Review

Translation and Translational Control in Dinoflagellates

Sougata Roy [1], Rosemary Jagus [2] and David Morse [1,*] (ID)

[1]   Institut de Recherche en Biologie Végétale, Département de Sciences Biologiques, Université de Montréal, 4101 Sherbrooke East, Montréal, QC H1X 2B2, Canada; sougatairbv20@gmail.com

[2]   Institute of Marine & Environmental Technology, University of Maryland Center for Environmental Science701 E. Pratt St., Baltimore, MD 21202, USA; jagus@umces.edu

*   Correspondence: david.morse@umontreal.ca; Tel.: +1-514-343-2133; Fax: +1-514-343-2288

Abstract: Dinoflagellates are unicellular protists that feature a multitude of unusual nuclear features, including large genomes, packaging of DNA without histones, and multiple gene copies organized as tandem gene arrays. Furthermore, all dinoflagellate mRNAs experience trans-splicing with a common 22-nucleotide splice leader (SL) sequence. These features challenge some of the concepts and assumptions about the regulation of gene expression derived from work on model eukaryotes such as yeasts and mammals. Translational control in the dinoflagellates, based on extensive study of circadian bioluminescence and by more recent microarray and transcriptome analyses, is now understood to be a crucial element in regulating gene expression. A picture of the translation machinery of dinoflagellates is emerging from the recent availability of transcriptomes of multiple dinoflagellate species and the first complete genome sequences. The components comprising the translational control toolkit of dinoflagellates are beginning to take shape and are outlined here.

Keywords: dinoflagellate; post-transcriptional control; gene expression; translation factors

1. Introduction

Dinoflagellates are a diverse group of unicellular protists that exhibit a great range of size, form and lifestyle. These organisms are notable for their symbioses with coral, the production of harmful algal blooms (HAB), and circadian bioluminescence. The free-living species are important primary producers in the marine environment, yet toxins released by HAB species can result in massive fish kills, human and marine mammal intoxications, as well as economic losses in fisheries and tourism. Dinoflagellates are also known for their diverse types of plastids; since plastids have been gained and lost multiple times, considerable insight has been gained into the process of endosymbiosis.

Dinoflagellates are members of the alveolata and are a sister group to the parasitic apicomplexans such as *Toxoplasma gondii* and *Plasmodium falciparum* as well as to the ciliates such as *Tetrahymena* and *Paramecium* [1–3]. While the major outlines of the alveolate lineages were established with rRNA phylogenies, protein-coding genes have been used to refine the branching patterns within the dinoflagellates. Protein-based phylogenies indicate the heterotroph *Oxyrrhis marina* and the perkinsosoan parasite *Perkinsus marinus* represent early branching members from the dinoflagellate lineage [4]. Within the dinoflagellates, there are two major clades, the syndinians and what have been termed the "core dinoflagellates" (formerly the Dinophyceae) [5,6]. Roughly half the marine dinoflagellates can grow autotrophically (photosynthetic) and are thought to be the most important eukaryotic primary producers after diatoms. The autotrophic dinoflagellates are found mostly in the core dinoflagellate group and can be either free-living, or associated with a broad range of hosts as endosymbionts, although many species are now recognized to be heterotrophic or mixotrophic [7]. The closely related syndinian dinoflagellates are primarily parasitic [8].

The dinoflagellates all contain alveoli, flattened membrane bound compartments of unknown function lying underneath the plasma membrane, that has given rise to the name alveolata. They also share a number of unusual nuclear characteristics including generally high DNA content, numerous chromosomes that remain condensed during interphase, DNA in a liquid crystal structure without observable nucleosomes and DNA in which 5-hydroxymethyluracil replaces some of the thymine [9–11]. Based on these features, the term dinokaryon was coined to designate nuclei in which the chromosomes appear fibrillar in appearance and histones are absent [12–14]. Core dinoflagellates have often undergone large scale gene duplication which has given rise to tandem gene arrays [15–17]. Tandemly-arrayed gene copies have no identifiable promoters in the intergenic spacers [15,18] and have been proposed to be transcribed into polycistronic mRNAs [19], although this has recently been challenged [20] and remains open to debate. Genomic sequence has now provided evidence that the upstream regions of dinoflagellate genes contain a conserved TTTT motif [17], of particular interest since the dinoflagellates have replaced the TATA-box binding protein (TBP) found in other eukaryotes with a TATA-box like protein (TLF) that binds TTTT instead [21]. It was previously thought that dinoflagellates lacked histone proteins [22] and instead used histone-like proteins for DNA organization [23], but recent work has refined this original view. First, transcriptome studies have revealed the presence of expressed genes encoding four core nucleosomal histones, H2A, H2B, H3, and H4 [24,25]. Conservation of these sequences suggests they are used, albeit not globally since protein levels are extremely hard to detect. Second, dinoflagellates express a dinoflagellate/viral nucleoprotein (DVNP) that appears to have displaced major histone functions in dinoflagellates [26]. DVNP is present in the core dinoflagellates, the syndinian *Hematodinium* sp. and *O. marina*, but is not found in *P. marinus* [27]. Since ribosomal protein phylogeny now firmly places *O. marina* as a lineage basal to both the syndinians and the core dinoflagellates, it seems likely that remodeling of the nucleus occurred after divergence of *P. marinus* and prior to the emergence of *O. marina*. Lastly, and subsequently to acquisition of DVNP, members of the core dinoflagellates acquired one of two distinct types of bacterial HU–like proteins, HLP-1 or HLP-2 [27]. Interestingly, the timing of this acquisition coincides with the appearance of the liquid crystalline structure with arched fibrils that are one of the hallmarks of dinoflagellate chromosomes. Presumably as a consequence of these latter substitutions, the protein-DNA ratio in core dinoflagellates is much lower than in syndinians, such as *Hematodinium* spp, which were previously considered more like typical eukaryotes because of their lower chromosome number, smaller genomes and no obvious gene amplification [16,28].

The unique physical features of dinoflagellate chromosomes are likely to affect both transcription and replication. Dinoflagellate DNA is packaged at a protein:DNA ratio of 1:10, unlike the equimolar ratios found in other eukaryotes [29]. Dinoflagellates possess some of the largest nuclear genomes known among eukaryotes (1500–245,000 Mbp), and have been predicted to contain 38,000–87,000 protein-coding genes [30]. Until recently, large genome sizes had deterred whole-genome sequencing. However, a complete and a partial genome sequence have now been reported for *Symbiodinium kawagutii* and *S. minutum* (which so far have the smallest dinoflagellate genomes) [17,31], with roughly 37,000 and 42,000 genes identified, respectively. A large number of genes (up to 68% in *S. kawagutii*) are found in gene families, suggestive of lineage-specific gene expansion by duplication. However, both transcriptomic [20] and genomic [17] analyses have shown there is a paucity of sequence-specific transcription factors in dinoflagellates, consistent with constant, steady transcription of most genes with fewer genes under sequence-specific transcriptional control.

2. Post-Transcriptional Regulation of Gene Expression in Dinoflagellates

Dinoflagellates are the most common and dramatic source of bioluminescence in the ocean [32] and have attracted keen scientific interest from the early pioneers in microscopy in the eighteenth century to modern molecular biologists. Although many dinoflagellates species exhibit bioluminescence, *Lingulodinium polyedra* (formerly *Gonyaulax polyedra*) has a brief but intense luminescent flash that is visible to the naked eye and easily measurable [33]. Bioluminescence by

L. polyedra, along with photosynthesis [34], cell division [35] and diurnal vertical migration [36] are regulated by an endogenous circadian (daily) clock [37]. Indeed, *L. polyedra* has been used for over 60 years as a model system in which to address the biochemical mechanisms underlying the observed rhythms, and the circadian clock has been shown to control the synthesis of the proteins involved in bioluminescence [38].

Studies on the clock-controlled rhythm of bioluminescence in *L. polyedra* began in the 1950s when it was noticed that light production is 40 to 60 times greater at night than during the day [39]. Light emission in dinoflagellates depends on luciferase-catalyzed oxidation of a high-energy luciferin substrate [40], normally sequestered and protected from oxidation by a luciferin binding protein (LBP) [41]. Both LBP and dinoflagellate luciferase show circadian rhythms in activity and abundance corresponding to the bioluminescence rhythm [42,43] and LBP has a measured peak of synthesis at the time when protein levels are increasing fastest in the cell [42]. Despite the rapid synthesis of LBP during the first several hours of the night phase, LBP mRNA remains constant over a 24-h period indicating that transcription is unlikely to regulate protein synthesis rates. There is a high amplitude to the bioluminescence rhythm because LBP is specifically degraded at the end of the night, by an as yet unknown mechanism.

In *L. polyedra*, the luminescence-related proteins are not the only ones to be circadian-controlled at the translational level. Two-dimensional gel analysis of *L. polyedra* extracts, from cells incorporating radiolabeled amino acids in vivo during the day and night phases, showed that many proteins exhibit circadian changes in synthesis rates [44,45]. However, no change in the amounts of protein synthesized by in vitro translation from poly(A)$^+$ RNA extracted from cells during day and night phases was observed, indicating that the mRNA itself is present and translatable at all times of the circadian cycle [44]. Consistent with this, analysis of mRNA recruitment by investigation of polysome profiles showed that during the subjective day phase (that corresponding to the day phase in constant conditions) less than half the ribosomes were present as polysomes [46]. In contrast, following the light to dark transition, polysome size increased and more than 90% of the ribosomes were found as polysomes, supporting the existence of translational control through changes in mRNA recruitment.

Other examples of translational control of specific transcripts in dinoflagellates have been reported. The NADP-dependent isocitrate dehydrogenase (NADP-ICDH) in the tricarboxylic acid (TCA) metabolic cycle also exhibits circadian changes of protein abundance and enzyme activity in *L. polyedra*, whereas its mRNA level remains constant throughout the diel cycle [47]. Circadian changes in protein synthesis, but not mRNA levels, have also been reported for glyceraldehyde 3-phosphate dehydrogenase (GAPDH) [48], peridinin-chlorophyll *a*-binding protein (PCP) [49], and superoxide dismutase [50]. In fact, most studies in this species show post-transcriptional control of gene expression. This includes genes typically under transcriptional control in other eukaryotes, such as cell cycle genes. In plants, fungi and metazoans, the transcription of several conserved S-phase genes such as proliferating cell nuclear antigen (PCNA), ribonucleotide reductase 2, replication factor C, and replication protein A is activated at S-phase entry by the E2F transcription factor, ensuring their timely availability for DNA synthesis (reviewed in [51,52]). In contrast, in *Karenia brevis*, transcript levels for these genes are unchanged over the cell cycle [53]. However, their proteins are maximally expressed during S-phase, suggesting their cell-cycle-dependent expression may be achieved at the level of translation. Microarray studies using *Karenia brevis* showed only 4% of the genes queried had transcript levels that changed by more than 1.7-fold in response to nitrogen or phosphorus limitation [54], and only 9.8% of the genes differed between day and night [55]. Among these latter, less than 1% of the genes queried varied both during growth under light dark (LD) cycles and in constant light [55]. Similarly, a microarray analysis of 3500 genes from *Pyrocystis lunula* revealed that fewer than 10% of gene transcripts increased in abundance by 2-fold and only ~1% of gene transcripts increased in abundance by 4-fold in response to treatment with 1 mM sodium nitrite or 0.5 mM paraquat [56]. In the dinoflagellate *P. lunula,* only 3% of the genes on the microarray were found to exhibit changes in transcript abundance (between 2- and 2.5-fold) during a diel cycle [57]. A similar

low magnitude transcriptome restructuring is found concurrent with the entry into stationary phase in *K. brevis* [53] and *Alexandrium minutum* [58].

Next-generation sequencing technologies have recently made genome-scale analyses of these organisms possible, and studies exploiting them confirm a paucity of transcriptional control. For example, a massively parallel signature sequencing (MPSS) analysis of the dinoflagellate *Alexandrium tamarense* transcriptome has shown that of a total of 40,029 signatures, only 18, 2, and 12 signatures were found exclusively in the nutrient-replete, nitrogen-depleted, and phosphate-depleted cultures, respectively [59]. Signature sequencing delivers several million 17 nucleotide transcript-derived fragments adjacent to a common restriction site in the cDNA and provides a quantitative overview of the transcriptome. The presence of bacteria had the most significant impact on the transcriptome, although changes were observed with only ~1% of the total number of transcribed genes, and only ~1.3% signatures were transcriptionally regulated under any condition [59]. This is lower than was found for the related species, *A. fundyense*, where MPSS showed that only 10% of the signatures were differentially expressed under nitrate- and phosphate-limited conditions [60]. Similarly, 454 pyrosequencing in the basal dinoflagellate, *O. marina*, revealed only 9 and 21 transcripts to be up- and down-regulated by saline stress, respectively, a total of 30 out of 7398 cDNA contigs [61]. Only 11 of these 30 sequences varied by more than 2-fold. Lastly, RNA-Seq analysis of *L. polyedra* showed that no transcripts varied significantly over the course of a light-dark or a circadian cycle, indicating no genes were regulated by either light or the circadian clock [62]. Overall, transcriptome studies have shown that although dinoflagellates do employ some transcriptional gene regulation, post-transcriptional regulation is a dominant and well-documented feature of the regulation of gene expression in these organisms. The importance of translational control in *L. polyedra* is underscored by the poor agreement between the amount of a protein and the level of its transcript, as measured for 3199 proteins by MS-based sequencing and by RNA-Seq for their transcripts [63].

3. Translational Control of Gene Expression in Eukaryotes

Gene expression is a multistep process that involves the transcription, translation and turnover of messenger RNAs and proteins. Throughout the eukaryotic tree of life, gene regulation at the translational level is widespread and significant. However, only a few studies have quantified the entire cascade on a genome-wide scale. The extent of gene regulation at the translational level has been demonstrated during early *Drosophila* embryogenesis on a genome-wide basis by determining ribosomal density and ribosomal occupancy of over 10,000 transcripts during the first ten hours after egg laying [64]. The diversity of the translation profiles suggests that multiple mechanisms are responsible for modulating transcript-specific translation. Furthermore, mRNAs involved in some biological processes are co-regulated at the translational level at certain developmental stages. In mammalian fibroblasts, absolute measurements of mRNA and protein abundance, along with turnover measures by parallel metabolic pulse labeling, have been investigated for more than 5000 genes [65]. Using this genome-scale investigation it was found that the cellular abundance of proteins is predominantly controlled at the level of translation. Translational control offers a rapid and specific response to the environmental and physiological changes, and transient alterations of genome-wide translational states are frequently observed in non-steady states, such as cell differentiation and stress responses (reviewed in [66,67]). Even in steady state situations, transcripts with the highest abundance are usually not those found in the highest levels on ribosomes [68,69].

