
30. Deprost D, Yao L, Sormani R, Moreau M, Leterreux G, Nicolai M, Bedu M, Robaglia C, Meyer C. The Arabidopsis TOR kinase links plant growth, yield, stress resistance and mRNA translation. EMBO Rep. 2007;8:864–70. doi:10.1038/sj.embor.7401043.
31. Caldana C, Li Y, Leisse A, Zhang Y, Bartholomaeus L, Fernie AR, Willmitzer L, Giavalisco P. Systemic analysis of inducible target of rapamycin mutants reveal a general metabolic switch controlling growth in Arabidopsis thaliana. Plant J. 2013;73:897–909. doi:10.1111/tpj.12080.
32. Dobrenel T, Marchive C, Azzopardi M, Clement G, Moreau M, Sormani R, Robaglia C, Meyer C. Sugar metabolism and the plant target of rapamycin kinase: a sweet operaTOR? Front Plant Sci. 2013;4:93. doi:10.3389/fpls.2013.00093.
33. Robaglia C, Thomas M, Meyer C. Sensing nutrient and energy status by SnRK1 and TOR kinases. Curr Opin Plant Biol. 2012;15:301–7. doi:10.1016/j.pbi.2012.01.012.
34. Xiong Y, Sheen J. Rapamycin and glucose-target of rapamycin (TOR) protein signaling in plants. J Biol Chem. 2012;287:2836–42. doi:10.1074/jbc.M111.300749.
35. Richter JD, Sonenberg N. Regulation of cap-dependent translation by eIF4E inhibitory proteins. Nature. 2005;433:477–80. doi:10.1038/nature03205.

36. Vardy L, Orr-Weaver TL. Regulating translation of maternal messages: multiple repression mechanisms. Trends Cell Biol. 2007;17:547–54. doi:10.1016/j.tcb.2007.09.002.
37. Stebbins-Boaz B, Cao Q, de Moor CH, Mendez R, Richter JD. Maskin is a CPEB-associated factor that transiently interacts with elF-4E. Mol Cell. 1999;4:1017–27.
38. Richter JD. CPEB: a life in translation. Trends Biochem Sci. 2007;32:279–85. doi:10.1016/j.tibs.2007.04.004.
39. Wells DG. RNA-binding proteins: a lesson in repression. J Neurosci Official J Soc Neurosci. 2006;26:7135–8. doi:10.1523/JNEUROSCI.1795-06.2006.
40. Jung MY, Lorenz L, Richter JD. Translational control by neuroguidin, a eukaryotic initiation factor 4E and CPEB binding protein. Mol Cell Biol. 2006;26:4277–87. doi:10.1128/MCB.02470-05.
41. Nelson MR, Leidal AM, Smibert CA. Drosophila Cup is an eIF4E-binding protein that functions in Smaug-mediated translational repression. The EMBO J. 2004;23:150–9. doi:10.1038/sj.emboj.7600026.
42. Nakamura A, Sato K, Hanyu-Nakamura K. Drosophila cup is an eIF4E binding protein that associates with Bruno and regulates oskar mRNA translation in oogenesis. Dev Cell. 2004;6:69–78.
43. Napoli I, Mercaldo V, Boyl PP, Eleuteri B, Zalfa F, De Rubeis S, Di Marino D, Mohr E, Massimi M, Falconi M, Witke W, Costa-Mattioli M, Sonenberg N, Achsel T, Bagni C. The fragile X syndrome protein represses activity-dependent translation through CYFIP1, a new 4E-BP. Cell. 2008;134:1042–54. doi:10.1016/j.cell.2008.07.031.
44. Rom E, Kim HC, Gingras AC, Marcotrigiano J, Favre D, Olsen H, Burley SK, Sonenberg N. Cloning and characterization of 4EHP, a novel mammalian eIF4E-related cap-binding protein. J Biol Chem. 1998;273:13104–9.
45. Cho PF, Poulin F, Cho-Park YA, Cho-Park IB, Chicoine JD, Lasko P, Sonenberg N. A new paradigm for translational control: inhibition via 5'-3' mRNA tethering by Bicoid and the eIF4E cognate 4EHP. Cell. 2005;121:411–23. doi:10.1016/j.cell.2005.02.024.
46. Cho PF, Gamberi C, Cho-Park YA, Cho-Park IB, Lasko P, Sonenberg N. Cap-dependent translational inhibition establishes two opposing morphogen gradients in Drosophila embryos. Current Biol CB. 2006;16:2035–41. doi:10.1016/j.cub.2006.08.093.
47. Kim HS, Abbasi N, Choi SB. Bruno-like proteins modulate flowering time via 3' UTR-dependent decay of SOC1 mRNA. New Phytol. 2013;198:747–56. doi:10.1111/nph.12181.
48. Francischini CW, Quaggio RB. Molecular characterization of Arabidopsis thaliana PUF proteins–binding specificity and target candidates. The FEBS J. 2009;276:5456–70. doi:10.1111/j.1742-4658.2009.07230.x.
49. Ruud KA, Kuhlow C, Goss DJ, Browning KS. Identification and characterization of a novel cap-binding protein from Arabidopsis thaliana. J Biol Chem. 1998;273:10325–30.
50. Zhang P, McGrath BC, Reinert J, Olsen DS, Lei L, Gill S, Wek SA, Vattem KM, Wek RC, Kimball SR, Jefferson LS, Cavener DR. The GCN2 eIF2alpha kinase is required for adaptation to amino acid deprivation in mice. Mol Cell Biol. 2002;22:6681–8.
51. Harding HP, Calfon M, Urano F, Novoa I, Ron D. Transcriptional and translational control in the Mammalian unfolded protein response. Annu Rev Cell Dev Biol. 2002;18:575–99. doi:10.1146/annurev.cellbio.18.011402.160624.
52. Clemens MJ. PKR–a protein kinase regulated by double-stranded RNA. Int J Biochem Cell Biol. 1997; 29:945–949. doi:S1357-2725(96)00169-0
53. Chen JJ. Regulation of protein synthesis by the heme-regulated eIF2alpha kinase: relevance to anemias. Blood. 2007;109:2693–9. doi:10.1182/blood-2006-08-041830.
54. Zhan K, Narasimhan J, Wek RC. Differential activation of eIF2 kinases in response to cellular stresses in Schizosaccharomyces pombe. Genetics. 2004;168:1867–75. doi:10.1534/genetics.104.031443.
55. Lageix S, Lanet E, Pouch-Pelissier MN, Espagnol MC, Robaglia C, Deragon JM, Pelissier T. Arabidopsis eIF2alpha kinase GCN2 is essential for growth in stress conditions and is activated by wounding. BMC Plant Biol. 2008;8:134. doi:10.1186/1471-2229-8-134.

56. Zhang Y, Wang Y, Kanyuka K, Parry MA, Powers SJ, Halford NG. GCN2-dependent phosphorylation of eukaryotic translation initiation factor-2alpha in Arabidopsis. J Exp Bot. 2008;59:3131–41. doi:10.1093/jxb/ern169.
57. Shaikhin SM, Smailov SK, Lee AV, Kozhanov EV, Iskakov BK. Interaction of wheat germ translation initiation factor 2 with GDP and GTP. Biochimie. 1992;74:447–54.
58. Gilks N, Kedersha N, Ayodele M, Shen L, Stoecklin G, Dember LM, Anderson P. Stress Granule Assembly Is Mediated by Prion-like Aggregation of TIA-1. Mol Biol Cell. 2004;15:5383–98. doi:10.1091/mbc.E04-08-0715.
59. Juntawong P, Sorenson R, Bailey-Serres J. Cold shock protein 1 chaperones mRNAs during translation in Arabidopsis thaliana. Plant J. 2013;74:1016–28. doi:10.1111/tpj.12187.
60. Li CH, Ohn T, Ivanov P, Tisdale S, Anderson P. eIF5A promotes translation elongation, polysome disassembly and stress granule assembly. PLoS ONE. 2010;5:e9942. doi:10.1371/journal.pone.0009942.
61. Thomas A, Goumans H, Amesz H, Benne R, Voorma HO. A Comparison of the Initiation Factors of Eukaryotic Protein Synthesis from Ribosomes and from the Postribosomal Supernatant. Eur J Biochem. 1979;98:329–37. doi:10.1111/j.1432-1033.1979.tb13192.x.
62. Saini P, Eyler DE, Green R, Dever TE. Hypusine-containing protein eIF5A promotes translation elongation. Nature. 2009;459:118–21.
63. Dever TE, Gutierrez E, Shin B-S. The hypusine-containing translation factor eIF5A. Crit Rev Biochem Mol Biol. 2014;49:413–25. doi:10.3109/10409238.2014.939608.
64. Park M. The post-translational synthesis of a polyamine-derived amino acid, hypusine, in the eukaryotic translation initiation factor 5A (eIF5A). J Biochem. 2006;139:161–9.
65. Shiba T, Mizote H, Kaneko T, Nakajima T, Yasuo K, sano I. Hypusine, a new amino acid occurring in bovine brain: Isolation and structural determination. Biochim Biophys Acta (BBA)- General Subjects. 1971; 244:523–531.
66. Park MH, Wolff EC, Folk JE. Hypusine—its posttranslational formation in eukaryotic initiation factor-5A and its potential role in cellular-regulation. BioFactors. 1993;4:95–104.
67. Chattopadhyay MK, Park MH, Tabor H. Hypusine modification for growth is the major function of spermidine in Saccharomyces cerevisiae polyamine auxotrophs grown in limiting spermidine. Proc Natl Acad Sci. 2008;105:6554–9. doi:10.1073/pnas.0710970105.
68. Pällmann N, Braig M, Sievert H, Preukschas M, Hermans-Borgmeyer I, Schweizer M, Nagel CH, Neumann M, Wild P, Haralambieva E, Hagel C, Bokemeyer C, Hauber J, Balabanov S. Biological relevance and therapeutic potential of the hypusine modification system. J Biol Chem. 2015;290:18343–60. doi:10.1074/jbc.M115.664490.
69. Nishimura K, Lee S, Park J, Park M. Essential role of eIF5A-1 and deoxyhypusine synthase in mouse embryonic development. Amino Acids. 2012;42:703–10. doi:10.1007/s00726-011-0986-z.
70. Pagnussat GC, Yu HJ, Ngo QA, Rajani S, Mayalagu S, Johnson CS, Capron A, Xie LF, Ye D, Sundaresan V. Genetic and molecular identification of genes required for female gametophyte development and function in Arabidopsis. Development. 2005;132:603–14.
71. Imai A, Matsuyama T, Hanzawa Y, Akiyama T, Tamaoki M, Saji H, Shirano Y, Kato T, Hayashi H, Shibata D, Tabata S, Komeda Y, Takahashi T. Spermidine synthase genes are essential for survival of arabidopsis. Plant Physiol. 2004;135:1565–73.
72. Wolff EC, Kang KR, Kim YS, Park MH. Posttranslational synthesis of hypusine: evolutionary progression and specificity of the hypusine modification. Amino Acids. 2007;33:341–50.
73. Navarre WW, Zou SB, Roy H, Xie JL, Savchenko A, Singer A, Edvokimova E, Prost LR, Kumar R, Ibba M, Fang FC. PoxA, YjeK, and elongation factor P coordinately modulate virulence and drug resistance in salmonella enterica. Mol Cell. 2010;39:209–21.
74. Lassak J, Keilhauer EC, Fürst M, Wuichet K, Gödeke J, Starosta AL, Chen J-M, Søgaard-Andersen L, Rohr J, Wilson DN, Häussler S, Mann M, Jung K. Arginine-rhamnosylation as new strategy to activate translation elongation factor P. Nat Chem Biol. 2015;11:266–70. doi:10.1038/nchembio.1751.

75. Bullwinkle TJ, Zou SB, Rajkovic A, Hersch SJ, Elgamal S, Robinson N, Smil D, Bolshan Y, Navarre WW, Ibba M. (R)-β-Lysine-modified elongation factor P functions in translation elongation. J Biol Chem. 2013;288:4416–23. doi:10.1074/jbc.M112.438879.
76. Blaha G, Stanley RE, Steitz TA. Formation of the first peptide bond: the structure of EF-P bound to the 70S ribosome. Science. 2009;325:966–70. doi:10.1126/science.1175800.
77. Gutierrez E, Shin B-S, Woolstenhulme Christopher J, Kim J-R, Saini P, Buskirk Allen R, Dever Thomas E. eIF5A Promotes Translation of Polyproline Motifs. Mol Cell. 2013; 51:35-45. doi:10.1016/j.molcel.2013.04.021
78. Doerfel LK, Wohlgemuth I, Kothe C, Peske F, Urlaub H, Rodnina MV. EF-P Is essential for rapid synthesis of proteins containing consecutive proline residues. Science. 2013;339:85–8. doi:10.1126/science.1229017.
79. Ude S, Lassak J, Starosta AL, Kraxenberger T, Wilson DN, Jung K. Translation elongation factor EF-P alleviates ribosome stalling at polyproline stretches. Science. 2013;339:82–5. doi:10.1126/science.1228985.
80. Li T, Belda-Palazon B, Ferrando A, Alepuz P. Fertility and polarized cell growth depends on eIF5A for translation of polyproline-rich formins in saccharomyces cerevisiae. Genetics. 2014;197:1191–200. doi:10.1534/genetics.114.166926.
81. Mandal A, Mandal S, Park MH. Genome-wide analyses and functional classification of proline repeat-rich proteins: potential role of eIF5A in eukaryotic evolution. PLoS ONE. 2014;9: e111800.
82. Ingolia NT, Ghaemmaghami S, Newman JRS, Weissman JS. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science. 2009;324:218–23. doi:10.1126/science.1168978.
83. Elgamal S, Katz A, Hersch SJ, Newsom D, White P, Navarre WW, Ibba M. EF-P dependent pauses integrate proximal and distal signals during translation. PLoS Genet. 2014;10: e1004553.
84. Woolstenhulme C, Guydosh N, Green R, Buskirk A. High-precision analysis of translational pausing by ribosome profiling in bacteria lacking EFP. Cell Reports. 2015;11:13–21. doi:10.1016/j.celrep.2015.03.014.
85. Duguay J, Jamal S, Liu Z, Wang TW, Thompson JE. Leaf-specific suppression of deoxyhypusine synthase in Arabidopsis thaliana enhances growth without negative pleiotropic effects. J Plant Physiol. 2007;164:408–20.
86. Feng H, Chen Q, Feng J, Zhang J, Yang X, Zuo J. Functional characterization of the arabidopsis eukaryotic translation initiation factor 5A-2 that plays a crucial role in plant growth and development by regulating cell division, cell growth, and cell death. Plant Physiol. 2007;144:1531–45.
87. Liu Z, Duguay J, Ma F, Wang TW, Tshin R, Hopkins MT, McNamara L, Thompson JE. Modulation of eIF5A1 expression alters xylem abundance in Arabidopsis thaliana. JExpBot. 2008;59:939–50.
88. Ma F, Liu Z, Wang TW, Hopkins MT, Peterson CA, Thompson JE. Arabidopsis eIF5A3 influences growth and the response to osmotic and nutrient stress. Plant Cell Environ. 2010;33:1682–96. doi:10.1111/j.1365-3040.2010.02173.x.
89. Belda-Palazón B, Nohales MA, Rambla JL, Aceña JL, Delgado O, Fustero S, Martínez MC, Granell A, Carbonell J, Ferrando A. Biochemical quantitation of the eIF5A hypusination in Arabidopsis thaliana uncovers ABA-dependent regulation. Front Plant Sci. 2014; 5. doi:10.3389/fpls.2014.00202.
90. Zuk D, Jacobson A. A single amino acid substitution in yeast eIF-5A results in mRNA stabilization. The EMBO J. 1998;17:2914–25. doi:10.1093/emboj/17.10.2914.
91. Ren B, Chen Q, Hong S, Zhao W, Feng J, Feng H, Zuo J. The arabidopsis eukaryotic translation initiation factor eIF5A-2 regulates root protoxylem development by modulating cytokinin signaling. Plant Cell. 2013;25:3841–57. doi:10.1105/tpc.113.116236.

Ribonucleoprotein Foci in Eukaryotes: How to Translate the Silence

Layana Carla, Gonzalo H. Corujo and Rolando V. Rivera Pomar

1 Introduction

Regulation of gene expression is a crucial point to control the cellular metabolism. Protein synthesis is an expensive process in energetic terms; thus, tight regulation is essential for the proper cellular energetic balance. There are multiple regulatory points at different levels: synthesis of the RNA primary transcript (at transcription step), post-transcriptional processing of RNA, RNA decay, translation regulation, post-translation modification, and protein degradation, among others. Transcription initiation is a key aspect of gene expression regulation; this step has been studied quite extensively in different organisms for many years [1–5]. Control at this stage is very efficient because it is located at the beginning of the gene expression process. However, regulation of the already synthesized mRNA allows both rapid and local changes in the synthesis of specific proteins to occur. Over the last few years, it has been determined that the regulation of gene expression involves several processes (splicing, polyadenylation, mRNA transport, or localization) in a more complex way than expected, for instance, new mechanisms of regulation as micro RNAs (miRNAs) have various functions [6, 7]. Recent work has shown that genome transcription occurs in its complete extension, including non-coding

L. Carla (✉) · G.H. Corujo · R.V. Rivera Pomar
Centro Regional de Estudios Genómicos, Facultad Cs. Exactas,
UNLP, La Plata, Argentina
e-mail: clayana@conicet.gov.ar

L. Carla · G.H. Corujo
Universidad Nacional de Avellaneda, Buenos Aires, Argentina

R.V. Rivera Pomar
Centro de Bioinvestigaciones (CeBio) and Centro de Investigación y Transferencia
del Noroeste de Buenos Aires (CITNOBA), Universidad Nacional del NO de BA,
CIC-BA, CONICET, Buenos Aires, Argentina

G. Hernández and R. Jagus (eds.), *Evolution of the Protein Synthesis Machinery and Its Regulation*, DOI 10.1007/978-3-319-39468-8_19

regions. This process, termed pervasive transcription, is regulated by miRNAs, which promote degradation of invalid transcripts [8]. This novel approach generates new and multiple functions for the common regulator factors. In eukaryotes, non-translating mRNAs often assemble together into visible cytoplasmic membraneless structures, called mRNP granules or cytoplasmic foci. The different types of mRNP granules are processing bodies (PBs), which occurs in all cells type, stress granules (SGs), which are evident in stressed cells, germ granules (GGs), and neuronal granules (NGs), among others [9–12]. PBs are constitutive mRNP granules composed of proteins involved in translational control and mRNA degradation and thus play a role in the control of cytoplasmic mRNA. In this review, we present the discussion about PB formation, their composition, function, dynamics, and relationship with other cellular mRNP granules and with other cellular structures and processes.

2 Processing Bodies: From Cytoplasmic Foci to Regulated mRNP Granules

Processing bodies were discovered in 1997 by Bashkirov et al. They showed that the mouse ortholog of the Xrn1 exonuclease (named mXrn1) was found in discrete cytoplasmic foci; they called these structures Xrn1 foci [13]. A few years later, it was shown in yeast that the decapping enzyme, Dcp2, aggregates in discrete cytoplasmic granules termed DCP foci [14]. Contemporarily, an unknown protein (glycine-tryptophan (GW) repeat-containing protein of 182 kDa) was also observed in cytoplasmic foci using autoimmune sera from patients with sensory and motor neuropathy; the protein and the immunoreactive cytoplasmic granules were called GW182 and GW foci, respectively [15]. Later on, Seth and Parker, working in yeast, brought up the novel concept that cytoplasmic foci contain RNA decay machinery and termed them as processing bodies or P bodies (PBs) [16]. Other laboratories have validated the occurrence of PBs in different model cells and organisms (Fig. 1) [9, 17–22].

The composition of PBs in yeast and mammalian cells includes the decapping enzyme (Dcp1p/Dcp2p), activators of decapping (Edc3p, RAP55/Scd6p, Lsm1-7 complex), the 5′–3′ exonuclease Xrn1/Pcm, translation repressors (Me31B/RCK/Dhh1p, Pat1p), and components of RNA inducing silencing complex, RISC (Ago 1, Ago 2), among a large list of other regulatory factors (summarized in Table 1 and references therein).

A variety of cellular processes are known to alter PB dynamics, such as the cell cycle, embryonic stages, and circadian cycles [12, 63]. Accumulation of specific mRNA in cytoplasmic foci regulates cellular translation levels at different stages

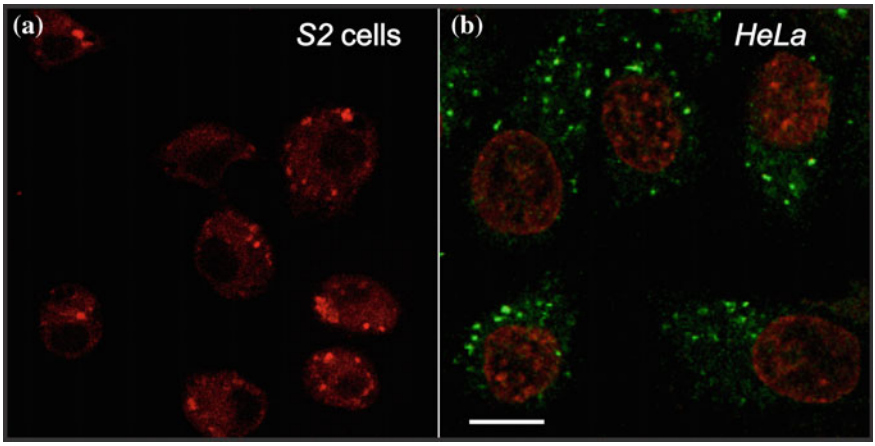


Fig. 1 Processing bodies in flies, mammals, and yeasts. **a** Transfection for Me31B in *Drosophila* Schneider S2R + cells. **b** Transfection for eIF4E in *HeLa* cells; the nucleus was stained with SYTO 17 dye. *Bars* 10 μm

[18, 58, 64]. In PBs, mRNA decay is likely to occur when mRNA degradation factors as well as the components of the RNAi machinery are concentrated [65]. Since the discovery of PBs, several studies have described them as mRNA degradation sites; however, the steps of mRNPs remodeling, leading to an inactive mRNA in PBs, are still unknown. This is mostly due to the lack of a method for the purification of PBs to allow biochemical studies. Moreover, is it still unclear whether miRNA-mediated translational inhibition is the cause or consequence of mRNAs targeted to PBs. Brengues et al. extended the role of PBs, which are not only degradation places, but also a place for reversible storage, as the mRNAs can go in/out from PBs, establishing a dynamic flow between translated mRNAs in polysomes and silenced ones in PBs [17]. This opened a more complex view of PBs with several steps of mRNP remodeling that can, perhaps, tag a specific mRNA for degradation (irreversible) or storage (reversible). An important point to consider is that the recruitment of mRNA into PBs is not only due to inhibition of translation, but the action of an active silencing mechanism is also necessary. This idea has emerged from the work of the Izaurralde group, which demonstrated that inhibition of translation with puromycin induces the PB assembly only in the presence of active RNAi or miRNA silencing pathways [66].

The proteins that compose PBs are involved in mRNA silencing and degradation; therefore, they contribute to controlling the level of translationally active mRNA in the cytoplasm. However, the diversity of the foci and their components leads to a major open question: Are they different foci with different compositions and functions, or are they intermediate forms of the same foci type?

Table 1 List of protein factors present in processing bodies

Function	Factor	Organism	Reference
Deadenylation	CCR4-NOT complex	Mammals, Sc	[23]
Decapping	DCP1/DCAP-1	Mammals, Sc, Ce, Dm	[16, 19, 22, 24]
	DCP2/DCAP-2	Mammals, Sc, Ce, Dm	[16, 19, 22, 24]
	Edc1-2	Sc	[25]
	Edc3	Mammals, Sc, Dm	[26, 27]
	Helds/Ge1	Mammals, Dm	[28–30]
	Lsm1-7	Mammals, Sc, Ce	[16, 22, 31]
miRNA	Ago	Mammals	[32]
	ALG-1	Ce	[33]
	GW182/AIN-1	Mammals, Ce, Dm	[15, 28, 33]
	TNRC6B	Mammals	[34]
mRNA decay	BRF1	Mammals	[35]
	PMR1	Mammals	[36]
	TTP	Mammals	[35]
mRNA stability	Pub1	Mammals	[37]
	Roquin	Mammals	[38]
	Vts1	Sc	[39]
	Pan2/3	Mammals	[40]
mRNA Transport	Ded1	Sc	[41]
	Gbp2	Sc	[37]
	Hrp1	Sc	[37]
	NXF2	Sc	[42]
	Staufen	Dm	[43]
NMD	SMG7	Mammals	[44]
	Upf1	Mammals, Sc	[45, 46]
	Upf2	Mammals, Sc	[45, 47]
	Upf3	Mammals, Sc	[45, 47]
Translation initiation factor	eIF3	Sc	[48]
	eIF4E/cdc33	Mammals, Sc	[23, 49, 50]
	eIF4G	Sc	[50]
Translation regulation	CPEB	Mammals	[24]
	eRF1	Sc	[37]
	eRF3	Sc	[37]
	hMex3A	Mammals	[51]
	hMex3B	Mammals	[52]
	hRNPA3	Mammals	[53]
	Ngr1	Sc	[37]
	PCBP2	Mammals	[54]
	Lin28	Mammals	[55]
	Pbp1	Sc	[37]
	hnRNPQ	Mammals	[56, 57]

(continued)

Table 1 (continued)

Function	Factor	Organism	Reference
Translation repression	eIF4ET	Mammals	[23]
	CG5208	Sc, Ce, Dm	[16, 22, 31]
	Rap55/Scd6/TraI/Dcp5/CAR-1	Mammals, Sc, Ce, Dm	[22, 36, 43, 58]
	Rck-p54/Me31B/CGH-1/Dhh1	Mammals, Sc, Ce, Dm	[16, 22, 24, 43]
Other functions	Sbp1	Sc	[59]
	Xrn1	Mammals, Sc	[35]
	Ppb4	Sc	[60]
	Rpm2	Sc	[61]
	APOBEC3G, APOBCE3F	Mammals	[62]

Summary of the most important PB components, grouped according to their biological function
Sc: *Saccharomyces cerevisiae*, Ce: *Caenorhabditis elegans*, Dm: *Drosophila melanogaster*

3 PBs and mRNA Decay

In eukaryotic cells, mRNA decay initiates with the removal of poly A tail by deadenylase. In these cells CAF1-CCR4-NOT complex is the most important deadenylase, and in yeast and mammals Pan2/Pan3 are also relevant [67]. Following this first rate-limiting step, mRNAs can take two different paths, undergoing 3′ → 5′ exonucleolytic decay or cap removal by decapping enzyme DCP2, rendering the mRNA susceptible to 5′ → 3′ digestion by Xrn1 (Fig. 2). mRNA decapping is a highly regulated process that requires the participation of several proteins. These proteins are collectively known as decapping co-activators (Dcp1, Lsm1-7 complex, rck/p54, and Pat 1), which are located in PBs with other proteins that function in the 5′ → 3′ mRNA decay pathway (Table 1 and references therein). In contrast, 3′ → 5′ exonucleolytic decay is catalyzed by the exosome, which is regulated by the SKI complex [68, 69]. Exosome components and the SKI complex are not present in PBs.

PB assembly is a crucial step for mRNA silencing and decay; in fact, translation inhibitors, such as actinomycin D, prevent PB assembly, suggesting that this process is RNA dependent. Therefore, mRNA is also required for the accumulation of mRNA degradation factors in PBs. In contrast, knockdown of the decapping enzyme Dcp2, which initiates the actual 5′ → 3′ mRNA degradation, did not abolish PB formation, indicating that its relevant functions after mRNA have been targeted to these cytoplasmic foci [16, 23]. Translation initiation competes with decapping; actually, some decapping factors can act as translation initiation inhibitors too [70, 71]. Taken together, these observations suggest that PBs could have an important role in mRNA degradation and that this process depends on the presence of mRNA destined to degradation [16, 17, 19, 23].

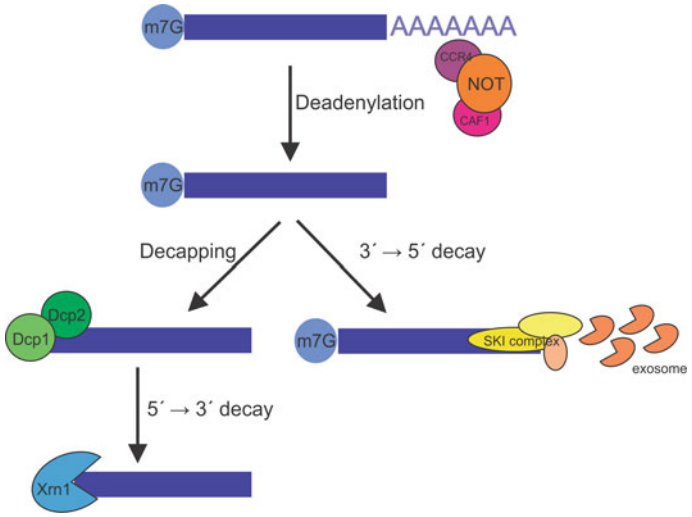


Fig. 2 mRNA decay. mRNA degradation is initiated by deadenylation. After deadenylation, mRNAs are degraded by exonucleolytic digestion at both ends. The 5′ → 3′ mRNA decay required the cleavage of the cap structure by decapping enzymes and exonucleolytic degradation by the exonuclease Xrn1. The 3′ → 5′ decay pathways involved the SKI complex and the exosome components. Most proteins involved in mRNA decay are present in PBs in eukaryotic cells

4 PBs Are Involved in mRNA Quality Control

The mRNA quality control mechanisms have evolved to ensure that only fully processed and error-free mRNAs are translated. Non-sense mediated mRNA decay (NMD) recognizes the non-sense mutation (premature stop codon) and degrades the mRNA, limiting the synthesis of truncated proteins [72, 73]. When NMD detects a premature stop codon, some factors (NMD effectors: UPF1, UPF2, and UPF3 in mammals and four additional effectors, SMG1 and SMG5-7, which do not have orthologs in yeast are recruited and assembled on mRNA, stimulating translation stop prematurely [74–76]. In general, the surveillance complex assembly leads to the recruitment of enzymes that participate in mRNA degradation (Dcp2, Xrn1, or exosome). Particularly, in yeast, the generation of this complex leads to deadenylation-independent rapid decapping; in contrast, the mRNA can be subject to accelerated deadenylation, decapping, and endonuclease activity in metazoans [68, 77–81].

In mammals, some NMD effectors (UPF2, UPF3, and SMG7) are found in PBs; however, in yeasts, NMD factors are absent from PBs, but can be found after the cells have been stimulated, for instance, when the dcp1 gene has been silenced [16, 45]. Moreover, it has also been shown that the Upf1 factor is involved in targeting the mRNA with a premature stop codon to PBs [45]. The mechanism to recruit NMD factors to PBs remains to be determined.

5 PBs and Gene Silencing

Two small RNA types regulate gene expression after transcription, siRNA and miRNA; both are important mediators of posttranscriptional gene expression regulation in higher eukaryotes. siRNAs are fully complementary to their targets and guide the Argonaute proteins to cleave the mRNA in the region that is base-paired with the siRNA. Following this endonucleolytic cleavage, the resulting mRNA fragments are captured by the general mRNA-decay machinery [82]. In contrast, most animal miRNAs are only partially complementary to their targets, so silence gene expression is mediated by repressing translation and/or by promoting mRNA decay (Fig. 3). miRNAs function as ribonucleoprotein complexes, miRISCs (miRNA-induced silencing complexes) [83]. Although their signaling pathways are different, miRNA and RNAi are executed by a member of the Argonaute family proteins and the miRISCs.