The degree of translational control in dinoflagellates is anticipated to be greater than in metazoans since so far the transcriptional responses to different stimuli appear modest (reviewed [70,71]) and there are few sequence-specific transcription factors [20]. However, despite a wealth of data pointing to the importance of regulation of the translation in dinoflagellates little is known about underlying mechanisms. Given the significance of translational regulation in dinoflagellates, it is important to first identify the components of their translational machinery and then to determine what regulatory mechanisms are in place. Data mining of the increasing number of dinoflagellate transcriptomes

suggests that translation and its control in these organisms may show some important differences with that of model organisms such as yeast, the model plant *Arabidopsis* and mammals.

3.1. Eukaryotic Translation

Protein synthesis is an ancient, conserved and complex multicomponent process in which a mRNA template is decoded to form a protein sequence. Of the three fundamental life processes—transcription, translation and DNA replication—translation is the most conserved. Amino acid polymerization occurs on the ribosome, a dynamic nucleoprotein machine that has inherent peptidyl-transferase activity and utilizes aminoacyl-tRNAs and a range of protein factors. The basic mechanism of protein synthesis is similar in all organisms and can be divided into four stages: initiation, elongation, termination and recycling. Although translational modulation can occur at all of these stages, translational kinetics have revealed that the major regulatory step is initiation which serves as a target for a multitude of regulatory cues [72,73].

3.2. Initiation of Translation

Translation initiation refers to the formation of translation-competent ribosomes, where the initiator tRNA (Met-tRNA$_i$) in the ribosomal P-site is base-paired to the mRNA start codon. Bacterial ribosomes, such as in *E. coli*, bind to the mRNA by sequence complementarity between the 3′ terminus of the 16S ribosomal RNA (rRNA) and a Shine–Dalgarno (SD) sequence in the 5′ end of the mRNA. Only three initiation factors, IF1, IF2, and IF3 are required to locate the start codon (reviewed in [74]), and the mRNAs are predominantly polycistronic (>60%). However, alternatives to the SD sequence-dependent mechanism are quite widespread, as only 50–70% of all bacterial transcripts contain a SD sequence [75]. In the archaea, most mRNAs either have leader sequences lacking SD motifs, or are lacking a 5′-UTR entirely [76]. In some archaeal species, leaderless mRNAs represent over 50% of all mRNAs [77–79] and the mechanism of recruitment is poorly understood [76]. Archaea possess a larger number of translation initiation factors than bacteria (reviewed in [80–82]). At least six translation initiation factors, which contain up to three subunits, are described for archaea [81]. They include orthologs to the bacterial/eukaryotic factors IF1/eIF1A and IF2/eIF5B, which are thus universally conserved in all three domains of life. Archaea also contain orthologs to eukaryotic factors that are absent in bacteria; aIF1, aIF2, and eIF4A-like (reviewed in [81,82]).

In contrast to the prokaryotes, eukaryotic translation initiation relies on the 5′-cap structure of mRNA. Ribosomes bind to the 5′-end and travel in a 5′–3′ direction to locate the start codon. Furthermore, eukaryotic initiation involves 13 core initiation factors, some of which are large, multimeric complexes (reviewed in [73]). The eukaryotic translational machinery clearly originated in the archaea, or the common ancestor with archaea, but in addition includes eIF3, eIF4E, eIF4B (some eukaryotes), eIF4G, eIF4H, poly(A) binding protein (PABP) and eIF5. With the exception of eIF5, all eukaryotic-specific initiation factors are involved in mRNA recruitment. For the purposes of this review, eIF6 is considered to be a ribosome anti-association factor and has not been included with the initiation factors. Similarly, eIF2C should not be included as a translation factor since it functions in RNA silencing by binding to microRNAs and repressing the translation of mRNAs complementary to them [83]. Conversely, poly(A) binding protein (PABP), also involved in mRNA recruitment, is generally considered to be a translation initiation factor. With the exception of eIF5, all the eukaryotic-specific initiation factors (eIF4E, eIF4G, eIF4B, eIF4H, eIF3, and PABP) play roles in mRNA recruitment.

During translation initiation in eukaryotes, the small ribosomal subunit first binds an mRNA at its 5′-end and scans in a 5′ to 3′ direction to locate the initiation codon (reviewed in [84]). At this point, the large subunit joins to form the 80S initiation complex, which is now ready for protein synthesis. This entire process is guided by eukaryotic initiation factors (eIFs) (Figure 1). During eukaryotic cap-dependent translation initiation, a preassembled 43S preinitiation complex (PIC) is formed at the capped 5′-end of the mRNA through the interaction of the cap binding protein eIF4E, the scaffold

protein eIF4G, the DExD/H-box helicase eIF4A, and eIF4B, (also eIF4H in mammals) (reviewed in [85–88]). The PIC itself is made up of the small ribosomal subunit, the initiator tRNA (Met-tRNAi) and the initiation factors eIF1, eIF1A, eIF2, and eIF3. eIF2 is a GTPase, and in its guanosine triphosphate (GTP) bound form facilitates recruitment of the eIF4E-bound RNA to the 43 PIC (reviewed in [85]), while eIF3 recruits the small ribosomal subunit (reviewed in [84,87]). The helicase eIF4A, bound to the scaffold eIF4G, is thought to expose a single-stranded region in mRNA for interaction with the ribosome. 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 at the 5′ end of mRNA and promotes circularization of mRNA to improve initiation efficiency in repeated rounds of translation [89].
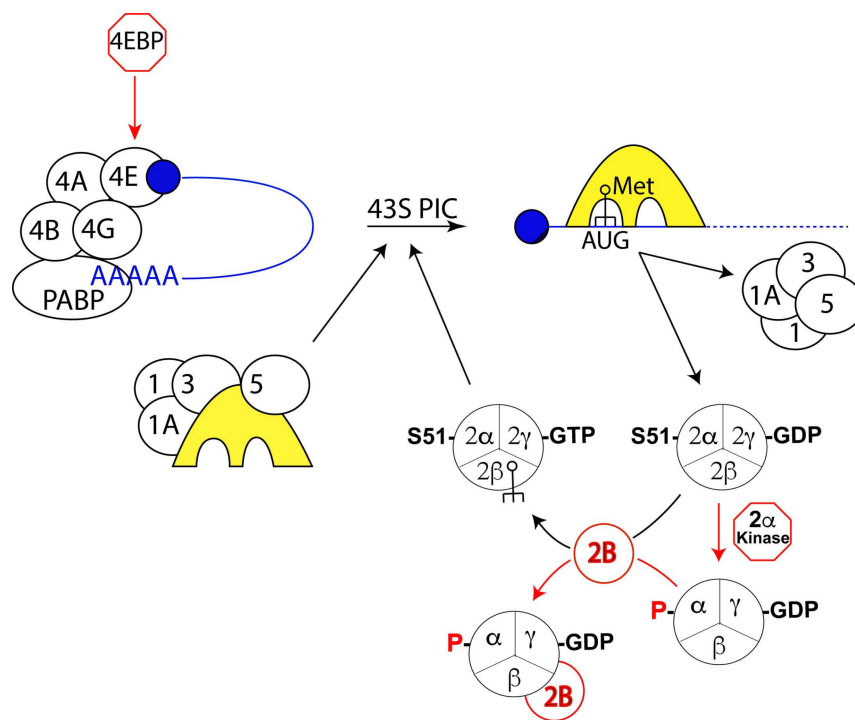

Figure 1. A schematic view of the factors involved in eukaryotic translation initiation. Transcripts (in blue) are present in the cytoplasm as a complex with the poly(A) binding protein (PABP), the cap-binding protein eIF4E, the scaffold protein eIF4G (linking PABP with the cap-binding protein) and two other initiation factors eIF4A and eIF4B. The 43S pre-initiation complex (43S PIC) is formed when the small ribosome subunit (yellow) in a complex with initiation factors eIF1, 1A, 3 and 5 binds the trimeric eIF2, which contains a tRNA charged with methionine (Met-tRNAi). The small ribosome subunit moves down the RNA sequence until the anticodon on the Met-tRNAi recognizes and binds the AUG start codon. This recognition allows hydrolysis of the GTP bound to eIF2, causing release of the eIF2 from the PIC along with the eIF1, eIF1A, eIF3 and eIF5. The GTP-bound form of eIF2 is regenerated by the guanine exchange factor eIF2B. Translation can be globally downregulated by phosphorylation of eIF2α, which then tightly binds eIF2B thus impeding recycling of eIF2. Orthologues for a GCN-type eIF2α kinase and the cap-binding regulatory protein 4E-BP have not been identified in the dinoflagellate transcriptome sequence assembly (TSA).

The 43S preinitiation complex (PIC) scans the mRNA leader in the 5′ to 3′ direction for an AUG codon in a suitable sequence context. During scanning, the 5′ UTR of mRNA is held within a channel on the 40S ribosome formed by the body and the head of the 40S subunit, along with eIF1 and eIF1A. mRNA slides through this channel. Scanning processivity is ensured by keeping the mRNA unstructured and oriented appropriately for "inspection" by the anticodon on the initiator tRNA

held in the partial P-site of the 43S PIC [90]. The "inspection" by the initiator tRNA mechanistically involves attempting to establish Watson–Crick base pairing between its anticodon and a nucleotide triplet in the mRNA as it moves through the partial P-site. Base pairing between the anticodon of Met-tRNAi and the AUG in the peptidyl-tRNA (P-) site of the 43S PIC is the initial event in start codon recognition [84,90–92]. AUG recognition stops the scanning process and triggers hydrolysis of eIF2-bound GTP to guanosine diphosphate (GDP). Hydrolysis of eIF2-bound GTP is the key step that commits the 43S PIC to initiation at a particular start codon; the two proteins that regulate the G-protein cycle are eIF5, which has both GTPase-activating protein (GAP) and GDP dissociation inhibitor (GDI) functions, and eIF2B, the guanine nucleotide exchange factor (GEF) for eIF2. Upon recognition of the initiation codon, eIF1 is displaced from the P-site, leading to the release of Pi and dissociation of eIF2. The dissociation of eIF1 and eIF2 from the 40S subunit results in a stable 48S mRNA-containing PIC. eIF5B binds to this complex and orients the acceptor stem of Met-tRNAi towards the P-site on the 60S subunit to facilitate subunit joining. eIF5B thus aids release of eIF2-GDP from the 48S PIC which promotes formation of the 80S initiation complex. The 80S complex contains Met-tRNAi base-paired to AUG in the P-site and is ready to begin the elongation phase of protein synthesis [93].

4. The Translation Machinery in Dinoflagellates

The translation machinery of dinoflagellates is reviewed from the perspective of its mRNAs, ribosomes and translation initiation factors.

4.1. Trans-Splicing of Mrna in Dinoflagellates

In most eukaryotes, the 5′ end of a transcript contains a monomethylated m^{7}GTP cap, co-transcriptionally added by a series of enzymatic reactions (reviewed in [94]). However, most if not all of the transcripts in dinoflagellates are post-transcriptionally modified at their 5′ end by *trans*-splicing. The splice leader (SL) RNA [95,96] is thus the source of the cap on the mRNA. SL *trans*-splicing was once considered an anomaly of the kinetoplastids, but subsequent identification of *trans*-splicing in dinoflagellates, euglenozoans and several major invertebrate phyla suggests that this form of RNA processing may represent an evolutionarily important form of regulating gene expression [97,98].

Trans-splicing has been found in all dinoflagellate species examined, including species in the basal lineages such as *Oxyrrhis* and *Amoebophrya* [16,61] and in *P. marinus*, indicating that this processing arose early in dinoflagellate evolution [96,99,100]. Because of the addition of the SL, all mRNA share a common 22 nucleotide sequence (5′-DCCGUAGCCATUUUGGCUCAAG-3′, with D = U, A, or G) at their 5′ end [96]. The identity of the cap structure for dinoflagellate transcripts needs to be determined, but there are more possibilities than the monomethylated m^{7}GTP typically found on mRNA or the trimethylguanosine found on small nuclear RNAs. For example, the SL in kinetoplastids contributes four consecutive modified nucleotides referred to as a cap-4 structure [101,102]. Independent of the nature of the cap, it is expected to play an important role in translation initiation. In *Trypanosoma brucei*, a minimal level of mRNA cap ribose methylation has been shown to be essential for viability, and hypermethylation of the cap-4 maximizes translation rates [103]. An unusual structure of the dinoflagellate cap would suggest that the initiation factors recognizing the dinoflagellate cap structure might also show some unique features.

4.2. mRNA Codon Flanking Regions

Analysis of the nucleotide frequency distribution in the flanking positions of the initiation codon, AUG, in eukaryotes has established a consensus sequence that is optimal for recognition by the partial P-site in the small ribosomal subunit [104]. This sequence is termed the Kozak sequence after the investigator who first recognized its importance [104,105]. In vertebrates, this sequence GCC<u>R</u>CC**AUG**<u>G</u> (where R = A or G) has conserved nucleotides at −3 (mostly A) and +4 (G) with respect to the +1 adenine of the start codon AUG (101). This consensus sequence differs between different phylogenetic groups. The consensus sequences of yeast (AAA<u>A</u>AA**AUG**UCU)

and plants (A<u>A</u>CA**AUG**<u>G</u>C) are similar, but not identical [106–108]. Protists such as Amoebozoa (AAA<u>A</u>AA**AUG**RNA), *Plasmodium* (UAAA<u>A</u>AA**AUG**AAN), *Toxoplasma* (GNCAAA**AUG**<u>G</u>) show similar conservation at the −3 and +4 positions. A compilation of transcript flanking sequences from two *Symbiodinium kawagutii* libraries generated <u>R</u>CC**AUG**<u>G</u>CN (R=A, G; N=A, G, C, T) [109], a sequence equivalent to that found in vertebrates and other protists. This sequence is also found in *L. polyedra* superoxide dismutase [110] and in the P-type protein pump in *Symbiodinium* [111].

4.3. Ribosomes in Dinoflagellates

At present, little is known about dinoflagellate ribosomes. The *Lingulodinium* transcriptome contains ribosomal proteins comparable to those of higher plant and mammalian genome [17,20,31]. However, when compared with eukaryotic and prokaryotic 5.8S rRNAs, *Prorocentrum micans* 5.8S rRNA revealed several dinoflagellate specific nucleotides. These distinctive nucleotides of 5.8S rRNA are found to be located in specific loops that have the potential to play a structural role in ribosome organization [112]. The size of the small subunit rRNA is similar to what is normally found, but the large subunit rRNA appears slightly smaller than the range of values typically found in eukaryotes [113].

4.4. Translation Initiation Factors in Dinoflagellates

A plethora of transcriptome data from multiple dinoflagellate species combined with the recently available genome sequence from *S. kawagutii* [17] have allowed the identification of translation initiation factors in a wide range of core dinoflagellates. In sharp contrast to the components of the transcription apparatus, the translational machinery in dinoflagellates resembles that of other eukaryotic organisms. This is consistent with the appreciable conservation of translation initiation factors evident for apicomplexans (reviewed in [114]), kinetoplasts (reviewed in [115,116]) and diatoms. Annotation of the *L. polyedra* transcriptome [20] first revealed almost all components of the eukaryotic translational machinery, and these results are confirmed by a more recent survey of the GenBank transcriptome sequence assembly (TSA) using eIF sequences from the vascular plant *Arabidopsis thaliana* [117] (Figure 2). Translational initiation factors not recognized so far include eIF4B and four of the five eIF-2B subunits. eIF4B also appears to be lacking in the genomes of apicomplexa and kinetoplasts, although the sequence is poorly conserved between distant eukaryotes and so may be difficult to recognize. The eIF4G-like sequences found so far seem incomplete, and some of the subunits of eIF3 core complex have not yet been definitively identified.
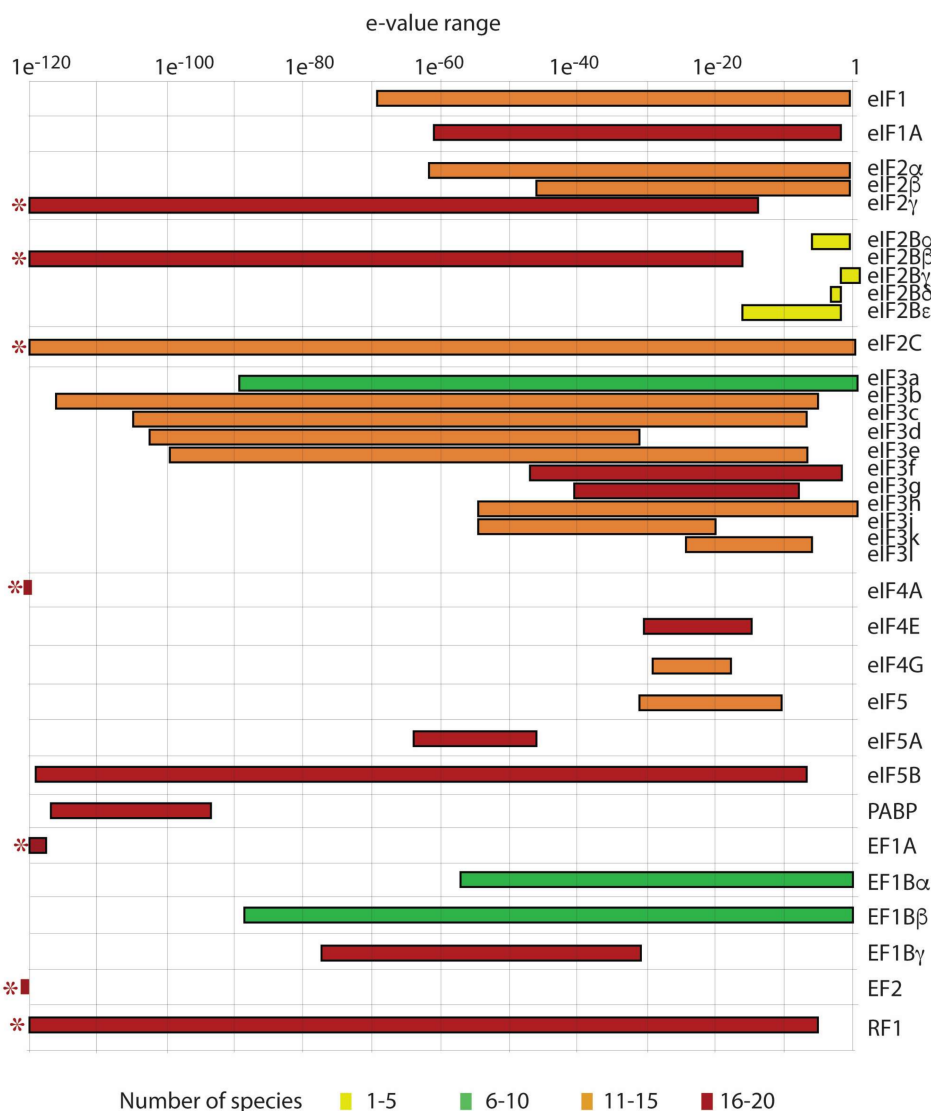
Figure 2. A list of eukaryotic initiation factors from *Arabidopsis thaliana* was used to query the dinoflagellate TSA using tBLASTn. Query sequences are listed at right, and the range of e-values (from e^{-120} to 1) are shown as horizontal red bars; an asterisk at left indicates an e-value of 0. The number of species in which hits were recovered is indicated by the color in the bar. Most of the eIF2B subunits are not found in dinoflagellates. Factors eIF4E. eIF4G and eIF5 are incomplete, suggesting their normal function may be complemented by different proteins in dinoflagellates.

4.5. eIF4E

The cap binding protein eIF4E is a central component in the recruitment of mRNA and regulation of translation (reviewed in [85,86,118,119]). The 5′-cap structure is a feature specific to eukaryotic transcripts, so unsurprisingly, eIF4E is an essential translational initiation factor found only in eukaryotes. It has a unique alpha/beta fold that is considered to have no homologs outside the eukaryotes as determined by sequence comparison or structural analyses [120]. The three-dimensional structure of eIF4E bound to a cap analog resembles a cupped hand in which the cap-structure is sandwiched between two conserved tryptophan residues (Trp-56 and Trp-102 of *H. sapiens* eIF4E) [121,122]. A third conserved Trp residue (Trp-166 of *H. sapiens* eIF4E) and an aspartate (Asp-103) also bind the cap-structure (Figure 3). Arg-112, Arg-157 and Lys-162 make direct contacts with the ribose and phosphate moieties. These basic amino acids generate the positive electrostatic potential that partially neutralizes the

charge on the two phosphate groups of m^7-GDP. Lastly, the motif (S/T)VxxFW-73 is a conserved region involved in binding eIF4G and a range of eIF4E interactive proteins [123].

　　Duplication of the genes encoding eIF4E seems to have taken place very early during eukaryotic evolution and has given rise to a family of proteins that function not only as prototypical initiation factors but also as regulators of mRNA recruitment and localization. Multiple eIF4E family members have been identified in a wide range of organisms including plants, flies, mammals, frogs, birds, nematodes, fish, and various protists (reviewed in [124]). These multiple family members seem to have resulted from a series of duplications of a single early eIF4E gene. The eIF4E family is made up of structurally related proteins within an organism, all containing a conserved core region of 160 to 170 amino acid residues [124]. Not all family members function as prototypical initiation factors, however. The prototypical eIF4E is considered to be eIF4E-1A of mammals, eIF4E and eIF (iso) 4E of plants, and eIF4E of *S. cerevisiae*. eIF4Es from plants, fungi and metazoans can be grouped into one of three classes [124], and different organisms express between one and eight eIF4E isoforms [124,125].
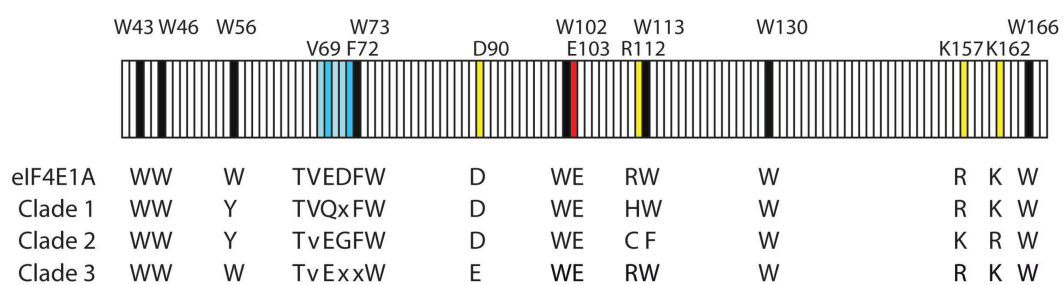

	W43 W46	W56	W73 V69 F72	D90	W102 E103	W113 R112	W130	W166 K157 K162
eIF4E1A	WW	W	TVEDFW	D	WE	RW	W	R K W
Clade 1	WW	Y	TVQxFW	D	WE	HW	W	R K W
Clade 2	WW	Y	TvEGFW	D	WE	C F	W	K R W
Clade 3	WW	W	TvExxW	E	WE	RW	W	R K W

Figure 3. A 130-amino acid conserved core of the mammalian eIF4E1A (subclade1a) illustrating the important binding regions for the m^{7}G cap and eIF4G. The eight conserved tryptophan residues (numbered as for the mammalian sequence) are shown in black along with the residues that are found at the corresponding positions in representative sequences from the three dinoflagellate clades. Residues in yellow are important for binding the phosphates of the m^{7}G (the aspartate at position 90 coordinates binding by arginine-157), while the eIF4G binding motif (S/TVxxFW) is shown in blue. The small v in the eIF4G binding motif could also be T (for clade 2) or A (for clade 3). The glutamate at position 103 (red) is involved in hydrogen bonding to the m7G. Note that dinoflagellate clade 1 eIF4Es have an insertion of 12–13 amino acids between W73 and W102, as well as an insertion of 7–9 amino acids between W130 and W166. Figure based on [126].

　　Dinoflagellates also encode multiple eIF4Es that cluster into three distinct clades, although none of these correspond to the previously described eIF4E classes [126]. However, because the cap structure in dinoflagellates is potentially different from that of other eukaryotes, it was anticipated that their eIF4E orthologs could show atypical features. Indeed, recent work has shown that the eIF4E family members of protists are distinct in many ways from those of non-protist eukaryotes [127]. Analysis of the relationship of selected eIF4E-family members from eleven dinoflagellate species shows clear evidence for three distinct clades [127]. Members of each clade differ significantly from each other, but all bear distinctive features of a cap-binding protein. The three clades can be divided into 9 subclasses, six of which were found in all eleven species [126]. There is considerable complexity in this gene family, with divergence seen at critical amino acids (Figure 3). This suggests that the different family members may have different roles.

　　eIF4Es from clade 1 are represented in all alveolate lineages, and are the only eIF4Es found in apicomplexans, suggesting the authentic translation initiation factor may be in this clade [126]. However, dinoflagellate clade 1 eIF4Es contain extended sequences between W73 and W102, and W130 to W166 compared to eIF4E-2 and -3, a feature not seen in any plant, fungi or metazoan eIF4E family member. The function of the extension is not known, but the different eIF4E-1 sub-clades show marked heterogeneity at this region. Sequence alignments of eIF4E family members from

nine core dinoflagellates [126] show that clade 1 and clade 2 eIF4Es have a Tyr substitution at the position equivalent to Trp-56 in mammalian Class I eIF4Es (Figure 3). A positively charged arginine at position 112 has been substituted with a positively charged histidine in clade 1 eIF4Es and with a non-conservative substitution of cysteine in clade 3 eIF4Es. Clade 1 and clade 3 eIF4Es have the conserved positively charged amino acids, equivalent to Arg-157 and Lys-162 of mammalian eIF4E1As, that are involved in binding the pyrophosphate bridge in metazoan eIF4E-1 [121,126]. The Clade 2 eIF4Es have Lys and Arg at positions equivalent to Arg-157 and Lys-162, respectively.

In non-protist eIF4E family members capable of binding eIF4G, the consensus sequence of the recognition motif is S/TVxxFW ending at W73 [121,122], and similar motifs are conserved in eIF4E sequences from core dinoflagellates [126]. However, there are subtle variations between the different eIF4E-1 sub-clades in the consensus sequence of the binding domain. Clade 2 eIF4Es have a tyrosine substitution at W73, and the region can be polar (TVQEFW), as in eIF4E-1a and -1b, acidic (TVEEFW) as in eIF4E-1c, or basic (TVKGFW), as in eIF4E-1d, suggesting different binding partners. The role of these motifs in dinoflagellates is unclear, since the dinoflagellate version of eIF4G lacks an eIF4E binding region (see below). However, it is to be expected that a range of other eIF4E interacting proteins (4E-IPs) may interact with these sites to play roles in mRNA recruitment or its regulation. An increasing number of proteins, generically known as eIF4E-interacting proteins, modulate the eIF4G–eIF4E interaction [128,129]. These include proteins such as maskin, cup, 4E-T, neuroguidin and others [128–131]. Sequences with similarity to mammalian neuroguidin, which has three eIF4E binding motifs, are found in both *Lingulodinium* and *Symbiodinium* (e-values of e^{-15}). In the axons and dendrites of the mammalian nervous system neuroguidin binds to eIF4E-1A and represses translation of cytoplasmic polyadenylation element (CPE) containing mRNAs [132], but it is not known what role, if any, they may play in dinoflagellates. No sequences encoding the regulatory eIF4E binding proteins (4E-BPs) have been identified in dinoflagellates to date. However, 4E-BPs are not consistently found throughout the eukaryotic tree of life. They are absent from plants and nematodes, but are present in some protists such as *Acanthamoeba castellani*, *Dictyostelium discoideum* and *Galucocystis nostochinearum* [123].

It has been suggested that the eIF4E family may contribute to the extensive translational control over gene expression seen in dinoflagellates [126]. Such a role would be greatly supported by differential transcript binding by the eIF4E family members and/or differential interaction with other components of the translational machinery.