The key components of miRISCs are proteins from the Argonaute family. These factors contain three conserved domains, which interact with the 3′ and 5′ ends of the miRNA. GW182 proteins are another group of factors, crucial for the

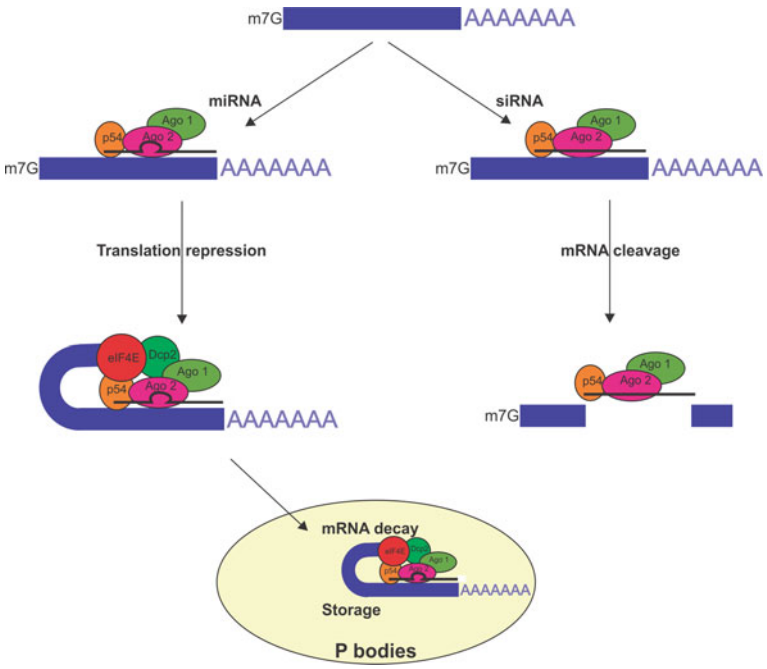


Fig. 3 PBs and gene silencing. Small RNA required for the silencing process. miRNAs are processed from highly structured genome-encoded transcripts; in animals they are partially complementary to their targets. In contrast, siRNAs are originated by the processing of long double-stranded RNA molecules and are fully complementary to their targets. Argonaute proteins cleave mRNA in both mechanisms; with the other factors necessary to silencing, they are found in PBs

miRNA-induced repression [84]. They interact directly with AGOs, and they also act downstream. GW182 proteins contain GW repeats in the N terminal portion followed by a glutamine (Q-rich region). These GW repeats are responsible for the interaction with the AGO proteins, and the Q-rich region is responsible for targeting to PBs [15, 85].

All the proteins involved in the decay of small RNA targets localize to P bodies, which strongly indicates a role of these bodies in RNA silencing. Given that target mRNA and RISC components have been co-localized in P-bodies [65, 86], it is possible that PBs are the bona fide sites for RISC-induced target cleavage or repression of translation. However, the mechanism to target mRNA and miRISC to PBs and how translation is repressed by miRISC are still unknown.

6 Do P Bodies Have a Role in Delivering mRNA and Proteins to Cellular Compartments?

Several lines of evidence suggest that a part of the PBs is associated to different cellular organelles. In *Drosophila* oocytes a subset of PBs were observed associated with the endoplasmic reticulum (ER) exit sites [87]. The same association was demonstrated in yeast by electron microscopy and density gradients studies [88]. In mammals, PBs have also recently been described to associate with mitochondria; however, the function of this link is not fully understood [89]. The significance of the association of PBs with various cellular organelles might imply an active participation in mRNA transport. According to this, PBs could either have a role in delivering protein factors to various cellular compartments or may play a role in the regulated translation of specific mRNAs that encode (specific, abundant, or necessary) factors required in those organelles.

In order to play a role in intracellular transport, PBs must associate with cytoskeleton factors. An update survey indicates that PBs are associated with actin bundles and microtubules [35]. Furthermore, some cytoskeleton factors are present in PBs, for instance, Myo 2 was visualized in PBs in yeast [90], and the microtubule motor protein dynein was observed in mammal cells under stress conditions [91].

Although these results clearly establish a link among mRNA transport, PBs, and cellular organelles, further research and more sophisticated methods are necessary to obtain an accurate conclusion.

7 Diversity Among Different mRNP Granule Types

The cytoplasmic foci are dynamic structures whose function and protein composition can vary. The diversity of mRNP granules depends on the different type of cells and the cellular context. For example, in neurons, the neuronal granules have a

transport function along dendrites and axons [10, 92]. These granules often contain RNA-binding proteins such as Staufen and in some cases can attach to ribosomes [43, 93, 94]. In germ cells, in contrast, the cytoplasmic mRNP foci are implicated in the storage and localization of mRNA, and they usually contain the RNA helicase Vasa as a common factor in different organisms [95]. Germ granules share components with PBs of somatic cells, and also contain proteins and RNAs uniquely required for germ cell development, but germ granules never attach to ribosomes [12]. Different subtypes of these granules have been described: P-granules, nuage granules, chromatoid bodies, sponge bodies, and mitochondrial cement [12]. Their sub-classification comprises the cellular localization, developmental stage, composition, and organism studied.

The PB dynamics is unclear; the number, size, and composition change in response to different cellular processes or different cell cycle phases. So, in dendritic cells, the dendritic stimulation generates variation in PB composition. In the same manner, during development and germline, maternal mRNA can be stored as translationally repressed mRNPs in *C. elegans* and *Drosophila* [18, 96], changing the number and PB composition along the different stages. One interesting example is the protein Dhh1, a decapping activator in yeast, whose orthologs in *C. elegans* (CGH-1), *Xenopus* (p54), and *Drosophila* (Me31B) are involved in the translational regulation of maternal mRNA [97–100].

Other evidence for the diversity of cytoplasmic foci and their components results from immunohistochemistry and co-localization studies [20, 35, 101]. A common factor present in most cytoplasmic mRNPs is the cap-binding protein eIF4E, the canonical initiation factor in eukaryotes. In active polysomes, eIF4E acts as a translation initiation factor, in SGs as part of the stalled initiation complex, and in PBs as a target of RNA-binding protein repressors. It is the unique translation factor that can be found in all PBs of eukaryotes. Co-localization studies in *Drosophila* S2

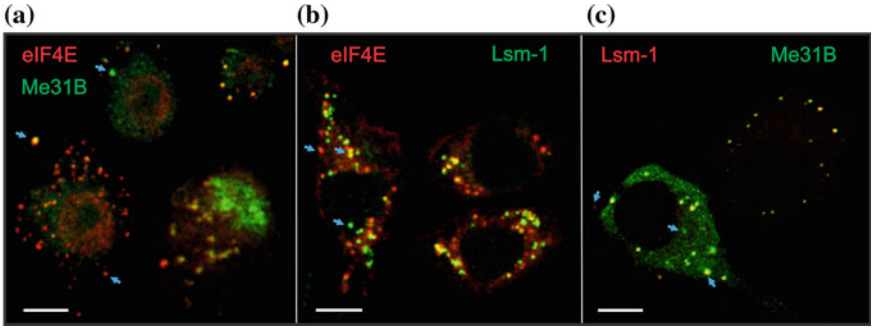


Fig. 4 Heterogeneous mRNA granules in *Drosophila* S2 and *HeLa* cells. **a** Transfection for eIF4E-YFP (*red*) and Me31B-CFP (*green*) in *Drosophila* cells. In some cells double-stained foci are more frequent than single-stained foci, but in other cells the opposite occurs. **b** and **c** Transfection of *HeLa* cells shows the same heterogeneity in granules; double-stained foci are highly frequent (**b** eIF4E-YFP: *red* and Lsm-1-CFP: *green*; **c** Lsm-1-YFP: *red* and Me31B-CFP: *green*). *Blue arrows* show single- and double-stained foci

cells and mammalian *HeLa* cells revealed the overlap of eIF4E with different factors in PBs (Fig. 4). Notably, the co-localization in PBs is not observed in all foci. Some granules contain only eIF4E, others contain only the factor, and in other cases both of them co-localize. This, which has been overlooked in many studies, indicates that the foci are not a homogeneous group. In the same line, a closer examination of PBs in mammalian neurons described specific granules that they have such as DCP2 and GW182, but Ago2 and rck/p54 are not always present [20, 102]. Dendritic foci that contain Lsm1 and exclude Dcp1a were also described [103].

In this way, the heterogeneity of granules may be indicative of a maturation process where distinct factors are recruited progressively. The foci are motile; they may come into close contact and even dock with each other, thus generating a way to exchange or incorporate different molecules [104]. As we mentioned above, the in vivo methods still lack the complement of reliable PB purification methods, and other foci/intermediates have prevented the detailed biochemical studies required to understand mRNA-protein complexes.

Similarly, the mRNP foci are highly dynamic and constantly exchange RNA and proteins with the cytosol, the polysomes, or other foci. Fluorescent Recovery After Photobleaching (FRAP) analysis of several protein components reveals a wide range of exchange rates; however, some protein components are almost static, like DCP2 in PBs [105]. Work from numerous laboratories yielded this idea about the dynamic exchange of components between different foci [20, 35, 106, 107]. In certain conditions, mammalian SGs and PBs may fuse, forming a hybrid granule containing components of both structures [20]. SG-PB interactions analyzed in real time using time-lapse microscopy showed that the interactions are variable. In some of them, PBs stably bind to SGs; in others, PBs appear intermittently attached to SGs or free PBs move freely in the cytoplasm [35].

In conclusion, SGs and PBs are diverse in composition and morphology, which indicates that the granule dynamics can generate intermediate forms, which can lose or acquire different components. However, one cannot rule out that they represent different foci committed to the regulation of specific mRNAs, including the interconversion of the different foci. The understanding of the dynamics of cytoplasmic foci is far from clear, and unpredictable paths remain to be discovered.

7.1 Drosophila Melanogaster *mRNPs in the Germline*

In *Drosophila*, several cytoplasmic granules have been recently described. Normally, *Drosophila S2* culture cells contain the classical PBs described previously, which comprise decapping enzyme complex Dcp1-Dcp2, decapping activator Lsm 1-7, Me31B, and many other factors involved in regulation (Table 1). They are sensitive to cycloheximide treatment [66, 108] and function in the silencing, storage, and decay of unnecessary or abortive mRNAs in normal conditions of the cell.

Similarly, in *Drosophila* nurse cells there are germ granules, which share many fundamental components with PBs; the oocytes, though distinct from previously characterized PBs, still acquire the full complement of PB components upon deposition of maternal dDcp1 in early stage embryos. This suggests that a developmentally regulated conversion occurs between maternal RNA granules and PBs from the oogenesis to embryogenesis transition. The presence of common components among different RNA granules implies that the interchange of components among them might be possible.

Another type of regulatory granule in *Drosophila* is the polar granule; these are present at the posterior pole of the early *Drosophila* embryo. They are electron-dense structures enriched in RNA and ribosomes [109]. After egg fertilization, maternally synthesized RNAs are protected from degradation in the germplasm. This implies that polar granules store RNA in a translationally repressed state until needed for primordial germ cell development [110]. After germ cell formation, most RNA is lost from polar granules [111]. The composition of polar granules (Me31B, eIF4A, Aubergine, and Transitional Endoplasmic Reticulum 94, and others) suggests their involvement in post-transcriptional regulation at two levels: mRNA recruitment and, through miRNA, silencing pathways. The RNA-rich nature of early polar granules supports the idea that germline-specific mRNAs are stored in polar granules in a translationally repressed state. Subsequently, these RNAs are translated, and their function may be required for germ cell formation and further development.

8 Oogenesis

Drosophila oogenesis provides a very good model system to study translational regulation of localized mRNA. The *Drosophila* egg chamber is made up of 16 germline cells surrounded by a layer of somatic follicle cells [112]. The 16 cells are produced from a single germline cell by four consecutive mitotic divisions. As cytokinesis is not complete during these divisions, the cells stay connected to each other through specialized cytoplasmic bridges named ring canals. The oocyte is determined from 1 of 16 germline cells, while the remaining 15 cells undergo differentiation into nurse cells. These last cells generate large quantities of RNAs and proteins, which are delivered to the developing oocyte through the ring canals. The mRNAs encoding determinants for embryonic polarity such as *bicoid* (*bcd*), *oskar* (*osk*), and *nanos* (*nos*) are synthesized in nurse cells, transported to the oocyte, and localized within the oocyte during oogenesis [113]. So, translation of these mRNAs is silenced during their localization and is activated when and where the protein is required [113, 114]. Although the translational repression of non-localized mRNAs in oocytes and embryos has been analyzed in depth, little is known as to how these mRNAs are silenced during their transport from nurse cells to the oocyte. During early oogenesis, the loss of Me31B causes premature translation of at least two mRNAs in the particles; the fact that *osk* and *Bicaudal-D* (*BicD)* mRNAs are

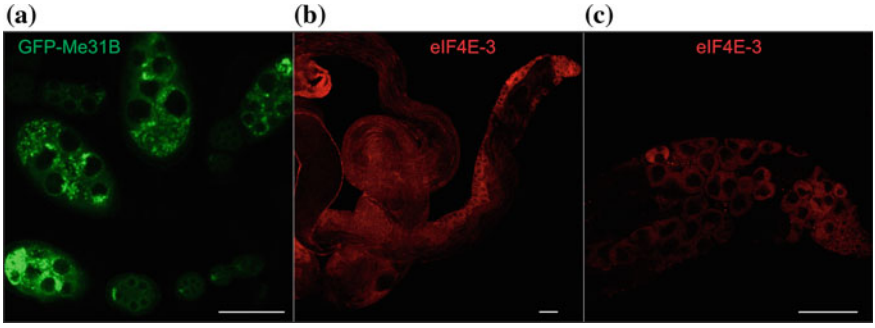


Fig. 5 **a** Confocal microscopy image of *Drosophila* egg chamber. Ovaries expressing GFP-Me31B (Bloomington flies: stock number 51530, genotype Y1 W*; P {PTT-GB} me31B^{CB05282}); endogenous Me31B also distributed in a granular pattern in the cytoplasm of nurse cells and oocytes. **b**, **c** Distribution of eIF4E-3 in the testes. Whole-mount immunostaining of adult testes using antibodies that recognize specifically eIF4E-3. *Bars* 50 μm

prematurely translated in nurse cells implies that Me31B plays an essential role in translational silencing of oocyte-localizing mRNAs during their transport to the oocyte [99]. Figure 5a shows Me31B localization in *Drosophila melanogaster* ovaries. GFP-Me31B signals show cytoplasmic particles in both nurse cells and the oocyte.

Translation can be modulated by affecting the recruitment of mRNAs to the ribosome, which involves recognition of the 5′ cap structure by the cap-binding protein eIF4E and recruits eIF4G, which interacts with eIF4A and PABP. *Drosophila* has several genes encoding eIF4E-related proteins, but the biological role of most of them remains unknown. The eIF4E-binding protein 4E-BP competitively inhibits the interaction of eIF4G with eIF4E, thus acting as a translation repressor. Interestingly, both eIF4G and 4E-BP could not be found in the granules. In oocytes the crucial role of eIF4E-1 during development was demonstrated by the regulation of protein expression with one eIF4E-related protein called CUP. eIF4E-1/CUP complex silences the *oskar* mRNA during oogenesis [115]. In the same way, the embryonic anterior-posterior (A-P) axis is formed through localization of *bcd* mRNA at the anterior as well as *nos* mRNAs to the posterior area of the oocyte. Bcd and Nos proteins regulate the translation of other morphogens, *caudal* and *hunchback*, respectively [114]. *Bicoid* silencing *caudal* mRNA at the embryo anterior pole is generated by the union to 3′UTR. d4E-HP binds to the cad mRNA 5′cap structure, while Bcd binds to BBR in its 3′UTR [115]. The interaction between d4E-HP and Bcd is mediated through a sequence motif in Bcd that resembles, but is distinct from, the consensus eIF4E binding domain present in classical eIF4E binding proteins such as 4E-BPs and eIF4G [116]. Similarly, d4E-HP binds to the hunchback 3′UTR mRNA region by interacting with other proteins such as Nanos. This generates the translation inhibition in the posterior pole [116].

9 Spermatogenesis

Spermatogenesis implies an intricate developmental cycle that generates changes in the cell cycle dynamics as well as severe cellular remodeling. Translational control is crucial for spermatogenesis in *Drosophila* as many mRNAs synthesized in the spermatocytes are translated only much later during spermatid differentiation [117]. In *Drosophila* testes, the sequential stages of spermatogenesis are disposed in a linear way. The apical region of the testes contains the 'hub' cells, which act as a niche that maintains the populations of the germline stem cell (GSC) and the somatic cyst progenitor cell (CPC). The GSC undergoes mitotic division to produce a spermatogonium, which is encapsulated by two cyst cells to form a cyst. Then, the spermatogonium divides mitotically four times with incomplete cytokinesis to generate 16 spermatocytes; two meiotic divisions follow, and 64 haploid onion stage spermatids are produced. Spermatogenesis involves a dramatic cellular transformation event that includes the formation of the elongated flagellar axoneme structure, nuclear shaping and condensation, and individualization to generate the mature sperm with a needle-like nucleus [117].

Similarly to oogenesis, a mechanism of a translational silencing process of messengers is necessary to assure the timely expression of localized mRNA during spermatogenesis. Recently, it was reported that *Drosophila* spermatogenesis requires specific initiation factors that differ from the canonical initiators eIF4E-1 and eIF4G [118, 119]. eIF4E-3 is required particularly during spermatogenesis. Males lacking eIF4E-3 are sterile, showing defects in meiotic chromosome segregation, cytokinesis, nuclear shaping, and individualization [118]. eIF4E-3 has been found in PBs in S2 cells [108]. Figure 5b, c shows the localization of eIF4E-3 in the *Drosophila* testes. eIF4E-3 is cytoplasmic and expressed in spermatocytes, round spermatids, and early differentiating spermatids.

The eIF4E-1 function in early germ cells and the surrounding somatic cells is critical for spermatogenesis. Both eIF4E-1 and eIF4E-3 are required in spermatocytes for chromosome condensation and cytokinesis during the meiotic stages [118]. Interestingly, eIF4G knockdown did not affect male fertility, whereas eIF4G2 has distinct functions during spermatogenesis. It is required in early germ cells for proper meiotic divisions and spermatid elongation, while its abrogation in spermatocytes caused meiotic arrest. In addition, double knockdown of eIF4G and eIF4G2 shows that these proteins act redundantly during the early stages of spermatogenesis [119]. This evidence supports the idea of the existence of specialized eIF4F complexes, which include eIF4E-3- and eIF4E-related proteins, specific for spermatocytes and post-meiotic spermatids, as occurs in the female germline. The occurrence of eIF4E-3 in PBs during spermatogenesis is an interesting hypothesis that remains to be tested.

10 Perspectives and Open Questions

Taken together, the above observations suggest that translation initiation is in dynamic equilibrium with an active process of translational silencing. RNA granules harbor translationally silenced mRNA. Although the content of each type of RNA granule is distinct, many proteins are found in more than one type of granule. Little is known about how RNA and proteins interact at the molecular level to create these dynamic cytoplasmic domains. Different models for PB assembly have been proposed [11, 95, 106, 107, 120]. They imply that silenced mRNPs are aggregated by specific dimerization or oligomerization domains, which direct the formation of distinct macromolecular complexes, likely corresponding to distinct silencing pathways. Therefore, interactions between different protein molecules present in separate silenced mRNPs may aggregate in different foci. This is in accordance with the evidence of specific eIF4F complexes in the *Drosophila* male and female germline that regulate translation initiation and silencing pathways in these cell lines. A tempting hypothesis to be challenged is whether foci loaded with different proteins correspond to different silencing pathways. Although currently classical biochemistry is not possible, the terrific development of in vivo cellular analysis by fluorescent microscopy and quantitative image analysis will allow addressing these questions in the future.

Here we propose a hypothesis for a dynamic switch among translational activation, translational repression, and degradation of mRNAs (Fig. 6). In this model the recruitment of mRNA from active polysomes to PBs implies the removal of the translation factors by translational repressors, which destabilize active polysomes and recruit PBs components, some of which interact with the translation initiation factor (eIF4E). All the intermediate steps of this process can represent different populations of granules coexisting in the cell, visible with different morphologies that might reflect a diversity of components. The exchange of protein components and mRNA with other foci, such as stress granules, might be possible. Likewise, evidence of connections with other cellular pathways and compartments is emerging.

Although considerable advances have been made in recent years, there are several key questions that need to be addressed. Obtaining knowledge of the complete protein and RNA composition of mRNP foci in their various cellular contexts will be a challenging goal for understanding the control of gene expression. It will be crucial to understand which mRNP remodeling events generate the selective silencing of an mRNA and not of others. We expect that the combination of biochemical, biophysical, and cell biological approaches will provide answers and surely exciting new questions.

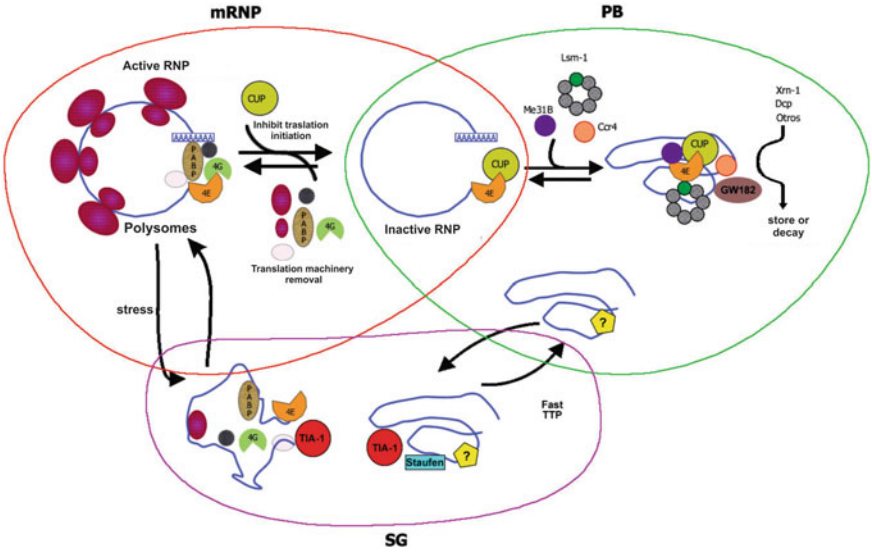


Fig. 6 Model of mRNP granule assembly and disassembly. The recruitment of mRNA from active polysomes to PB implies the removal of the translation factors by translational repressors. A putative mechanism proposes that one protein factor, probably CUP, destabilizes active polysomes and recruits PB components, some of which interact with eIF4E. All the intermediate steps of this process can represent different populations of granules coexisting in the cell and visible with different morphologies that might reflect a variety of components. Besides, stress conditions generate SG assembly in the cell. The exchange of protein components and mRNA between these foci might be possible

References

1. Smale ST, Baltimore D. The "initiator" as a transcription control element. Cell. 1989; 57:103–113. doi:http://dx.doi.org/10.1016/0092-8674(89)90176-1.
2. Weis L, Reinberg D. Accurate positioning of RNA polymerase II on a natural TATA-less promoter is independent of TATA-binding-protein-associated factors and initiator-binding proteins. Mol Cell Biol. 1997;17:2973–84. doi:10.1128/mcb.17.6.2973.
3. Struhl K. Fundamentally different logic of gene regulation in eukaryotes and prokaryotes. Cell. 1999;98:1–4. doi:10.1016/S0092-8674(00)80599-1.
4. Butler JE, Kadonaga JT. The RNA polymerase II core promoter: a key component in the regulation of gene expression. Genes Dev. 2002;16:2583–92. doi:10.1101/gad.1026202.
5. Juven-Gershon T, Kadonaga JT. Regulation of gene expression via the core promoter and the basal transcriptional machinery. Dev Biol. 2010;339:225–9. doi:10.1016/j.ydbio.2009.08.009.
6. Filipowicz W, Bhattacharyya SN, Sonenberg N. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? Nat Rev Genet. 2008;9:102–14. doi:10.1038/nrg2290.
7. Jackson RJ, Standart N. How Do MicroRNAs Regulate Gene Expression? Sci STKE. 2007; 2007:re1-. doi:10.1126/stke.3672007re1.
8. Kapranov P, Cheng J, Dike S, Nix DA, Duttagupta R, Willingham AT, Stadler PF, Hertel J, Hackermuller J, Hofacker IL, Bell I, Cheung E, Drenkow J, Dumais E, Patel S, Helt G,

Ganesh M, Ghosh S, Piccolboni A, Sementchenko V, Tammana H, Gingeras TR. RNA maps reveal new RNA classes and a possible function for pervasive transcription. Science. 2007;316:1484–8. doi:10.1126/science.1138341.
9. Eulalio A, Behm-Ansmant I, Izaurralde E. P bodies: at the crossroads of post-transcriptional pathways. Nat Rev Mol Cell Biol. 2007;8:9–22. doi:10.1038/nrm2080.
10. Hillebrand J, Barbee SA, Ramaswami M. P-body components, microRNA regulation, and synaptic plasticity. Sci World J. 2007;7:178–90. doi:10.1100/tsw.2007.206.
11. Anderson P, Kedersha N. Stress granules: the Tao of RNA triage. Trends Biochem Sci. 2008;33:141–50. doi:10.1016/j.tibs.2007.12.003.
12. Voronina E, Seydoux G, Sassone-Corsi P, Nagamori I. RNA granules in germ cells. Cold Spring Harb perspectives in biology. 2011; 3. doi:10.1101/cshperspect.a002774.
13. Bashkirov VI, Scherthan H, Solinger JA, Buerstedde JM, Heyer WD. A mouse cytoplasmic exoribonuclease (mXRN1p) with preference for G4 tetraplex substrates. J Cell Biol. 1997;136:761–73.
14. van Dijk E, Cougot N, Meyer S, Babajko S, Wahle E, Seraphin B. Human Dcp2: a catalytically active mRNA decapping enzyme located in specific cytoplasmic structures. EMBO J. 2002;21:6915–24.
15. Eystathioy T, Jakymiw A, Chan EK, Seraphin B, Cougot N, Fritzler MJ. The GW182 protein colocalizes with mRNA degradation associated proteins hDcp1 and hLSm4 in cytoplasmic GW bodies. RNA. 2003;9:1171–3. doi:10.1261/rna.5810203.
16. Sheth U, Parker R. Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. Science. 2003;300:805–8. doi:10.1126/science.1082320.
17. Brengues M, Teixeira D, Parker R. Movement of eukaryotic mRNAs between polysomes and cytoplasmic processing bodies. Science. 2005;310:486–9. doi:10.1126/science.1115791.
18. Anderson P, Kedersha N. RNA granules. J Cell Biol. 2006;172:803–8. doi:10.1083/jcb.200512082.
19. Cougot N, Babajko S, Seraphin B. Cytoplasmic foci are sites of mRNA decay in human cells. J Cell Biol. 2004;165:31–40. doi:10.1083/jcb.200309008.
20. Thomas MG, Loschi M, Desbats MA, Boccaccio GL. RNA granules: the good, the bad and the ugly. Cell Signal. 2011;23:324–34. doi:10.1016/j.cellsig.2010.08.011.
21. Cassola A. RNA granules living a post-transcriptional life: the trypanosomes' case. Curr Chem Biol. 2011;5:108–17.
22. Gallo CM, Munro E, Rasoloson D, Merritt C, Seydoux G. Processing bodies and germ granules are distinct RNA granules that interact in C. elegans embryos. Dev Biol. 2008;323:76–87. doi:10.1016/j.ydbio.2008.07.008.
23. Andrei MA, Ingelfinger D, Heintzmann R, Achsel T, Rivera-Pomar R, Luhrmann R. A role for eIF4E and eIF4E-transporter in targeting mRNPs to mammalian processing bodies. RNA. 2005;11:717–27. doi:10.1261/rna.2340405.
24. Wilczynska A, Aigueperse C, Kress M, Dautry F, Weil D. The translational regulator CPEB1 provides a link between dcp1 bodies and stress granules. J Cell Sci. 2005;118:981–92. doi:10.1242/jcs.01692.
25. Neef DW, Thiele DJ. Enhancer of decapping proteins 1 and 2 are important for translation during heat stress in Saccharomyces cerevisiae. Mol Microbiol. 2009;73:1032–42. doi:10.1111/j.1365-2958.2009.06827.x.
26. Kshirsagar M, Parker R. Identification of Edc3p as an enhancer of mRNA decapping in Saccharomyces cerevisiae. Genetics. 2004;166:729–39.
27. Tritschler F, Eulalio A, Truffault V, Hartmann MD, Helms S, Schmidt S, Coles M, Izaurralde E, Weichenrieder O. A divergent Sm fold in EDC3 proteins mediates DCP1 binding and P-body targeting. Mol Cell Biol. 2007;27:8600–11. doi:10.1128/MCB.01506-07.
28. Behm-Ansmant I, Rehwinkel J, Izaurralde E. MicroRNAs silence gene expression by repressing protein expression and/or by promoting mRNA decay. Cold Spring Harb Symp Quant Biol. 2006;71:523–30. doi:10.1101/sqb.2006.71.013.

29. Fenger-Gron M, Fillman C, Norrild B, Lykke-Andersen J. Multiple processing body factors and the ARE binding protein TTP activate mRNA decapping. Mol Cell. 2005;20:905–15. doi:10.1016/j.molcel.2005.10.031.
30. Yu JH, Yang WH, Gulick T, Bloch KD, Bloch DB. Ge-1 is a central component of the mammalian cytoplasmic mRNA processing body. RNA. 2005;11:1795–802. doi:10.1261/rna.2142405.
31. Scheller N, Resa-Infante P, de la Luna S, Galao RP, Albrecht M, Kaestner L, Lipp P, Lengauer T, Meyerhans A, Diez J. Identification of PatL1, a human homolog to yeast P body component Pat1. Biochim Biophys Acta. 2007;1773:1786–92. doi:10.1016/j.bbamcr.2007.08.009.
32. Leung AK, Calabrese JM, Sharp PA. Quantitative analysis of Argonaute protein reveals microRNA-dependent localization to stress granules. Proc Natl Acad Sci U S A. 2006;103:18125–30. doi:10.1073/pnas.0608845103.
33. Ding L, Spencer A, Morita K, Han M. The developmental timing regulator AIN-1 interacts with miRISCs and may target the argonaute protein ALG-1 to cytoplasmic P bodies in C. elegans. Mol Cell. 2005;19:437–47. doi:10.1016/j.molcel.2005.07.013.
34. Meister G, Landthaler M, Peters L, Chen PY, Urlaub H, Luhrmann R, Tuschl T. Identification of novel argonaute-associated proteins. Curr Biol CB. 2005;15:2149–55. doi:10.1016/j.cub.2005.10.048.
35. Kedersha N, Stoecklin G, Ayodele M, Yacono P, Lykke-Andersen J, Fritzler MJ, Scheuner D, Kaufman RJ, Golan DE, Anderson P. Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. J Cell Biol. 2005;169:871–84. doi:10.1083/jcb.200502088.
36. Yang WH, Yu JH, Gulick T, Bloch KD, Bloch DB. RNA-associated protein 55 (RAP55) localizes to mRNA processing bodies and stress granules. RNA. 2006;12:547–54. doi:10.1261/rna.2302706.
37. Buchan JR, Muhlrad D, Parker R. P bodies promote stress granule assembly in Saccharomyces cerevisiae. J Cell Biol. 2008;183:441–55. doi:10.1083/jcb.200807043.
38. Yu D, Tan AH, Hu X, Athanasopoulos V, Simpson N, Silva DG, Hutloff A, Giles KM, Leedman PJ, Lam KP, Goodnow CC, Vinuesa CG. Roquin represses autoimmunity by limiting inducible T-cell co-stimulator messenger RNA. Nature. 2007;450:299–303. doi:10.1038/nature06253.
39. Rendl LM, Bieman MA, Smibert CAS. Cerevisiae Vts1p induces deadenylation-dependent transcript degradation and interacts with the Ccr4p-Pop2p-Not deadenylase complex. RNA. 2008;14:1328–36. doi:10.1261/rna.955508.
40. Zheng D, Ezzeddine N, Chen CY, Zhu W, He X, Shyu AB. Deadenylation is prerequisite for P-body formation and mRNA decay in mammalian cells. J Cell Biol. 2008;182:89–101. doi:10.1083/jcb.200801196.
41. Beckham C, Hilliker A, Cziko AM, Noueiry A, Ramaswami M, Parker R. The DEAD-box RNA helicase Ded1p affects and accumulates in Saccharomyces cerevisiae P-bodies. Mol Biol Cell. 2008;19:984–93. doi:10.1091/mbc.E07-09-0954.
42. Takano K, Miki T, Katahira J, Yoneda Y. NXF2 is involved in cytoplasmic mRNA dynamics through interactions with motor proteins. Nucleic Acids Res. 2007;35:2513–21. doi:10.1093/nar/gkm125.
43. Barbee SA, Estes PS, Cziko AM, Hillebrand J, Luedeman RA, Coller JM, Johnson N, Howlett IC, Geng C, Ueda R, Brand AH, Newbury SF, Wilhelm JE, Levine RB, Nakamura A, Parker R, Ramaswami M. Staufen- and FMRP-containing neuronal RNPs are structurally and functionally related to somatic P bodies. Neuron. 2006;52:997–1009. doi:10.1016/j.neuron.2006.10.028.
44. Unterholzner L, Izaurralde E. SMG7 acts as a molecular link between mRNA surveillance and mRNA decay. Mol Cell. 2004;16:587–96. doi:10.1016/j.molcel.2004.10.013.
45. Sheth U, Parker R. Targeting of aberrant mRNAs to cytoplasmic processing bodies. Cell. 2006;125:1095–109. doi:10.1016/j.cell.2006.04.037.