4.6. eIF4G

In plants and metazoans, recruitment of capped mRNAs requires, in addition to eIF4E, the large scaffold protein eIF4G that serves to coordinate the assembly of the various translation initiation factors and the 40S ribosomal subunit. However, dinoflagellate transcriptomes do not contain sequences that encode a full-length eIF4G. Instead, all dinoflagellate eIF4G transcripts found to date encode only a portion of the mammalian eIF4G equivalent to the MIF4G/HEAT1 domain. This form is expected to allow interactions with eIF4A, eIF3 and RNA. However, the domains at the amino-terminal of eIF4G, responsible for binding PABP and eIF4E, as well as at the carboxyl-terminal region, responsible for binding the eIF4E kinase along with additional eIF4A binding sites, are missing in the dinoflagellate sequences. The binding domains for eIF4E and PABP allow mammalian eIF4G to coordinate independent interactions with mRNA via the cap and poly(A) tail which promotes circularization of the mRNA and a more efficient re-utilization of recruited mRNAs [133]. The dinoflagellate eIF4G is thus unlikely to be able to fulfill these functions, although it must be noted that multiple truncated atypical eIF4Gs have been found in trypanosomatids, and they have been shown to associate with different eIF4Es, eIF4As and PABPs to support translation [116,134,135]. Furthermore, in *S. cerevisiae*, eIF3 and eIF2 are more critical than eIF4G for stable binding of mRNA to native preinitiation complexes, and this appears to require the same RNA-binding domain found in the dinoflagellate eIF4G [88]. In *Leishmania*, eIF4G-3 contains the domain found in the dinoflagellate eIF4G as well as an eIF4E binding domain close to the amino terminal. It is usually assumed that a PABP binding domain is also

required to allow the "closed-loop" conformation which is crucial for efficient recruitment of the 43S PIC. However, eliminating the PABP–eIF4G interaction by deletion or mutation of the PABP-binding domain in eIF4G is not lethal in yeast and deletion of the PABP-binding domain has no effect on yeast cell growth provided that the RNA-binding region in the amino terminus of eIF4G1 (RNA1) is intact [136]. eIF3 contains subunits that bind RNA as isolated proteins, such that eIF3 could interact directly with mRNA in the initiation complex [137].

4.7. eIF2

All subunits of the translation initiation factor eIF2 are found in dinoflagellate transcriptomes. GTP-bound eIF2 also binds Met-tRNAi in what is termed a ternary complex, and this complex is essential in delivering the Met-tRNAi to the translational machinery during the initiation step of protein synthesis (reviewed in [84,86,138]). The α-subunit of eIF2 functions in a regulatory capacity, as phosphorylation of eIF2α reduces eIF2 activity as well as the overall rate of translation [138–140]. In fungi, metazoans, apicomplexans and trypanosomes, phosphorylation of eIF2α has been shown to regulate mRNA recruitment in response to environmental stress, latency and during entry into S-phase [138]. In vertebrates, five eIF2 kinases have been described, each of which is activated by unique stressors (reviewed in [141,142]). Regulation of translation by eIF2α-kinases has been shown to be critical in the life cycle of apicomplexans such as *Toxoplasma gondii* and *P. falciparum* [143–147]. Phosphorylation of eIF2α plays a key role in sustaining parasite viability when they are in their vulnerable extracellular form and when they transition from the proliferating tachyzoites to encysted bradyzoites forms [144–148]. Furthermore, when growing in glutamine-free media, knockouts of a GCN2-like eIF2α-kinase reduce the levels of eIF2α phosphorylation and cause a severe growth defect [149]. GCN2 kinases bind uncharged tRNAs and are thus activated when levels of amino acids are low. Since dinoflagellates are known to respond to a wide range of stressors such as oxidative stress, nitrogen depletion, phosphate limitation, and osmotic stress, it might be expected that phosphorylation of eIF2α would form part of their translational regulatory repertoire. However, only one sequence in the dinoflagellate transcriptome shotgun assembly shows similarity to the *T. gondii* GCN2-like eIF2α-kinase (9% coverage, e-value $9e^{-16}$) making its identity as an eIF2α-kinase uncertain. Another *T. gondii* eIF2α-kinase, TgIF2K-A, has significant hits with many dinoflagellate species (best blast hit 6% coverage, e-value $1e^{-31}$). TgIF2K-A has been proposed to regulate response to unfolded proteins in the ER [144].

The inhibitory effects of the eIF2α-kinases on translation are a consequence of reduced recycling of eIF2-GDP to eIF2-GTP by the guanine exchange factor (GEF) eIF2B (reviewed in [138,139]). eIF2B is a decameric protein, with two copies of the regulatory α, β, and δ subunits, and two copies of the catalytic γ and ε subunits. The αβδ regulatory subcomplex contributes to eIF2-GDP binding through interactions with eIF2α, and it is these interactions that are enhanced by phosphorylation of eIF2α [150]. However, of the eIF2B subunits, only eIF2Bβ has a significant presence in the dinoflagellate transcriptome shotgun assembly (Figure 2). Could eIF2Bβ function alone? A recent study in plants has discovered a new role for eIF2Bβ in mediating viral resistance [151], although it is not known if this requires the other eIF2B subunits to function. Dinoflagellate sequences have been found that have weak homology to the ε subunit, which has been shown to be able to act alone in guanine nucleotide exchange [152].

5. Translational Regulation by RNA-Binding Proteins

RNA in cells is bound to a number of RNA-binding proteins (RBP) forming ribonucleoprotein complexes that affect splicing, transport to the cytoplasm, translation, stability and subcellular localization [153]. The classical view of RBP is that they bind RNA using specialized RNA binding domains. Thus, one approach to the analysis of RBP in dinoflagellates is to compare sequence homology in the dinoflagellate transcriptome sequencing assembly with a number of known RNA binding domains. A recent publication on mammalian cells has provided a list of 15 RBP containing

a variety of different RNA binding domains [154]. When these were used to query eight different dinoflagellate species, four RBPs were found to have the greatest degree of sequence similarity to the dinoflagellates, RRM, the OB NTP fold, S1 and Pumilio (Figure 4).
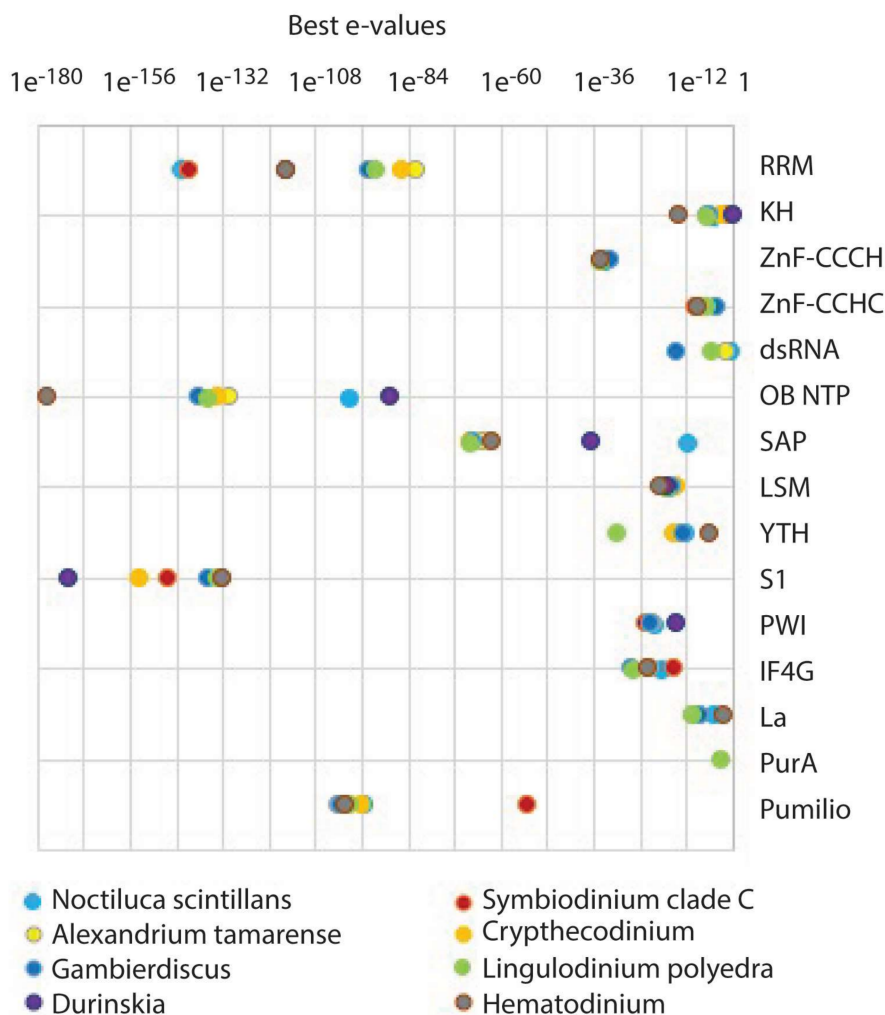

Figure 4. A list of RNA-binding proteins from mammalian cells was used to query the dinoflagellate TSA using tBLASTn, Individual species were queried one-by-one, and the best e-value recorded for each species. Mouse proteins (except where indicated) selected as representatives for the different domains are RRM, polyadenylate-binding protein 1; KH, *Plasmodium* PF3D7_0605100; ZnF-CCCH, splicing factor U2AF 35 kDa subunit; ZnF-CCHC, protein lin-28 homolog; dsRNA, *E. coli* Rnase III; OB NTP-binding, ATP-dependent RNA helicase A; SAP, heterogeneous nuclear ribonucleoprotein U; LSM, LSM14 homolog; YTH, YTH domain-containing family protein; S1, ATP-dependent RNA helicase DHX8; PWI, serine/arginine repetitive matrix protein; eIF4G, eukaryotic translation initiation factor 4 gamma; La, lupus La.

The poor sequence similarity to members of the other RNA binding protein suggests that the dinoflagellates may have evolved different RNA binding modules. Indeed, new RNA "interactome" studies, in which cross-linking RNA-protein complexes with UV light in vivo is followed by high-throughput proteomics, are providing a quantum shift in how RBP are viewed. Importantly, these proteomic studies, based on function, have revealed that many RBP lack a recognizable RNA binding domain [155]. It is thus possible that many RBPs still await discovery in the dinoflagellates.

The circadian regulation of lbp translation in *L. polyedra* has been correlated with binding of a protein factor to the 3′-UTR of the transcript [156]. Unfortunately, the protein factor was not identified at the time, and subsequent work to repeat the experiments using a different strain of *L. polyedra* have not shown evidence for protein binding by a variety of techniques [157].

6. Translational Regulation by Small RNAs

RNA interference (RNAi) is a mechanism whereby a few molecules of double stranded RNA introduced into a cell can interfere with gene expression [158]. RNAi is caused by small RNA molecules, termed small interfering RNAs (siRNAs) or microRNAs (miRNAs), and these can regulate translation in both plants and animals [159–162]. siRNAs, either synthetic or produced naturally, are short double stranded RNA molecules, 20–25 base pairs in length with two nucleotide overhangs on the 3′ ends of each strand. The natural siRNAs are formed by cleavage of longer double-stranded RNAs by an RNase III-like enzyme named Dicer or Dicer-like. The siRNA then bind to an RNA binding protein of the Argonaute/Piwi family, where they are denatured into single-stranded guide RNA. These guide RNAs, usually as part of a larger RNA-induced silencing complex called RISC, target RISC to mRNAs with a sequence complementary to the guide. miRNA differ from siRNAs in that after transcription they form a pre-miRNA molecule which is then processed. Processing involves first the formation of a roughly 60 base hairpin loop structure then, after export to the cytoplasm, cleavage by Dicer. In many eukaryotes, an RNA-dependent RNA polymerase (RdRP) generates double-stranded RNA from single stranded transcripts, thus acting to amplify the RNAi response [163].

The miRNA pathway is widespread but not ubiquitous, as it is not found in phylogenetically diverse organisms such as *Saccharomyces cerevisiae*, *Trypanosoma cruzi*, *Leishmania major* and *Cyanidioschyzon merolae* [164]. The classical pathway is also absent in the apicomplexan *Plasmodium*, a sister group to the dinoflagellates [165], although an abundance of antisense RNAs in the *Plasmodium* transcriptome suggests that these organisms may employ an alternative mechanism [166]. Surprisingly, ciliates (which belong to the same lineage as apicomplexans and dinoflagellates) as well as the heterokonts *Phytophthera*, contain all the components required for miRNA biogenesis and formation of the RISC complex [164]. Other heterokonts, such as the diatoms, lack canonical Dicer/Dicer-like or RdRP genes despite having mi/siRNAs and Argonaute family proteins [167]. This suggests that these organisms must have an alternative to Dicer in order to generate their small RNAs. BLAST searches have identified Dicer-like proteins in *Perkinsus marinus*, a basal lineage in the dinoflagellate clade, as well as in transcriptomes from the core dinoflagellates *Symbiodinium* and *Lingulodinium* [168]. This indicates that Dicer-like enzymes and Argonaute-like proteins are present, although RdRp related proteins have not yet been detected. Searches of the *Lingulodinium* and *Symbiodinium* transcriptomes also contain a protein with RdRp domain as well as several piwi domain-containing proteins (four in *L. polyedra* and two in *Symbiodinium*). The poorly named eIF2C found in dinoflagellates (Figure 2) is a member of the Argonaute family [169]. In mice and humans, this protein binds to miRNAs or siRNAs and represses translation of mRNAs with complementary sequences. It is also involved in transcriptional gene silencing of promoter regions that are complementary to short antisense RNAs bound to the promoter, as well as in the degradation of miRNA-bound mRNA targets [169].

Recent computational approaches identified 21 and 18 predicted small RNAs with miRNA characteristics in *Symbiodinium microadriaticum* and *Alexandrium tamarense*, respectively [170,171]. These predicted miRNAs have both animal (partial complementarity) and plant (perfect complementarity) properties with respect to their targets. In *S. kawagutii*, 102 miRNAs were identified using small RNA sequencing, and 99 of these were also found among the 367 miRNAs predicted from the genome [17]. Of these, the vast majority had partial complementarity to their targets with only one having perfect complementarity. Interestingly, confirmation of miRNA expression was made using Northern blots, and these experiments showed that some of the miRNAs were downregulated by an increase in the temperature from 25 to 35 °C [17]. One of these downregulated miRNAs had heat shock proteins 70 and 90 as predicted targets, thus suggesting a role in regulating gene expression. The fact that

miRNAs can in some cases mediate translation rates without affecting mRNA levels is potentially applicable to understanding circadian regulation of translation in *Lingulodinium* where changes in protein synthesis rates are not accompanied by changes in RNA levels. This was tested explicitly using small RNA sequencing to detect miRNAs in this species [168]. However, the results indicate miRNAs do not regulate circadian translation, as out of 9787 sequence reads corresponding to LBP only six were found in an antisense orientation, and none of the 5749 reads corresponding to luciferase were in an antisense orientation. Furthermore, attempts to introduce antisense LBP RNA by biolistic bombardment had no effect on LBP translation.

7. Post-Translational Regulation

It is important to note that control over the synthesis of a protein determines the amount but not necessarily the activity of a given protein, as several types of post-translational modifications can intervene to fine tune the levels of active protein. Phosphorylation is the most frequently found post-translational modification (PTM) although many other PTM are known. These include addition of small molecules (acetate, amides, hydroxyls or methyl groups), addition of carbohydrate groups (the process of glycosylation), addition of ubiquitin or small ubiquitin-like modifiers (SUMO), addition of hydrophobic groups such as palmitic acid, myristic acid or prenyl groups), or proteolytic cleavage [172]. Indeed, while protein abundance has long been considered as the most important step in regulating protein function, PTM can affect the structure and function of proteins and can induce effects that can be just as profound [173].

A recent survey of post-translational modifications showed that phosphorylation accounted for roughly two-thirds of the 72,430 experimentally characterized modifications [174]. Furthermore, almost one third of proteins can undergo phosphorylation in eukaryotes [175,176], though for the most part their significance is still unknown. Among those phosphorylated proteins that have been characterized, effects have been found on protein folding, enzyme activity, interactions between proteins, degradation rates and sub-cellular localization [175]. In terms of cellular activities, phosphorylation can be used to regulate signaling, growth, proliferation, differentiation and cell death [177]. Importantly, recent advances in phosphoprotein or phosphopeptide enrichment techniques have made analysis of phosphorylation amenable to analysis by high throughput mass spectroscopy sequencing. This technique requires either genomic or transcriptomic databases, but fortunately, the growing availability of dinoflagellate transcriptome sequence databases is opening the door to studies of proteomics and phosphoproteomics. There are now a sufficiently large number of species represented so that even species without a transcriptome can be analyzed by homology [178].

Phosphorylation is also part of the circadian clock mechanism in animals, plants, fungi and cyanobacterial [179–182]. Furthermore, in *Arabidopsis*, mRNA levels of several kinases and phosphatases are regulated by circadian clock [183,184], and these in turn regulate different rhythmic activities. In *Lingulodinium*, inhibitors of both serine/threonine kinases and phosphatases affect the timing of the bioluminescence rhythm [185,186], supporting a role for phosphorylation in the dinoflagellate clock mechanism. As yet, the full kinase repertoire in dinoflagellates is unknown, although all the kinases found in the transcriptome of *L. polyedra* have been cataloged [187]. It will clearly be of interest to determine the different kinase classes, their abundance and most importantly, their substrates at different times during the circadian cycle. These studies are likely to reveal considerable functional insight. For example, the variation of PCNA abundance with cell cycle stages in *K. brevis* was also accompanied by a shift in electrophoretic mobility consistent with the fact that post-translational modifications of PCNA, such as phosphorylation, are known to control its activity [188]. 2-D polyacrylamide gel electrophoresis (PAGE), phosphoprotein staining and LC-MS/MS after phosphopeptide enrichment were used to assess the phosphoproteome of *L. polyedra* at mid-day and mid-night [189], and the level of 8 among the 45 phosphoproteins varied more than 2-fold between the two times studied. Interestingly, three of these phosphoproteins contained RNA

binding domains [189], a finding pertinent to the predominant role of translational regulation in this species.

Lastly, it must also be noted that degradation rates also affect protein abundance in the cell. Unfortunately, protein degradation has not yet been studied in dinoflagellates. In most eukaryotic cells, a ubiquitous protein degradation mechanism involving addition of a small 76 amino acid protein called ubiquitin (Ub) to a target protein controls access to the proteasome, an organelle specialized for protein degradation [190]. Ubiquitination involves a complex of ubiquitin-activating (E1) enzymes, ubiquitin-conjugating (E2) enzymes and E3 ubiquitin ligases, with the choice among many different E2-E3 complexes determining the specificity of the ubiquitination reaction [191]. In both *Symbiodinium* and *Lingulodinium* most of the ubiquitin pathway elements are conserved [24]. BLAST searches have shown that many components to the ubiquitin-proteasome system are found in the transcriptome. In comparison to mammals, however, *Lingulodinium* contains 40% of the E1, E2 and E3 components, similar to what is found in *Arabidopsis*.

While the most well-studied role for ubiquitination lies in mediating protein degradation, the cellular consequences of the PTM depend on where, and how many, ubiquitin moieties are added to a target protein [192]. Target proteins can be mono-ubiquitinated or polyubiquitinated, and in the latter case, ubiquitin groups can be added to different amino acids in the target or to the same amino acid in a series or branched format. Thus, in addition to controlling the amount of protein, ubiquitination can influence protein-protein interactions, in particular by allowing binding to proteins with ubiquitin-binding domains, and this can subsequently regulate the activity and cellular localization of the target protein.

8. Conclusions

Dinoflagellates have often found novel ways of carrying out fundamental processes within the cell, making them useful systems in which to study variability in a highly-conserved process such as protein synthesis. Most of the expected players in the initiation of protein synthesis are found in dinoflagellates, at least with respect to their counterparts among the Alveolata. The exceptions are eIF2Bε and γ subunits, eIF4B and eIF4G. In addition, dinoflagellate eIF4G sequences are truncated and atypical, resembling those found in trypanosomatids. It seems likely that variability in mRNA recruitment may explain these differences. Further studies using biochemical approaches should allow functions of the novel eIF4E family members to be uncovered and the apparently "missing" components to be identified in the dinoflagellates, thus providing a unique evolutionary insight into the translational machinery.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1.	Cavalier-Smith, T. *The Biology of Free-Living Heterotrophic Flagellates*; Oxford Universty Press: Oxford, UK, 1991.
2.	Taylor, F.J. Ultrastructure as a control for protistan molecular phylogeny. *Am. Nat.* **1999**, *154*, S125–S136. [CrossRef] [PubMed]
3.	Fast, N.M.; Xue, L.; Bingham, S.; Keeling, P.J. Re-examining alveolate evolution using multiple protein molecular phylogenies. *J. Eukaryot. Microbiol.* **2002**, *49*, 30–37. [CrossRef] [PubMed]
4.	Saldarriaga, J.F.; McEwan, M.L.; Fast, N.M.; Taylor, F.J.; Keeling, P.J. Multiple protein phylogenies show that *Oxyrrhis marina* and *Perkinsus marinus* are early branches of the dinoflagellate lineage. *Int. J. Syst. Evol. Microbiol.* **2003**, *53*, 355–365. [CrossRef] [PubMed]

5.	Adl, S.M.; Simpson, A.G.; Farmer, M.A.; Andersen, R.A.; Anderson, O.R.; Barta, J.R.; Bowser, S.S.; Brugerolle, G.; Fensome, R.A.; Fredericq, S.; et al. The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *J. Eukaryot. Microbiol.* **2005**, *52*, 399–451. [CrossRef] [PubMed]

6.	Bachvaroff, T.R.; Gornik, S.G.; Concepcion, G.T.; Waller, R.F.; Mendez, G.S.; Lippmeier, J.C.; Delwiche, C.F. Dinoflagellate phylogeny revisited: Using ribosomal proteins to resolve deep branching dinoflagellate clades. *Mol. Phylogenet. Evol.* **2014**, *70*, 314–322. [CrossRef] [PubMed]

7.	Dagenais-Bellefeuille, S.; Morse, D. Putting the N in dinoflagellates. *Front. Microbiol.* **2013**, *4*, 369. [CrossRef] [PubMed]

8.	Taylor, F.J.R.; Hoppenrath, M.; Saldarriaga, J.F. Dinoflagellate diversity and distribution. *Biodivers. Conserv.* **2008**, *17*, 407–418. [CrossRef]

9.	Rae, P.M. Hydroxymethyluracil in eukaryote DNA: A natural feature of the pyrrophyta (dinoflagellates). *Science* **1976**, *194*, 1062–1064. [CrossRef] [PubMed]

10.	Steele, R.E.; Rae, P.M. Ordered distribution of modified bases in the DNA of a dinoflagellate. *Nucleic Acids Res.* **1980**, *8*, 4709–4725. [CrossRef] [PubMed]

11.	Moreno Diaz de la Espina, S.; Alverca, E.; Cuadrado, A.; Franca, S. Organization of the genome and gene expression in a nuclear environment lacking histones and nucleosomes: The amazing dinoflagellates. *Eur. J. Cell Biol.* **2005**, *84*, 137–149. [CrossRef] [PubMed]

12.	Spector, D.L. Dinoflagellate Nuclei. In *Dinoflagellates*; Spector, D.L., Ed.; Academic Press: New York, NY, USA, 1984; pp. 107–147.

13.	Bouligand, Y.; Norris, V. Chromosome separation and segregation in dinoflagellates and bacteria may depend on liquid crystalline states. *Biochimie* **2001**, *83*, 187–192. [CrossRef]

14.	Lin, S. Genomic understanding of dinoflagellates. *Res. Microbiol.* **2011**, *162*, 551–569. [CrossRef] [PubMed]

15.	Bachvaroff, T.R.; Place, A.R. From stop to start: Tandem gene arrangement, copy number and trans-splicing sites in the dinoflagellate *Amphidinium carterae*. *PLoS ONE* **2008**, *3*, e2929. [CrossRef] [PubMed]

16.	Bachvaroff, T.R.; Place, A.R.; Coats, D.W. Expressed sequence tags from *Amoebophrya* sp. infecting *Karlodinium veneficum*: Comparing host and parasite sequences. *J. Eukaryot. Microbiol.* **2009**, *56*, 531–541. [CrossRef] [PubMed]

17.	Lin, S.; Cheng, S.; Song, B.; Zhong, X.; Lin, X.; Li, W.; Li, L.; Zhang, Y.; Zhang, H.; Ji, Z.; et al. The *Symbiodinium kawagutii* genome illuminates dinoflagellate gene expression and coral symbiosis. *Science* **2015**, *350*, 691–694. [CrossRef] [PubMed]

18.	Le, Q.H.; Markovic, P.; Jovine, R.; Hastings, J.; Morse, D. Structure and organization of the peridinin-chlorophyll a-protein gene in *Gonyaulax polyedra*. *Mol. Gen. Genet.* **1997**, *255*, 595–604. [PubMed]

19.	Lukes, J.; Leander, B.S.; Keeling, P.J. Cascades of convergent evolution: The corresponding evolutionary histories of euglenozoans and dinoflagellates. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 9963–9970. [CrossRef] [PubMed]

20.	Beauchemin, M.; Roy, S.; Daoust, P.; Dagenais-Bellefeuille, S.; Bertomeu, T.; Letourneau, L.; Lang, B.F.; Morse, D. Dinoflagellate tandem array gene transcripts are highly conserved and not polycistronic. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 15793–15798. [CrossRef] [PubMed]

21.	Guillebault, D.; Sasorith, S.; Derelle, E.; Wurtz, J.M.; Lozano, J.C.; Bingham, S.; Tora, L.; Moreau, H. A new class of transcription initiation factor, intermediate between TBPs and TLFs is present in the marine unicellular organism: The dinoflagellate *Crypthecodinium cohnii*. *J. Biol. Chem.* **2002**, *277*, 40881–40886. [CrossRef] [PubMed]

22.	Rizzo, P. Biochemistry of the dinoflagellate nucleus. In *The Biology of Dinoflagellates*; Taylor, F.J.R., Ed.; Blackwell Scientific: Oxford, UK, 1987; pp. 143–173.

23.	Chan, Y.H.; Wong, J.T. Concentration-dependent organization of DNA by the dinoflagellate histone-like protein HCc3. *Nucleic Acids Res.* **2007**, *35*, 2573–2583. [CrossRef] [PubMed]

24.	Bayer, T.; Aranda, M.; Sunagawa, S.; Yum, L.K.; Desalvo, M.K.; Lindquist, E.; Coffroth, M.A.; Voolstra, C.R.; Medina, M. *Symbiodinium* transcriptomes: Genome insights into the dinoflagellate symbionts of reef-building corals. *PLoS ONE* **2012**, *7*, e35269. [CrossRef] [PubMed]

25.	Roy, S.; Morse, D. A full suite of histone and histone modifying genes are transcribed in the dinoflagellate *Lingulodinium*. *PLoS ONE* **2012**, *7*, e34340. [CrossRef] [PubMed]

26. Gornik, S.G.; Ford, K.L.; Mulhern, T.D.; Bacic, A.; McFadden, G.I.; Waller, R.F. Loss of nucleosomal DNA condensation coincides with appearance of a novel nuclear protein in dinoflagellates. *Curr. Biol.* **2012**, *22*, 2303–2312. [CrossRef] [PubMed]

27. Janouskovec, J.; Gavelis, G.S.; Burki, F.; Dinh, D.; Bachvaroff, T.R.; Gornik, S.G.; Bright, K.J.; Imanian, B.; Strom, S.L.; Delwiche, C.F.; et al. Major transitions in dinoflagellate evolution unveiled by phylotranscriptomics. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E171–E180. [CrossRef] [PubMed]

28. Cachon, J.; Cachon, M. Botanical Monographs. In *The Biology of Dinoflagellates*; Taylor, F.J.R., Ed.; Blackwell Scientific Publications: Hoboken, NJ, USA, 1987; Volume 21.