46. Durand S, Cougot N, Mahuteau-Betzer F, Nguyen CH, Grierson DS, Bertrand E, Tazi J, Lejeune F. Inhibition of nonsense-mediated mRNA decay (NMD) by a new chemical molecule reveals the dynamic of NMD factors in P-bodies. J Cell Biol. 2007;178:1145–60. doi:10.1083/jcb.200611086.
47. Stalder L, Muhlemann O. Processing bodies are not required for mammalian nonsense-mediated mRNA decay. RNA. 2009;15:1265–73. doi:10.1261/rna.1672509.
48. Grousl T, Ivanov P, Frydlova I, Vasicova P, Janda F, Vojtova J, Malinska K, Malcova I, Novakova L, Janoskova D, Valasek L, Hasek J. Robust heat shock induces eIF2alpha-phosphorylation-independent assembly of stress granules containing eIF3 and 40S ribosomal subunits in budding yeast, Saccharomyces cerevisiae. J Cell Sci. 2009;122:2078–88. doi:10.1242/jcs.045104.
49. Ferraiuolo MA, Basak S, Dostie J, Murray EL, Schoenberg DR, Sonenberg N. A role for the eIF4E-binding protein 4E-T in P-body formation and mRNA decay. J Cell Biol. 2005;170:913–24. doi:10.1083/jcb.200504039.
50. Hoyle NP, Castelli LM, Campbell SG, Holmes LE, Ashe MP. Stress-dependent relocalization of translationally primed mRNPs to cytoplasmic granules that are kinetically and spatially distinct from P-bodies. J Cell Biol. 2007;179:65–74. doi:10.1083/jcb.200707010.
51. Buchet-Poyau K, Courchet J, Le Hir H, Seraphin B, Scoazec JY, Duret L, Domon-Dell C, Freund JN, Billaud M. Identification and characterization of human Mex-3 proteins, a novel family of evolutionarily conserved RNA-binding proteins differentially localized to processing bodies. Nucleic Acids Res. 2007;35:1289–300. doi:10.1093/nar/gkm016.
52. Courchet J, Buchet-Poyau K, Potemski A, Bres A, Jariel-Encontre I, Billaud M. Interaction with 14-3-3 adaptors regulates the sorting of hMex-3B RNA-binding protein to distinct classes of RNA granules. J Biol Chem. 2008;283:32131–42. doi:10.1074/jbc.M802927200.
53. Katahira J, Miki T, Takano K, Maruhashi M, Uchikawa M, Tachibana T, Yoneda Y. Nuclear RNA export factor 7 is localized in processing bodies and neuronal RNA granules through interactions with shuttling hnRNPs. Nucleic Acids Res. 2008;36:616–28. doi:10.1093/nar/gkm556.
54. Fujimura K, Kano F, Murata M. Identification of PCBP2, a facilitator of IRES-mediated translation, as a novel constituent of stress granules and processing bodies. RNA. 2008;14:425–31. doi:10.1261/rna.780708.
55. Balzer E, Moss EG. Localization of the developmental timing regulator Lin28 to mRNP complexes, P-bodies and stress granules. RNA Biol. 2007;4:16–25.
56. Tadauchi T, Inada T, Matsumoto K, Irie K. Posttranscriptional regulation of HO expression by the Mkt1-Pbp1 complex. Mol Cell Biol. 2004;24:3670–81.
57. Quaresma AJ, Bressan GC, Gava LM, Lanza DC, Ramos CH, Kobarg J. Human hnRNP Q re-localizes to cytoplasmic granules upon PMA, thapsigargin, arsenite and heat-shock treatments. Exp Cell Res. 2009;315:968–80. doi:10.1016/j.yexcr.2009.01.012.
58. Teixeira D, Sheth U, Valencia-Sanchez MA, Brengues M, Parker R. Processing bodies require RNA for assembly and contain nontranslating mRNAs. RNA. 2005;11:371–82. doi:10.1261/rna.7258505.
59. Segal SP, Dunckley T, Parker R. Sbp1p affects translational repression and decapping in Saccharomyces cerevisiae. Mol Cell Biol. 2006;26:5120–30. doi:10.1128/MCB.01913-05.
60. Lotan R, Bar-On VG, Harel-Sharvit L, Duek L, Melamed D, Choder M. The RNA polymerase II subunit Rpb4p mediates decay of a specific class of mRNAs. Genes Dev. 2005;19:3004–16. doi:10.1101/gad.353205.
61. Stribinskis V, Ramos KS. Rpm2p, a protein subunit of mitochondrial RNase P, physically and genetically interacts with cytoplasmic processing bodies. Nucleic Acids Res. 2007;35:1301–11. doi:10.1093/nar/gkm023.
62. Wichroski MJ, Robb GB, Rana TM. Human retroviral host restriction factors APOBEC3G and APOBEC3F localize to mRNA processing bodies. PLoS Pathog. 2006;2:e41. doi:10.1371/journal.ppat.0020041.

63. Peruquetti RL, de Mateo S, Sassone-Corsi P. Circadian proteins CLOCK and BMAL1 in the chromatoid body, a RNA processing granule of male germ cells. PLoS ONE. 2012;7:e42695. doi:10.1371/journal.pone.0042695.
64. Parker R, Sheth U. P bodies and the control of mRNA translation and degradation. Mol Cell. 2007;25:635–46. doi:10.1016/j.molcel.2007.02.011.
65. Rossi JJ. RNAi and the P-body connection. Nat Cell Biol. 2005;7:643–4.
66. Eulalio A, Behm-Ansmant I, Schweizer D, Izaurralde E. P-body formation is a consequence, not the cause, of RNA-mediated gene silencing. Mol Cell Biol. 2007;27:3970–81. doi:10.1128/MCB.00128-07.
67. Parker R, Song H. The enzymes and control of eukaryotic mRNA turnover. Nat Struct Mol Biol. 2004;11:121–7. doi:10.1038/nsmb724.
68. Lejeune F, Li X, Maquat LE. Nonsense-mediated mRNA decay in mammalian cells involves decapping, deadenylating, and exonucleolytic activities. Mol Cell. 2003;12:675–87.
69. Houseley J, LaCava J, Tollervey D. RNA-quality control by the exosome. Nat Rev Mol Cell Biol. 2006;7:529–39.
70. Coller J, Parker R. General translational repression by activators of mRNA decapping. Cell. 2005;122:875–86. doi:10.1016/j.cell.2005.07.012.
71. Nissan T, Rajyaguru P, She M, Song H, Parker R. Decapping activators in Saccharomyces cerevisiae act by multiple mechanisms. Mol Cell. 2010;39:773–83. doi:10.1016/j.molcel.2010.08.025.
72. Conti E, Izaurralde E. Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species. Curr Opin Cell Biol. 2005;17:316–25. doi:10.1016/j.ceb.2005.04.005.
73. Amrani N, Sachs MS, Jacobson A. Early nonsense: mRNA decay solves a translational problem. Nat Rev Mol Cell Biol. 2006;7:415–25. doi:10.1038/nrm1942.
74. Conti E, Izaurralde E. Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species. Curr Opin Cell Biol. 2005;17:316–25. doi:10.1016/j.ceb.2005.04.005.
75. Lejeune F, Maquat LE. Mechanistic links between nonsense-mediated mRNA decay and pre-mRNA splicing in mammalian cells. Curr Opin Cell Biol. 2005;17:309–15. doi:10.1016/j.ceb.2005.03.002.
76. Amrani N, Sachs MS, Jacobson A. Early nonsense: mRNA decay solves a translational problem. Nat Rev Mol Cell Biol. 2006;7:415–25.
77. Cao D, Parker R. Computational modeling and experimental analysis of nonsense-mediated decay in yeast. Cell. 2003;113:533–45.
78. Chen CY, Shyu AB. Rapid deadenylation triggered by a nonsense codon precedes decay of the RNA body in a mammalian cytoplasmic nonsense-mediated decay pathway. Mol Cell Biol. 2003;23:4805–13.
79. Couttet P, Grange T. Premature termination codons enhance mRNA decapping in human cells. Nucleic Acids Res. 2004;32:488–94. doi:10.1093/nar/gkh218.
80. Gatfield D, Izaurralde E. Nonsense-mediated messenger RNA decay is initiated by endonucleolytic cleavage in Drosophila. Nature. 2004; 429:575–578. doi:http://www.nature.com/nature/journal/v429/n6991/suppinfo/nature02559_S1.html.
81. Mitchell P, Tollervey D. An NMD pathway in yeast involving accelerated deadenylation and exosome-mediated 3′– > 5′ degradation. Mol Cell. 2003;11:1405–13.
82. Filipowicz W. RNAi: the nuts and bolts of the RISC machine. Cell. 2005;122:17–20. doi:10.1016/j.cell.2005.06.023.
83. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell. 2004;116:281–97.
84. Behm-Ansmant I, Rehwinkel J, Doerks T, Stark A, Bork P, Izaurralde E. mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. Genes Dev. 2006;20:1885–98. doi:10.1101/gad.1424106.

85. Eystathioy T, Chan EK, Tenenbaum SA, Keene JD, Griffith K, Fritzler MJ. A phosphorylated cytoplasmic autoantigen, GW182, associates with a unique population of human mRNAs within novel cytoplasmic speckles. Mol Biol Cell. 2002;13:1338–51. doi:10.1091/mbc.01-11-0544.
86. Liu J, Valencia-Sanchez MA, Hannon GJ, Parker R. MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. Nat Cell Biol. 2005;7:719–23. doi:10.1038/ncb1274.
87. Wilhelm JE, Buszczak M, Sayles S. Efficient protein trafficking requires trailer hitch, a component of a ribonucleoprotein complex localized to the ER in Drosophila. Dev Cell. 2005;9:675–85. doi:10.1016/j.devcel.2005.09.015.
88. Kilchert C, Weidner J, Prescianotto-Baschong C, Spang A. Defects in the secretory pathway and high Ca2 + induce multiple P-bodies. Mol Biol Cell. 2010;21:2624–38. doi:10.1091/mbc.E10-02-0099.
89. Huang L, Mollet S, Souquere S, Le Roy F, Ernoult-Lange M, Pierron G, Dautry F, Weil D. Mitochondria associate with P-bodies and modulate microRNA-mediated RNA interference. J Biol Chem. 2011;286:24219–30. doi:10.1074/jbc.M111.240259.
90. Chang W, Zaarour RF, Reck-Peterson S, Rinn J, Singer RH, Snyder M, Novick P, Mooseker MS. Myo2p, a class V myosin in budding yeast, associates with a large ribonucleic acid-protein complex that contains mRNAs and subunits of the RNA-processing body. RNA. 2008;14:491–502. doi:10.1261/rna.665008.
91. Loschi M, Leishman CC, Berardone N, Boccaccio GL. Dynein and kinesin regulate stress-granule and P-body dynamics. J Cell Sci. 2009;122:3973–82. doi:10.1242/jcs.051383.
92. Kiebler MA, Bassell GJ. Neuronal RNA granules: movers and makers. Neuron. 2006;51:685–90. doi:10.1016/j.neuron.2006.08.021.
93. Thomas MG. Martinez Tosar LJ, Desbats MA, Leishman CC, Boccaccio GL. Mammalian Staufen 1 is recruited to stress granules and impairs their assembly. J Cell Sci. 2009;122:563–73. doi:10.1242/jcs.038208.
94. Pimentel J, Boccaccio GL. Translation and silencing in RNA granules: a tale of sand grains. Front Mol Neurosci. 2014;7:68. doi:10.3389/fnmol.2014.00068.
95. Buchan JR. mRNP granules. Assembly, function, and connections with disease. RNA Biol. 2014; 11:1019–1030. doi:10.4161/15476286.2014.972208.
96. Kloc M, Etkin LD. RNA localization mechanisms in oocytes. J Cell Sci. 2005;118:269–82. doi:10.1242/jcs.01637.
97. Coller JM, Tucker M, Sheth U, Valencia-Sanchez MA, Parker R. The DEAD box helicase, Dhh1p, functions in mRNA decapping and interacts with both the decapping and deadenylase complexes. RNA. 2001;7:1717–27.
98. Ladomery M, Wade E, Sommerville J. Xp54, the Xenopus homologue of human RNA helicase p54, is an integral component of stored mRNP particles in oocytes. Nucleic Acids Res. 1997;25:965–73.
99. Nakamura A, Amikura R, Hanyu K, Kobayashi S. Me31B silences translation of oocyte-localizing RNAs through the formation of cytoplasmic RNP complex during Drosophila oogenesis. Development. 2001;128:3233–42.
100. Navarro RE, Shim EY, Kohara Y, Singson A, Blackwell TK. cgh-1, a conserved predicted RNA helicase required for gametogenesis and protection from physiological germline apoptosis in C. elegans. Development. 2001; 128:3221–3232.
101. Layana C, Ferrero P, Rivera-Pomar R. Cytoplasmic Ribonucleoprotein Foci in Eukaryotes: Hotspots of Bio(chemical)Diversity. Comp Funct Genomics. 2012;2012:504292. doi:10.1155/2012/504292.
102. Cougot N, Bhattacharyya SN, Tapia-Arancibia L, Bordonne R, Filipowicz W, Bertrand E, Rage F. Dendrites of mammalian neurons contain specialized P-body-like structures that respond to neuronal activation. J Neurosci Official J Soc Neurosci. 2008;28:13793–804. doi:10.1523/JNEUROSCI.4155-08.2008.

103. di Penta A, Mercaldo V, Florenzano F, Munck S, Ciotti MT, Zalfa F, Mercanti D, Molinari M, Bagni C, Achsel T. Dendritic LSm1/CBP80-mRNPs mark the early steps of transport commitment and translational control. J Cell Biol. 2009;184:423–35. doi:10.1083/jcb.200807033.
104. Aizer A, Brody Y, Ler LW, Sonenberg N, Singer RH, Shav-Tal Y. The dynamics of mammalian P body transport, assembly, and disassembly in vivo. Mol Biol Cell. 2008;19:4154–66. doi:10.1091/mbc.E08-05-0513.
105. Mollet S, Cougot N, Wilczynska A, Dautry F, Kress M, Bertrand E, Weil D. Translationally repressed mRNA transiently cycles through stress granules during stress. Mol Biol Cell. 2008;19:4469–79. doi:10.1091/mbc.E08-05-0499.
106. Balagopal V, Parker R. Polysomes, P bodies and stress granules: states and fates of eukaryotic mRNAs. Curr Opin Cell Biol. 2009;21:403–8. doi:10.1016/j.ceb.2009.03.005.
107. Buchan JR, Parker R. Eukaryotic stress granules: the ins and outs of translation. Mol Cell. 2009;36:932–41. doi:10.1016/j.molcel.2009.11.020.
108. Ferrero PV, Layana C, Paulucci E, Gutierrez P, Hernandez G, Rivera-Pomar RV. Cap binding-independent recruitment of eIF4E to cytoplasmic foci. Biochim Biophys Acta. 2012;1823:1217–24. doi:10.1016/j.bbamcr.2012.03.013.
109. Mahowald AP. Polar granules of drosophila. IV. Cytochemical studies showing loss of RNA from polar granules during early stages of embryogenesis. The. J Exp Zool. 1971;176:345–52. doi:10.1002/jez.1401760309.
110. Tadros W, Lipshitz HD. Setting the stage for development: mRNA translation and stability during oocyte maturation and egg activation in Drosophila. Dev Dyn Official Pub Am Assoc Anat. 2005;232:593–608. doi:10.1002/dvdy.20297.
111. Mahowald AP. Assembly of the drosophila germ plasm. Int Rev Cytol. 2001;203:187–213.
112. Spradling AC. Germline cysts: communes that work. Cell. 1993;72:649–51.
113. Nakamura A, Amikura R, Hanyu K, Kobayashi S. Me31B silences translation of oocyte-localizing RNAs through the formation of cytoplasmic RNP complex during Drosophila oogenesis. Development. 2001;128:3233–42.
114. Lasko P. mRNA Localization and Translational Control in Drosophila Oogenesis. Cold Spring Harb Perspect Biol. 2012;4:a012294. doi:10.1101/cshperspect.a012294.
115. Nakamura A, Sato K, Hanyu-Nakamura K. Drosophila cup is an eIF4E binding protein that associates with Bruno and regulates oskar mRNA translation in oogenesis. Dev Cell. 2004;6:69–78.
116. Cho PF, Poulin F, Cho-Park YA, Cho-Park IB, Chicoine JD, Lasko P, Sonenberg N. A new paradigm for translational control: inhibition via 5′-3′ mRNA tethering by Bicoid and the eIF4E cognate 4EHP. Cell. 2005;121:411–23. doi:10.1016/j.cell.2005.02.024.
117. White-Cooper H. Tissue, cell type and stage-specific ectopic gene expression and RNAi induction in the Drosophila testis. Spermatogenesis. 2012;2:11–22. doi:10.4161/spmg.19088.
118. Hernandez G, Han H, Gandin V, Fabian L, Ferreira T, Zuberek J, Sonenberg N, Brill JA, Lasko P. Eukaryotic initiation factor 4E-3 is essential for meiotic chromosome segregation, cytokinesis and male fertility in Drosophila. Development. 2012;139:3211–20. doi:10.1242/dev.073122.
119. Ghosh S, Lasko P. Loss-of-function analysis reveals distinct requirements of the translation initiation factors eIF4E, eIF4E-3, eIF4G and eIF4G2 in Drosophila spermatogenesis. PLoS ONE. 2015;10:e0122519. doi:10.1371/journal.pone.0122519.
120. Franks TM, Lykke-Andersen J. The control of mRNA decapping and P-body formation. Mol Cell. 2008;32:605–15. doi:10.1016/j.molcel.2008.11.001.

RNA-Mediated Silencing in Eukaryotes: Evolution of Protein Components and Biological Roles

J. Armando Casas-Mollano, Ericka Zacarias, Xinrong Ma, Eun-Jeong Kim and Heriberto Cerutti

1 Introduction

RNA-mediated silencing is an evolutionarily conserved process in eukaryotes by which small RNAs (~20–30 nt in length) induce the inactivation of cognate sequences through a variety of mechanisms, including translation repression, RNA degradation, transcriptional inhibition and/or, in a few organisms, DNA elimination [1–9]. The role of double-stranded RNAs (dsRNAs) in triggering repression, acting as precursors for processing into small RNAs (sRNAs), was initially characterized in *Caenorhabditis elegans* and termed RNA interference (RNAi) [10]. Interestingly, recent studies indicate that these non-coding RNAs may also participate in transcriptional or translational activation [11–14]. Despite the mechanistic diversity of these RNA-mediated processes, in most characterized pathways sRNAs are incorporated into effector complexes containing at their core Argonaute-PIWI (AGO-PIWI) proteins, which include two major families of polypeptides in eukaryotes, named after *Arabidopsis thaliana* ARGONAUTE1 (AGO1) and *Drosophila melanogaster* P-element induced wimpy testis (PIWI) [2, 3, 5, 7, 15–17]. Some AGO-PIWI proteins function as sRNA-guided endonucleases (slicers) that cleave complementary transcripts, whereas others lack endonucleolytic activity and repress their targets through other mechanisms [3, 5, 17–23]. From a cellular/organismal perspective, RNA-mediated silencing appears to play two main roles: protecting cells from selfish genetic elements such as viruses and transposons

J.A. Casas-Mollano (✉) · E. Zacarias
Departamento de Bioquímica, Instituto de Química,
Universidade de São Paulo, São Paulo, Brazil
e-mail: acasasmollano@gmail.com

X. Ma · E.-J. Kim · H. Cerutti (✉)
School of Biological Sciences and Center for Plant Science Innovation,
University of Nebraska-Lincoln, Lincoln, NE, USA
e-mail: hcerutti1@unl.edu

G. Hernández and R. Jagus (eds.), *Evolution of the Protein Synthesis Machinery and Its Regulation*, DOI 10.1007/978-3-319-39468-8_20

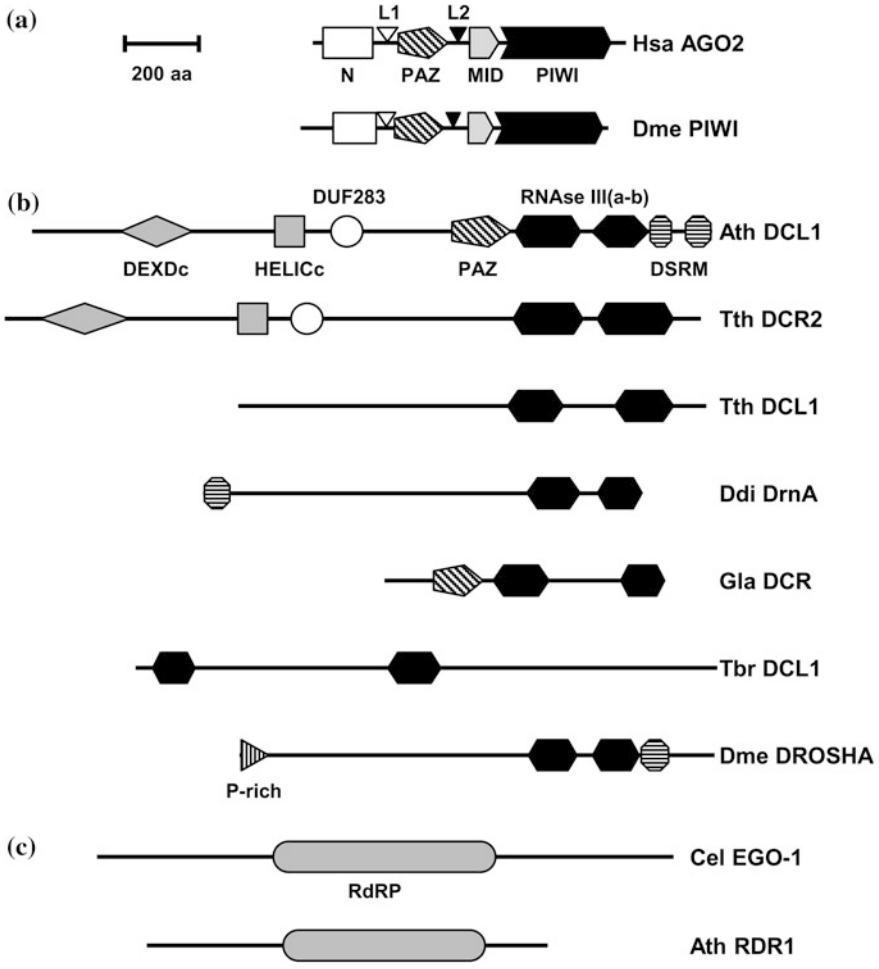
(possibly with an extended role in heterochromatin organization in some eukaryotes) and/or regulating the expression of endogenous genes [3, 4, 24–32].

Three major classes of sRNAs have been recognized in metazoans: microRNAs (miRNAs), PIWI-interacting RNAs (piRNAs) and small interfering RNAs (siRNAs) [15–17, 26, 33–36]. Land plants and green algae lack PIWI proteins and contain only miRNAs and siRNAs that associate with members of the AGO-like family [1, 2, 17, 36]. MicroRNAs generally originate from single-stranded non-coding RNAs or introns, which fold into imperfectly paired hairpin structures [15, 26, 34, 36]. They usually modulate gene expression, affecting many biological processes such as development, metabolism or stress responses [4, 5, 8, 17, 26, 36]. siRNAs are produced from long, near-perfect complementarity dsRNAs of diverse origins, including the products of convergent transcription or RNA-dependent RNA polymerase (RdRP) activity, viral or transposon RNAs, or dsRNAs experimentally introduced into cells [4, 5, 15, 17, 24, 30, 37]. In plants and algae, these siRNAs play various roles in the suppression of viruses and transposable elements, post-transcriptional regulation of gene expression, DNA double-strand break repair, DNA methylation and/or heterochromatin formation [4, 6, 15, 17, 24, 25, 32]. However, there is a growing realization that, despite their differences, distinct small RNA pathways (which include multiple other variants in diverse eukaryotes) often interact, compete for and share effector proteins and cross-regulate each other.

In addition to the AGO-PIWI superfamily, the core protein components of the RNAi machinery (Fig. 1) include an RNaseIII-like endonuclease (named Dicer), involved in the processing of long dsRNAs into small RNAs, and an RdRP protein, involved in the generation of dsRNA from single-stranded transcripts and/or in the production of secondary small RNAs [15, 25, 38]. The eukaryotic RNAi apparatus is remarkable in that it seems to have been assembled from three distinct prokaryotic sources, archaeal (AGO-PIWIs), bacterial (the RNaseIII domains of Dicer) and viral (RdRP) [15, 38]. Interestingly, the phyletic distribution patterns of AGO-PIWIs, Dicer-like proteins and RdRPs suggest that these components came together as a functional unit in the last common ancestor of eukaryotes [15, 25, 38, 39]. Subsequently, each of these proteins evolved through a series of duplications, some antedating the divergence of the major eukaryotic lineages and some being lineage specific. However, as discussed in more detail below, AGO-PIWI proteins appear to be the primary players in the eukaryotic RNAi system, and they can function independently of other core components. Moreover, the functional role of AGO-PIWIs (and of the RNAi machinery) appears to have changed during evolution, from relatively simple host-defense proteins to key players in complex multiprotein regulatory pathways in multicellular organisms.

2 Evolution of Argonaute-PIWI Proteins

In the most extensively characterized RNAi pathways, short dsRNA molecules are loaded into eukaryotic AGO-PIWI proteins, one strand (guide) is selectively retained and the other strand (passenger) removed, and then AGO-PIWIs use the retained strand as a guide to bind to complementary RNAs for inactivation [2, 3, 5, 17–19]. Typical eukaryotic AGO-PIWI polypeptides are characterized by a conserved structure (Fig. 1a), comprised of the following domains: (1) an N (N-terminal) domain, which plays a role in target RNA cleavage and in the dissociation of the passenger strand (i.e., the complementary strand in the short dsRNA precursor of a miRNA/siRNA); (2) the PAZ (PIWI-Argonaute-Zwille) domain, which binds the 3′ end of the guide strand (i.e., the miRNA/siRNA proper);



◀ **Fig. 1** Domain composition and architecture of AGO-PIWI, Dicer and RdRP proteins from representative eukaryotes. **a** Schematic illustration of human AGO2 (AGO-like family) and Drosophila PIWI (PIWI-like family) proteins. *N*, N-terminal domain; *PAZ*, PIWI-Argonaute-Zwille domain; *MID*, middle domain; *PIWI*, PIWI domain; L1 and L2, linker motifs. **b** Protein domain diagrams demonstrating the structural diversity of Dicer-like polypeptides. *DEXDc*, DEAD/DEAH-like helicase superfamily domain; *HELICc*, helicase superfamily C-terminal domain. *DUF283*, divergent dsRNA-binding fold and heterodimerization domain. *RNAseIII* (a-b), ribonuclease III catalytic domains a and b; *DSRM*, double-stranded RNA binding motif; *P-rich*, proline-rich domain. DEXDc and HELICc are referred to in the text as the superfamily II helicase domain. **c** Representation of two RdRP proteins, *C. elegans* EGO-1 and Arabidopsis RDR1. *RdRP*, RNA-dependent RNA polymerase domain. Domains present in the individual proteins were identified using the SMART [40] and PFAM [41] databases. Accession numbers of the proteins represented in the figures are as follows: Ath DCL1, *Arabidopsis thaliana* DCL1 (Q9SP32); Ath RDR1, *A. thaliana* RNA-dependent RNA polymerase 1 (NP_172932); Cel EGO-1, *Caenorhabditis elegans* Enhancer of Glp-One (NP_492132); Ddi DrnA, *Dictyostelium discoideum* DrnA (CAC41976); Dme DROSHA, *Drosophila melanogaster* DROSHA (AAF59169); Dme PIWI, *D. melanogaster* PIWI (AAD08704); Gla DCR, *Giardia lamblia* Dicer (XP_001705536); Hsa AGO2, *Homo sapiens* Argonaute-2 isoform 1 (NP_036286); Tbr DCL1, *Trypanosoma brucei* DCL1 (AAX69739); Tth DCL1 *Tetrahymena thermophila* DCL1 (XP_001018276); Tth DCR2, *T. thermophila* DCR2 (EDK31487). Proteins are drawn to scale. The *scale bar* represents 200 amino acids

(3) the MID (middle) domain, which binds the 5′ end of the guide strand; (4) the PIWI domain, an RNaseH-like fold domain, which accommodates the target RNA [2, 3, 18–23]. Two linker motifs, L1 and L2, are located between the N and PAZ and the PAZ and MID domains, respectively [3]. In cleavage active AGO-PIWIs (slicers), the PIWI domain contains a catalytic pocket (the DEDX motif) consisting of a DDX triad (where D is aspartic acid and X is aspartic acid, histidine or, in rare cases, lysine) and a glutamate (E) that resides on a loop (the so-called glutamate finger) [3, 18, 19, 23, 42].