29. Rizzo, P.J.; Jones, M.; Ray, S.M. Isolation and properties of isolated nuclei from the Florida red tide dinoflagellate *Gymnodinium breve* (Davis). *J. Protozool.* **1982**, *29*, 217–222. [CrossRef] [PubMed]

30. Hou, Y.; Lin, S. Distinct gene number-genome size relationships for eukaryotes and non-eukaryotes: Gene content estimation for dinoflagellate genomes. *PLoS ONE* **2009**, *4*, e6978. [CrossRef] [PubMed]

31. Shoguchi, E.; Shinzato, C.; Kawashima, T.; Gyoja, F.; Mungpakdee, S.; Koyanagi, R.; Takeuchi, T.; Hisata, K.; Tanaka, M.; Fujiwara, M.; et al. Draft assembly of the *Symbiodinium minutum* nuclear genome reveals dinoflagellate gene structure. *Curr. Biol.* **2013**, *23*, 1399–1408. [CrossRef] [PubMed]

32. Wilson, T.; Hastings, J.W. Bioluminescence. *Annu. Rev. Cell Dev. Biol.* **1998**, *14*, 197–230. [CrossRef] [PubMed]

33. Hastings, J.; Sweeney, B.M. The luminescent reaction in extracts of the marine dinoflagellate *Gonyaulax polyedra*. *J. Cell. Physiol.* **1957**, *49*, 209–218. [CrossRef]

34. Hastings, J.W.; Astrachan, L.; Sweeney, B.M. A persistent daily rhythm in photosynthesis. *J. Gen. Physiol.* **1961**, *45*, 69–76. [CrossRef] [PubMed]

35. Sweeney, B.M.; Hastings, J.W. Effects of temperature upon diurnal rhythms. *Cold Spring Harb. Symp. Quant. Biol.* **1960**, *25*, 87–104. [CrossRef] [PubMed]

36. Eppley, R.W.; Holm-Hansen, O.; Strickland, J.D.H. Some observations on the vertical migration of dinoflagellates. *J. Phycol.* **1968**, *4*, 333–340. [CrossRef] [PubMed]

37. Morse, D.S.; Fritz, L.; Hastings, J.W. What is the clock? Translational regulation of circadian bioluminescence. *Trends biochem. Sci.* **1990**, *15*, 262–265. [CrossRef]

38. Hastings, J.W. The *Gonyaulax* clock at 50: Translational control of circadian expression. *Cold Spring Harb. Symp. Quant. Biol.* **2007**, *72*, 141–144. [CrossRef] [PubMed]

39. Sweeney, B.M.; Hastings, J.W. Characteristics of the diurnal rhythm of luminescence in *Gonyaulax polyedra*. *J. Cell. Physiol.* **1957**, *49*, 115–128. [CrossRef]

40. Nakamura, H.; Kishi, Y.; Shimomura, O.; Morse, D.; Hastings, J.W. Structure of dinoflagellate luciferin and its enzymatic and non-enzymatic air oxidation products. *J. Am. Chem. Soc.* **1989**, *111*, 7607–7611. [CrossRef]

41. Morse, D.; Pappenheimer, A.M., Jr.; Hastings, J.W. Role of a luciferin-binding protein in the circadian bioluminescent reaction of *Gonyaulax polyedra*. *J. Biol. Chem.* **1989**, *264*, 11822–11826. [PubMed]

42. Morse, D.; Milos, P.M.; Roux, E.; Hastings, J.W. Circadian regulation of bioluminescence in *Gonyaulax* involves translational control. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 172–176. [CrossRef] [PubMed]

43. Johnson, C.H.; Roeber, J.F.; Hastings, J.W. Circadian changes in enzyme concentration account for rhythm of enzyme activity in *Gonyaulax*. *Science* **1984**, *223*, 1428–1430. [CrossRef] [PubMed]

44. Milos, P.; Morse, D.; Hastings, J.W. Circadian control over synthesis of many *Gonyaulax* proteins is at a translational level. *Naturwiss* **1990**, *77*, 87–89. [CrossRef] [PubMed]

45. Markovic, P.; Roenneberg, T.; Morse, D. Phased protein synthesis at several circadian times does not change protein levels in *Gonyaulax*. *J. Biol. Rhythm.* **1996**, *11*, 57–67. [CrossRef] [PubMed]

46. Schroder-Lorenz, A.; Rensing, L. Circadian changes in protein-synthesis rate and protein phosphorylation in cell-free extracts of *Gonyaulax polyedra*. *Planta* **1987**, *170*, 7–13. [CrossRef] [PubMed]

47. Akimoto, H.; Kinumi, T.; Ohmiya, Y. Circadian rhythm of a TCA cycle enzyme is apparently regulated at the translational level in the dinoflagellate *Lingulodinium polyedrum*. *J. Biol. Rhythm.* **2005**, *20*, 479–489. [CrossRef] [PubMed]

48. Fagan, T.; Morse, D.; Hastings, J. Circadian synthesis of a nuclear encoded chloroplast Glyceraldehyde-3-phosphate dehydrogenase in the dinoflagellate *Gonyaulax polyedra* is translationally controlled. *Biochemistry* **1999**, *38*, 7689–7695. [CrossRef] [PubMed]

49. Ten Lohuis, M.R.; Miller, D.J. Light-regulated transcription of genes encoding peridinin chlorophyll a proteins and the major intrinsic light-harvesting complex proteins in the dinoflagellate *Amphidinium carterae* Hulburt (Dinophycae). Changes In cytosine methylation accompany photoadaptation. *Plant Physiol.* **1998**, *117*, 189–196. [PubMed]

50. Okamoto, O.K.; Robertson, D.L.; Fagan, T.F.; Hastings, J.W.; Colepicolo, P. Different regulatory mechanisms modulate the expression of a dinoflagellate iron-superoxide dismutase. *J. Biol. Chem.* **2001**, *276*, 19989–19993. [CrossRef] [PubMed]

51. Black, A.R.; Azizkhan-Clifford, J. Regulation of E2F: A family of transcription factors involved in proliferation control. *Gene* **1999**, *237*, 281–302. [CrossRef]

52. Lavia, P.; Jansen-Durr, P. E2F target genes and cell-cycle checkpoint control. *Bioessays* **1999**, *21*, 221–230. [CrossRef]

53. Brunelle, S.A.; van Dolah, F.M. Post-transcriptional regulation of S-phase genes in the dinoflagellate, *Karenia brevis*. *J. Eukaryot. Microbiol.* **2011**, *58*, 373–382. [CrossRef] [PubMed]

54. Morey, J.S.; Monroe, E.A.; Kinney, A.L.; Beal, M.; Johnson, J.G.; Hitchcock, G.L.; van Dolah, F.M. Transcriptomic response of the red tide dinoflagellate, *Karenia brevis*, to nitrogen and phosphorus depletion and addition. *BMC Genom.* **2011**, *12*, 346. [CrossRef] [PubMed]

55. Van Dolah, F.M.; Lidie, K.B.; Morey, J.S.; Brunelle, S.A.; Ryan, J.C.; Monroe, E.A.; Haynes, B.L. Microarray analysis of diurnal and circadian regulated genes in the florida red-tide dinoflagellate *Karenia brevis* (Dinophyceae). *J. Phycol.* **2007**, *43*, 741–752. [CrossRef]

56. Okamoto, O.K.; Hastings, J.W. Genome-wide analysis of redox-regulated genes in a dinoflagellate. *Gene* **2003**, *321*, 73–81. [CrossRef] [PubMed]

57. Okamoto, O.K.; Hastings, J.W. Novel dinoflagellate circadian-clock genes identified through microarray analysis of a phase shifted clock. *J. Phycol.* **2003**, *39*, 1–9. [CrossRef]

58. Yang, I.; Beszteri, S.; Tillmann, U.; John, U. Growth and nutrient-dependent gene expression in the toxigenic marine dinoflagellate *Alexandrium minutum. Harmful Algae* **2013**, *12*, 15–25. [CrossRef]

59. Moustafa, A.; Evans, A.N.; Kulis, D.M.; Hackett, J.D.; Erdner, D.L.; Anderson, D.M.; Bhattacharya, D. Transcriptome profiling of a toxic dinoflagellate reveals a gene-rich protist and a potential impact on gene expression due to bacterial presence. *PLoS ONE* **2010**, *5*, e9688. [CrossRef] [PubMed]

60. Erdner, D.L.; Anderson, D.M. Global transcriptional profiling of the toxic dinoflagellate *Alexandrium fundyense* using Massively Parallel Signature Sequencing. *BMC Genom.* **2006**, *7*, 88. [CrossRef] [PubMed]

61. Lowe, C.D.; Mello, L.V.; Samatar, N.; Martin, L.E.; Montagnes, D.J.; Watts, P.C. The transcriptome of the novel dinoflagellate *Oxyrrhis marina* (Alveolata: Dinophyceae): Response to salinity examined by 454 sequencing. *BMC Genom.* **2011**, *12*, 519. [CrossRef] [PubMed]

62. Roy, S.; Beauchemin, M.; Dagenais-Bellefeuille, S.; Letourneau, L.; Cappadocia, M.; Morse, D. The *Lingulodinium* circadian system lacks rhythmic changes in transcript abundance. *BMC Biol.* **2014**, *12*, 107. [CrossRef] [PubMed]

63. Tse, S.P.K.; Beauchemin, M.; Morse, D.; Lo, S.C.L. Refining transcriptome gene catalogs by MS-validation of expressed proteins. *Proteomics* **2018**, *18*. [CrossRef] [PubMed]

64. Qin, X.; Ahn, S.; Speed, T.P.; Rubin, G.M. Global analyses of mRNA translational control during early Drosophila embryogenesis. *Genome Biol.* **2007**, *8*, R63. [CrossRef] [PubMed]

65. Schwanhausser, B.; Busse, D.; Li, N.; Dittmar, G.; Schuchhardt, J.; Wolf, J.; Chen, W.; Selbach, M. Global quantification of mammalian gene expression control. *Nature* **2011**, *473*, 337–342. [CrossRef] [PubMed]

66. Beilharz, T.H.; Preiss, T. Translational profiling: The genome-wide measure of the nascent proteome. *Brief. Funct. Genom. Proteom.* **2004**, *3*, 103–111. [CrossRef]

67. Pradet-Balade, B.; Boulme, F.; Beug, H.; Mullner, E.W.; Garcia-Sanz, J.A. Translation control: Bridging the gap between genomics and proteomics? *Trends Biochem. Sci.* **2001**, *26*, 225–229. [CrossRef]

68. Maier, T.; Guell, M.; Serrano, L. Correlation of mRNA and protein in complex biological samples. *FEBS Lett.* **2009**, *583*, 3966–3973. [CrossRef] [PubMed]

69. Thomas, A.; Lee, P.J.; Dalton, J.E.; Nomie, K.J.; Stoica, L.; Costa-Mattioli, M.; Chang, P.; Nuzhdin, S.; Arbeitman, M.N.; Dierick, H.A. A versatile method for cell-specific profiling of translated mRNAs in *Drosophila. PLoS ONE* **2012**, *7*, e40276. [CrossRef] [PubMed]

70. Wisecaver, J.H.; Hackett, J.D. Dinoflagellate genome evolution. *Annu. Rev. Microbiol.* **2011**, *65*, 369–387. [CrossRef] [PubMed]

71. Roy, S.; Morse, D. Transcription and Maturation of mRNA in Dinoflagellates. *Microorganisms* **2013**, *1*, 71–99. [CrossRef] [PubMed]

72. McCarthy, J.E. Posttranscriptional control of gene expression in yeast. *Microbiol. Mol. Biol. Rev.* **1998**, *62*, 1492–1553. [PubMed]

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

74. Laursen, B.S.; Sorensen, H.P.; Mortensen, K.K.; Sperling-Petersen, H.U. Initiation of protein synthesis in bacteria. *Microbiol. Mol. Biol. Rev.* **2005**, *69*, 101–123. [CrossRef] [PubMed]

75. Chang, B.; Halgamuge, S.; Tang, S.L. Analysis of SD sequences in completed microbial genomes: Non-SD-led genes are as common as SD-led genes. *Gene* **2006**, *373*, 90–99. [CrossRef] [PubMed]

76. Hering, O.; Brenneis, M.; Beer, J.; Suess, B.; Soppa, J. A novel mechanism for translation initiation operates in haloarchaea. *Mol. Microbiol.* **2009**, *71*, 1451–1463. [CrossRef] [PubMed]

77. Tolstrup, N.; Sensen, C.W.; Garrett, R.A.; Clausen, I.G. Two different and highly organized mechanisms of translation initiation in the archaeon *Sulfolobus solfataricus*. *Extremophiles* **2000**, *4*, 175–179. [CrossRef] [PubMed]

78. Slupska, M.M.; King, A.G.; Fitz-Gibbon, S.; Besemer, J.; Borodovsky, M.; Miller, J.H. Leaderless transcripts of the crenarchaeal hyperthermophile *Pyrobaculum aerophilum*. *J. Mol. Biol.* **2001**, *309*, 347–360. [CrossRef] [PubMed]

79. Zheng, X.; Hu, G.Q.; She, Z.S.; Zhu, H. Leaderless genes in bacteria: Clue to the evolution of translation initiation mechanisms in prokaryotes. *BMC Genom.* **2011**, *12*, 361. [CrossRef] [PubMed]

80. Benelli, D.; Londei, P. Begin at the beginning: Evolution of translational initiation. *Res. Microbiol.* **2009**, *160*, 493–501. [CrossRef] [PubMed]

81. Benelli, D.; Londei, P. Translation initiation in Archaea: Conserved and domain-specific features. *Biochem. Soc. Trans.* **2011**, *39*, 89–93. [CrossRef] [PubMed]

82. Londei, P. Evolution of translational initiation: New insights from the archaea. *FEMS Microbiol. Rev.* **2005**, *29*, 185–200. [CrossRef] [PubMed]

83. Doi, N.; Zenno, S.; Ueda, R.; Ohki-Hamazaki, H.; Ui-Tei, K.; Saigo, K. Short-interfering-RNA-mediated gene silencing in mammalian cells requires Dicer and eIF2C translation initiation factors. *Curr. Biol.* **2003**, *13*, 41–46. [CrossRef]

84. Hinnebusch, A.G. Molecular mechanism of scanning and start codon selection in eukaryotes. *Microbiol. Mol. Biol. Rev.* **2011**, *75*, 434–467. [CrossRef] [PubMed]

85. Aitken, C.E.; Lorsch, J.R. A mechanistic overview of translation initiation in eukaryotes. *Nat. Struct. Mol. Biol.* **2012**, *19*, 568–576. [CrossRef] [PubMed]

86. Hinnebusch, A.G.; Lorsch, J.R. The mechanism of eukaryotic translation initiation: New insights and challenges. *Cold Spring Harb. Perspect. Biol.* **2012**, *4*. [CrossRef] [PubMed]

87. Valasek, L.S. 'Ribozoomin'—Translation initiation from the perspective of the ribosome-bound eukaryotic initiation factors (eIFs). *Curr. Protein Pept. Sci.* **2012**, *13*, 305–330. [CrossRef] [PubMed]

88. Yanagiya, A.; Svitkin, Y.V.; 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*, 1661–1669. [CrossRef] [PubMed]

89. Kahvejian, A.; Svitkin, Y.V.; Sukarieh, R.; M'Boutchou, M.N.; Sonenberg, N. Mammalian poly(A)-binding protein is a eukaryotic translation initiation factor, which acts via multiple mechanisms. *Genes Dev.* **2005**, *19*, 104–113. [CrossRef] [PubMed]

90. Lomakin, I.B.; Steitz, T.A. The initiation of mammalian protein synthesis and mRNA scanning mechanism. *Nature* **2013**, *500*, 307–311. [CrossRef] [PubMed]

91. Lomakin, I.B.; Shirokikh, N.E.; Yusupov, M.M.; Hellen, C.U.; Pestova, T.V. The fidelity of translation initiation: Reciprocal activities of eIF1, IF3 and YciH. *EMBO J.* **2006**, *25*, 196–210. [CrossRef] [PubMed]

92. Kolitz, S.E.; Takacs, J.E.; Lorsch, J.R. Kinetic and thermodynamic analysis of the role of start codon/anticodon base pairing during eukaryotic translation initiation. *RNA* **2009**, *15*, 138–152. [CrossRef] [PubMed]

93. Pestova, T.V.; Lorsch, J.R.; Hellen, C.U.T. The mechanism of translation initiation in eukaryotes. In *Translational Control in Biology and Medicine*; Cold Spring Harbor Laboratory Press: Suffolk County, NY, USA, 2007; pp. 87–128.