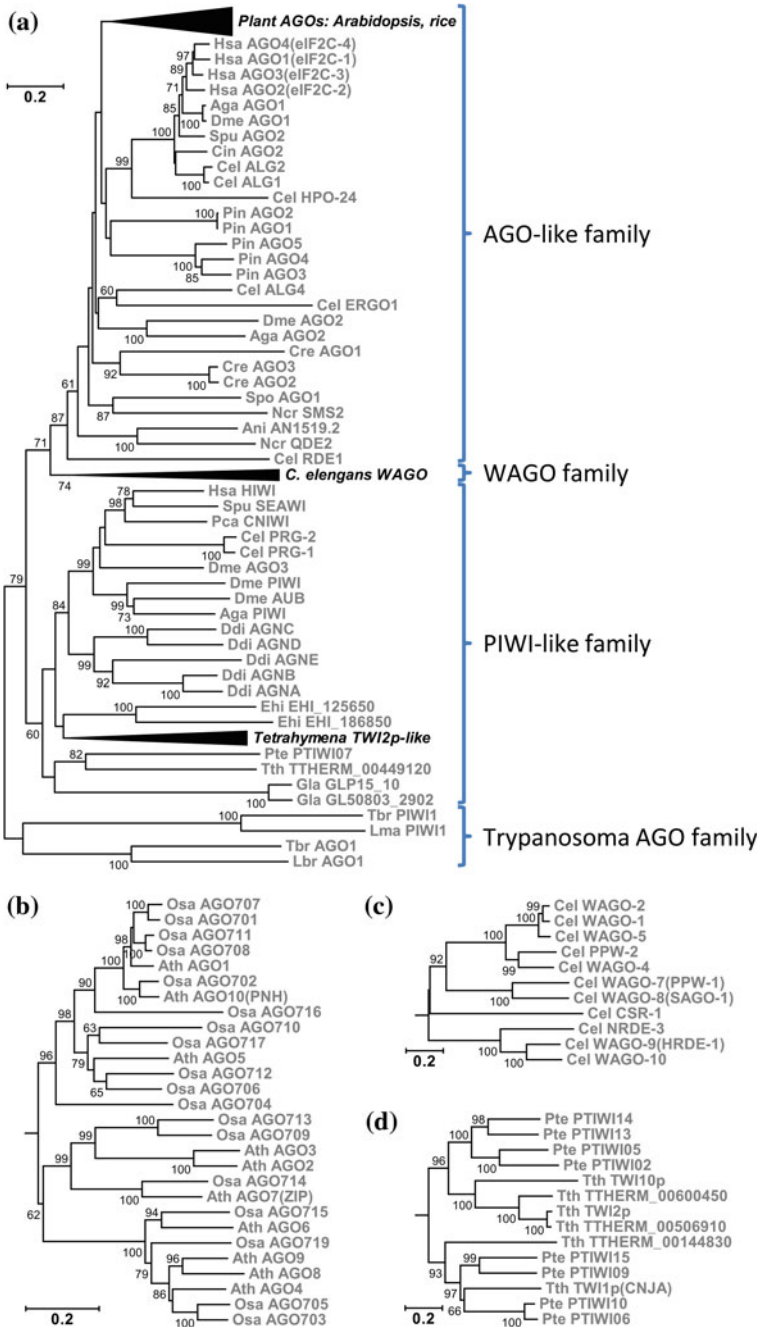
AGO-PIWI proteins are widespread in eukaryotes, being found in all eukaryotic supergroups [3, 15, 25]. However, members of the AGO-PIWI superfamily, defined by the presence of a PIWI domain, are also present in many prokaryotes [3, 15, 43]. The PIWI RNaseH fold domain is most closely related to the endonuclease V (EndoV) family of deoxyribonucleases [44] and likely derived from an EndoV-like precursor originally involved in DNA repair in prokaryotes [15]. Prokaryotic AGO-PIWI genes, like most prokaryotic defense genes [45], show a patchy taxonomic distribution, with at most 70 % representation in any bacterial or archaeal phylum [3]. These prokaryotic AGO-PIWIs fall into three major clades: short PIWIs (class I), consisting of the MID and PIWI domains; PIWI-REs, similar to short PIWIs but with characteristically conserved arginine (R) and glutamic acid (E) residues [43]; long PIWIs (class II), consisting of the classical multidomain AGO-PIWI architecture as well as some truncated variants [3, 15, 46]. Interestingly, eukaryotic AGO-PIWIs cluster in phylogenetic analyses with prokaryotic long PIWIs from several euryarchaeal species, mostly thermophiles, suggesting that eukaryotic AGO-PIWIs originated from an archaeal PIWI subgroup [3, 46].

Moreover, the ancestral eukaryotic AGO-PIWI was likely cleavage active, and the catalytic function was subsequently lost in a subset of proteins [3, 15].

The evolutionary history of the eukaryotic AGO-PIWI superfamily is characterized by an initial bifurcation that resulted in the classical AGO-PIWI proteins and the MedPIWI polypeptides [15, 43]. The latter proteins have only the MID domain and a catalytically inactive PIWI domain and are typified by Med13, a subunit of the transcription regulatory Mediator complex in mammals [43]. To infer the evolutionary path taken by the classical AGO-PIWIs, we have used sequences of the PAZ and PIWI domains to build phylogenetic trees using a representative set of eukaryotic AGO-PIWI proteins (Fig. 2). As previously reported [3, 47], classical eukaryotic AGO-PIWIs cluster into four major groups: the Trypanosoma AGO family, typified by *Trypanosoma brucei* TbAGO1; the WAGO family, typified by *Caenorhabditis elegans*-specific AGO-PIWIs; the AGO-like family, typified by *A. thaliana* AGO1; the PIWI family, typified by *D. melanogaster* PIWI. The AGO-like and PIWI families are represented in several eukaryotic supergroups [15, 25], suggesting that at least one duplication of the classical AGO-PIWIs may have occurred early during eukaryotic evolution. In contrast, the Trypanosoma AGO and the WAGO families appear to have emerged as a result of additional, lineage-specific duplications accompanied by fast evolutionary divergence (Fig. 2). For instance, the WAGO family seems to function uniquely in several distinct silencing mechanisms in *C. elegans* involving secondary sRNAs generated by RdRP activity [30, 48–50].

Fungi, green algae, land plants and *Phytophthora infestans* appear to encode exclusively AGO-like proteins in their genomes (Fig. 2). In contrast, Amoebozoa, *Giardia lamblia* and ciliates (*Paramecium tetraurelia* and *Tetrahymena thermophila*) seem to encode exclusively PIWI proteins. Animals have representatives of both AGO-PIWI families (Fig. 2). A parsimonious interpretation of these data suggests that the last common ancestor of eukaryotes (or at least the ancestor of the 'crown' group) contained both classes of polypeptides and that specific lineages independently lost one or the other [3, 15, 25]. Only animals appear to have retained both classes of proteins, where the PIWI family became functionally linked primarily with the emerging piRNA class of small RNAs and the protection of germline integrity against transposable elements [16, 30, 33] and the AGO-like family mainly maintained roles typically attributed to the ancestral AGO-PIWI proteins in other eukaryotic lineages [3, 15, 54]. The AGO-like and PIWI families have also undergone marked expansion in certain eukaryotic lineages, most notably ciliates, plants and metazoans (Fig. 2), frequently accompanied by functional specialization and binding to distinct classes of sRNAs [2, 4, 5, 30, 36, 51, 52].

Based on the conservation of the four active site residues, the majority of the eukaryotic AGO-like and PIWI proteins are predicted to be catalytically active [3]. However, as a caveat, certain AGOs with archetypal DEDX tetrads are slicing inactive, such as human AGO3, which harbors all four residues but cannot cleave RNA targets in vitro [20, 21]. In addition, some guide-target interactions are characterized by limited, imperfect base pairing that is incompatible with target RNA cleavage (as is usually the case for metazoan miRNA interactions) [3, 5, 26,


(a)
Plant AGOs: Arabidopsis, rice
Hsa AGO4(eIF2C-4)
Hsa AGO1(eIF2C-1)
Hsa AGO3(eIF2C-3)
Hsa AGO2(eIF2C-2)
Aga AGO1
Dme AGO1
Spu AGO2
Cin AGO2
Cel ALG2
Cel ALG1
Cel HPO-24
Pin AGO2
Pin AGO1
Pin AGO5
Pin AGO4
Pin AGO3
Cel ALG4
Cel ERGO1
Dme AGO2
Aga AGO2
Cre AGO1
Cre AGO3
Cre AGO2
Spo AGO1
Ncr SMS2
Ani AN1519.2
Ncr QDE2
Cel RDE1
C. elengans WAGO
Hsa HIWI
Spu SEAWI
Pca CNIWI
Cel PRG-2
Cel PRG-1
Dme AGO3
Dme PIWI
Dme AUB
Aga PIWI
Ddi AGNC
Ddi AGND
Ddi AGNE
Ddi AGNB
Ddi AGNA
Ehi EHI_125650
Ehi EHI_186850
Tetrahymena TWI2p-like
Pte PTIWI07
Tth TTHERM_00449120
Gla GLP15_10
Gla GL50803_2902
Tbr PIWI1
Lma PIWI1
Tbr AGO1
Lbr AGO1
AGO-like family
WAGO family
PIWI-like family
Trypanosoma AGO family
0.2
(b)
Osa AGO707
Osa AGO701
Osa AGO711
Osa AGO708
Ath AGO1
Osa AGO702
Ath AGO10(PNH)
Osa AGO716
Osa AGO710
Osa AGO717
Ath AGO5
Osa AGO712
Osa AGO706
Osa AGO704
Osa AGO713
Osa AGO709
Ath AGO3
Ath AGO2
Osa AGO714
Ath AGO7(ZIP)
Osa AGO715
Ath AGO6
Osa AGO719
Ath AGO9
Ath AGO8
Ath AGO4
Osa AGO705
Osa AGO703
0.2
(c)
Cel WAGO-2
Cel WAGO-1
Cel WAGO-5
Cel PPW-2
Cel WAGO-4
Cel WAGO-7(PPW-1)
Cel WAGO-8(SAGO-1)
Cel CSR-1
Cel NRDE-3
Cel WAGO-9(HRDE-1)
Cel WAGO-10
0.2
(d)
Pte PTIWI14
Pte PTIWI13
Pte PTIWI05
Pte PTIWI02
Tth TWI10p
Tth TTHERM_00600450
Tth TWI2p
Tth TTHERM_00506910
Tth TTHERM_00144830
Pte PTIWI15
Pte PTIWI09
Tth TWI1p(CNJA)
Pte PTIWI10
Pte PTIWI06
0.2

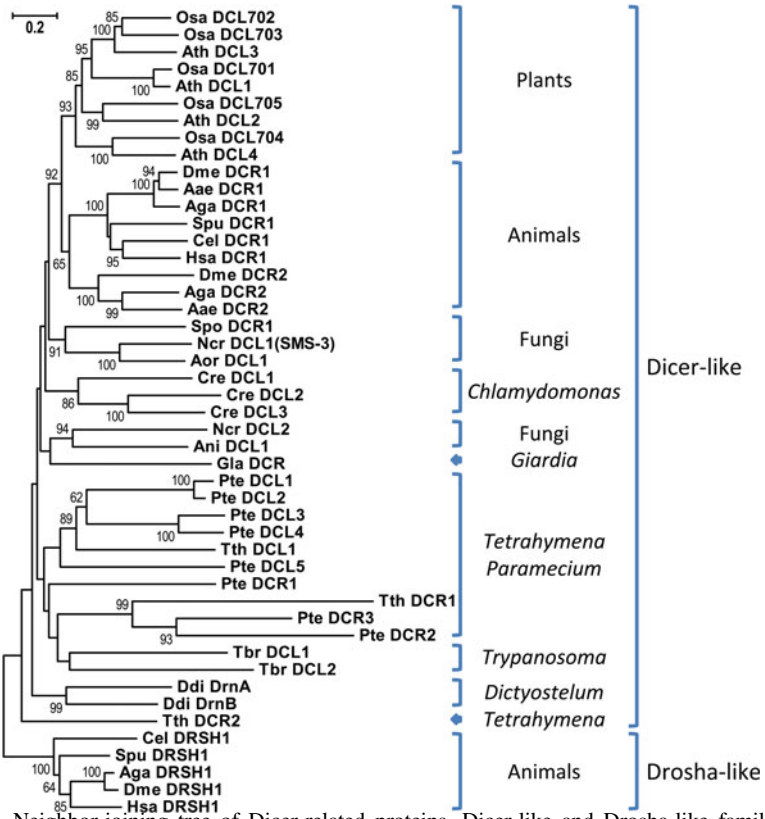
◀ **Fig. 2** Phylogenetic analysis of AGO-PIWI proteins from a representative set of eukaryotic organisms. **a** Neighbor-joining tree of 105 AGO-PIWI proteins from diverse eukaryotes. The four major families of AGO-PIWI proteins are indicated to the right of the tree. Branches corresponding to plant AGOs, *Caenorhabditis elegans* WAGOs and TWI2p-like proteins have been collapsed for presentation purposes. **b** Uncollapsed plant AGO subtree containing AGO proteins from *Arabidopsis thaliana* and *Oryza sativa*. **c** WAGO subtree showing the relationship between different WAGO proteins from *C. elegans*. **d** TWI2p-like subtree showing clustering of PIWI-like proteins from *Tetrahymena thermophila* and *Paramecium tetraurelia*. To construct the tree, sequences corresponding to the PAZ and PIWI domains were obtained using SMART [40] and aligned with the ClustalX program [53]. The alignment generated was then used to construct a neighbor-joining tree [54] using MEGA version 4.0 [55]. Tree reliability was assessed by conducting a bootstrap test based on 1000 pseudoreplicates. Numbers on the nodes correspond to bootstrap values bigger than 60 %. Protein names are preceded by an abbreviation of the species they correspond to, which is indicated by a three-letter code consisting of the initial of the genus followed by the first two letters of the species as follow: Aga, *Anopheles gambiae*; Ani, *Aspergillus nidulans*; Ath, *Arabidopsis thaliana*; Cel, *Caenorhabditis elegans*; Cin, *Ciona intestinalis*; Cre, *Chlamydomonas reinhardtii*; Ddi, *Dictyostelium discoideum*; Dme, *Drosophila melanogaster*; Ehi, *Entamoeba histolytica*; Gla, *Giardia lamblia*; Hsa, *Homo sapiens*; Lbr, *Leishmania braziliensis*; Lma, *Leishmania major*; Ncr, *Neurospora crassa*; Osa, *Oryza sativa*; Pca, *Podocoryna carnea*; Pin, *Phytophthora infestans*; Pte, *Paramecium tetraurelia*; Spo, *Schizosaccharomyces pombe*; Spu, *Strongylocentrotus purpuratus*; Tbr, *Trypanosoma brucei*; Tth, *Tetrahymena thermophila*. Accession numbers of proteins used to draw this tree are: Aga AGO1, EAA00062; Aga AGO2, EAL41436; Aga PIWI, EAA05900; Ani AN1519.2, EAA63775; Ath AGO1, AAC18440; Ath AGO2, NP_174413; Ath AGO3, NP_174414; Ath AGO4, NP_565633; Ath AGO5, NP_850110; Ath AGO6, NP_180853; Ath AGO7(ZIP), NP_177103; Ath AGO8, NP_197602; Ath AGO9, NP_197613; Ath AGO10(PNH), CAA11429; Cel ALG1, NP_001257238; Cel ALG2, NP_871992; Cel ALG4, NP_499192; Cel CSR-1, NP_001040939; Cel ERGO1, O61931; Cel HPO-24, NP_492643; Cel NRDE-3, NP_508092; Cel PPW-2, NP_491535; Cel PRG-1, CAA98113; Cel PRG-2, NP_500994; Cel RDE1, NP_741611; Cel WAGO-1, NP_492045; Cel WAGO-2, NP_491579; Cel WAGO-4, NP_496751; Cel WAGO-5, NP_495151; Cel WAGO-7 (PPW-1), NP_740835; Cel WAGO-8(SAGO-1), NP_504610; Cel WAGO-9(HRDE-1), NP_497834; Cel WAGO-10, NP_503177; Cin AGO2, XP_009857554; Cre AGO1, Cre02.g141050*; Cre AGO2, Cre04.g214250*; Cre AGO3, Cre16.g689647*; Ddi AGNA, EAL69296; Ddi AGNB, EAL62204; Ddi AGNC, EAL71514; Ddi AGND, XP_001134555; DdiAGNE, EAL62770; Dme AGO1, BAA88078; Dme AGO2, Q9VUQ5; Dme AGO3, NP_001036627; Dme AUB, CAA64320; Dme PIWI, AAD08704; Ehi EHI_125650, EAL51127; Ehi EHI_186850, XP_656436; Gla GL50803_2902, XP_001707926; Gla GLP15_10, EFO65454; Hsa AGO1 (eIF2C-1), AAH63275; Hsa AGO2(eIF2C-2), NP_036286; Hsa AGO3(eIF2C-3), NP_079128; Hsa AGO4(eIF2C-4), BAB13393; Has HIWI, AAC97371; Lbr AGO1, ACI22628; Lma PIWI1, XP_001682974; Ncr QDE2, AAF43641; Ncr SMS2, AAN32951; Osa AGO701, NP_001048621; Osa AGO702, Q69VD5; Osa AGO703, NP_001042722; Osa AGO704, EEE67140; Osa AGO705, NP_001052115; Osa AGO706, Q851R2; Osa AGO707, NP_001058648; Osa AGO708, Q7XSA2; Osa AGO709, NP_001053871; Osa AGO710, NP_001050911; Osa AGO711, NP_001047704; Osa AGO712, NP_001059079; Osa AGO713, EAZ31978; Osa AGO714, EAZ27470; Osa AGO715, Q6YSJ5; Osa AGO716, NP_001046021; Osa AGO717, Q7Y001; Osa AGO719, NP_001042721; Pca CNIWI, AAS01181; Pin AGO1, XP_002906080; Pin AGO2, XP_002906081; Pin AGO3, XP_002908068; Pin AGO4, XP_002908108; Pin AGO5, XP_002908109; Pte PTIWI02, CAI44470; Pte PTIWI05, CAI44468; Pte PTIWI06, CAI39075; Pte PTIWI07, CAI39074; Pte PTIWI09, CAI39072; Pte TIWI10, CAI39070; Pte PTIWI13, CAI39067; Pte PTIWI14, CAI39066; Pte PTIWI15, CAI39065; Spo AGO1, O74957; Spu AGO2, XP_011668809; Spu SEAWI, AAG42533; Tbr AGO1, XP_823303; Tbr PIWI1, AAR10811; Tth TTHERM_00144830, XP_001011123; Tth TTHERM_00449120, XP_001013312; Tth TTHERM_00506910, XP_001015193; Tth TTHERM_00600450, XP_001032516; Tth TWI1p (CNJA), AAM77972; Tth TWI2p, AAQ74967; Tth TWI10p, ABP68415. *Accession numbers correspond to those given in the *Chlamydomonas reinhardtii* genome project (https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Creinhardtii)

34, 54]. Thus, some classical eukaryotic AGO-PIWIs may act purely as sequence-specific RNA binding proteins, either because they are catalytically inactive or because cleavage is prevented by the interaction context. Moreover, in various multicellular eukaryotes, AGO-PIWIs appear to function in larger protein networks and have evolved to interact with a multitude of polypeptides involved, for instance, in guide processing, guide loading, activity regulation, translation repression, transcription inhibition and/or recruitment of accessory polypeptides [2, 3, 6, 7, 17].

3 Evolution of Dicer and RdRP Proteins

As already mentioned, the core protein components of the RNAi machinery also include an RNaseIII-like endonuclease (named Dicer) and an RdRP polypeptide. Dicer-like proteins are relatively well conserved among eukaryotes containing classical AGO-PIWI proteins, albeit with significant variability in their primary sequence and domain organization (Fig. 1b) [25, 39]. In *G. lamblia* the sole Dicer-like protein, consisting of a PAZ domain and tandem RNaseIII motifs, is fully functional in the processing of long dsRNAs to short duplexes [56]. Indeed, a minimal Dicer may consist of only duplicated RNaseIII catalytic motifs [25, 57], which form a single dsRNA-processing center through intramolecular dimerization [58]. However, the multidomain Dicer variants of 'crown' eukaryotes also contain superfamily II helicase, PAZ and dsRNA binding domains (Fig. 1b) [25, 38, 39]. All eukaryotic RNaseIII domains seem to be of common origin, and it has been speculated that the RNaseIII and the dsRNA binding motifs of Dicer were acquired from an ancient bacterial source with subsequent duplication of the RNaseIII domain to generate an ancestral Dicer [15, 38]. Support for this idea came from budding yeast, lacking a canonical Dicer, in which a single RNaseIII domain protein, acting as a homodimer, is able to provide dsRNA dicing activity [59, 60]. The multidomain Dicers of 'crown' eukaryotes may have appeared gradually during early lineage evolution through the subsequent acquisition of PAZ and helicase domains [15, 38].

Yet, the evolutionary history of the eukaryotic Dicer enzymes has been unclear because the topology of phylogenetic trees (built using sequences from the RNaseIII domains) does not follow expected eukaryotic phylogeny (Fig. 3) and a monophyletic relationship cannot be unequivocally confirmed [25]. This may be partly due to rapid sequence variation among duplicated genes within specific lineages as well as to the high divergence of RNAi-active RNaseIII superfamily nucleases in certain basal lineages, such as those recently characterized in *T. brucei* [15, 57, 61]. Nonetheless, current evidence suggests that the Dicer family originated early in eukaryotic evolution and independently expanded in animals, plants, fungi and ciliates [25, 39, 62, 63]. Interestingly, Drosha, another endonuclease with tandem RNaseIII domains (Fig. 1b) involved in the processing of miRNA precursors in animals, forms an outgroup with respect to the Dicer polypeptides (Fig. 3). This type



◀ **Fig. 3** Neighbor-joining tree of Dicer-related proteins. Dicer-like and Drosha-like families are indicated to the right of the tree. Sequences corresponding to the RNAseIII domains (a and b) were obtained using the SMART database [40] and aligned with the ClustalX program [53]. This alignment was then used to infer a tree using MEGA 4.0 [55]. The numbers on the branches indicate bootstrap values (based on 1000 pseudoreplicates) bigger than 60 %. Species are designated by a three-letter abbreviation preceding the name of each protein. The species are the same as in Fig. 2, except for the addition of *Aedes aegypti* (Aae). Accession numbers of the proteins used to draw the tree are: Aae DCR1, AAW48724; Aae DCR2, AAW48725; Aga DCR1, AAO73809; Aga DCR2, EAA00264; Aga DRSH1, EAL39656; Ani DCL1, XP_660793; Aor DCL1, XP_001824024; Ath DCL1, NP_171612; Ath DCL2, NP_566199; Ath DCL3, NP_189978; Ath DCL4, NP_197532; Cel DCR1, P34529; Cel DRSH1, AAD31170; Cre DCL1, Cre02.g141000*; Cre DCL2, Cre16.g684715*; Cre DCL3, Cre07.g345900*; Ddi DrnA, CAC41976; Ddi DrnB, XP_647462; Dme DCR1, Q9VCU9; Dme DCR2, BAB69959; Dme DRSH1, AAF59169; Gla DCR, XP_001705536; Hsa DCR1, NP_803187; Hsa DRSH1, Q9NRR4; Ncr DCL1(SMS-3), XP_961898; Ncr DCL2, Q7SCC1; Osa DCL701, Q8LMR2; Osa DCL702, NP_001045148; Osa DCL703, BAT11379; Osa DCL704, BAF80150; Osa DCL705, NP_001050564; Pte DCL1, XP_001444844; Pte DCL2, XP_001439412; Pte DCL3, XP_001462500; Pte DCL4, XP_001429424; Pte DCL5, XP_001455480; Pte DCR1, XP_001455980; Pte DCR2, XP_001430660; Pte DCR3, XP_001461156; Spo DCR1, Q09884; Spu DCR1, XP_011668948; Spu DRSH1, XP_800324; Tbr DCL1, AAX69739; Tbr DCL2, AAX69562; Tth DCL1, XP_001018276; Tth DCR1, XP_001009465; Tth DCR2, EDK31487. *Accession numbers correspond to those given in the *Chlamydomonas reinhardtii* genome project (https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Creinhardtii)

of protein may have evolved independently from Dicer within the animal lineage or, as recently argued [15], it may have evolved from an earlier duplication of an ancestral tandem RNaseIII protein within the main eukaryotic lineage (with subsequent losses in several taxonomic groups).

Eukaryotic RdRPs (Fig. 1c) appear to have evolved from enzymes that functioned as DNA-dependent RNA polymerases in the transcription of certain bacteriophages or primer synthesis of certain distinctive bacterial mobile elements [15, 38]. Interestingly, the eukaryotic ancestor may have had three distinct RdRPs [4, 64], perhaps reflecting ancient functional diversification, which seem to have been inherited vertically during eukaryotic evolution in a complex pattern of duplications and losses [25, 64]. Indeed, RdRPs are not as widely distributed among extant eukaryotes as AGO-PIWI and Dicer-like proteins [25, 38, 64], and there is compelling evidence for lineage-specific losses of some of these proteins [65]. This is consistent with a postulated ancillary role of RdRPs in generating precursor dsRNAs and/or in amplifying sRNA levels [24, 25]. Thus, if small RNAs can be produced by other means, RdRPs may not be needed for RNAi, explaining their pervasive loss from certain eukaryotic lineages.

4 Ancestral Role(s) of the RNAi Machinery

The taxonomic distribution patterns of AGO-PIWIs, Dicer-like proteins and RdRPs suggest that they came together as a functional unit in the last common ancestor of eukaryotes [15, 25, 38, 39]. Moreover, the fact that the eukaryotic RNAi machinery seems to have been assembled from prokaryotic proteins apparently involved in diverse DNA repair and RNA synthesis/processing pathways points toward an independent origin of RNAi in eukaryotes. Indeed, although both prokaryotes and eukaryotes possess multiple RNA-mediated defense/regulatory mechanisms, these systems appear to have evolved independently [15].

AGO-PIWI proteins seem to be the key players in the development of the eukaryotic RNAi system. They are the most phylogenetically conserved RNAi component [15, 25, 38], and there is evidence that they can function independently of the other core factors. For instance, in the PIWI interacting small RNA system, different PIWI variants work with non-RNaseIII nucleases both in piRNA production and as functional effectors [35, 37]. Additionally, a Dicer-independent miRNA-processing pathway exists by which a catalytically active AGO-PIWI cleaves a precursor hairpin, independently of an RNaseIII-like enzyme, to yield mature miRNAs [66, 67]. This suggests that basic versions of RNA interference could have emerged with just a nuclease-active PIWI protein [3, 15]. Interestingly, nearly all major classes of prokaryotic AGO-PIWIs are combined in operons with a diverse range of deoxyribonucleases, indicating that the ancestral role of the PIWI domain may have been in RNA- or DNA-guided DNA restriction systems as defense responses against invasive DNAs [3, 15, 46]. Several prokaryotic long PIWIs work as RNA- or DNA-guided DNA-targeting systems [68–70], but few are

predicted to function in RNA-targeting [15]. Eukaryotes may have built their RNAi system primarily upon these RNA-targeting PIWIs, acquired early in their evolution from archaea [15].

Acquisition of the other core components resulted in the development of a fairly complex RNAi machinery in the last common ancestor of eukaryotes [25, 38, 64]. As previously proposed, the 'raison d'être' of this machinery may have been in defense responses against selfish genetic elements such as viruses and transposons, conceivably through repression at both the transcriptional (heterochromatin organization and/or cytosine DNA methylation) and post-transcriptional (RNA destabilization and/or translation inhibition) levels [15, 24, 25, 38, 71]. The diversification of function, with transcriptional (nuclear) and post-transcriptional (mostly cytoplasmic) mechanisms allowing the inactivation of genome-integrated as well as extra-chromosomal parasitic sequences, could account for the ancient conservation of duplicated RNAi components [15, 25]. Consistent with this hypothesis, taxonomically diverse RNAi-defective mutants commonly show reactivation of transposons and/or certain repetitive sequences [4, 29, 32, 72–75].

In contrast, an ancestral role of RNAi in the regulation of crucial cellular functions seems unlikely since the RNAi machinery does not appear to be essential for unicellular life, and it has been independently lost in several divergent eukaryotic lineages [15, 25, 28, 29, 76–78]. Yet, nearly all these lineages never had a paucity of transcription factors or chromatin-level regulators, suggesting that as a regulatory mechanism RNAi is probably lower in the hierarchy than the former two nuclear systems [15]. Supporting this interpretation, in organisms that contain a single Dicer gene such as *Schizosaccharomyces pombe* and vertebrates, Dicer null-mutants are RNAi-defective but viable at the cellular level [72, 73]. Likewise, mutants of the unicellular green alga *Chlamydomonas reinhardtii*, defective in small RNA biogenesis and lacking virtually all tested sRNAs, have no obvious abnormal phenotypes under standard laboratory conditions [79]. Moreover, miRNAs/endogenous siRNAs appear to have a very limited role in gene regulation in this unicellular species [79, 80] as well as in *S. pombe* [81].

Despite several hypotheses [28, 29, 78], the exact reason for the loss of RNAi in some eukaryotic lineages remains unclear, particularly because of the existence of sister species with similar lifestyles that differ in its presence or absence (e.g., *Naumovozyma castellii* vs. *Saccharomyces cerevisiae* or *T. brucei* vs. *T. cruzi*) [15]. Conversely, there is a strong tendency to retain the RNAi machinery in certain lineages (e.g., plants and animals), possibly associated with greater organizational complexity [31, 77, 82–84]. Indeed, miRNA-mediated regulation plays key roles in maintaining cell/tissue differentiation in multicellular plants and animals, and disruption in miRNA production leads to abnormal development or is lethal to these organisms [2, 4, 31, 36, 85]. It is therefore tempting to speculate that novel classes of endogenous regulatory sRNAs (such as animal and plant miRNAs) appeared later during eukaryotic evolution but became essential for controlling newly arisen processes such as cell differentiation and development in specific lineages.

As previously proposed [38, 77, 83, 86], inverted repeats, transcribed into RNA hairpins, may serve as a pool of novel miRNA genes that have the potential to

acquire regulatory function. The RNAi machinery would process transcribed RNA hairpins generating endogenous sRNAs and potential miRNAs. As eukaryotic genomes evolve, inverted repeats may randomly arise, but these hairpins would evolve neutrally unless a processed sRNA confers a selective advantage or disadvantage to the cell. As a result, miRNA precursors could arise, evolve accumulating mutations over time and eventually be lost from the genome. Under this scenario, conserved animal and plant miRNAs would represent versions that became 'fixed' as crucial gene regulators involved in unique evolutionary adaptations in each lineage, also making the RNAi machinery essential for organismal development.

At present at least five eukaryotic lineages (animals, slime molds, land plants, chlorophyte green algae and brown algae) are known to possess bona fide but unrelated miRNAs, suggesting that miRNAs have evolved independently through co-option of their shared inherited RNAi machinery [26, 34, 77, 79, 84, 86–88]. Additionally, within land plants or animals, only a minority of annotated miRNAs are conserved among all species, while the majority are family- or species-specific, suggesting that most known miRNA genes arose relatively recently in evolutionary time [77, 83, 86, 87]. MicroRNAs and multiple kinds of endogenous sRNAs may be flexible innovations that allowed gathering the selectivity of RNAi for the regulation of gene expression and other processes in specific eukaryotic lineages. Remarkably, a considerable degree of functional diversification of RNAi components and pathways seems to have occurred during eukaryotic evolution [1, 4, 5, 25, 27, 30, 32, 36, 63]. The archaeal AGO-PIWIs, most closely related to the putative ancestral eukaryotic AGO-PIWI, appear to function as stand-alone proteins involved in host defense [3]. In contrast, the eukaryotic AGO-PIWIs from multicellular plants and animals have now evolved to associate with multiple proteins in a variety of sRNA pathways and regulatory networks and contain unique insertion segments that likely provide binding surfaces for complex subunit interactions [2, 3, 6, 7, 17].

5 Perspective

A relatively complex RNAi machinery appears to have evolved in the last common ancestor of eukaryotes [25]. The original role of these components was likely in defense responses against selfish genetic elements such as viruses and transposable elements. From a mechanistic perspective, the ancestral RNAi machinery may have been capable of both transcriptional and post-transcriptional repression, roles widespread among extant eukaryotes, and this diversification of function could account for the evolutionary conservation of duplicated RNAi components, particularly AGO-PIWI proteins. The existence of a proto-miRNA system at an early stage of evolution (presumably based on transcribed hairpins processed by an siRNA-like pathway) cannot be ruled out [38], but it seems unlikely to have played a key role in the regulation of vital genes since RNAi is not essential for unicellular life. MicroRNAs as well as many other endogenous sRNAs may have evolved and

become ‘fixed’ more recently as important regulators in lineage-specific biological processes. Indeed, the introduction of miRNAs as post-transcriptional regulators of gene expression may have been a major innovation at the dawn of multicellular life [77, 88]. RNAi operates in a myriad of eukaryotes, and recent advances in sequencing technology have allowed the identification of numerous distinct small RNAs and pathways. However, except for a select group of well-studied organisms, the biological role(s) of the majority of these small RNAs remains largely uncharacterized. Particularly the impact of non-conserved, lineage-specific and lowly expressed miRNAs is virtually unknown. Are most of them generated from transient, neutrally evolving genes without a biological role? Do the miRNA-like sequences identified in several unicellular protists have a biological function? If so, how do they integrate within gene regulatory networks? Solving these questions may provide crucial information toward understanding the origin and evolution of sRNA-dependent regulatory systems.