94. Ghosh, A.; Lima, C.D. Enzymology of RNA cap synthesis. *Wiley Interdiscip. Rev. RNA* **2010**, *1*, 152–172. [CrossRef] [PubMed]

95. Lidie, K.B.; van Dolah, F.M. Spliced leader RNA-mediated trans-splicing in a dinoflagellate, *Karenia brevis*. *J. Eukaryot. Microbiol.* **2007**, *54*, 427–435. [CrossRef] [PubMed]

96. Zhang, H.; Hou, Y.; Miranda, L.; Campbell, D.A.; Sturm, N.R.; Gaasterland, T.; Lin, S. Spliced leader RNA *trans*-splicing in dinoflagellates. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 4618–4623. [CrossRef] [PubMed]

97. Hastings, K.E. SL trans-splicing: Easy come or easy go? *Trends Genet.* **2005**, *21*, 240–247. [CrossRef] [PubMed]

98. Lasda, E.L.; Blumenthal, T. Trans-splicing. *WIREs RNA* **2011**, *2*, 417–434. [CrossRef] [PubMed]

99. Zhang, H.; Campbell, D.A.; Sturm, N.R.; Dungan, C.F.; Lin, S. Spliced leader RNAs, mitochondrial gene frameshifts and multi-protein phylogeny expand support for the genus *Perkinsus* as a unique group of alveolates. *PLoS ONE* **2011**, *6*, e19933. [CrossRef] [PubMed]

100. Hearne, J.L.; Pitula, J.S. Identification of two spliced leader RNA transcripts from *Perkinsus marinus*. *J. Eukaryot. Microbiol.* **2011**, *58*, 266–268. [CrossRef] [PubMed]

101. Perry, K.L.; Watkins, K.P.; Agabian, N. Trypanosome mRNAs have unusual "cap 4" structures acquired by addition of a spliced leader. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 8190–8194. [CrossRef] [PubMed]

102. Bangs, J.D.; Crain, P.F.; Hashizume, T.; McCloskey, J.A.; Boothroyd, J.C. Mass spectrometry of mRNA cap 4 from trypanosomatids reveals two novel nucleosides. *J. Biol. Chem.* **1992**, *267*, 9805–9815. [PubMed]

103. Zamudio, J.R.; Mittra, B.; Campbell, D.A.; Sturm, N.R. Hypermethylated cap 4 maximizes *Trypanosoma brucei* translation. *Mol. Microbiol.* **2009**, *72*, 1100–1110. [CrossRef] [PubMed]

104. Kozak, M. Regulation of translation via mRNA structure in prokaryotes and eukaryotes. *Gene* **2005**, *361*, 13–37. [CrossRef] [PubMed]

105. Kozak, M. Mechanism of mRNA recognition by eukaryotic ribosomes during initiation of protein synthesis. *Curr. Top. Microbiol. Immunol.* **1981**, *93*, 81–123. [PubMed]

106. Hamilton, R.; Watanabe, C.K.; de Boer, H.A. Compilation and comparison of the sequence context around the AUG startcodons in *Saccharomyces cerevisiae* mRNAs. *Nucleic Acids Res.* **1987**, *15*, 3581–3593. [CrossRef] [PubMed]

107. Cigan, A.M.; Donahue, T.F. Sequence and structural features associated with translational initiator regions in yeast—A review. *Gene* **1987**, *59*, 1–18. [CrossRef]

108. Joshi, C.P.; Zhou, H.; Huang, X.; Chiang, V.L. Context sequences of translation initiation codon in plants. *Plant Mol. Biol.* **1997**, *35*, 993–1001. [CrossRef] [PubMed]

109. Zhang, H.; Zhuang, Y.; Gill, J.; Lin, S. Proof that dinoflagellate spliced leader (DinoSL) is a useful hook for fishing dinoflagellate transcripts from mixed microbial samples: *Symbiodinium kawagutii* as a case study. *Protist* **2013**, *164*, 510–527. [CrossRef] [PubMed]

110. Bodyl, A.; Mackiewicz, P. Analysis of the targeting sequences of an iron-containing superoxide dismutase (SOD) of the dinoflagellate *Lingulodinium polyedrum* suggests function in multiple cellular compartments. *Arch. Microbiol.* **2007**, *187*, 281–296. [CrossRef] [PubMed]

111. Bertucci, A.; Tambutte, E.; Tambutte, S.; Allemand, D.; Zoccola, D. Symbiosis-dependent gene expression in coral-dinoflagellate association: Cloning and characterization of a P-type H+-ATPase gene. *Proc. Biol. Sci.* **2010**, *277*, 87–95. [CrossRef] [PubMed]

112. Maroteaux, L.; Herzog, M.; Soyer-Gobillard, M.O. Molecular organization of dinoflagellate ribosomal DNA: evolutionary implications of the deduced 5.8 S rRNA secondary structure. *Biosystems* **1985**, *18*, 307–319. [CrossRef]

113. Gressel, J.; Berman, T.; Cohen, N. Dinoflagellate ribosomal RNA; an evolutionary relic? *J. Mol. Evol.* **1975**, *5*, 307–313. [CrossRef] [PubMed]

114. Jackson, K.E.; Habib, S.; Frugier, M.; Hoen, R.; Khan, S.; Pham, J.S.; Ribas de Pouplana, L.; Royo, M.; Santos, M.A.; Sharma, A.; et al. Protein translation in *Plasmodium* parasites. *Trends Parasitol.* **2011**, *27*, 467–476. [CrossRef] [PubMed]

115. Zinoviev, A.; Shapira, M. Evolutionary conservation and diversification of the translation initiation apparatus in trypanosomatids. *Comp. Funct. Genom.* **2012**, *2012*, 813718. [CrossRef] [PubMed]

116. De Melo Neto, O.P.; Reis, C.R.S.; Moura, D.M.N.; Freire, E.R.; Carrington, M. Unique and Conserved Features of the Protein Synthesis Apparatus in Parasitic Trypanosomatid (Trypanosoma and Leishmania) Species. In *Evolution of the Protein Synthesis Machinery and Its Regulation, Hernández, G., Jagus, R., Eds.*; Springer International Publishing: Cham, Switzerland, 2016; pp. 435–475.

117. Eukaryotic Initiation Factors Gene Family. Available online: https://www.arabidopsis.org/browse/genefamily/eIF.jsp (accessed on 3 April 2018).

118. Gingras, A.C.; Raught, B.; Sonenberg, N. eIF4 initiation factors: Effectors of mRNA recruitment to ribosomes and regulators of translation. *Annu. Rev. Biochem.* **1999**, *68*, 913–963. [CrossRef] [PubMed]

119. Von der Haar, T.; Gross, J.D.; Wagner, G.; McCarthy, J.E. The mRNA cap-binding protein eIF4E in post-transcriptional gene expression. *Nat. Struct. Mol. Biol.* **2004**, *11*, 503–511. [CrossRef] [PubMed]

120. Aravind, L.; Koonin, E.V. Eukaryote-specific domains in translation initiation factors: Implications for translation regulation and evolution of the translation system. *Genome Res.* **2000**, *10*, 1172–1184. [CrossRef] [PubMed]

121. Marcotrigiano, J.; Gingras, A.C.; Sonenberg, N.; Burley, S.K. Cocrystal structure of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP. *Cell* **1997**, *89*, 951–961. [CrossRef]

122. Matsuo, H.; Li, H.; McGuire, A.M.; Fletcher, C.M.; Gingras, A.C.; Sonenberg, N.; Wagner, G. Structure of translation factor eIF4E bound to m7GDP and interaction with 4E-binding protein. *Nat. Struct. Biol.* **1997**, *4*, 717–724. [CrossRef] [PubMed]

123. Hernández, G.; Gillespie, K.M.; Bachvaroff, T.R.; Jagus, R.; Igreja, C.; Peter, D.; Bulfoni, M.; Cosson, B. Evolution of eIF4E-interacting proteins. In *Evolution of the Protein Synthetic Machinery and Its Regulation*; Hernandez, G., Jagus, R., Eds.; Spinger: Cham, Switzerland, 2016; pp. 165–186.

124. Joshi, B.; Lee, K.; Maeder, D.L.; Jagus, R. Phylogenetic analysis of eIF4E-family members. *BMC Evol. Biol.* **2005**, *5*, 48. [CrossRef] [PubMed]

125. Hernandez, G.; Vazquez-Pianzola, P. Functional diversity of the eukaryotic translation initiation factors belonging to eIF4 families. *Mech. Dev.* **2005**, *122*, 865–876. [CrossRef] [PubMed]

126. Jones, G.D.; Williams, E.P.; Place, A.R.; Jagus, R.; Bachvaroff, T.R. The alveolate translation initiation factor 4E family reveals a custom toolkit for translational control in core dinoflagellates. *BMC Evol. Biol.* **2015**, *15*, 14. [CrossRef] [PubMed]

127. Jagus, R.; Bachvaroff, T.R.; Joshi, B.; Place, A.R. Diversity of Eukaryotic Translational Initiation Factor eIF4E in Protists. *Comp. Funct. Genom.* **2012**, *2012*, 134839. [CrossRef] [PubMed]

128. Rhoads, R.E.; Dinkova, T.D.; Jagus, R. Approaches for analyzing the differential activities and functions of eIF4E family members. *Methods Enzymol.* **2007**, *429*, 261–297. [CrossRef] [PubMed]

129. Rhoads, R.E. eIF4E: New family members, new binding partners, new roles. *J. Biol. Chem.* **2009**, *284*, 16711–16715. [CrossRef] [PubMed]

130. Kamenska, A.; Simpson, C.; Standart, N. eIF4E-binding proteins: New factors, new locations, new roles. *Biochem. Soc. Trans.* **2014**, *42*, 1238–1245. [CrossRef] [PubMed]

131. Lasko, P. Posttranscriptional regulation in Drosophila oocytes and early embryos. *Wiley Interdiscip. Rev. RNA* **2011**, *2*, 408–416. [CrossRef] [PubMed]

132. Jung, M.Y.; Lorenz, L.; Richter, J.D. Translational control by neuroguidin, a eukaryotic initiation factor 4E and CPEB binding protein. *Mol. Cell. Biol.* **2006**, *26*, 4277–4287. [CrossRef] [PubMed]

133. Wells, S.E.; Hillner, P.E.; Vale, R.D.; Sachs, A.B. Circularization of mRNA by eukaryotic translation initiation factors. *Mol. Cell* **1998**, *2*, 135–140. [CrossRef]

134. Freire, E.R.; Sturm, N.R.; Campbell, D.A.; de Melo Neto, O.P. The Role of Cytoplasmic mRNA Cap-Binding Protein Complexes in Trypanosoma brucei and Other Trypanosomatids. *Pathogens* **2017**, *6*, 55. [CrossRef] [PubMed]

135. Yoffe, Y.; Leger, M.; Zinoviev, A.; Zuberek, J.; Darzynkiewicz, E.; Wagner, G.; Shapira, M. Evolutionary changes in the Leishmania eIF4F complex involve variations in the eIF4E-eIF4G interactions. *Nucleic Acids Res.* **2009**, *37*, 3243–3253. [CrossRef] [PubMed]

136. Park, E.H.; Walker, S.E.; Lee, J.M.; Rothenburg, S.; Lorsch, J.R.; Hinnebusch, A.G. Multiple elements in the eIF4G1 N-terminus promote assembly of eIF4G1*PABP mRNPs in vivo. *EMBO J.* **2011**, *30*, 302–316. [CrossRef] [PubMed]

137. Jivotovskaya, A.V.; Valasek, L.; Hinnebusch, A.G.; Nielsen, K.H. 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*, 1355–1372. [CrossRef] [PubMed]

138. Wek, R.C.; Jiang, H.Y.; Anthony, T.G. Coping with stress: EIF2 kinases and translational control. *Biochem. Soc. Trans.* **2006**, *34*, 7–11. [CrossRef] [PubMed]

139. Sonenberg, N.; Hinnebusch, A.G. Regulation of translation initiation in eukaryotes: Mechanisms and biological targets. *Cell* **2009**, *136*, 731–745. [CrossRef] [PubMed]

140. Jennings, M.D.; Kershaw, C.J.; Adomavicius, T.; Pavitt, G.D. Fail-safe control of translation initiation by dissociation of eIF2 alpha phosphorylated ternary complexes. *Elife* **2017**, *6*. [CrossRef] [PubMed]