Acknowledgments JAC-M is supported by a Young Investigator grant from the São Paulo Research Foundation (FAPESP 2011/50483-2). This work was supported in part by a grant from the National Science Foundation (to HC). We apologize to all researchers whose contributions could not be cited because of space limitations.

References

1. Cerutti H, Ma X, Msanne J, Repas T. RNA-mediated silencing in Algae: biological roles and tools for analysis of gene function. Eukaryot Cell. 2011;10:1164–72. doi:10.1128/EC.05106-11.
2. Poulsen C, Vaucheret H, Brodersen P. Lessons on RNA silencing mechanisms in plants from eukaryotic argonaute structures. Plant Cell. 2013;25:22–37. doi:10.1105/tpc.112.105643.
3. Swarts DC, Makarova K, Wang Y, Nakanishi K, Ketting RF, Koonin EV, Patel DJ, van der Oost J. The evolutionary journey of Argonaute proteins. Nat Struct Mol Biol. 2014;21:743–53. doi:10.1038/nsmb.2879.
4. Borges F, Martienssen RA. The expanding world of small RNAs in plants. Nat Rev Mol Cell Biol. 2015;16:727–41. doi:10.1038/nrm4085.
5. Ipsaro JJ, Joshua-Tor L. From guide to target: molecular insights into eukaryotic RNA-interference machinery. Nat Struct Mol Biol. 2015;22:20–8. doi:10.1038/nsmb.2931.
6. Iwakawa HO, Tomari Y. The functions of microRNAs: mRNA decay and translational repression. Trends Cell Biol. 2015;25:651–65. doi:10.1016/j.tcb.2015.07.011.
7. Jonas S, Izaurralde E. Towards a molecular understanding of microRNA-mediated gene silencing. Nat Rev Genet. 2015;16:421–33. doi:10.1038/nrg3965.
8. Leung AK. The whereabouts of microRNA actions: cytoplasm and beyond. Trends Cell Biol. 2015;25:601–10. doi:10.1016/j.tcb.2015.07.005.
9. Wilczynska A, Bushell M. The complexity of miRNA-mediated repression. Cell Death Differ. 2015;22:22–33. doi:10.1038/cdd.2014.112.
10. Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature. 1998;391:806–11. doi:10.1038/35888.
11. Chu Y, Yue X, Younger ST, Janowski BA, Corey DR. Involvement of argonaute proteins in gene silencing and activation by RNAs complementary to a non-coding transcript at the

progesterone receptor promoter. Nucleic Acids Res. 2010;38:7736–48. doi:10.1093/nar/gkq648.
12. Mortensen RD, Serra M, Steitz JA, Vasudevan S. Posttranscriptional activation of gene expression in Xenopus laevis oocytes by microRNA-protein complexes (microRNPs). Proc Natl Acad Sci USA. 2011;108:8281–6. doi:10.1073/pnas.1105401108.
13. Gagnon KT, Li L, Chu Y, Janowski BA, Corey DR. RNAi factors are present and active in human cell nuclei. Cell Rep. 2014;6:211–21. doi:10.1016/j.celrep.2013.12.013.
14. Guo D, Barry L, Lin SS, Huang V, Li LC. RNAa in action: from the exception to the norm. RNA Biol. 2014;11:1221–5. doi:10.4161/15476286.2014.972853.
15. Burroughs AM, Ando Y, Aravind L. New perspectives on the diversification of the RNA interference system: insights from comparative genomics and small RNA sequencing. Wiley Interdiscip Rev RNA. 2014;5:141–81. doi:10.1002/wrna.1210.
16. Peng JC, Lin H. Beyond transposons: the epigenetic and somatic functions of the Piwi-piRNA mechanism. Curr Opin Cell Biol. 2013;25:190–4. doi:10.1016/j.ceb.2013.01.010.
17. Dueck A, Meister G. Assembly and function of small RNA—argonaute protein complexes. Biol Chem. 2014;395:611–29. doi:10.1515/hsz-2014-0116.
18. Schirle NT, MacRae IJ. The crystal structure of human Argonaute2. Science. 2012;336:1037–40. doi:10.1126/science.1221551.
19. Sasaki HM, Tomari Y. The true core of RNA silencing revealed. Nat Struct Mol Biol. 2012;19:657–60. doi:10.1038/nsmb.2302.
20. Faehnle CR, Elkayam E, Haase AD, Hannon GJ, Joshua-Tor L. The making of a slicer: activation of human Argonaute-1. Cell Rep. 2013;3:1901–9. doi:10.1016/j.celrep.2013.05.033.
21. Hauptmann J, Dueck A, Harlander S, Pfaff J, Merkl R, Meister G. Turning catalytically inactive human Argonaute proteins into active slicer enzymes. Nat Struct Mol Biol. 2013;20:814–7. doi:10.1038/nsmb.2577.
22. Nakanishi K, Ascano M, Gogakos T, Ishibe-Murakami S, Serganov AA, Briskin D, Morozov P, Tuschl T, Patel DJ. Eukaryote-specific insertion elements control human Argonaute slicer activity. Cell Rep. 2013;3:1893–900. doi:10.1016/j.celrep.2013.06.010.
23. Sheng G, Zhao H, Wang J, Rao Y, Tian W, Swarts DC, van der Oost J, Patel DJ, Wang Y. Structure-based cleavage mechanism of Thermus thermophilus Argonaute DNA guide strand-mediated DNA target cleavage. Proc Natl Acad Sci USA. 2014;111:652–7. doi:10.1073/pnas.1321032111.
24. Baulcombe D. RNA silencing in plants. Nature. 2004;431:356–63. doi:10.1038/nature02874.
25. Cerutti H, Casas-Mollano JA. On the origin and functions of RNA-mediated silencing: from protists to man. Curr Genet. 2006;50:81–99. doi:10.1007/s00294-006-0078-x.
26. Bartel DP. MicroRNAs: target recognition and regulatory functions. Cell. 2009;136:215–33. doi:10.1016/j.cell.2009.01.002.
27. Dang Y, Yang Q, Xue Z, Liu Y. RNA interference in fungi: pathways, functions, and applications. Eukaryot Cell. 2011;10:1148–55. doi:10.1128/EC.05109-11.
28. Billmyre RB, Calo S, Feretzaki M, Wang X, Heitman J. RNAi function, diversity, and loss in the fungal kingdom. Chromosome Res. 2013;21:561–72. doi:10.1007/s10577-013-9388-2.
29. Nicolas FE, Torres-Martinez S, Ruiz-Vazquez RM. Loss and retention of RNA interference in fungi and parasites. PLoS Pathog. 2013;9:e1003089. doi:10.1371/journal.ppat.1003089.
30. Hoogstrate SW, Volkers RJ, Sterken MG, Kammenga JE, Snoek LB. Nematode endogenous small RNA pathways. Worm. 2014;3:e28234. doi:10.4161/worm.28234.
31. Posadas DM, Carthew RW. MicroRNAs and their roles in developmental canalization. Curr Opin Genet Dev. 2014;27:1–6. doi:10.1016/j.gde.2014.03.005.
32. Gebert D, Rosenkranz D. RNA-based regulation of transposon expression. Wiley Interdiscip Rev RNA. 2015;6:687–708. doi:10.1002/wrna.1310.
33. Malone CD, Hannon GJ. Small RNAs as guardians of the genome. Cell. 2009;136:656–68. doi:10.1016/j.cell.2009.01.045.
34. Axtell MJ, Westholm JO, Lai EC. Vive la difference: biogenesis and evolution of microRNAs in plants and animals. Genome Biol. 2011;12:221. doi:10.1186/gb-2011-12-4-221.

35. Ishizu H, Siomi H, Siomi MC. Biology of PIWI-interacting RNAs: new insights into biogenesis and function inside and outside of germlines. Genes Dev. 2012;26:2361–73. doi:10.1101/gad.203786.112.
36. Rogers K, Chen X. Biogenesis, turnover, and mode of action of plant microRNAs. Plant Cell. 2013;25:2383–99. doi:10.1105/tpc.113.113159.
37. Ketting RF. The many faces of RNAi. Dev Cell. 2011;20:148–61. doi:10.1016/j.devcel.2011.01.012.
38. Shabalina SA, Koonin EV. Origins and evolution of eukaryotic RNA interference. Trends Ecol Evol. 2008;23:578–87. doi:10.1016/j.tree.2008.06.005.
39. Mukherjee K, Campos H, Kolaczkowski B. Evolution of animal and plant dicers: early parallel duplications and recurrent adaptation of antiviral RNA binding in plants. Mol Biol Evol. 2013;30:627–41. doi:10.1093/molbev/mss263.
40. Letunic I, Doerks T, Bork P. SMART: recent updates, new developments and status in 2015. Nucleic Acids Res. 2015;43:D257–60. doi:10.1093/nar/gku949.
41. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, Potter SC, Punta M, Qureshi M, Sangrador-Vegas A, Salazar GA, Tate J, Bateman A. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res. 2016;44:D279–85. doi:10.1093/nar/gkv1344.
42. Nakanishi K, Weinberg DE, Bartel DP, Patel DJ. Structure of yeast Argonaute with guide RNA. Nature. 2012;486:368–74. doi:10.1038/nature11211.
43. Burroughs AM, Iyer LM, Aravind L. Two novel PIWI families: roles in inter-genomic conflicts in bacteria and Mediator-dependent modulation of transcription in eukaryotes. Biol Direct. 2013;8:13. doi:10.1186/1745-6150-8-13.
44. Aravind L, Walker DR, Koonin EV. Conserved domains in DNA repair proteins and evolution of repair systems. Nucleic Acids Res. 1999;27:1223–42.
45. Makarova KS, Wolf YI, Koonin EV. Comparative genomics of defense systems in archaea and bacteria. Nucleic Acids Res. 2013;41:4360–77. doi:10.1093/nar/gkt157.
46. Makarova KS, Wolf YI, van der Oost J, Koonin EV. Prokaryotic homologs of Argonaute proteins are predicted to function as key components of a novel system of defense against mobile genetic elements. Biol Direct. 2009;4:29. doi:10.1186/1745-6150-4-29.
47. Hock J, Meister G. The Argonaute protein family. Genome Biol. 2008;9:210. doi:10.1186/gb-2008-9-2-210.
48. Pak J, Fire A. Distinct populations of primary and secondary effectors during RNAi in C. elegans. Science. 2007; 315:241–244. doi:10.1126/science.1132839.
49. Sijen T, Steiner FA, Thijssen KL, Plasterk RH. Secondary siRNAs result from unprimed RNA synthesis and form a distinct class. Science. 2007;315:244–7. doi:10.1126/science.1136699.
50. Wedeles CJ, Wu MZ, Claycomb JM. Protection of germline gene expression by the C. elegans Argonaute CSR-1. Dev Cell. 2013;27:664–71. doi:10.1016/j.devcel.2013.11.016.
51. Meister G. Argonaute proteins: functional insights and emerging roles. Nat Rev Genet. 2013;14:447–59. doi:10.1038/nrg3462.
52. Swart EC, Nowacki M. The eukaryotic way to defend and edit genomes by sRNA-targeted DNA deletion. Ann NY Acad Sci. 2015;1341:106–14. doi:10.1111/nyas.12636.
53. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. Clustal W and Clustal X version 2.0. Bioinformatics. 2007;23:2947–8. doi:10.1093/bioinformatics/btm404.
54. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 1987;4:406–25.
55. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol. 2007;24:1596–9. doi:10.1093/molbev/msm092.
56. Macrae IJ, Zhou K, Li F, Repic A, Brooks AN, Cande WZ, Adams PD, Doudna JA. Structural basis for double-stranded RNA processing by Dicer. Science. 2006;311:195–8. doi:10.1126/science.1121638.

57. Patrick KL, Shi H, Kolev NG, Ersfeld K, Tschudi C, Ullu E. Distinct and overlapping roles for two Dicer-like proteins in the RNA interference pathways of the ancient eukaryote Trypanosoma brucei. Proc Natl Acad Sci USA. 2009;106:17933–8. doi:10.1073/pnas.0907766106.
58. Zhang H, Kolb FA, Jaskiewicz L, Westhof E, Filipowicz W. Single processing center models for human Dicer and bacterial RNase III. Cell. 2004;118:57–68. doi:10.1016/j.cell.2004.06.017.
59. Drinnenberg IA, Weinberg DE, Xie KT, Mower JP, Wolfe KH, Fink GR, Bartel DP. RNAi in budding yeast. Science. 2009;326:544–50. doi:10.1126/science.1176945.
60. Weinberg DE, Nakanishi K, Patel DJ, Bartel DP. The inside-out mechanism of Dicers from budding yeasts. Cell. 2011;146:262–76. doi:10.1016/j.cell.2011.06.021.
61. Shi H, Tschudi C, Ullu E. An unusual Dicer-like1 protein fuels the RNA interference pathway in Trypanosoma brucei. RNA. 2006;12:2063–72. doi:10.1261/rna.246906.
62. Hu Y, Stenlid J, Elfstrand M, Olson A. Evolution of RNA interference proteins dicer and argonaute in Basidiomycota. Mycologia. 2013;105:1489–98. doi:10.3852/13-171.
63. Sandoval PY, Swart EC, Arambasic M, Nowacki M. Functional diversification of Dicer-like proteins and small RNAs required for genome sculpting. Dev Cell. 2014;28:174–88. doi:10.1016/j.devcel.2013.12.010.
64. Zong J, Yao X, Yin J, Zhang D, Ma H. Evolution of the RNA-dependent RNA polymerase (RdRP) genes: duplications and possible losses before and after the divergence of major eukaryotic groups. Gene. 2009;447:29–39. doi:10.1016/j.gene.2009.07.004.
65. Hammond TM, Keller NP. RNA silencing in Aspergillus nidulans is independent of RNA-dependent RNA polymerases. Genetics. 2005;169:607–17. doi:10.1534/genetics.104.035964.
66. Yang JS, Lai EC. Dicer-independent, Ago2-mediated microRNA biogenesis in vertebrates. Cell Cycle. 2010;9:4455–60.
67. Yang JS, Maurin T, Lai EC. Functional parameters of Dicer-independent microRNA biogenesis. RNA. 2012;18:945–57. doi:10.1261/rna.032938.112.
68. Olovnikov I, Chan K, Sachidanandam R, Newman DK, Aravin AA. Bacterial argonaute samples the transcriptome to identify foreign DNA. Mol Cell. 2013;51:594–605. doi:10.1016/j.molcel.2013.08.014.
69. Swarts DC, Jore MM, Westra ER, Zhu Y, Janssen JH, Snijders AP, Wang Y, Patel DJ, Berenguer J, Brouns SJ, van der Oost J. DNA-guided DNA interference by a prokaryotic Argonaute. Nature. 2014;507:258–61. doi:10.1038/nature12971.
70. Zander A, Holzmeister P, Klose D, Tinnefeld P, Grohmann D. Single-molecule FRET supports the two-state model of Argonaute action. RNA Biol. 2014;11:45–56. doi:10.4161/rna.27446.
71. Ma X, Kim EJ, Kook I, Ma F, Voshall A, Moriyama E, Cerutti H. Small interfering RNA-mediated translation repression alters ribosome sensitivity to inhibition by cycloheximide in Chlamydomonas reinhardtii. Plant Cell. 2013;25:985–98. doi:10.1105/tpc.113.109256.
72. Kanellopoulou C, Muljo SA, Kung AL, Ganesan S, Drapkin R, Jenuwein T, Livingston DM, Rajewsky K. Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. Genes Dev. 2005;19:489–501. doi:10.1101/gad.1248505.
73. Martienssen RA, Zaratiegui M, Goto DB. RNA interference and heterochromatin in the fission yeast Schizosaccharomyces pombe. Trends Genet. 2005;21:450–6. doi:10.1016/j.tig.2005.06.005.
74. Casas-Mollano JA, Rohr J, Kim EJ, Balassa E, van Dijk K, Cerutti H. Diversification of the core RNA interference machinery in Chlamydomonas reinhardtii and the role of DCL1 in transposon silencing. Genetics. 2008;179:69–81. doi:10.1534/genetics.107.086546.
75. Wang X, Hsueh YP, Li W, Floyd A, Skalsky R, Heitman J. Sex-induced silencing defends the genome of Cryptococcus neoformans via RNAi. Genes Dev. 2010;24:2566–82. doi:10.1101/gad.1970910.

76. Nakayashiki H, Kadotani N, Mayama S. Evolution and diversification of RNA silencing proteins in fungi. J Mol Evol. 2006;63:127–35. doi:10.1007/s00239-005-0257-2.
77. Berezikov E. Evolution of microRNA diversity and regulation in animals. Nat Rev Genet. 2011;12:846–60. doi:10.1038/nrg3079.
78. Drinnenberg IA, Fink GR, Bartel DP. Compatibility with killer explains the rise of RNAi-deficient fungi. Science. 2011;333:1592. doi:10.1126/science.1209575.
79. Voshall A, Kim EJ, Ma X, Moriyama EN, Cerutti H. Identification of AGO3-associated miRNAs and computational prediction of their targets in the green alga Chlamydomonas reinhardtii. Genetics. 2015;200:105–21. doi:10.1534/genetics.115.174797.
80. Yamasaki T, Kim EJ, Cerutti H, Ohama T. Argonaute3 is a key player in miRNA-mediated target cleavage and translational repression in Chlamydomonas. Plant J. 2015;. doi:10.1111/tpj.13107.
81. Smialowska A, Djupedal I, Wang J, Kylsten P, Swoboda P, Ekwall K. RNAi mediates post-transcriptional repression of gene expression in fission yeast Schizosaccharomyces pombe. Biochem Biophys Res Commun. 2014;444:254–9. doi:10.1016/j.bbrc.2014.01.057.
82. Grimson A, Srivastava M, Fahey B, Woodcroft BJ, Chiang HR, King N, Degnan BM, Rokhsar DS, Bartel DP. Early origins and evolution of microRNAs and Piwi-interacting RNAs in animals. Nature. 2008;455:1193–7. doi:10.1038/nature07415.
83. Cuperus JT, Fahlgren N, Carrington JC. Evolution and functional diversification of MIRNA genes. Plant Cell. 2011;23:431–42. doi:10.1105/tpc.110.082784.
84. Lenz D, May P, Walther D. Comparative analysis of miRNAs and their targets across four plant species. BMC Res Notes. 2011;4:483. doi:10.1186/1756-0500-4-483.
85. Sheng Y, Previti C. Genomic features and computational identification of human microRNAs under long-range developmental regulation. BMC Genom. 2011;12:270. doi:10.1186/1471-2164-12-270.
86. Nozawa M, Miura S, Nei M. Origins and evolution of microRNA genes in plant species. Genome Biol Evol. 2012;4:230–9. doi:10.1093/gbe/evs002.
87. Tarver JE, Donoghue PC, Peterson KJ. Do miRNAs have a deep evolutionary history? BioEssays. 2012;34:857–66. doi:10.1002/bies.201200055.
88. Tarver JE, Cormier A, Pinzon N, Taylor RS, Carre W, Strittmatter M, Seitz H, Coelho SM, Cock JM. microRNAs and the evolution of complex multicellularity: identification of a large, diverse complement of microRNAs in the brown alga Ectocarpus. Nucleic Acids Res. 2015;43:6384–98. doi:10.1093/nar/gkv578.

Evolution of the Molecules Coupling mRNA Transport with Translational Control in Metazoans

Paula Vazquez-Pianzola, Beat Suter and Greco Hernández

1 Introduction

Eukaryotes arose from ancestral prokaryotes as a result of profound evolutionary changes at the molecular, metabolic, and morphological levels. These changes resulted in the emergence of novel and more sophisticated levels of cellular architecture. An essential structure of eukaryotes is the cytoskeleton, whose evolution from prokaryotic cytoskeleton proteins allowed novel and fundamental processes such as mitosis, meiosis, inheritance of genetic material, and cellular motility to evolve [1–7]. The emergence of the cytoskeleton also led to the evolution of motors driving intracellular transport to discrete regions of a cell, and these motors are capable of transporting an amazing variety of different cargos, ranging from vesicles and organelles to a plethora of proteins and RNAs required for most cellular processes [1, 2, 7–10]. mRNA transport coupled with translation emerged as a key process of gene expression that targets protein synthesis to specific compartments of cells. In this process, motors act in concert with the cytoskeleton to assemble, stabilize, and transport mRNAs, and this process is also coupled with the control of translation. During their journey translation of mRNAs is repressed, and it is only activated once the mRNAs reach their final destination [11–15].

In this chapter, we will review recent findings that shed new light on the evolution of the molecules involved in translational control of transported mRNAs. To date, regulation of gene expression involving this phenomenon is known for diverse transcripts, and the transport motors as well as diverse proteins involved in this

P. Vazquez-Pianzola (✉) · B. Suter
Institute of Cell Biology, University of Bern, Baltzerstrasse 4, 3012 Bern, Switzerland
e-mail: paula.vazquez@izb.unibe.ch

G. Hernández (✉)
Division of Basic Research, National Institute of Cancer (INCan),
22 San Fernando Ave., Tlalpan, 14080 Mexico City, Mexico
e-mail: greco.hernandez@gmail.com

G. Hernández and R. Jagus (eds.), *Evolution of the Protein Synthesis Machinery and Its Regulation*, DOI 10.1007/978-3-319-39468-8_21

process have been characterized to a good extent. While progress has been made across eukaryotes, we will put special emphasis on the processes described in metazoans.

2 mRNA Localization Is Coupled with Local Protein Synthesis in Metazoans

Translational control of asymmetrically localized mRNAs allows cells to determine the precise time and place when a protein is synthesized. Upon translational activation, the proteins can be synthesized rapidly because transcription is not required anymore. This posttranscriptional gene expression control underlies many biological processes in metazoans, such as germline development, embryonic axis specification, and embryonic patterning. Additionally, it contributes to various cell differentiation processes, including neurogenesis and synaptic transmission. Some of the first examples of gene expression regulation involving translational control of localized mRNAs were described while studying the embryonic development of *Drosophila* and *Xenopus* [11–17]. The *Drosophila* processes take place during oogenesis and embryogenesis and include *bicoid* (*bcd*), *oskar* (*osk*), and *nanos* (*nos*) mRNAs, which encode the maternal polarity determinants that localize to the anterior and to the posterior cortex of the oocyte, respectively. Their correct localization and translation are crucial for the antero-posterior axis specification of the embryo. Similarly, localization of *gurken* (*grk*) mRNA to the dorso-anterior corner of the oocyte is essential for egg chamber and embryo dorso-ventral axis specification [11–13, 16–18]. Likewise, the early examples from *Xenopus* described that mRNAs encoding the T-box transcription factor VegT and transforming growth factor-beta (TGF-ß) family member, Vg1, localize to the vegetal pole cortex of oocytes and play critical roles in endodermal and mesodermal specification during early embryogenesis [19].

Localization of mRNAs appears to be a generalized phenomenon in metazoans, since a growing number of mRNAs have also been reported to localize in oocytes, eggs, and cleaving embryos of diverse species of vertebrates, cnidarians, and arthropods [13, 20–23]. Indeed, two high-throughput in situ screens in *Drosophila* ovaries and embryos covering 1/4 − 1/3 of the transcriptome revealed that ∼35 and ∼70 % of mRNAs exhibit a differential localization pattern in the developing ovary and embryo, respectively [20, 22, 24]. An extensive correlation between mRNA localization and protein distribution was also observed in embryos, indicating that translation control is tightly regulated during mRNA transport. The high abundance of mRNA localization strongly suggests that most cellular processes are somehow impacted by mRNA localization coupled to translational control [20, 24]. The evolutionary widespread occurrence of transport and translational control of many different mRNAs also illustrates the crucial role of this process in metazoans. For example, in mammalian mesenchymal-like cells, the establishment of front-back

polarity at the proteome level is maintained by localized translation of mRNAs [25]. Surprisingly, although many mRNAs are asymmetrically localized between the cell body and the protrusions, in this study no correlation was detected between the distribution of mRNAs and the corresponding proteins, suggesting that mRNA localization alone is not a significant predictor of protein localization. Differential distribution of mRNAs in polarized cells may be a mean to store repressed mRNAs in order to rapidly activate translation on site upon specific stimuli [25].

Many of the factors involved in coupling transport with translation of homologous mRNAs are conserved in different species. For example, a 54-nucleotide cytoplasmic localization element of the 3′-UTR of *β-actin* mRNA (termed a zipcode) is recognized by zipcode-binding protein 1 (ZBP1), a step that is required for carrying and translating mRNA to lamellipodia of chicken fibroblasts. This event produces an enrichment of actin at the leading edge of cells, which is required for cell motility [26]. Similar phenomena of localized *β-actin* mRNA have been observed for different cell lines from several vertebrates, including developing neurons of rat hippocampus and the *Xenopus* retinal axons where its translation might also be regulated by ZBP1 [14, 27–30]. Moreover, it has been found that ZBP1 inhibits mRNA translation by preventing 80S ribosomal complex formation [31]. In *Drosophila*, Fragile X Mental Retardation Protein (FMR1) also exerts translational control on localized mRNAs. FMR1 forms a complex with Argonaute 2 (AGO2), an essential component of the RNA-induced silencing complex (RISC) [32], and with the ribosome to directly block translation by inhibiting tRNA association [33]. FMR1 is also able to function as a translational activator [34]. dFMR1 not only regulates translation but also controls the efficacy of mRNA transport in neurons [35] In mammalian neurons, FMRP colocalizes and coimmunoprecipitates with subsets of dendritically localized mRNAs [36]. FMRP knockdown enhances protein synthesis of some localized mRNAs in mice and interferes with DHPG trafficking of specific mRNAs in neurons, indicating that FMRP promotes transport and regulation of local translation of mRNAs at the synapses [35, 36].

3 Origin of Cytoskeleton and Molecular Motors

The highly sophisticated organization of eukaryotic cells was made possible by the evolutionary emergence of protein motors that facilitate trafficking between different cellular compartments. Molecular motors carry a plethora of cargoes such as RNAs, proteins, organelles, and diverse macromolecular complexes to a variety of destinations within the cytoplasm. To do so, motors travel directionally along the tracks of a dynamic and extremely elaborate system of intracellular polymers termed a cytoskeleton, which is also responsible for maintaining the shape and the mechanical dynamics of the cell [37].

In all extant eukaryotes, cytoskeletal elements involved in the transport of cargo consist of two major types of structural components: tubulins form microtubules

(MT) [37, 38], whereas actins form actin filaments (AF) [37, 39]. Unlike MT and AF, a third class of cytoskeletal elements, the intermediate filaments (IFs), lack structural directionality and cytomotiliy, and no motor proteins have been found associated with them [37, 40].

MT and AF evolved from prokaryotic homolog filaments. Indeed, both bacteria and archaea are endowed with cytomotive cytoskeletons that can function as motors because of the kinetics of polymerization/depolymerization itself [41]. Bacteria and archaea possess genes encoding clear homologs of tubulin and actin, namely FtsZ, TubZ, and RepX for tubulin (being FtsZ the nearest extant relative), and MreB and FtsA families for actin, the latter playing critical roles in prokaryotic plasmid segregation and cell shape and septation [1–4, 6, 7, 42, 43]. Regarding the origin of cytoskeleton proteins, on one hand highly conserved orthologs of tubulins have only been found in the genomes of archaeal species of the *phylum* Thaumarchaeota [1, 5, 7, 44, 45]. On the other hand, actin and its prokaryotic homologs MreB and FtsA belong to a large superfamily of ATPases present in all three domains of life. Recently, phylogenomic analyses have discovered proteins with high similarity to eukaryotic actins in archaeal species of the *phylum* Crenarchaeota. Accordingly, they are dubbed "crenactins." Altogether, these findings support the emerging view that the two major components of the eukaryotic cytoskeleton have archaeal origins [5, 43] and that the *l*ast *c*ommon *e*ukaryotic *a*ncestor (LECA) possessed an established, complex cytoskeleton composed of multiple paralogs of the tubulin (FtsZ/TubZ) and actin (MreB/crenactin) families of proteins [1–7, 42, 43].

It appears that the ability of cargo-carrying molecules was strongly augmented in eukaryotes by the mergence in early eukaryotic evolution of molecular motors that function in coordination with the cytoskeleton [1, 3, 7, 8, 46]. However, it is intriguing that prokaryotes possess only cytoskeletal cytomotive polymers while no good candidate motor protein has been found yet. Eukaryotes evolved three major superfamilies of motors that drive transport of mRNA cargoes. These are kinesins and dyneins, which work along MTs, and myosins that work along AFs. Thus, numerous kinesins, myosins, and dyneins have evolved to cope with the much more sophisticated needs that have arisen during eukaryotic evolution. Even though we do not know the origin of motors, kinesins and myosins share a common ancestor. Dyneins belong to the large AAA+ superfamily of proteins and most likely evolved from multiple duplication events of a single AAA+ domain before LECA. Some evidence suggests that the closest relative prokaryotic protein is MoxR, but it does not possess any motor activity. The ubiquitous distribution of different paralogous proteins of all three motors across eukaryotes supports the notion that LECA already possessed several families of all three motor types working along with an established cytoskeleton. However, multiple losses of paralogs of the three families of motors happened during eukaryotic diversification. After eukaryotes emerged, the ancient "toolbox" of motors expanded into a wide battery of motors coupled with different and additional cargo-bound "receptor" proteins, each designed to carry distinct and specific cargoes [1–4, 7, 9, 10].

4 The *Drosophila* BicD/Egl/Dynein Machinery Paradigm

The *Drosophila* BicD/Egl/Dyn complex is arguably one of the best-studied mRNA transport machineries. It plays a key role in oogenesis and embryogenesis by localizing a plethora of mRNAs required for cell determination, differentiation, and formation of the anterior-posterior and dorsal-ventral axes. This machinery is composed of BicD and Egalitarian (Egl) proteins, which interact with the motor dynein (Dyn)/dynactin to transport mRNA cargoes along the microtubule cytoskeleton to specific cellular compartments. To form the complex, Egl interacts directly with both BicD and Dyn, as well as with transported mRNAs [11, 13, 47–50]. BicD/Egl/dynein complex may work in conjunction with additional proteins that confer specificity and, at the same time, translational control.

Genetic and biochemical studies have provided evidence of the *Drosophila* BicD/Egl function in mRNA transport. During oogenesis, a single germline cell produces a cluster of 16 interconnected cells of which one differentiates into an oocyte. In parallel, the remaining 15 germline cells differentiate into nurse cells that provide the oocyte with all the material required for growth and differentiation. This process includes the transport into the oocyte of a subset of mRNAs produced in the nurse cells. *BicD* loss-of-function mutant females produce a germline that is composed only of cells with nurse cell appearance, indicating that *BicD* is essential for oocyte differentiation. *BicD* mutant egg chambers also fail to accumulate oocyte-specific mRNAs [such as *osk*, *orb, BicD* and *fs(1)K10*] in the oocyte. Thus, it is suggested that the loss of oocyte differentiation may be due to a failure in the transport of oocyte-specific proteins and mRNAs from the nurse cells into the oocyte [51, 52]. Ovaries mutant for *egl* as well as wild-type ovaries treated with microtubule disrupting drugs show the same 16-nurse-cell phenotype as *BicD* mutants [53, 54]. Studies using fluorescently labeled *grk* and *bcd* mRNAs injected into the nurse cells have shown that BicD and Egl are recruited to these mRNAs and that these genes are required for *grk* transport into the oocyte [55]. This study also revealed that transport along MTs requires Dyn for efficient localization of *grk*, *bcd*, and *osk* mRNAs from the nurse cells into the oocyte [55]. Moreover, the BicD/Egl/Dyn machinery is not only active in the germline, but is also used for the apical localization of *inscuteable* mRNA in neurons [56] and for apical localization of mRNAs from several segmentation genes in blastoderm embryos [57].

A recent NMR study on the *K10* mRNA localization signal showed that it folds in a special A'-form RNA conformation that is also found in the stem loops responsible for localization of other BicD/Egl targets, namely, *ftz, h, grk, wg, bcd, I-factor,* and *osk* mRNAs, suggesting that they are all recognized directly by Egl [58, 59]. However, whether Egl is a general link for all mRNAs transported by the BicD/Egl/Dyn machinery or whether other proteins are required for cargo specificity is not known.

5 The Importance of Being *Oskar*

Drosophila osk gene expression has been one of the most studied models of translation control during mRNA transport, becoming a paradigm for this phenomenon. Localization of *osk* mRNA to the posterior of the oocyte proceeds by the action of the BicD/Egl/Dyn transport motor that imports the mRNA from nurse cells into the oocyte [55, 57, 60–64]. Then, *osk* mRNA switches to a kinesin-based motor that transports it towards the posterior cortex. Kinesin heavy chain (KHC) and the kinesin light chain (KLC)-like protein PAT1 are required for this process. While kinesin is involved in the long-range MT-based transport of *osk* mRNA throughout the oocyte, there is evidence that *osk* mRNA localization is followed by a myosin-V-dependent short-range actomyosin translocation of *osk* mRNA at the posterior cortex [65].

During its journey, the translation of *osk* mRNA is repressed until it reaches its final destination at the oocyte posterior cortex after stage 8 of oogenesis. Mutants in *armitage* (*armi*), *aubergine* (*aub*), *spindle-E* (*spn-E*), *maelstrom* (*mael*) [66], *zucchini* (*zuc*), *squah* (*squ*) [67], and *krimper* (*krimp*) [68] show premature translation of *osk* mRNA in the oocyte during early oogenesis. It is therefore possible that translational silencing of *osk* mRNA during these stages is driven by piRNA-Piwi-Argonaute complexes interacting with *osk* mRNA. Alternatively, the reduced activity of any of these proteins coupled with the higher expression of mobile genetic elements might titrate the repressors of *osk* mRNA translation. Other proteins are also involved in exerting *osk* mRNA translational repression. As opposed to wild types, egg chambers mutant for the *Maternal expression at 31B* (*Me31B*) gene show ectopic Osk protein accumulation in the nurse cells rather than in the oocyte during early oogenesis, indicating that Me31Bs normally repress *osk* translation during its transport through the nurse cell into the oocyte [69].

During oogenesis, polypyrimidine tract-binding protein (PTB) mediates assembly of high-order complexes containing multiple *osk* RNAs, and this causes translational silencing [70]. A complex made up by Bruno (Bru) and Cup represses cap-dependent translation of *osk* mRNA from stage 5-6 onwards [71]. Bru binds simultaneously to Bru-response elements (BRE) in *osk* 3′-UTR and to Cup, which in turn binds eIF4E, thereby inhibiting recruitment of the small ribosomal subunit to *osk* mRNA [71]. Accordingly, egg chambers expressing mutant Cup unable to bind eIF4E show precocious expression of *osk* mRNA in stages 6–9 as well as increased expression in stage 9 oocytes. Another mechanism is independent of the Cup-eIF4E interaction, but still depends on Bru. This one causes translation repression during mid oogenesis, and it also involves the formation of Bru-dependent *osk* mRNA oligomers, which, bound to Bru, form large silencing complexes that cannot be accessed by ribosomes [72]. Finally, the *Drosophila* hnRNPA/B homolog (hrp48) binds sequences in the *osk* 5′- and 3′-UTRs, being involved in localization and translational repression of *osk* mRNA after stage 9 of oogenesis [73].

Interestingly, Cup is also involved in translational repression of *grk* mRNA, which is also transported by the BicD/Egl complex. A model for translation

regulation of *grk* mRNA during its transport has been put forward in which both Cup and Bru also function in complex with Sqd, out, and Hrb27C/Hrp48 [74]. It was shown that that before *grk* RNA reaches its final destination at the dorsal-anterior region of the oocyte, a well-established translation factor, poly(A)-binding protein (PABP), functions with Encore (Enc) to facilitate translational activation of *grk* mRNA [74].

Our research group has reported that *Drosophila Pabp* interacts genetically and biochemically with *BicD* and that the biochemical interaction depends on RNA [75]. *Pabp* mutants show both reduced stability and mislocalization of *osk* mRNA during early oogenesis, demonstrating that PABP plays a key role in *osk* mRNA localization [75]. Although there is no evidence for PABP involvement in *osk* mRNA translational control during early oogenesis, it might be possible that PABP activates *osk* mRNA translation after it has reached its final destination during late oogenesis. All in all, Cup, Me31B, PTB, PABP, IMP, Bru, and Hrp48 are factors that can associate with the BicD/Egl/Dyn motor to regulate the fate and translation of *osk* and of other transported mRNAs as well.

6 Evolution of the BicD/Egl/Dyn Complex

Recent studies in the wasp *Nasonia vitripennis* have shown a conserved role of BicD in mRNA localization and organization of a polarized microtubule network during oogenesis in non-dipteran insects [76]. *Drosophila* and *Nasonia* share a similar germline development, even though they diverged over 200 million years ago. Although a role of BicD in mRNA transport in other *phyla* has not been described yet, BicD are coiled-coil protein adaptors linking the Dyn/dynactin minus-end-directed motor complex with different cargos [13, 47, 49, 77]. Because of this versatility *Drosophila* BicD does not only perform mRNA localization, but is also involved in the transport of other cargoes, such as clathrin, synaptic vesicles at the neuromuscular junction, lipid droplets, and even nuclei of photoreceptor cells and oocytes, [13, 47, 49, 77, 78].

A conserved role of BicD in neuronal development in other species is supported by several recent findings. Like *D. melanogaster BicD*, *C. elegans* BicD is also involved in nuclear migration and in neuron branching [79, 80], while the mammalian BicD1/Rab6 complex regulates COPI-independent Golgi-ER transport as well as retrograde membrane transport in human neurons [47, 49, 78]. Furthermore, *BicD2*-deficient mice show impaired radial neuronal migration [81], suggesting that *BicD2* is linked to cargo trafficking also in glial cells. In a similar way, mouse BicD1 was recently shown to modulate endosomal trafficking and signaling of ligand-activated neurotrophin receptors in motor neurons [82]. Furthermore, mutations in human *BicD2* have been shown to cause congenital autosomal-dominant spinal muscular atrophy and hereditary spastic paraplegia in humans [83–85]. These mutations cause BicD to bind more strongly to dynein/dynactin complexes and to produce Golgi fragmentation, which may result in defects in neuronal cargo

trafficking and impairment of neuron outgrowth. Altogether, these findings highlight the essential and conserved role of *BicD* in nervous system development and physiology. It appears to perform the same function across metazoans by regulating different cargo trafficking needed for polarizing nerve and glial cells.

The *BicD* gene is conserved throughout metazoans, but is not present in other eukaryotes. Like in *Drosophila*, mammalian orthologs of BicD bind directly to components of the Dyn and dynactin complexes [86]. While there is only one gene encoding BicD in insects, *C. elegans*, and some ascidians, the gene is duplicated in various vertebrates including humans. In the amphibian *Xenopus,* one *BicD1* and two *BicD2* homologs are present. Interestingly, the fishes *Danio rerio, Gasterosteus aculeatus, Oryzias latipes, Takifugu rubripes,* and *Tetraodon nigroviridis* have two homologs of the *BicD1* gene and two homologs of *Bic-D2*. In addition, in fishes there is also a third, deeply divergent gene, probably representing an ancestral version of the *BicD* gen. The sea lamprey *Petromyzon marinus* has also two *BicD* genes, one *BicD1* ortholog and one that could also be close to the ancestral *BicD* gene [13, 47, 87–89]. Two shorter BicD-related genes, BicDR1 and BicDR2, contain only two coiled coil regions and have been described in mammals and other vertebrates. These cognate genes are involved in neural development in *Zebrafish* [90]. Despite both genes being conserved in vertebrates, only BicDR1 is present in flies. Like BicD1/2, BicDR1 also binds Rab6. The highest degree of similarity to BicD is in the cargo-binding domain at the C-terminus of BicDR.

Egl is present in many arthropods and in *C. elegans*. As in *Drosophila*, studies on the giant tiger shrimp (*Penaeus monodon*) *egl* ortholog gene have suggested an involvement of Egl in ovary development as well [91]. In contrast to BicD, a clear Egl homolog has not been identified in mammals. Thus, it is possible that different, so-far unidentified adaptor proteins not related to Egl might link the BicD/Dyn localization motor to localized mRNAs in other *phyla* [13].

Several RNA-binding proteins present in *Drosophila* BicD complexes, such as PABP [75], FMRP [92], and the insulin-like growth factor II mRNA binding proteins (IMPs; Vazquez-Pianzola, Bullock and Suter, unpublished), are highly conserved across eukaryotes. Cytoplasmic PABP is a translation factor that is present in all eukaryotes and that has diversified in multiple gene families. Indeed, cytoplasmic PABP proteins are involved in different processes of RNA metabolism, including mRNA stability, transport, and translation [93–98]. To date, most functional studies have been focused on the prototype PABP1. However, the versatility and high number of genes encoding PABPs in different species point to the possibility that distinct PABPs might regulate the localization and/or translation of different localized mRNAs. Interestingly, PABPs have been found to bind not only to the poly(A) tail of mRNAs, but also to A-rich sequences in the UTRs of *osk*, *bcd*, and *Vasopressin* mRNAs, and their binding is critical for proper mRNA localization in *Drosophila* oocytes and mammalian neurons, respectively [75, 99–102]. These additional binding sites further contribute to the versatility of PABPs.

In addition to Egl, another RNA-binding protein has been reported to link mRNAs with BicD and FMR1 [92], and this function was implicated in branching

of the dendritic arbor [92]. FMRP regulates mRNA transport and also functions as a negative regulator of translation [36]. While the three vertebrate paralogs, FMR1, FXR1, and FXR2, share a conserved gene structure derived from a common ancestral gene, *Drosophila* and most invertebrates possess a single ortholog with high overall similarity to human FXR2 [103–105]. Thus, as opposed to Egl, FMRP is conserved in vertebrates and invertebrates opening the possibility that FMRP/BicD complexes regulate RNA transport and translation in higher eukaryotes.

IMPs form a family of RNA-binding proteins highly conserved across the animal kingdom. In *Drosophila*, IMP is required for translational control of localized *osk* and *grk* mRNAs [106, 107]. Chicken IMP1, also known as ZBP-1, is required for *beta-actin* mRNA localization and translational repression during transport to the leading edge of motile fibroblasts and neurons. In *Xenopus*, IMP is required for localization of *Vg1* mRNAs to the oocyte vegetal pole during maturation [108, 109]. Preliminary results from our laboratory indicate that *Drosophila* IMP also forms a complex with BicD/Egl during specific developmental stages (Vazquez-Pianzola, Bullock and Suter, unpublished). Because most invertebrates, including *D. melanogaster, C. elegans*, and different ascidian species, are endowed with only one *IMP* gene, whereas most vertebrates possess more than one paralog, it appears that the vertebrate IMP family originated from repeated gene duplications shortly after the divergence of these two lineages. Most vertebrates (i.e., humans, rats, mice, birds, and reptiles) contain three IMP paralogs, namely IMP1, IMP2, and IMP3. Interestingly, *Gorilla* and the fish *D. rerio* have four orthologous *IMP* genes, the additional one being most closely related to mammalian IMP2. On the other hand, the frog *Xenopus tropicalis* contains only one *IMP* gene, an ortholog of mammalian IMP3 [87–89].

Dyneins are microtubule-based motor complexes consisting of a core of heavy chains (HCs) that contain the motor domains, associated with a variety of smaller subunits termed intermediate, light intermediate, and light chains, which can interact with diverse cargoes [2, 10, 77, 110]. In comparison to lower eukaryotes, metazoans have expanded the number of multifunctional adaptors associated to dyneins, including dynactin, nuclear distribution protein E, lissencephaly 1, Spindly, and BicD, among others. This has led metazoan dyneins to play a role in a large diversity of activities, such as mitotic spindle assembly, apoptosis, centrosomal protein transport, chromosome segregation, and the transport of diverse cargoes such as mRNAs, viruses, organelles, signaling molecules, and intermediate filaments.

Dynein HCs comprise a large eukaryotic family of proteins [2, 10, 77, 110]. Phylogenomic analyses of hundreds of genomes have established the notion that LECA was endowed with at least nine distinct types of dynein HCs [46, 111] and that further diversification of eukaryotes led to multiple duplication of the dynein repertory in most phyla but also to lineage-specific losses in some others [2, 46, 111, 112]. For example, higher plants are devoid of dynein genes and use primarily myosin motors; *Entamoeba* and red algae also have independently lost all dyneins. In contrast, the unicellular parasite *Giardia* contains many dynein and kinesin

genes, but no myosins [2, 46, 113], and ciliates encode more dynein HC genes than most eukaryotes thus far analyzed [114]. The *chlamydomonas*, sea urchin, and human genomes possess between 14 and 16 dynein HC genes.

While all sequenced species of arthropods [112], including 12 *Drosophila* species, contain only one copy for each gene encoding a dynactin subunit, they contain a highly variable repertory of dynein heavy chains, and different numbers of light chains, which allows these species to form a large variety of dynein complexes for many cargoes [112]. All *Drosophila* species have the largest number and most divergent set of light chains [112].

The high conservation of BicD proteins and the associated dynein/dynactin motors in the animal kingdom suggests that BicD orthologs have played a conserved role in the transport of diverse cargoes, including mRNAs. Other proteins that are needed for mRNA localization have been found associated with BicD, and they are conserved throughout evolution. The future will show whether the human BicD orthologs have lost their ability to transport mRNAs or whether this function and its mRNA adaptor have simply not been discovered yet.

7 Concluding Remarks

To cope with the high sophistication of cell architecture, eukaryotes evolved two cytoskeletons that also serve as tracks for molecular motors. These are filamentous actin and microtubules. A limited number of molecular motors, myosins, kinesins, and dyneins, associate with a wide array of adaptors to gain specificity for many different cargoes. This provides the cells with the opportunity to evolve cargo-specific regulatory controls [2]. One example is the BicD/Egl complex in *Drosophila*, which exerts mRNA localization coupled with translational control of various mRNAs and which is crucial for oogenesis and embryogenesis. However, despite their importance, only a few cargo-specific adaptors for dynein have been studied so far [77]. Moreover, studies on which proteins regulate the translation of the majority of localized mRNAs are still missing. Since dynein-based motors function in the transport of a plethora of disparate cargoes, many more dynein adaptors as well as additional proteins controlling translation may still await their discovery.

Acknowledgments P. V.-P. and B.S. were supported by the Swiss National Science Foundation and the Canton of Bern. G.H. was supported by the National Institute of Cancer (INCan), Mexico, and the National Council of Science and Technology (CONACyT, grant no. 168154 to G.H.), Mexico.

References

1. Jekely G. Origin and evolution of self-organizing cytoskeleton in the network of eukaryotic organelles. In: Keeling PJ, Koonin EV, editors. The origin and evolution of eukaryotes. Cold Spring Harbor, New York, U.S.A.: Cold Spring Harbor Laboratory Press; 2014. p. 199–218.
2. Vale RD. The molecular motor toolbox for intracellular transport. Cell. 2003;112:467–80.
3. Wickstead B, Gull K, Richards TA. Patterns of kinesin evolution reveal a complex ancestral eukaryote with a multifunctional cytoskeleton. BMC Evol Biol. 2010;10:110.
4. Wickstead B, Gull K. The evolution of the cytoskeleton. J Cell Biol. 2011;194:513–25.
5. Koonin EV, Yutin N. The dispersed archaeal eukaryome and the complex archaeal ancestor of eukaryotes. Cold Spring Harb Perspect Biol. 2014;6:a016188.
6. Williams TA, Foster PG, Cox CJ, Embley M. An archaeal origin of eukaryotes supports only two primary domains of life. Nature. 2013;504:231–6.
7. Koumandou VL, Wickstead B, Ginger ML, van der Giezen M, Dacks JB, Field MC. Molecular paleontology and complexity in the last eukaryotic common ancestor. Crit Rev Biochem Mol Biol. 2013;48:373–96.
8. Gibbons BH, Asai DJ, Tang WJ, Hays TS, Gibbons IR. Phylogeny and expression of axonemal and cytoplasmic dynein genes in sea urchins. Mol Biol Cell. 1994;5:57–70.
9. Kamal A, Goldstein LSB. Principles of cargo attachment to cytoplasmic motor proteins. Curr Opin Cell Biol. 2002;14:63–8.
10. Karcher RL, Deacon SW, Geldfand VI. Motor–cargo interactions: the key to transport specificity. Trends Cell Biol. 2002;12:21–7.
11. Czaplinski K, Singer R. Pathways for mRNA localization in the cytoplasm. Trends Biochem Sci. 2006;31:687–93.
12. Kong J, Lasko P. Translational control in cellular and developmental processes. Nat Rev Genet. 2012;13:383–94.
13. Vazquez-Pianzola P, Suter B. Conservation of the RNA transport machineries and their coupling to translation control across eukaryotes. Comp Funct Genom. 2012;2012:287852.
14. Rodriguez AJ, Czaplinski K, Condeelis JS, Singer RH. Mechanisms and cellular roles of local protein synthesis in mammalian cells. Curr Opin Cell Biol. 2008;20:144–9.
15. Weis BL, Schleiff E, Zerges W. Protein targeting to subcellular organelles via MRNA localization. Biochim Biophys Acta. 2013;1833:260–73.
16. Lasko P. Translational control during early development. In: Hershey JWB, editors. Progress in Molecular Biology and Translational Science vol 90. Burlington: Academic Press; 2009. p. 211–254.
17. Richter JD, Lasko P. Translational control in oocyte development. Cold Spring Harb Perspect Biol. 2011;3:a002758.
18. McLaughlin JM, Bratu DP. *Drosophila melanogaster* oogenesis: an overview. Methods Mol Biol. 2015;1328:1–20.
19. King ML, Messitt TJ, Mowry KL. Putting RNAs in the right place at the right time: RNA localization in the frog oocyte. Biol Cell. 2005;97:19–33.
20. Lécuyer E, Yoshida H, Parthasarathy N, Alm C, Babak T, Cerovina T, Hughes TR, Tomancak P, Krause HM. Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. Cell. 2007;131:174–87.
21. Kumano G. Polarizing animal cells via mRNA localization in oogenesis and early development. Dev Growth Differ. 2011;54:1–18.
22. Jambor H, Surendranath V, Kalinka AT, Mejstrik P, Saalfeld S, Tomancak P. Systematic imaging reveals features and changing localization of mRNAs in *Drosophila* development. eLife. 2015;4.
23. Taliaferro JM, Wang ET, Burge CB. Genomic analysis of RNA localization. RNA Biol. 2014;11:1040–50.
24. Lécuyer E, Yoshida H, Krause HM. Global implications of mRNA localization pathways in cellular organization. Curr Opin Cell Biol. 2009;21:409–15.

25. Mardakheh FK, Paul A, Kümper S, Sadok A, Paterson H, Mccarthy A, Yuan Y, Marshall CJ. Global analysis of mRNA, translation, and protein localization: local translation is a key regulator of cell protrusions. Dev Cell. 2015;35:344–57.
26. Condeelis J, Singer RH. How and why does *beta-actin* mRNA target? Biol Cell. 2005;97:97–110.
27. Eom T, Antar LN, Singer RH, Bassell GJ. Localization of a beta-actin messenger ribonucleoprotein complex with zipcode-binding protein modulates the density of dendritic filopodia and filopodial synapses. J Neurosci. 2003;23:10433–44.
28. Zhang HL, Eom T, Oleynikov Y, Shenoy SM, Liebelt DA, Dictenberg JB, Singer RH, Bassell GJ. Neurotrophin-induced transport of a *beta-actin* mRNP complex increases beta-actin levels and stimulates growth cone motility. Neuron. 2001;31:261–75.
29. Leung KM, van Horck FP, Lin AC, Allison R, Standart N, Holt CE. Asymmetrical *beta-actin* mRNA translation in growth cones mediates attractive turning to netrin-1. Nat Neurosci. 2006;9:1247–56.
30. Donnelly CJ, Fainzilber M, Twiss JL. Subcellular communication through RNA transport and localized protein synthesis. Traffic. 2010;11:1498–505.
31. Hüttelmaier S, Zenklusen D, Lederer M, Dictenberg J, Lorenz M, Meng X, Bassell GJ, Condeelis J, Singer RH. Spatial regulation of beta-actin translation by Src-dependent phosphorylation of ZBP1. Nature. 2005;438:512–5.
32. Ishizuka A, Siomi MC, Siomi H. A *Drosophila* fragile X protein interacts with components of RNAi and ribosomal proteins. Genes Dev. 2002;16:2497–508.
33. Chen E, Sharma MR, Shi X, Agrawal RK, Joseph S. Fragile X mental retardation protein regulates translation by binding directly to the ribosome. Mol Cell. 2014;54:407–17.
34. Bechara EG, Didiot MC, Melko M, Davidovic L, Bensaid M, Martin P, Castets M, Pognonec P, Khandjian EW, Moine H. A novel function for fragile X mental retardation protein in translational activation. PLoS Biol. 2009;7:e16.
35. Dictenberg JB, Swanger SA, Antar LN, Singer RH, Bassell GJ. A direct role for FMRP in activity-dependent dendritic mRNA transport links filopodial-spine morphogenesis to fragile X syndrome. Dev Cell. 2008;14:926–39.
36. Bassell GJ, Warren ST. Fragile X syndrome: loss of local mRNA regulation alters synaptic development and function. Neuron. 2008;60:201–14.
37. Huber F, Boire A, López MP, Koenderink GH. Cytoskeletal crosstalk: when three different personalities team up. Curr Opin Cell Biol. 2015;32:39–47.
38. Alfaro-Aco R, Petry S. Building the microtubule cytoskeleton piece by piece. J Biol Chem. 2015;290:17154–62.
39. De la Cruz EM, Gardel ML. Actin mechanics and fragmentation. J Biol Chem. 2015;290:17137–44.
40. Lowery J, Kuczmarski ER, Herrmann H, Goldman RD. Intermediate filaments play a pivotal role in regulating cell architecture and function. J Biol Chem. 2015;290:17145–53.
41. Shih YL, Rothfield L. The bacterial cytoskeleton. Microbiol Mol Biol Rev. 2006;70:729–54.
42. Löwe J, Amos LA. Evolution of cytomotive filaments: the cytoskeleton from prokaryotes to eukaryotes. Int J Biochem Cell Biol. 2009;41:323–9.
43. Koonin EV. Origin of eukaryotes from within archaea, archaeal eukaryome and bursts of gene gain: eukaryogenesis just made easier? Philos Trans R Soc Lond B Biol Sci. 2015;370:20140333.
44. Makarova KS, Yutin N, Bell SD, Koonin EV. Evolution of diverse cell division and vesicle formation systems in Archaea. Nat Rev Microbiol. 2010;8:731–41.
45. Ettema TJ, Lindås AC, Bernander R. An actin-based cytoskeleon in Archaea. Mol Microbiol. 2011;80:1052–6.
46. Wickstead B, Gull K. Dyneins across eukaryotes: a comparative genomic analysis. Traffic. 2007;8:1708–21.
47. Dienstbier M, Li X. Bicaudal-D and its role in cargo sorting by microtubule-based motors. Biochem Soc Trans. 2009;37:1066–71.

48. Bullock SL. Translocation of mRNAs by molecular motors: think complex. Semin Cell Dev Biol. 2007;18:194–201.
49. Claußen M, Suter B. BicD-dependent localization processes: from *Drosophilia* development to human cell biology. Ann Anat. 2005;187:539–53.
50. Dienstbier M, Boehl F, Li X, Bullock SL. Egalitarian is a selective RNA-binding protein linking mRNA localization signals to the dynein motor. Genes Dev. 2009;23:1546–58.
51. Ran B, Bopp R, Suter B. Null alleles reveal novel requirement for *Bic-D* during *Drosophila* oogenesis and zygotic development. Development. 1994;120:1233–42.
52. Suter B, Steward R. Requirement for phosphorylation and localization of the Bicaudal-D protein in *Drosophila* oocyte differentiation. Cell. 1991;67:917–26.
53. Schüpbach T, Wieschaus E. Female sterile mutations on the second chromosome of *Drosophila* melanogaster: II Mutations blocking oogenesis or altering egg morphology. Genetics. 1991;129:1119–36.
54. Theurkauf WE, Alberts BM, Jan YN, Jongens TA. A central role for microtubules in the differentiation of *Drosophila* oocytes. Development. 1993;118:1169–80.
55. Clark A, Meignin C, Davis I. A Dynein-dependent shortcut rapidly delivers axis determination transcripts into the *Drosophila* oocyte. Development. 2007;134:1955–65.
56. Hughes JR, Bullock SL, Ish-Horowicz D. Inscuteable mRNA localization is dynein-dependent and regulates apicobasal polarity and spindle length in *Drosophila* neuroblasts. Curr Biol. 2004;14:1950–6.
57. Bullock SL, Ish-Horowicz D. Conserved signals and machinery for RNA transport in *Drosophila* oogenesis and embryogenesis. Nature. 2001;414:611–6.
58. Bullock SL, Ringel I, Ish-Horowicz D, Lukavsky PJ. A-form RNA helices are required for cytoplasmic mRNA transport in *Drosophila*. Nat Struct Mol Biol. 2010;17:703–9.
59. Jambor H, Mueller S, Bullock SL, Ephrussi A. A stem-loop structure directs *oskar* mRNA to microtubule minus ends. RNA. 2014;20:429–39.
60. Hachet O, Ephrussi A. Splicing of *oskar* RNA in the nucleus is coupled to its cytoplasmic localization. Nature. 2004;428:959–63.
61. Jambor H, Brunel C, Ephrussi A. Dimerization of *oskar* 3′ UTRs promotes hitchhiking for RNA localization in the *Drosophila* oocyte. RNA. 2011;17:2049–57.
62. Tanaka T, Kato Y, Matsuda K, Hanyu-Nakamura K, Nakamura A. *Drosophila* Mon2 couples Oskar induced endocytosis with actin remodeling for cortical anchorage of the germ plasm. Development. 2011;138:2523–32.
63. Vanzo N, Oprins A, Xanthakis D, Ephrussi A, Rabouille C. Stimulation of endocytosis and actin dynamics by Oskar polarizes the *Drosophila* oocyte. Dev Cell. 2007;12:543–55.
64. Zimyanin VL. *In vivo* imaging of *oskar* mRNA transport reveals the mechanism of posterior localization. Cell. 2008;134:843–53.
65. Krauss J. Lopez de Quinto S, Nusslein-Volhard C, Ephrussi A. Myosin-V regulates *oskar* mRNA localization in the *Drosophila* oocyte. Curr Biol. 2009;19:1058–63.
66. Cook HA, Koppetsch BS, Wu J, Theurkauf WE. The *Drosophila* SDE3 homolog armitage is required for *oskar* mRNA silencing and embryonic axis specification. Cell. 2004;116:817–29.
67. Pane A, Wehr K, Schüpbach T. *zucchini* and *squash* encode two putative nucleases required for rasiRNA production in the *Drosophila* germline. Dev Cell. 2007;12:851–62.
68. Lim AK, Kai T. Unique germ-line organelle, nuage, functions to repress selfish genetic elements in *Drosophila melanogaster*. Proc Natl Acad Sci USA. 2007;104:6714–9.
69. Nakamura A, Amikura R, Hanyu K, Kobayashi S. Me31B silences translation of oocyte-localizing RNAs through the formation of cytoplasmic RNP complex during *Drosophila* oogenesis. Development. 2001;128:3233–42.
70. Besse F. Lopez de Quinto S, Marchand V, Trucco A, Ephrussi A. *Drosophila* PTB promotes formation of high-order RNP particles and represses *oskar* translation. Genes Dev. 2009;23:195–207.

71. Nakamura A, Sato K, Hanyu-Nakamura K. *Drosophila* cup is an eIF4E binding protein that associates with Bruno and regulates *oskar* mRNA translation in oogenesis. Dev Cell. 2004;6:69–78.
72. Chekulaeva M, Hentze MW, Ephrussi A. Bruno acts as a dual repressor of *oskar* translation, promoting mRNA oligomerization and formation of silencing particles. Cell. 2006;124:521–33.
73. Yano T, Lopez de Quinto S, Matsui Y, Shevchenko A, Ephrussi A. Hrp48, a Drosophila hnRNPA/B homolog, binds and regulates translation of oskar mRNA. Dev Cell. 2004;6:637–648.
74. Clouse KN, Ferguson SB, Schupbach T. Squid, Cup, and PABP55B function together to regulate *gurken* translation in *Drosophila*. Dev Biol. 2008;313:713–24.
75. Vazquez-Pianzola P, Urlaub H, Suter B. Pabp binds to the osk 3′UTR and specifically contributes to osk mRNA stability and oocyte accumulation. Dev Biol. 2011;357:404–18.
76. Olesnicky EC, Brent AE, Tonnes L, Walker M, Pultz MA, Leaf D, Desplan C. A caudal mRNA gradient controls posterior development in the wasp *Nasonia*. Development. 2006;133:3973–82.
77. Kardon JR, Vale RD. Regulators of the cytoplasmic dynein motor. Nat Rev Mol Cell Biol. 2009;10:854–65.
78. Vazquez-Pianzola P, Adam J, Haldemann D, Hain D, Urlaub H, Suter B. Clathrin heavy chain plays multiple roles in polarizing the *Drosophila* oocyte downstream of Bic-D. Development. 2014;141:1915–26.
79. Aguirre-Chen C, Bulow HE, Kaprielian ZC. elegans bicd-1, homolog of the Drosophila dynein accessory factor Bicaudal D, regulates the branching of PVD sensory neuron dendrites. Development. 2011;138:507–18.
80. Fridolfsson HN, Starr DA. Kinesin-1 and dynein at the nuclear envelope mediate the bidirectional migrations of nuclei. J Cell Biol. 2010;191:115–28.
81. Jaarsma D, van den Berg R, Wulf PS, van Erp S, Keijzer N, Schlager MA, de Graaff E, De Zeeuw CI, Pasterkamp RJ, Akhmanova A, Hoogenraad CC. A role for Bicaudal-D2 in radial cerebellar granule cell migration. Nat Commun. 2014;5:3411.
82. Terenzio M, Golding M, Russell MR, Wicher KB, Rosewell I, Spencer-Dene B, Ish-Horowicz D, Schiavo G. Bicaudal-D1 regulates the intracellular sorting and signalling of neurotrophin receptors. EMBO J. 2014;33:1582–98.
83. Lipka J, Kuijpers M, Jaworski J, Hoogenraad CC. Mutations in cytoplasmic dynein and its regulators cause malformations of cortical development and neurodegenerative diseases. Biochem Soc Trans. 2013;41:1605–12.
84. Neveling K, Martinez-Carrera LA, Holker I, Heister A, Verrips A, Hosseini-Barkooie SM, Gilissen C, Vermeer S, Pennings M, Meijer R, te Riele M, Frijns CJ, Suchowersky O, MacLaren L, Rudnik-Schoneborn S, Sinke RJ, Zerres K, Lowry RB, Lemmink HH, Garbes L, Veltman JA, Schelhaas HJ, Scheffer H, Wirth B. Mutations in BICD2, which encodes a golgin and important motor adaptor, cause congenital autosomal-dominant spinal muscular atrophy. Am J Hum Genet. 2013;92:946–54.
85. Peeters K, Litvinenko I, Asselbergh B, Almeida-Souza L, Chamova T, Geuens T, Ydens E, Zimon M, Irobi J, De Vriendt E, De Winter V, Ooms T, Timmerman V, Tournev I, Jordanova A. Molecular defects in the motor adaptor BICD2 cause proximal spinal muscular atrophy with autosomal-dominant inheritance. Am J Hum Genet. 2013;92:955–64.
86. Hoogenraad CC, Akhmanova A, Howell SA, Dortland BR, De Zeeuw CI, Willemsen R, Visser P, Grosveld F, Galjart N. Mammalian Golgi-associated Bicaudal-D2 functions in the dynein-dynactin pathway by interacting with these complexes. EMBO J. 2001;20:4041–54.
87. Kriventseva EV, Tegenfeldt F, Petty TJ, Waterhouse RM, Simão FA, Pozdnyakov IA, Ioannidis P, Zdobnov EM. OrthoDB v8: update of the hierarchical catalog of orthologs and the underlying free software. Nucleic Acids Res. 2015;43:D250–6.
88. Waterhouse RM, Tegenfeldt F, Li J, Zdobnov EM, Kriventseva EV. OrthoDB: a hierarchical catalog of animal, fungal and bacterial orthologs. Nucleic Acids Res. 2013;41:D358–65.

89. Waterhouse RM, Zdobnov EM, Tegenfeldt F, Li J, Kriventseva EV. OrthoDB: the hierarchical catalog of eukaryotic orthologs in 2011. Nucleic Acids Res. 2011;39:D283–8.
90. Schlager MA, Kapitein LC, Grigoriev I, Burzynski GM, Wulf PS, Keijzer N, de Graaff E, Fukuda M, Shepherd IT, Akhmanova A, Hoogenraad CC. Pericentrosomal targeting of Rab6 secretory vesicles by Bicaudal-D-related protein 1 (BICDR-1) regulates neuritogenesis. EMBO J. 2010;29:1637–51.
91. Preechaphol R, Klinbunga S, Khamnamtong B, Menasveta P. Isolation and characterization of genes functionally involved in ovarian development of the giant tiger shrimp *Penaeus monodon* by suppression subtractive hybridization (SSH). Genet Mol Biol. 2010;33:676–85.
92. Bianco A, Dienstbier M, Salter HK, Gatto G, Bullock SL. Bicaudal-D regulates fragile X mental retardation protein levels, motility, and function during neuronal morphogenesis. Curr Biol. 2010;20:1487–92.
93. Guerra N, Vega-Sendino M, Pérez-Morgado MI, Ramos E, Soto M, Gonzalez VM, Martín ME. Identification and functional characterization of a poly(A)-binding protein from *Leishmania infantum* (LiPABP). FEBS Lett. 2011;585:193–8.
94. Burgess HM, Gray NK. mRNA-specific regulation of translation by poly(A)-binding proteins. Biochem Soc Trans. 2010;38:1517–22.
95. Mangus DA, Evans MC, Jacobson A. Poly(A)-binding proteins: multifunctional scaffolds for post-transcriptional control of gene expression. Genom Biol. 2003;4:223.
96. Gallie DR, Liu R. Phylogenetic analysis reveals dynamic evolution of the poly(A)-binding protein gene family in plants. BMC Evol Biol. 2014;14:238.
97. Hernández G, Altmann M, Lasko P. Origins and evolution of the mechanisms regulating translation initiation in eukaryotes. Trends Biochem Sci. 2010;35:63–73.
98. Jimenez-Lopez D, Bravo J, Guzman P. Evolutionary history exposes radical diversification among classes of interaction partners of the MLLE domain of plant poly(A)-binding proteins. BMC Evol Biol. 2015;15:195.
99. Arn EA, Cha BJ, Theurkauf WE, Macdonald PM. Recognition of a *bicoid* mRNA localization signal by a protein complex containing Swallow, Nod, and RNA binding proteins. Dev Cell. 2003;4:41–51.
100. Mohr E, Fuhrmann C, Richter D. VP-RBP, a protein enriched in brain tissue, specifically interacts with the dendritic localizer sequence of rat *vasopressin* mRNA. Eur J Neurosci. 2001;13:1107–12.
101. Mohr E, Prakash N, Vieluf K, Fuhrmann C, Buck F, Richter D. *Vasopressin* mRNA localization in nerve cells: characterization of cis-acting elements and trans-acting factors. Proc Natl Acad Sci USA. 2001;98:7072–9.
102. Mohr E, Richter D. Subcellular *vasopressin* mRNA trafficking and local translation in dendrites. J Neuroendocrinol. 2004;16:333–9.
103. Coffee RL, Tessier CR, Woodruff EA, Broadie K. Fragile X mental retardation protein has a unique, evolutionarily conserved neuronal function not shared with FXR1P or FXR2P. Dis Model Mech. 2010;3:471–85.
104. Tucker B, Richards R, Lardelli M. Expression of three zebrafish orthologs of human *FMR1*-related genes and their phylogenetic relationships. Dev Genes Evol. 2004;214:567–74.
105. Kirkpatrick LL, McIlwain KA, Nelson DL. Comparative genomic sequence analysis of the FXR gene family: *FMR1*, *FXR1*, and *FXR2*. Genomics. 2001;78:169–77.
106. Munro TP, Kwon S, Schnapp BJ. St Johnston D. A repeated IMP-binding motif controls *oskar* mRNA translation and anchoring independently of *Drosophila melanogaster* IMP. J Cell Biol. 2006;172:577–88.
107. Geng C, Macdonald PM. Imp associates with squid and Hrp48 and contributes to localized expression of *gurken* in the oocyte. Mol Cell Biol. 2006;26:9508–16.
108. Nielsen FC, Nielsen J, Christiansen J. A family of IGF-II mRNA binding proteins (IMP) involved in RNA trafficking. Scand J Clin Lab Invest Suppl. 2001;234:93–9.
109. Paquin N, Chartrand P. Local regulation of mRNA translation: new insights from the bud. Trends Cell Biol. 2008;18:105–11.
110. Asai DJ, Wilkes DE. The dynein heavy chain family. J Eukaryot Microbiol. 2004;51:23–9.

111. Wilkes DE, Watson HE, Mitchell DR, Asai DJ. Twenty-five dyneins in *Tetrahymena*: a re-examination of the multidynein hypothesis. Cell Motil Cytoskeleton. 2008;65:342–51.
112. Odronitz F, Becker S, Kollmar M. Reconstructing the phylogeny of 21 completely sequenced arthropod species based on their motor proteins. BMC Genom. 2009;10:173.
113. Lawrence CJ, Morris NR, Meagher RB, Dawe RK. Dyneins have run their curse in plant lineage. Traffic. 2001;2:362–3.
114. Rajagopalan V, Wilkes DE. Evolution of the dynein heavy chain family in ciliates. J Eukaryot Microbiol. 2015;63:138–41.

IRES Elements: Issues, Controversies and Evolutionary Perspectives

Rosario Francisco-Velilla, Gloria Lozano, Rosa Diaz-Toledano, Javier Fernandez-Chamorro, Azman M. Embarek and Encarnacion Martinez-Salas

1 Introduction

Translation control is a key step in gene expression regulation in all organisms. In eukaryotes, the vast majority of mRNAs initiate translation by a mechanism that depends on the recognition of the $m^7G(5')ppp(5')N$ structure (termed cap) located at the 5′ end of most mRNAs [1]. This process is achieved in several steps. The first step begins with the binding of the cap structure by the translation initiation factor (eIF)-4F complex, consisting of the cap-binding protein eIF4E, the scaffold eIF4G and the helicase eIF4A. This complex recruits the 40S ribosomal subunit bound to eIF3, eIF2 and the initiator tRNA and scans the 5′UTR until an AUG triplet is found in the appropriate context to start protein synthesis (for a review, see [1]). Joining of the 60S ribosomal subunit follows this step, producing a translation competent complex. Beyond this general manner to initiate translation, specific mRNA regions referred to as internal ribosome entry sites (IRES) can recruit the 40S ribosomal subunit through a cap-independent mechanism [2, 3]. Likewise, initiation of protein synthesis may also occur by other cap-independent mechanisms. Indeed, recent studies have shown that m^6A modification in mRNA promote translation initiation regardless of the 5′end N(7)-methyl guanosine cap [4, 5]. Additionally, RNA looping was proposed as the mechanism leading to the identification of the translation initiation codon regardless of the presence of cap-structure [6].

It is well established that strong cellular stresses, such as viral infection or apoptosis, compromise cap-dependent translation initiation. Yet, a specific type of

R. Francisco-Velilla · G. Lozano · R. Diaz-Toledano · J. Fernandez-Chamorro
A. M. Embarek · E. Martinez-Salas (✉)
Centro de Biología Molecular Severo Ochoa, Consejo Superior de Investigaciones Científicas - Universidad Autónoma de Madrid, Nicolas Cabrera 1, 28049 Madrid, Spain
e-mail: emartinez@cbm.csic.es

R. Francisco-Velilla
e-mail: rfrancisco@cbm.uam.es

G. Hernández and R. Jagus (eds.), *Evolution of the Protein Synthesis Machinery and Its Regulation*, DOI 10.1007/978-3-319-39468-8_22

mRNAs can overcome the inhibitory conditions, as exemplified by viral mRNAs that harbor IRES elements. Viral IRES elements hijack the translational machinery of the host cell, facilitating translation initiation by recruiting and actively manipulating the ribosome using a subset of canonical initiation factors [7]. IRES elements were first reported in the genomic RNA of picornaviruses [2, 3]. Many positive-strand RNA viruses are naturally uncapped and instead depend on IRES elements to express their genome. In full agreement with this, IRES elements also control protein synthesis in the genome of various genera of the RNA virus (such as hepacivirus, pestivirus, dicistrovirus, retrovirus and lentivirus, among others) [8–11]. Additionally, IRES elements have been reported to control protein synthesis in a diverse group of cellular mRNAs under conditions that compromise cap-dependent translation [12–16].

RNA structure determines the function of most viral IRES elements. Yet, IRES elements belonging to different families of RNA viruses lack overall conserved features. For instance, the IRES located in the intergenic region (IGR) of dicistroviruses and the IRES of hepatitis C virus (HCV) lack sequence homology, exhibit different structural organization and also have a different requirement of factors. Nonetheless, the natural selection pressure has evolved specialized three-dimensional structures in each family of RNA virus that finally render a functional element. For instance, the RNA architecture of the IGR is strongly conserved across different species of dicistroviruses [17, 18]. Also, in spite of the high genetic variability of RNA viruses exemplified by the HCV or the foot-and-mouth disease virus (FMDV) genomic RNA, the IRES structure is preserved by compensatory mutations, monitored by covariation analysis. In accordance with this, engineered disruption of the stems reduced IRES activity whereas compensatory mutations restoring the secondary structure recovered translation efficiency [8, 19, 20], further demonstrating the relevance of RNA structure for IRES function.

Regarding their position on the mRNA, most known IRES elements are located within the 5′UTR, but a few others are placed within the coding sequence. The location downstream of the main start codon gives rise to a shorter protein from the internal initiation codon that sometimes can have a function different from that of the main open reading frame (ORF). This characteristic, which is currently unpredictable, greatly increases the coding potential of the genome and, importantly, evidences the urgent need for a correct annotation of the eukaryotic genomes. Another characteristic of most IRES elements is that, although there is a distribution of functions among modular domains, they function as a single entity; this feature implies that discrete pieces do not exhibit the activity produced by the entire element. This observation, however, could be relevant to understand the evolutionary history of RNA structural motifs present in distinct types of IRES elements.

2 IRES Elements: Diversity of Internal Initiation Mechanisms

Internal initiation of translation can occur either by direct interaction of the IRES with the 40S ribosmal subunit, or binding of the IRES element to initiation factors, which then recruit the 40S ribosomal subunit. Thus, according to the minimal set of factors required for internal initiation, IRES elements can be grouped into different categories. The IGR of dicistroviruses represents the simplest category. This IRES element adopts a complex three-dimensional structure, consisting of a triple-pseudoknot (PK I, II and III), and initiates translation at a non-AUG triplet without the help of eIFs [9, 17, 21]. Hence, the IGR functionally substitutes for the initiator met-tRNA$_i$ during internal initiation. The near atomic resolution structure of IGR-ribosome complexes showed that the PKI of the IGR resembles a tRNA/mRNA interaction in the decoding center of the A site, mimicking a pre-translocation rather than initiation state of the ribosome [22]. Translocation of the IGR by elongation factor 2 (eEF2) is required to bring the first codon of the mRNA into the A site and to trigger translation; during this translocation event the IRES undergoes a structural change to a stretched conformation, demonstrating the active role of the mRNA structure in manipulating the ribosome to initiate protein synthesis.

A larger and more diverse category of IRES elements includes those that do need eIFs to recruit the ribosomal subunits in order to promote initiation of protein synthesis. Examples of this category are found in the genomic RNA of HCV, but also in pestivirus, picornavirus or retrovirus RNAs, among others [23]. All these IRES elements, however, differ in nucleotide sequence, RNA structure and the specific eIFs required for 48S complex assembly. The HCV IRES element, as well as the so-called HCV-like present in the genomic RNA of pestivirus and a few picornaviruses, only requires eIF3 and the ternary complex eIF2-GTP-tRNA$_i$ for the assembly in vitro of 48S initiation complexes [24]. The HCV IRES is organized into three domains, designated as II, III and IV, each one performing a distinct function. Domain II is involved in eIF2-catalyzed GTP hydrolysis and 60S ribosomal subunit joining; in addition, the apical loop of domain II contacts the ribosomal protein RPS5, stabilizing the ribosome in the proper conformation leading to translocation. Domain III binds the 40S ribosomal subunit and eIF3, while domain IV harbors the AUG initiation codon [25, 26]. It should be noted that the full length HCV IRES in solution shows a flexible RNA structure, which consists of an ensemble of conformers made of rigid parts that can move relative to each other [27]. This feature is consistent with the presence of four subpopulations for the 80S•HCV IRES complex, revealing dynamic conformational modes of the complex [28]. Evidence for RNA flexibility of the HCV RNA was also observed in long transcripts that harbor the IRES element in the natural viral RNA context, flanked by domain I at the 5′end, and domains V and VI at the 3′end [29].

Translation initiation promoted by the HCV-like IRES elements also relies on their interaction with the 40S ribosomal subunit. During initiation, the IRES-40S

complex places the start codon into the P site, base-pairing with eIF2-bound initiator met-tRNAi to form a 48S complex. Reconstitution of a 40S ribosomal complex containing eIF3 and a pestivirus (HCV-like) IRES shows that eIF3 is displaced from its ribosomal position in the 43S complex relative to the HCV IRES and instead interacts through its ribosome-binding surface with the apical region of domain III of the IRES [30]. These results suggest a role for the specific interaction of HCV-like IRES elements with eIF3 in preventing ribosomal association of eIF3, thus favoring translation of viral mRNAs.

In contrast to HCV, the IRES elements governing translation initiation in the genome of picornaviruses are heterogeneous in the nucleotide sequence, RNA structure and requirement of factors for ribosome recruitment [23]. Furthermore, the list of new types of IRES elements increases in correlation with the incessant discovery of new species of picornavirus with different nucleotide composition. The picornavirus IRES elements are currently classified into five types and several subclasses; notably, each type harbors a common RNA structure core maintained by evolutionary conserved covariant substitutions [19, 31–33]. In addition, each type exhibits a distinct requirement of factors for the assembly of the initiation complex. The IRES elements classified as types I and II require the C-terminal region of eIF4G, eIF4A, eIF2, and eIF3, but not eIF4E, to assemble 48S initiation complexes [34, 35]. In contrast, translation initiation driven by type III depends on the integrity of eIF4G [36]. Type IV (termed HCV-like because of its similarity with the HCV IRES element [37]) is eIF4G-independent but depends on eIF2 and eIF3. Conversely, the Aichi-like type (present in Aichi virus, AV) depends on eIF4G but, unlike type I and type II, it depends also on the RNA helicase DHX29 because of sequestration of its initiation codon in a stable hairpin [33]. A recently discovered subclass by phylogenetic analysis of RNAs detected in samples from domestic dog and wild animals revealed an Aichi-like IRES secondary structure [38]. It is worth noting that recombination events occurring during picornavirus coinfection can generate IRES elements with unique properties [39, 40], including novel tissue tropism and/or host-range spectrum. The latter could lead to crossing the species barrier. Importantly, the IRES element is essential for viral infection in all RNA viruses analyzed, and therefore it constitutes a specific target for antiviral compounds [41–45].

Three distinct IRES elements, HIV-1 5′UTR, HIV-1 gag and HIV-2 gag, have been reported in the genome of lentiviruses (a genus of the retroviridae family), although there are controversial results concerning these IRES elements because the viral mRNA is naturally capped. These IRES elements exhibit different structural features and also have a different requirement of factors than the IRES of HCV or picornavirus. It appears that the activity of HIV-1 5′UTR IRES tolerates point mutations, but it is strictly dependent on specific host factors, and its activity is linked to the cell cycle phase [10, 46]. On the other hand, and in contrast to the vast majority of viral IRES elements, the HIV-1 gag and HIV-2 gag IRES elements are positioned downstream of the start codon. Hence, a back-scanning mechanism was proposed for both HIV-1 gag and HIV-2 gag IRES elements [11, 47].

Further increasing the IRES diversity, the activity of a few viral IRES elements depends on specific single-stranded regions. For instance, the activity of the IRES placed upstream of the coat protein ORF of the crucifer-infecting tobamovirus (CrTMV) was mapped to A-rich regions [48] Conversely, U-rich single-stranded regions determine the function of the IRES located at the 5′ end of the *Rhopalosiphum padi* virus RNA [49]. Also, an unstructured region determines the function of the *Halastavi arva* virus, a positive strand RNA virus with a dicistronic genome [50]. In the latter, initiation of translation involves direct attachment of the 43S preinitiation complexes immediately downstream of the initiation codon; then, 43S complexes undergo a retrograde scanning dependent on eIF1 and eIF1A.

Due to the observation that viral IRES elements confer resistance to inhibitory conditions for cap-dependent translation, the IRES field has attracted research interests not only in the translation field, but also in other areas of RNA biology impacting on cell proliferation, differentiation, stress response, development, etc., leading to the discovery of a significant number of mRNAs able to be translated using a cap-independent mechanism under cellular stress. Hence, these RNA regions were designated as IRES elements. In line with the observations made on distinct types of viral IRES elements, structural and functional studies carried out in a few cellular IRES elements [51–54] suggest that these elements are defined by their local RNA structure flexibility, depending upon short sequence motifs and trans-acting factors for their function. The diversity of sequences and structures mediating cap-independent translation issued many caveats in the translation field, mainly due to the lack of appropriately controls in published data. Notably, and in spite of the controversy on some of these studies, recently published evidence supports distinct types of cap-independent translation mechanisms; the unanticipated diversity of translation initiation mechanisms suggests that various sorts of RNA regions governing distinct cap-independent mechanisms might have been erroneously designated as IRES elements.

3 Diversity of IRES-Ribosome Interactions: Insights for Evolutionarily Traits?

As mentioned earlier, IRES elements can recruit the ribosomal subunits directly, e.g., by direct contact with the 40S ribosomal subunit or, indirectly, by using functional bridges, generally eIFs and RNA-binding proteins, to capture the ribosome. Physical association of the IGR or the HCV IRES elements and the 40S ribosomal subunit has been shown by different experimental approaches [26, 55]. Evidence for direct interaction between the ribosomal RNA and the IRES element has been reported in studies of the HCV IRES element, the 9-nt element of the GTX mRNA 5′UTR (encoding a murine homeodomain transcription factor expressed specifically in glial cells of the brain and germ cells of testis) or the G-quadruplex of the vascular endothelial growth factor (VEGF) mRNA [56–58]. Several studies of

the HCV IRES reveal the dynamic nature of the RNA in recruiting the translation machinery; the HCV IRES contacts the backbone and bases of the CCC triplet in the 18S rRNA, inducing a rearrangement of 18S rRNA structure in the vicinity of the conserved nucleotide G1639 [59]. Conversely, formation of a kissing complex between the loop IIId of the HCV IRES and the 18S rRNA was also observed in binary complexes [60]. More recently, short sequences complementary to the 18S rRNA have been found in mRNAs selected in a high-throughput screening among a library of several thousand oligonucleotides (210 nt long) cloned in a lentiviral bicistronic plasmid between mRFP and eGFP reporters translated by a cap-independent mechanism in H1299 cells [61].

A different manner used by many IRES elements to recruit the ribosomal subunits relies on the involvement of proteins, both eIFs and RNA-binding proteins (RBPs). These factors serve as functional bridges between the IRES element on one side and the ribosomal subunit on the other, as exemplified by picornavirus IRES elements. As mentioned earlier, picornavirus IRES elements are rather heterogeneous, and thus they are classified into five types. The RNA structure of enterovirus IRES element (the representative member of type I) is organized in five domains, designated II to VI [32]. The involvement of these domains in IRES activity was demonstrated by mutational studies, and their interaction with host factors modulating internal initiation of translation was shown by RNA protein-binding analysis. For instance, domain IV harbors a C-rich loop that provides the binding site for the polyr(C)-binding protein PCBP2, while domain V provides the binding site for the polypyrimidine tract-binding protein (PTB) [62].

The RNA structure of type II IRES (represented by encephalomyocarditis virus (EMCV) and FMDV) is also arranged in five domains designated 2–5, or H to L, respectively [63, 64]. Domain 2 contains a conserved pyrimidine tract that provides a binding site for the PTB protein [65]. Domain 3 is a self-folding cruciform structure [66, 67]; the basal region of this domain consists of a long stem interrupted with bulges that includes several non-canonical base pairs and a helical structure essential for IRES activity. Domain 4 is organized into two hairpin loops, which contain the binding site for eIF4G, an essential factor for these IRES elements [34, 68, 69]. Domain 5 consists of a short hairpin followed by a single-stranded stretch of nucleotides on its 3′ end including a conserved pyrimidine tract; this domain provides the binding site for eIF4B, PTB and other RNA-binding proteins [70–73]. Notably, a construct that harbors domain 4 and 5 (hence, able to interact with eIF4G, eIF3, eIF4B, PTB and several other proteins) is not sufficient to promote IRES-dependent protein synthesis [74]. The partial recovery of IRES activity observed upon coexpression of two defective IRES elements classified as type I and type II, each one having a defect on a different domain [75, 76], suggested functional complementation. These data show that although there is a distribution of functions in each domain, the entire element is required for full IRES function.

It is well established that eIFs and auxiliary RBPs stimulate 48S complex formation in reconstitution assays with picornavirus IRES elements [33, 77–79]. Interestingly, picornaviruses classified as type I and type II IRES have two

polypyrimidine tracts placed at each end of the IRES region such that PTB binds to the IRES constraining the RNA structure in a unique orientation [80], possibly facilitating protein synthesis. Beyond PTB, multiple RBPs involved in distinct RNA processes have been identified by mass spectrometry analysis of affinity-purified factors associated with viral IRES elements. Most of these factors interact with multiple targets and show promiscuous binding to RNA. These characteristics represent an obstacle to elucidate whether the proteins recognizing distinct IRES elements reflect, or not, common evolutionary traits.

With the exception of short pyrimidine tracts, distinctive RNA motifs are lacking in cellular IRES elements [81, 82]. It is worth noting that the observation that artificial constructs containing (CCU)(n) motifs function as internal ribosome entry segments in the presence of PTB [83] led to suggest that PTB (or its interacting protein partners) could provide a bridge between the IRES and the ribosome, and thus it could be a general IRES transacting factor (ITAF). In support of this proposal, PTB stimulates translation driven by IRES elements placed in cellular mRNAs encoding proteins controlling cell growth, as well as factors involved in apoptosis, nutrient deprivation or cell proliferation [84, 85]. Nevertheless, this is not always the case, as there are examples where PTB inhibits translation [86].

On the other hand, the roles of 3′-poly(A) tail of the mRNA and the poly(A)-binding protein (PABP) in internal initiation vary depending on the type of IRES and the biological characteristics of the system [87–89]. In the case of the A-rich CrTMV IRES, mutations in the internal polypurine tract decrease both the IRES activity and binding of PABP. Furthermore, enhancement of IRES function in the presence of 3′-poly(A) and the absence of 5′-cap suggests a crosstalk among PABP, the CrTMV IRES and the 3′-poly(A) tail [90].

4 Lack of a Universal Motif Across IRES Elements: Evolutionary Perspectives

The evolutionary origin of IRES elements remains elusive. Nevertheless, the observation that IRES elements have been described not only in RNA viruses infecting animal cells, but also in cellular RNAs from mammals, insects, plants, parasites and yeast encoding proteins synthesized upon stress or related to cell survival [91] suggests that these elements may have an ancient origin. Indeed, current data support two different evolutionary origins of IRES elements. Early studies on the mechanism of action of IRES elements suggested similarities with prokaryote-like mechanisms [24]. In agreement with this hypothesis, the study of the insulin-like growth factor 1 receptor (IGF1R) IRES proposed a direct interaction with the 18S rRNA modulated by a polymorphic poly(U) loop, presumably operating as a Shine-Dalgarno-like (SD-like) site [92].

The possibility that some IRES elements could arise by divergent evolution from the prokaryote SD-like translation initiation regulatory element has gained support

by the discovery of a large number of short sequences that have the property of being complementary to the rRNA and behave as putative IRES elements [61]. Additionally, a work on land plant chloroplast RNAs described an operon that harbors two cistrons separated by a spacer region. The secondary structure of this intercistronic spacer consists of a loop projecting the SD-like site and a stem that interacts with the adjacent coding region and sequesters the start codon, suggesting that translation of the second cistron is regulated by a cis-acting mechanism comparable to prokaryotes [93]. Bicistronic RNAs are rarely found in eukaryotic organisms, but there are some examples in both plant and animal cells. Two examples are the WDR53-CesA8A transcript of the domestic apple, which contains a spliced intergenic sequence predicted to fold into a hairpin structure [94], and the colinear transcripts GPR40-GPR41 (G protein-coupled receptors activated by long or short chain fatty acids, respectively), which is expressed in pancreatic beta cells [95].

Given the great diversity of IRES elements, it is also conceivable that some type of IRES elements could have arisen by convergent evolution. Possible events resulting in the generation of novel functional IRES elements could be the assembly of discrete RNA modules derived from different molecules, presumably by RNA recombination, integrative events or RNA ligation of small fragments. These events, although infrequent, could eventually generate a regulatory structural motif with new functions, unrelated to that displayed by the RNA molecule, which harbors the combined RNA modules. In support of this possibility, there is proven evidence that nature uses thermodynamically stable modular motifs from natural RNA molecules to generate complex programmable structures [96]. In agreement with this, RNA molecules have unique structural attributes, which include the ability to self-assemble in a controlled manner through the arrangement of structural building blocks, as shown in riboswitches, ribozymes or ribosomal RNAs, among other RNA molecules.

An example of a modular structural motif that is found in various RNAs, yet performing different functions, is the tRNA-like motif. This motif is found in the 3′ UTR of plant viral RNAs promoting translation and replication of the viral RNA [97, 98]. Indeed, tRNA structure mimicry, experimentally shown in the case of the dicistrovirus IGR (see Sect. 2), was proposed as a candidate for a universal IRES motif. In favor of this proposal, the results of RNase P cleavage assays (a structure-dependent enzyme that recognizes and processes the tRNA precursor) suggested the presence of tRNA-like motifs in a few viral IRES elements, including the IGR, the HCV, and the EMCV and FMDV IRES elements [99–101]. This proposal has received direct support by near-atomic resolution structures only in the case of the discistrovirus IGR. Thus, the results obtained from other IRES elements could suggest that the regions of tRNA mimicry are remnants from an ancestor molecule, presumably unrelated to their function. Along this line, RNA molecules could arise by fusion, followed by gradual addition and evolutionary selection, from ancestral modules. Yet, as it occurs in proteins, domains representative of known folds whose similarity is indicative of a common descent can occur in non-homologous domains.

5 Prediction of IRES-Like Motifs in Genomes

Prediction of potential IRES elements genome wide remains a challenging task because currently known IRES elements lack conservation of primary sequence and RNA secondary structure. To our knowledge, a tool capable of using sequence and structure information, and superimposing inferences about the evolutionary behavior of individual residues, is not available for RNA molecules. In spite of this, and given that the RNA structure plays a fundamental role in most viral IRES-dependent translations [102], it is plausible that similar RNA motifs can be found in genomes, expanding gene regulatory elements and also providing hints about their evolutionary history.

In addition to the stress response, cap-independent translation also can occur during physiological cellular processes (for example, mitosis and apoptosis). Hence, it is unlikely that viruses developed the IRES-dependent translation initiation mechanism and remained solely used by viral RNAs. It should be noted that research in viruses has been the first step to discover critical post-transcriptional events. To name a few, studies on reovirus mRNAs led to the identification of mRNA capping, splicing was first reported in adenovirus, and polyadenylation was described in vaccinia virus. Viruses usually have small genomes and thus depend on the host cell to carry out most of the viral replication steps. This feature ensures that virus genomes evolved to increase their replication potential in the infected host. This evolutionary plasticity makes viruses an invaluable system to identify new mechanisms used not only by viruses but also by the host cell. Hence, regardless of the lack of a universal conserved structural motif unique to all viral IRES elements, a better understanding of the structural organization and function of a specific IRES domain (or subdomain) can help to predict IRES-like motifs in cellular genomes.

Functional annotation of short ORFs using non-conventional translation initiation, as well as noncoding RNAs, remains a current task of great biological importance. Evidence for the existence of non-annotated, yet translated short ORFs (<100 aa) is supported by experimental data derived from the ribosome profiling and mass spectrometry analysis [103]. The periodicity of ribosome movement on the mRNA defines actively translated ORFs by ribosome footprinting in eukaryotes, which are also supported by computational prediction of small ORFs from codon conservation patterns. Additionally, short peptides, which have emerged as important regulators of development and physiology, result from translation initiation events at non-canonical initiation codons. Hence, regulatory mechanisms different from the conventional cap-dependent scanning must be involved in their translation initiation. While we are not implying that internal initiation will solely be driven by IRES elements, it may occur that at least some of these short ORFs may be using IRES-like mechanisms to initiate translation. Moreover, the rapid advance in the sequencing methodologies has led to the discovery of novel RNA forms, such as the long noncoding RNAs (lncRNAs) and the circular RNAs (circRNAs). Increasing evidence supports that some of these RNAs encode short

peptides although they do not match the criteria of the cap-dependent initiation mechanism. The function of such molecules as well as the mechanism used to initiate their synthesis remains to be explored.

Attempts to explore the presence of putative IRES elements in genome-wide approaches have made use of in vitro selection strategies [104]. This cell-free method is based on the mRNA display of genomic fragments (about 150 nucleotides long), selection for initiation of translation and high-throughput deep sequencing. When translated in vitro, the tagged RNA sequences that mediate cap-independent initiation of translation become covalently linked to a peptide affinity tag encoded in the coding sequence. Formation of a 3′ -puromycin chemical bond between newly translated peptides and their encoding mRNA occurs via the natural peptidyl transferase activity of the ribosome, which recognizes puromycin as a tyrosyl-tRNA analog. This method reported several thousand translation-enhancing elements (TEEs) in the human genome associated with genes involved in signal transduction, cell communication and neurological system development pathways. Surprisingly, many of these RNAs harbor TEEs in intronic sequences flanked by regions that are conserved among species.

A radically different approach to search for putative IRES elements takes advantage of sequence comparative analysis using gene finder algorithms. Using as reference the cap-independent A-rich RNA sequences of yeast [105], a bioinformatic method based on a support vector machine was developed for the prediction of putative IRES elements using the 5′-UTR sequences of fungi [106]. Surprisingly, this method predicted over 6000 putative IRES elements. Although this attempt to develop informatic tools to identify potential IRES elements rendered promising results, no functional studies were reported for any of the identified RNAs.

The use of conserved structural motifs present in viral IRES elements would be advantageous for searching for regions putatively folding as IRES-like subdomains in genome sequences. Indeed, the relationship between RNA structure and biological function is generally inferred from the conservation of structural motifs. Although high-resolution structures of picornavirus IRES elements are still lacking, there is extensive information regarding the secondary structure of picornavirus IRES elements. For instance, several motifs are conserved in type II IRES elements [107]. Notably, some of them are placed in domain 3, a region that determines a three-dimensional folding of the IRES [108]. The apical region of this domain harbors conserved motifs involved in the structural organization of the IRES molecule, such as the GNRA (N stands for any nucleotide and R for purine) motif, which adopts a tetraloop conformation in several picornavirus IRES elements [66, 109–111]. Computational modeling of this domain generated a three-dimensional RNA structure that integrates experimental evidence for tertiary contacts between distant residues of the secondary structure [112]. This model also indicated that not only the motifs exposed on loops, but also the sequence of the junctions is conserved in natural isolates, implying that the secondary structure is evolutionary constrained to deliver its function.

Given the unusual combination of motifs that constrains domain 3 of the FMDV IRES [19, 113], it was hypothesized that a search for RNA sequences with the capacity to adopt an IRES-like subdomain fold could identify RNA regions potentially promoting IRES activity, undetectable by other approaches. The RNA inverse folding approach allows searching for sequences that are predicted to adopt an RNA structure similar to conserved RNA motifs. Indeed, several RNA sequences were predicted to adopt an IRES-like subdomain following application of computational filters and biological insight to prioritize the hits returned by RNAiFold [114]. More importantly, one of these sequences was proven to confer weak but positive internal initiation [115]. The weak activity of this region is not surprising, since viral IRES elements are active as an entity and not as short fragments or even individual domains. Despite the low level of activity of the selected candidate, the RNA inverse folding approach opens new avenues to search for other structurally conserved motifs present in well-characterized IRES elements. Future studies aimed to explore the presence of structural subdomains of other IRES elements, such as the dicistrovirus tRNA motif, may help to uncover new IRES-like motifs in eukaryotic genomes.

6 Concluding Remarks

A common property of many IRES elements is their RNA flexibility, which increases in correlation with the requirements of factors to assemble competent translation initiation complexes. The question that remains open is whether the lack of a universal structural motif in IRES elements indicates no common evolutionary traits in these regulatory elements. Even IRES elements from genetically distant viral RNAs lack overall conserved features. RNA flexibility, hence plasticity, is inherent to relative disordered regions of the RNA structure. In this regard, it is well established that many interactions are mediated by sequence motifs in the regulatory regions of genes and in the intrinsically disordered regions of proteins [116]. On the other hand, regulatory modules have a plasticity that facilitates a rapid acquisition, resulting in large networks that may respond in a coordinated manner to changes in the cell state. These elements, which are present in DNA and RNA, and proteins, suggest that coregulation and motif-driven regulatory programs are mechanisms that emerge from the use of dynamic modules. Thus, a hypothesis that accommodates the diversity of IRES elements should take into account that most of them exploit RNA structure flexibility, and therefore plasticity, as the core functional element.

Acknowledgments This work was supported by grant BFU2014-54564-P from MINECO and by an Institutional grant from Fundación Ramón Areces.

References

1. Sonenberg N, Hinnebusch AG. Regulation of translation initiation in eukaryotes: mechanisms and biological targets. Cell. 2009;136(4):731–45.
2. Pelletier J, Sonenberg N. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. Nature. 1988;334(6180):320–5.
3. Jang SK, Krausslich HG, Nicklin MJ, Duke GM, Palmenberg AC, Wimmer E. A segment of the 5′ nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. J Virol. 1988;62(8):2636–43.
4. Meyer KD, Patil DP, Zhou J, Zinoviev A, Skabkin MA, Elemento O, et al. 5′ UTR m(6)A promotes cap-independent translation. Cell. 2015;163(4):999–1010.
5. Zhou J, Wan J, Gao X, Zhang X, Jaffrey SR, Qian SB. Dynamic m(6)A mRNA methylation directs translational control of heat shock response. Nature. 2015;526(7574):591–4.
6. Paek KY, Hong KY, Ryu I, Park SM, Keum SJ, Kwon OS, et al. Translation initiation mediated by RNA looping. Proc Natl Acad Sci USA. 2015;112(4):1041–6.
7. Walsh D, Mohr I. Viral subversion of the host protein synthesis machinery. Nat Rev Microbiol. 2011;9(12):860–75.
8. Honda M, Ping LH, Rijnbrand RC, Amphlett E, Clarke B, Rowlands D, et al. Structural requirements for initiation of translation by internal ribosome entry within genome-length hepatitis C virus RNA. Virology. 1996;222(1):31–42.
9. Wilson JE, Pestova TV, Hellen CU, Sarnow P. Initiation of protein synthesis from the A site of the ribosome. Cell. 2000;102(4):511–20.
10. Vallejos M, Deforges J, Plank TD, Letelier A, Ramdohr P, Abraham CG, et al. Activity of the human immunodeficiency virus type 1 cell cycle-dependent internal ribosomal entry site is modulated by IRES trans-acting factors. Nucleic Acids Res. 2011;39(14):6186–200.
11. Herbreteau CH, Weill L, Decimo D, Prevot D, Darlix JL, Sargueil B, et al. HIV-2 genomic RNA contains a novel type of IRES located downstream of its initiation codon. Nat Struct Mol Biol. 2005;12(11):1001–7.
12. Xue S, Tian S, Fujii K, Kladwang W, Das R, Barna M. RNA regulons in Hox 5′ UTRs confer ribosome specificity to gene regulation. Nature. 2015;517(7532):33–8.
13. Du X, Wang J, Zhu H, Rinaldo L, Lamar KM, Palmenberg AC, et al. Second cistron in CACNA1A gene encodes a transcription factor mediating cerebellar development and SCA6. Cell. 2013;154(1):118–33.
14. Burkart C, Fan JB, Zhang DE. Two independent mechanisms promote expression of an N-terminal truncated USP18 isoform with higher DeISGylation activity in the nucleus. J Biol Chem. 2012;287(7):4883–93.
15. Henis-Korenblit S, Shani G, Sines T, Marash L, Shohat G, Kimchi A. The caspase-cleaved DAP5 protein supports internal ribosome entry site-mediated translation of death proteins. Proc Natl Acad Sci USA. 2002;99(8):5400–5.
16. Ul-Hussain M, Olk S, Schoenebeck B, Wasielewski B, Meier C, Prochnow N, et al. Internal ribosomal entry site (IRES) activity generates endogenous carboxyl-terminal domains of Cx43 and is responsive to hypoxic conditions. J Biol Chem. 2014;289(30):20979–90.
17. Sasaki J, Nakashima N. Translation initiation at the CUU codon is mediated by the internal ribosome entry site of an insect picorna-like virus in vitro. J Virol. 1999;73(2):1219–26.
18. Nakashima N, Uchiumi T. Functional analysis of structural motifs in dicistroviruses. Virus Res. 2009;139(2):137–47.
19. Fernandez N, Fernandez-Miragall O, Ramajo J, Garcia-Sacristan A, Bellora N, Eyras E, et al. Structural basis for the biological relevance of the invariant apical stem in IRES-mediated translation. Nucleic Acids Res. 2011;39(19):8572–85.
20. Fernandez N, Buddrus L, Pineiro D, Martinez-Salas E. Evolutionary conserved motifs constrain the RNA structure organization of picornavirus IRES. FEBS Lett. 2013;587(9):1353–8.

21. Au HH, Cornilescu G, Mouzakis KD, Ren Q, Burke JE, Lee S, et al. Global shape mimicry of tRNA within a viral internal ribosome entry site mediates translational reading frame selection. Proc Natl Acad Sci USA. 2015;112(47):E6446–55.
22. Fernandez IS, Bai XC, Murshudov G, Scheres SH, Ramakrishnan V. Initiation of translation by cricket paralysis virus IRES requires its translocation in the ribosome. Cell. 2014;157 (4):823–31.
23. Martinez-Salas E, Francisco-Velilla R, Fernandez-Chamorro J, Lozano G, Diaz-Toledano R. Picornavirus IRES elements: RNA structure and host protein interactions. Virus Res. 2015; 206:62–73.
24. Pestova TV, Shatsky IN, Fletcher SP, Jackson RJ, Hellen CU. A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initiation of hepatitis C and classical swine fever virus RNAs. Genes Dev. 1998;12(1):67–83.
25. Filbin ME, Vollmar BS, Shi D, Gonen T, Kieft JS. HCV IRES manipulates the ribosome to promote the switch from translation initiation to elongation. Nat Struct Mol Biol. 2013;20 (2):150–8.
26. Kieft JS, Zhou K, Jubin R, Doudna JA. Mechanism of ribosome recruitment by hepatitis C IRES RNA. RNA. 2001;7(2):194–206.
27. Perard J, Leyrat C, Baudin F, Drouet E, Jamin M. Structure of the full-length HCV IRES in solution. Nat Commun. 2013;4:1612.
28. Yamamoto H, Collier M, Loerke J, Ismer J, Schmidt A, Hilal T, et al. Molecular architecture of the ribosome-bound Hepatitis C Virus internal ribosomal entry site RNA. EMBO J. 2015;34(24):3042–58.
29. Garcia-Sacristan A, Moreno M, Ariza-Mateos A, Lopez-Camacho E, Jaudenes RM, Vazquez L, et al. A magnesium-induced RNA conformational switch at the internal ribosome entry site of hepatitis C virus genome visualized by atomic force microscopy. Nucleic Acids Res. 2015;43(1):565–80.
30. Hashem Y, des Georges A, Dhote V, Langlois R, Liao HY, Grassucci RA, et al. Hepatitis-C-virus-like internal ribosome entry sites displace eIF3 to gain access to the 40S subunit. Nature. 2013;503(7477):539–43.
31. Asnani M, Kumar P, Hellen CU. Widespread distribution and structural diversity of Type IV IRESs in members of Picornaviridae. Virology. 2015;478:61–74.
32. Bailey JM, Tapprich WE. Structure of the 5′ nontranslated region of the coxsackievirus b3 genome: chemical modification and comparative sequence analysis. J Virol. 2007;81(2): 650–68.
33. Yu Y, Sweeney TR, Kafasla P, Jackson RJ, Pestova TV, Hellen CU. The mechanism of translation initiation on Aichivirus RNA mediated by a novel type of picornavirus IRES. EMBO J. 2011;30(21):4423–36.
34. Kolupaeva VG, Pestova TV, Hellen CU, Shatsky IN. Translation eukaryotic initiation factor 4G recognizes a specific structural element within the internal ribosome entry site of encephalomyocarditis virus RNA. J Biol Chem. 1998;273(29):18599–604.
35. de Breyne S, Yu Y, Unbehaun A, Pestova TV, Hellen CU. Direct functional interaction of initiation factor eIF4G with type 1 internal ribosomal entry sites. Proc Natl Acad Sci USA. 2009;106(23):9197–202.
36. Ali IK, McKendrick L, Morley SJ, Jackson RJ. Activity of the hepatitis A virus IRES requires association between the cap-binding translation initiation factor (eIF4E) and eIF4G. J Virol. 2001;75(17):7854–63.
37. Pisarev AV, Chard LS, Kaku Y, Johns HL, Shatsky IN, Belsham GJ. Functional and structural similarities between the internal ribosome entry sites of hepatitis C virus and porcine teschovirus, a picornavirus. J Virol. 2004;78(9):4487–97.
38. Olarte-Castillo XA, Heeger F, Mazzoni CJ, Greenwood AD, Fyumagwa R, Moehlman PD, et al. Molecular characterization of canine kobuvirus in wild carnivores and the domestic dog in Africa. Virology. 2015;477:89–97.
39. Lau SK, Woo PC, Lai KK, Huang Y, Yip CC, Shek CT, et al. Complete genome analysis of three novel picornaviruses from diverse bat species. J Virol. 2011;85(17):8819–28.

40. Hellen CU, de Breyne S. A distinct group of hepacivirus/pestivirus-like internal ribosomal entry sites in members of diverse picornavirus genera: evidence for modular exchange of functional noncoding RNA elements by recombination. J Virol. 2007;81(11):5850–63.
41. Lozano G, Trapote A, Ramajo J, Elduque X, Grandas A, Robles J, et al. Local RNA flexibility perturbation of the IRES element induced by a novel ligand inhibits viral RNA translation. RNA Biol. 2015;12(5):555–68.
42. Boerneke MA, Dibrov SM, Gu J, Wyles DL, Hermann T. Functional conservation despite structural divergence in ligand-responsive RNA switches. Proc Natl Acad Sci USA. 2014;111(45):15952–7.
43. Laxton C, Brady K, Moschos S, Turnpenny P, Rawal J, Pryde DC, et al. Selection, optimization, and pharmacokinetic properties of a novel, potent antiviral locked nucleic acid-based antisense oligomer targeting hepatitis C virus internal ribosome entry site. Antimicrob Agents Chemother. 2011;55(7):3105–14.
44. Stone JK, Rijnbrand R, Stein DA, Ma Y, Yang Y, Iversen PL, et al. A morpholino oligomer targeting highly conserved internal ribosome entry site sequence is able to inhibit multiple species of picornavirus. Antimicrob Agents Chemother. 2008;52(6):1970–81.
45. Fajardo T Jr, Rosas MF, Sobrino F, Martinez-Salas E. Exploring IRES Region Accessibility by Interference of Foot-and-Mouth Disease Virus Infectivity. PLoS ONE. 2012;7(7):e41382.
46. Soto-Rifo R, Rubilar PS, Limousin T, de Breyne S, Decimo D, Ohlmann T. DEAD-box protein DDX3 associates with eIF4F to promote translation of selected mRNAs. EMBO J. 2012;31(18):3745–56.
47. Locker N, Chamond N, Sargueil B. A conserved structure within the HIV gag open reading frame that controls translation initiation directly recruits the 40S subunit and eIF3. Nucleic Acids Res. 2011;39(6):2367–77.
48. Dorokhov YL, Ivanov PA, Komarova TV, Skulachev MV, Atabekov JG. An internal ribosome entry site located upstream of the crucifer-infecting tobamovirus coat protein (CP) gene can be used for CP synthesis in vivo. J Gen Virol. 2006;87(Pt 9):2693–7.
49. Groppelli E, Belsham GJ, Roberts LO. Identification of minimal sequences of the Rhopalosiphum padi virus 5′ untranslated region required for internal initiation of protein synthesis in mammalian, plant and insect translation systems. J Gen Virol. 2007; 88(Pt 5):1583–8.
50. Abaeva IS, Pestova TV, Hellen CU. Attachment of ribosomal complexes and retrograde scanning during initiation on the Halastavi arva virus IRES. Nucleic Acids Res. 2016;44 (5):2362–77.
51. Le Quesne JP, Stoneley M, Fraser GA, Willis AE. Derivation of a structural model for the c-myc IRES. J Mol Biol. 2001;310(1):111–26.
52. Baird SD, Lewis SM, Turcotte M, Holcik M. A search for structurally similar cellular internal ribosome entry sites. Nucleic Acids Res. 2007;35(14):4664–77.
53. Jimenez-Gonzalez AS, Fernandez N, Martinez-Salas E. Sanchez de Jimenez E. Functional and structural analysis of maize hsp101 IRES. PLoS ONE. 2014;9(9):e107459.
54. Bisio A, Latorre E, Andreotti V, Bressac-de Paillerets B, Harland M, Scarra GB, et al. The 5′-untranslated region of p16INK4a melanoma tumor suppressor acts as a cellular IRES, controlling mRNA translation under hypoxia through YBX1 binding. Oncotarget. 2015;6 (37):39980–94.
55. Spahn CM, Jan E, Mulder A, Grassucci RA, Sarnow P, Frank J. Cryo-EM visualization of a viral internal ribosome entry site bound to human ribosomes: the IRES functions as an RNA-based translation factor. Cell. 2004;118(4):465–75.
56. Matsuda D, Mauro VP. Base pairing between hepatitis C virus RNA and 18S rRNA is required for IRES-dependent translation initiation in vivo. Proc Natl Acad Sci USA. 2014;111(43):15385–9.
57. Chappell SA, Edelman GM, Mauro VP. A 9-nt segment of a cellular mRNA can function as an internal ribosome entry site (IRES) and when present in linked multiple copies greatly enhances IRES activity. Proc Natl Acad Sci USA. 2000;97(4):1536–41.

58. Bhattacharyya D, Diamond P, Basu S. An Independently folding RNA G-quadruplex domain directly recruits the 40S ribosomal subunit. Biochemistry. 2015;54(10):1879–85.
59. Malygin AA, Kossinova OA, Shatsky IN, Karpova GG. HCV IRES interacts with the 18S rRNA to activate the 40S ribosome for subsequent steps of translation initiation. Nucleic Acids Res. 2013;41(18):8706–14.
60. Angulo J, Ulryck N, Deforges J, Chamond N, Lopez-Lastra M, Masquida B, et al. LOOP IIId of the HCV IRES is essential for the structural rearrangement of the 40S-HCV IRES complex. Nucleic Acids Res. 2016;44(3):1309–25.
61. Weingarten-Gabbay S, Elias-Kirma S, Nir R, Gritsenko AA, Stern-Ginossar N, Yakhini Z, et al. Comparative genetics. Systematic discovery of cap-independent translation sequences in human and viral genomes. Science. 2016;351(6270).
62. Sean P, Nguyen JH, Semler BL. Altered interactions between stem-loop IV within the 5′ noncoding region of coxsackievirus RNA and poly(rC) binding protein 2: effects on IRES-mediated translation and viral infectivity. Virology. 2009;389(1–2):45–58.
63. Fernandez N, Garcia-Sacristan A, Ramajo J, Briones C, Martinez-Salas E. Structural analysis provides insights into the modular organization of picornavirus IRES. Virology. 2011;409 (2):251–61.
64. Kaminski A, Belsham GJ, Jackson RJ. Translation of encephalomyocarditis virus RNA: parameters influencing the selection of the internal initiation site. EMBO J. 1994;13(7): 1673–81.
65. Jang SK, Wimmer E. Cap-independent translation of encephalomyocarditis virus RNA: structural elements of the internal ribosomal entry site and involvement of a cellular 57-kD RNA-binding protein. Genes Dev. 1990;4:1560–72.
66. Fernandez-Miragall O, Martinez-Salas E. Structural organization of a viral IRES depends on the integrity of the GNRA motif. RNA. 2003;9(11):1333–44.
67. Lozano G, Fernandez N, Martinez-Salas E. Magnesium-dependent folding of a picornavirus IRES element modulates RNA conformation and eIF4G interaction. FEBS J. 2014;281 (16):3685–700.
68. de Quinto Lopez. S, Martinez-Salas E. Interaction of the eIF4G initiation factor with the aphthovirus IRES is essential for internal translation initiation in vivo. RNA. 2000;6(10): 1380–92.
69. Clark AT, Robertson ME, Conn GL, Belsham GJ. Conserved nucleotides within the J domain of the encephalomyocarditis virus internal ribosome entry site are required for activity and for interaction with eIF4G. J Virol. 2003;77(23):12441–9.
70. de Quinto Lopez. S, Lafuente E, Martinez-Salas E. IRES interaction with translation initiation factors: functional characterization of novel RNA contacts with eIF3, eIF4B, and eIF4GII. RNA. 2001;7(9):1213–26.
71. Fernandez-Chamorro J, Pineiro D, Gordon JM, Ramajo J, Francisco-Velilla R, Macias MJ, et al. Identification of novel non-canonical RNA-binding sites in Gemin5 involved in internal initiation of translation. Nucleic Acids Res. 2014;42(9):5742–54.
72. Ochs K, Rust RC, Niepmann M. Translation initiation factor eIF4B interacts with a picornavirus internal ribosome entry site in both 48S and 80S initiation complexes independently of initiator AUG location. J Virol. 1999;73(9):7505–14.
73. Pineiro D, Fernandez N, Ramajo J, Martinez-Salas E. Gemin5 promotes IRES interaction and translation control through its C-terminal region. Nucleic Acids Res. 2013;41(2):1017–28.
74. Fernandez-Miragall O, Lopez de Quinto S, Martinez-Salas E. Relevance of RNA structure for the activity of picornavirus IRES elements. Virus Res. 2009;139(2):172–82.
75. Serrano P, Ramajo J, Martinez-Salas E. Rescue of internal initiation of translation by RNA complementation provides evidence for a distribution of functions between individual IRES domains. Virology. 2009;388(1):221–9.
76. Roberts LO, Belsham GJ. Complementation of defective picornavirus internal ribosome entry site (IRES) elements by the coexpression of fragments of the IRES. Virology. 1997;227 (1):53–62.

77. Andreev DE, Fernandez-Miragall O, Ramajo J, Dmitriev SE, Terenin IM, Martinez-Salas E, et al. Differential factor requirement to assemble translation initiation complexes at the alternative start codons of foot-and-mouth disease virus RNA. RNA. 2007;13(8):1366–74.
78. Sweeney TR, Abaeva IS, Pestova TV, Hellen CU. The mechanism of translation initiation on Type 1 picornavirus IRESs. EMBO J. 2014;33(1):76–92.
79. Yu Y, Abaeva IS, Marintchev A, Pestova TV, Hellen CU. Common conformational changes induced in type 2 picornavirus IRESs by cognate trans-acting factors. Nucleic Acids Res. 2011;39(11):4851–65.
80. Kafasla P, Morgner N, Poyry TA, Curry S, Robinson CV, Jackson RJ. Polypyrimidine tract binding protein stabilizes the encephalomyocarditis virus IRES structure via binding multiple sites in a unique orientation. Mol Cell. 2009;34(5):556–68.
81. Spriggs KA, Cobbold LC, Jopling CL, Cooper RE, Wilson LA, Stoneley M, et al. Canonical initiation factor requirements of the Myc family of internal ribosome entry segments. Mol Cell Biol. 2009;29(6):1565–74.
82. Majumder M, Yaman I, Gaccioli F, Zeenko VV, Wang C, Caprara MG, et al. The hnRNA-binding proteins hnRNP L and PTB are required for efficient translation of the Cat-1 arginine/lysine transporter mRNA during amino acid starvation. Mol Cell Biol. 2009;29 (10):2899–912.
83. Mitchell SA, Spriggs KA, Bushell M, Evans JR, Stoneley M, Le Quesne JP, et al. Identification of a motif that mediates polypyrimidine tract-binding protein-dependent internal ribosome entry. Genes Dev. 2005;19:1556–71.
84. Mitchell SA, Spriggs KA, Coldwell MJ, Jackson RJ, Willis AE. The Apaf-1 internal ribosome entry segment attains the correct structural conformation for function via interactions with PTB and unr. Mol Cell. 2003;11(3):757–71.
85. Yang B, Hu P, Lin X, Han W, Zhu L, Tan X, et al. PTBP1 induces ADAR1 p110 isoform expression through IRES-like dependent translation control and influences cell proliferation in gliomas. Cell Mol Life Sci. 2015;72(22):4383–97.
86. Ungureanu NH, Cloutier M, Lewis SM, de Silva N, Blais JD, Bell JC, et al. Internal ribosome entry site-mediated translation of Apaf-1, but not XIAP, is regulated during UV-induced cell death. J Biol Chem. 2006;281(22):15155–63.
87. de Quinto Lopez. S, Saiz M, de la Morena D, Sobrino F, Martinez-Salas E. IRES-driven translation is stimulated separately by the FMDV 3′-NCR and poly(A) sequences. Nucleic Acids Res. 2002;30(20):4398–405.
88. Song Y, Friebe P, Tzima E, Junemann C, Bartenschlager R, Niepmann M. The hepatitis C virus RNA 3′-untranslated region strongly enhances translation directed by the internal ribosome entry site. J Virol. 2006;80(23):11579–88.
89. Bradrick SS, Dobrikova EY, Kaiser C, Shveygert M, Gromeier M. Poly(A)-binding protein is differentially required for translation mediated by viral internal ribosome entry sites. RNA. 2007;13(9):1582–93.
90. Marom L, Hen-Avivi S, Pinchasi D, Chekanova JA, Belostotsky DA, Elroy-Stein O. Diverse poly(A) binding proteins mediate internal translational initiation by a plant viral IRES. RNA Biol. 2009;6(4):446–54.
91. Martinez-Salas E, Pineiro D, Fernandez N. Alternative Mechanisms to Initiate Translation in Eukaryotic mRNAs. Comp Funct Genomics. 2012;2012:391546.
92. Meng Z, Jackson NL, Shcherbakov OD, Choi H, Blume SW. The human IGF1R IRES likely operates through a Shine-Dalgarno-like interaction with the G961 loop (E-site) of the 18S rRNA and is kinetically modulated by a naturally polymorphic polyU loop. J Cell Biochem. 2010;110(2):531–44.
93. Peredo EL, Les DH, King UM, Benoit LK. Extreme conservation of the psaA/psaB intercistronic spacer reveals a translational motif coincident with the evolution of land plants. J Mol Evol. 2012;75(5–6):184–97.

94. Guerriero G, Spadiut O, Kerschbamer C, Giorno F, Baric S, Ezcurra I. Analysis of cellulose synthase genes from domesticated apple identifies collinear genes WDR53 and CesA8A: partial co-expression, bicistronic mRNA, and alternative splicing of CESA8A. J Exp Bot. 2012;63(16):6045–56.
95. Bahar Halpern K, Veprik A, Rubins N, Naaman O, Walker MD. GPR41 gene expression is mediated by internal ribosome entry site (IRES)-dependent translation of bicistronic mRNA encoding GPR40 and GPR41 proteins. J Biol Chem. 2012;287(24):20154–63.
96. Grabow WW, Jaeger L. RNA self-assembly and RNA nanotechnology. Acc Chem Res. 2014;47:1871–80.
97. Gao F, Gulay SP, Kasprzak W, Dinman JD, Shapiro BA, Simon AE. The kissing-loop T-shaped structure translational enhancer of Pea enation mosaic virus can bind simultaneously to ribosomes and a 5′ proximal hairpin. J Virol. 2013;87(22):11987–2002.
98. Dreher TW. Role of tRNA-like structures in controlling plant virus replication. Virus Res. 2009;139:217–29.
99. Lyons AJ, Robertson HD. Detection of tRNA-like structure through RNase P cleavage of viral internal ribosome entry site RNAs near the AUG start triplet. J Biol Chem. 2003;278 (29):26844–50.
100. Serrano P, Gomez J, Martinez-Salas E. Characterization of a cyanobacterial RNase P ribozyme recognition motif in the IRES of foot-and-mouth disease virus reveals a unique structural element. RNA. 2007;13(6):849–59.
101. Piron M, Beguiristain N, Nadal A, Martinez-Salas E, Gomez J. Characterizing the function and structural organization of the 5′ tRNA-like motif within the hepatitis C virus quasispecies. Nucleic Acids Res. 2005;33(5):1487–502.
102. Lozano G, Martinez-Salas E. Structural insights into viral IRES-dependent translation mechanisms. Curr Opin Virol. 2015;12:113–20.
103. Saghatelian A, Couso JP. Discovery and characterization of smORF-encoded bioactive polypeptides. Nature Chem Biol. 2015;11(12):909–16.
104. Wellensiek BP, Larsen AC, Stephens B, Kukurba K, Waern K, Briones N, et al. Genome-wide profiling of human cap-independent translation-enhancing elements. Nat Methods. 2013;10(8):747–50.
105. Gilbert WV, Zhou K, Butler TK, Doudna JA. Cap-independent translation is required for starvation-induced differentiation in yeast. Science. 2007;317(5842):1224–7.
106. Peguero-Sanchez E, Pardo-Lopez L, Merino E. IRES-dependent translated genes in fungi: computational prediction, phylogenetic conservation and functional association. BMC Genom. 2015;16(1):1059.
107. Lozano G, Fernandez N, Martinez-Salas E. Modeling three-dimensional structural motifs of viral IRES. J Mol Biol. 2016;428(5 Pt A):767–76.
108. Ramos R, Martinez-Salas E. Long-range RNA interactions between structural domains of the aphthovirus internal ribosome entry site (IRES). RNA. 1999;5(10):1374–83.
109. Phelan M, Banks RJ, Conn G, Ramesh V. NMR studies of the structure and Mg2+ binding properties of a conserved RNA motif of EMCV picornavirus IRES element. Nucleic Acids Res. 2004;32(16):4715–24.
110. Du Z, Ulyanov NB, Yu J, Andino R, James TL. NMR structures of loop B RNAs from the stem-loop IV domain of the enterovirus internal ribosome entry site: a single C to U substitution drastically changes the shape and flexibility of RNA. Biochemistry. 2004;43 (19):5757–71.
111. Bhattacharyya S, Das S. An apical GAGA loop within 5′ UTR of the Coxsackievirus B3 RNA maintains structural organization of the IRES element required for efficient ribosome entry. RNA Biol. 2006;3(2).
112. Jung S, Schlick T. Candidate RNA structures for domain 3 of the foot-and-mouth-disease virus internal ribosome entry site. Nucleic Acids Res. 2013;41(3):1483–95.
113. Fernandez-Miragall O, Ramos R, Ramajo J, Martinez-Salas E. Evidence of reciprocal tertiary interactions between conserved motifs involved in organizing RNA structure essential for internal initiation of translation. RNA. 2006;12(2):223–34.

114. Garcia-Martin JA, Dotu I, Clote P. RNAiFold 2.0: a web server and software to design custom and Rfam-based RNA molecules. Nucleic Acids Res. 2015;43(W1):W513–21.
115. Dotu I, Lozano G, Clote P, Martinez-Salas E. Using RNA inverse folding to identify IRES-like structural subdomains. RNA Biol. 2013;10(12):1842–52.
116. Oldfield CJ, Dunker AK. Intrinsically disordered proteins and intrinsically disordered protein regions. Annu Rev Biochem. 2014;83:553–84.

Erratum to: Evolution of TOR and Translation Control

Bruno D. Fonseca, Tyson E. Graber, Huy-Dung Hoang, Asier González, Alexander A. Soukas, Greco Hernández, Tommy Alain, Stephanie L. Swift, Ronit Weisman, Christian Meyer, Christophe Robaglia, Joseph Avruch and Michael N. Hall

Erratum to: "Evolution of TOR and Translation Control" in: G. Hernández and R. Jagus (eds.), *Evolution of the Protein Synthesis Machinery and Its Regulation*, DOI 10.1007/978-3-319-39468-8_15

The subjected book was published with the following errors:

- An incorrect spelling of the author's name as "Alexandre A. Soukas" in Chap.15 and in the table of contents. It should be "Alexander A. Soukas".

The updated original online version for this chapter can be found at 10.1007/978-3-319-39468-8_15

B.D. Fonseca (✉) · T.E. Graber · H.-D. Hoang · T. Alain · S.L. Swift
Children's Hospital of Eastern Ontario Research Institute, 401 Smyth Road, Ottawa, ON K1H 8L1, Canada
e-mail: brunodfonseca@gmail.com

A. González · M.N. Hall
Biozentrum, University of Basel, 4056 Basel, Switzerland

A.A. Soukas · J. Avruch
Massachusetts General Hospital, Boston, MA, USA

G. Hernández
Division of Basic Science, National Institute of Cancer, 22 San Fernando Avenue, Tlalpan, 14080 Mexico City, Mexico

R. Weisman
Department of Molecular Microbiology and Biotechnology, Tel Aviv University, Tel Aviv, Israel

R. Weisman
Department of Natural and Life Sciences, The Open University of Israel, Raanana, Israel

G. Hernández and R. Jagus (eds.), *Evolution of the Protein Synthesis Machinery and Its Regulation*, DOI 10.1007/978-3-319-39468-8_23

- A figure was placed incorrectly on p. 354 whereas it should have been placed on p. 366.
- A typographical error occurred on page 358 (line no 988). The spelling of a word was given as “nicotic” whereas it should be “nicotinic”.

The erratum has been updated in the book and chapters.

C. Meyer
Institut Jean-Pierre Bourgin, Unité Mixte de Recherche 1318, Institut National de la Recherche Agronomique, AgroParisTech, ERL CNRS 3559, Saclay Plant Sciences, 78026 Versailles Cedex, France

C. Robaglia
Aix-Marseille Université, Laboratoire de Génétique et Biophysique des Plantes, 13009 Marseille, France

C. Robaglia
CNRS, UMR 7265 Biologie Végétale & Microbiologie Environnementales, 13009 Marseille, France

C. Robaglia
CEA, IBEB, 13009 Marseille, France