141. Baird, T.D.; Wek, R.C. Eukaryotic initiation factor 2 phosphorylation and translational control in metabolism. *Adv. Nutr.* **2012**, *3*, 307–321. [CrossRef] [PubMed]

142. Donnelly, N.; Gorman, A.M.; Gupta, S.; Samali, A. The eIF2α kinases: Their structures and functions. *Cell. Mol. Life Sci.* **2013**, *70*, 3493–3511. [CrossRef] [PubMed]

143. Mohrle, J.J.; Zhao, Y.; Wernli, B.; Franklin, R.M.; Kappes, B. Molecular cloning, characterization and localization of PfPK4, an eIF2 α kinase-related enzyme from the malarial parasite *Plasmodium falciparum*. *Biochem. J.* **1997**, *328*, 677–687. [CrossRef] [PubMed]

144. Narasimhan, J.; Joyce, B.R.; Naguleswaran, A.; Smith, A.T.; Livingston, M.R.; Dixon, S.E.; Coppens, I.; Wek, R.C.; Sullivan, W.J., Jr. Translation regulation by eukaryotic initiation factor-2 kinases in the development of latent cysts in *Toxoplasma gondii*. *J. Biol. Chem.* **2008**, *283*, 16591–16601. [CrossRef] [PubMed]

145. Vonlaufen, N.; Kanzok, S.M.; Wek, R.C.; Sullivan, W.J., Jr. Stress response pathways in protozoan parasites. *Cell. Microbiol.* **2008**, *10*, 2387–2399. [CrossRef] [PubMed]

146. Joyce, B.R.; Konrad, C.; Wek, R.C.; Sullivan, W.J., Jr. Translation control is critical during acute and chronic stages of toxoplasmosis infection. *Expert Rev. Anti-Infect. Ther.* **2011**, *9*, 1–3. [CrossRef] [PubMed]

147. Zhang, M.; Joyce, B.R.; Sullivan, W.J., Jr.; Nussenzweig, V. Translational control in *Plasmodium* and *Toxoplasma* parasites. *Eukaryot. Cell* **2013**, *12*, 161–167. [CrossRef] [PubMed]

148. Konrad, C.; Wek, R.C.; Sullivan, W.J., Jr. A GCN2-like eukaryotic initiation factor 2 kinase increases the viability of extracellular *Toxoplasma gondii* parasites. *Eukaryot. Cell* **2011**, *10*, 1403–1412. [CrossRef] [PubMed]

149. Konrad, C.; Wek, R.C.; Sullivan, W.J., Jr. GCN2-like eIF2α kinase manages the amino acid starvation response in Toxoplasma gondii. *Int. J. Parasitol.* **2014**, *44*, 139–146. [CrossRef] [PubMed]

150. Dev, K.; Qiu, H.; Dong, J.; Zhang, F.; Barthlme, D.; Hinnebusch, A.G. The β/Gcd7 subunit of eukaryotic translation initiation factor 2B (eIF2B), a guanine nucleotide exchange factor, is crucial for binding eIF2 in vivo. *Mol. Cell. Biol.* **2010**, *30*, 5218–5233. [CrossRef] [PubMed]

151. Shopan, J.; Mou, H.; Zhang, L.; Zhang, C.; Ma, W.; Walsh, J.A.; Hu, Z.; Yang, J.; Zhang, M. Eukaryotic translation initiation factor 2B-beta (eIF2B β), a new class of plant virus resistance gene. *Plant J. Cell Mol. Biol.* **2017**, *90*, 929–940. [CrossRef] [PubMed]

152. Gomez, E.; Pavitt, G.D. Identification of domains and residues within the epsilon subunit of eukaryotic translation initiation factor 2B (eIF2B epsilon) required for guanine nucleotide exchange reveals a novel activation function promoted by eIF2B complex formation. *Mol. Cell. Biol.* **2000**, *20*, 3965–3976. [CrossRef] [PubMed]

153. Glisovic, T.; Bachorik, J.L.; Yong, J.; Dreyfuss, G. RNA-binding proteins and post-transcriptional gene regulation. *FEBS Lett.* **2008**, *582*, 1977–1986. [CrossRef] [PubMed]

154. Baltz, A.G.; Munschauer, M.; Schwanhausser, B.; Vasile, A.; Murakawa, Y.; Schueler, M.; Youngs, N.; Penfold-Brown, D.; Drew, K.; Milek, M.; et al. The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Mol. Cell* **2012**, *46*, 674–690. [CrossRef] [PubMed]

155. Hentze, M.W.; Castello, A.; Schwarzl, T.; Preiss, T. A brave new world of RNA-binding proteins. *Nat Rev Mol. Cell. Biol.* **2018**. [CrossRef] [PubMed]

156. Mittag, M.; Lee, D.-H.; Hastings, J.W. Circadian expression of the luciferin-binding protein correlates with the binding of a protein to the 3' untranslated region of its mRNA. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 5257–5261. [CrossRef] [PubMed]

157. Lapointe, M.; Morse, D. Reassessing the role of a 3'UTR-binding translational inhibitor in regulation of the circadian bioluminescence reaction of the dinoflagellate *Gonyaulax. Biol. Chem.* **2008**, *389*, 13–19. [CrossRef] [PubMed]

158. Fire, A.; Xu, S.; Montgomery, M.K.; Kostas, S.A.; Driver, S.E.; Mello, C.C. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature* **1998**, *391*, 806–811. [CrossRef] [PubMed]

159. Lee, R.C.; Ambros, V. An extensive class of small RNAs in Caenorhabditis elegans. *Science* **2001**, *294*, 862–864. [CrossRef] [PubMed]

160. Reinhart, B.J.; Weinstein, E.G.; Rhoades, M.W.; Bartel, B.; Bartel, D.P. MicroRNAs in plants. *Genes Dev.* **2002**, *16*, 1616–1626. [CrossRef] [PubMed]

161. Liu, J.; Carmell, M.A.; Rivas, F.V.; Marsden, C.G.; Thomson, J.M.; Song, J.J.; Hammond, S.M.; Joshua-Tor, L.; Hannon, G.J. Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **2004**, *305*, 1437–1441. [CrossRef] [PubMed]

162. Vasudevan, S.; Tong, Y.; Steitz, J.A. Switching from repression to activation: MicroRNAs can up-regulate translation. *Science* **2007**, *318*, 1931–1934. [CrossRef] [PubMed]

163. Sijen, T.; Fleenor, J.; Simmer, F.; Thijssen, K.L.; Parrish, S.; Timmons, L.; Plasterk, R.H.; Fire, A. On the role of RNA amplification in dsRNA-triggered gene silencing. *Cell* **2001**, *107*, 465–476. [CrossRef]

164. Cerutti, H.; Casas-Mollano, J.A. On the origin and functions of RNA-mediated silencing: From protists to man. *Curr. Genet.* **2006**, *50*, 81–99. [CrossRef] [PubMed]

165. Baum, J.; Papenfuss, A.T.; Mair, G.R.; Janse, C.J.; Vlachou, D.; Waters, A.P.; Cowman, A.F.; Crabb, B.S.; de Koning-Ward, T.F. Molecular genetics and comparative genomics reveal RNAi is not functional in malaria parasites. *Nucleic Acids Res.* **2009**, *37*, 3788–3798. [CrossRef] [PubMed]

166. Gunasekera, A.M.; Patankar, S.; Schug, J.; Eisen, G.; Kissinger, J.; Roos, D.; Wirth, D.F. Widespread distribution of antisense transcripts in the Plasmodium falciparum genome. *Mol. Biochem. Parasitol.* **2004**, *136*, 35–42. [CrossRef] [PubMed]

167. De Riso, V.; Raniello, R.; Maumus, F.; Rogato, A.; Bowler, C.; Falciatore, A. Gene silencing in the marine diatom Phaeodactylum tricornutum. *Nucleic Acids Res.* **2009**, *37*, e96. [CrossRef] [PubMed]

168. Dagenais-Bellefeuille, S.; Beauchemin, M.; Morse, D. miRNAs Do Not Regulate Circadian Protein Synthesis in the Dinoflagellate *Lingulodinium polyedrum*. *PLoS ONE* **2017**, *12*, e0168817. [CrossRef] [PubMed]

169. Hausser, J.; Landthaler, M.; Jaskiewicz, L.; Gaidatzis, D.; Zavolan, M. Relative contribution of sequence and structure features to the mRNA binding of Argonaute/EIF2C-miRNA complexes and the degradation of miRNA targets. *Genome Res.* **2009**, *19*, 2009–2020. [CrossRef] [PubMed]

170. Gao, D.; Qiu, L.; Hou, Z.; Zhang, Q.; Wu, J.; Gao, Q.; Song, L. Computational Identification of MicroRNAs from the Expressed Sequence Tags of Toxic Dinoflagellate *Alexandrium tamarense*. *Evol. Bioinform. Online* **2013**, *9*, 479–485. [CrossRef] [PubMed]

171. Baumgarten, S.; Bayer, T.; Aranda, M.; Liew, Y.J.; Carr, A.; Micklem, G.; Voolstra, C.R. Integrating microRNA and mRNA expression profiling in *Symbiodinium microadriaticum*, a dinoflagellate symbiont of reef-building corals. *BMC Genom.* **2013**, *14*, 704. [CrossRef] [PubMed]

172. Walsh, C.T.; Garneau-Tsodikova, S.; Gatto, G.J., Jr. Protein posttranslational modifications: The chemistry of proteome diversifications. *Angew. Chem. Int. Ed. Engl.* **2005**, *44*, 7342–7372. [CrossRef] [PubMed]

173. Karve, T.M.; Cheema, A.K. Small changes huge impact: The role of protein posttranslational modifications in cellular homeostasis and disease. *J. Amino Acids* **2011**, *2011*, 207691. [CrossRef] [PubMed]

174. Khoury, G.A.; Baliban, R.C.; Floudas, C.A. Proteome-wide post-translational modification statistics: Frequency analysis and curation of the swiss-prot database. *Sci Rep* **2011**, *1*. [CrossRef] [PubMed]

175. Cohen, P. The regulation of protein function by multisite phosphorylation—A 25 year update. *Trends Biochem. Sci.* **2000**, *25*, 596–601. [CrossRef]

176. Holt, L.J.; Tuch, B.B.; Villen, J.; Johnson, A.D.; Gygi, S.P.; Morgan, D.O. Global analysis of Cdk1 substrate phosphorylation sites provides insights into evolution. *Science* **2009**, *325*, 1682–1686. [CrossRef] [PubMed]

177. Kersten, B.; Agrawal, G.K.; Durek, P.; Neigenfind, J.; Schulze, W.; Walther, D.; Rakwal, R. Plant phosphoproteomics: An update. *Proteomics* **2009**, *9*, 964–988. [CrossRef] [PubMed]

178. Wang, D.Z.; Li, C.; Xie, Z.X.; Dong, H.P.; Lin, L.; Hong, H.S. Homology-Driven Proteomics of Dinoflagellates with Unsequenced Genomes Using MALDI-TOF/TOF and Automated De Novo Sequencing. *Evid. Based Complement. Alternat. Med.* **2011**, *2011*, 471020. [CrossRef] [PubMed]

179. Vanselow, K.; Kramer, A. Role of phosphorylation in the mammalian circadian clock. *Cold Spring Harb. Symp. Quant. Biol.* **2007**, *72*, 167–176. [CrossRef] [PubMed]

180. Kusakina, J.; Dodd, A.N. Phosphorylation in the plant circadian system. *Trends Plant Sci.* **2012**, *17*, 575–583. [CrossRef] [PubMed]

181. Schafmeier, T.; Kaldi, K.; Diernfellner, A.; Mohr, C.; Brunner, M. Phosphorylation-dependent maturation of Neurospora circadian clock protein from a nuclear repressor toward a cytoplasmic activator. *Genes Dev.* **2006**, *20*, 297–306. [CrossRef] [PubMed]

182. Kurosawa, G.; Aihara, K.; Iwasa, Y. A model for the circadian rhythm of cyanobacteria that maintains oscillation without gene expression. *Biophys. J.* **2006**, *91*, 2015–2023. [CrossRef] [PubMed]

183. Dodd, A.N.; Gardner, M.J.; Hotta, C.T.; Hubbard, K.E.; Dalchau, N.; Love, J.; Assie, J.M.; Robertson, F.C.; Jakobsen, M.K.; Goncalves, J.; et al. The Arabidopsis circadian clock incorporates a cADPR-based feedback loop. *Science* **2007**, *318*, 1789–1792. [CrossRef] [PubMed]

184. Edwards, K.D.; Anderson, P.E.; Hall, A.; Salathia, N.S.; Locke, J.C.; Lynn, J.R.; Straume, M.; Smith, J.Q.; Millar, A.J. FLOWERING LOCUS C mediates natural variation in the high-temperature response of the Arabidopsis circadian clock. *Plant Cell* **2006**, *18*, 639–650. [CrossRef] [PubMed]

185. Comolli, J.; Taylor, W.; Hastings, J.W. An inhibitor of protein phosphorylation stops the circadian oscillator and blocks light-induced phase shifting in *Gonyaulax polyedra*. *J. Biol. Rhythm.* **1994**, *9*, 13–26. [CrossRef] [PubMed]

186. Comolli, J.C.; Hastings, J.W. Novel effects on the *Gonyaulax* circadian system produced by the protein kinase inhibitor staurosporine. *J. Biol. Rhythm.* **1999**, *14*, 11–19. [CrossRef] [PubMed]

187. Roy, S.; Morse, D. The dinoflagellate *Lingulodinium* has predicted casein kinase 2 sites in many RNA binding proteins. *Protist* **2014**, *165*, 330–342. [CrossRef] [PubMed]

188. Prosperi, E.; Scovassi, A.I.; Stivala, L.A.; Bianchi, L. Proliferating cell nuclear antigen bound to DNA synthesis sites: Phosphorylation and association with cyclin D1 and cyclin A. *Exp. Cell Res.* **1994**, *215*, 257–262. [CrossRef] [PubMed]

189. Liu, B.; Lo, S.C.; Matton, D.P.; Lang, B.F.; Morse, D. Daily changes in the phosphoproteome of the dinoflagellate *Lingulodinium*. *Protist* **2012**, *163*, 746–754. [CrossRef] [PubMed]

190. Hochstrasser, M. Origin and function of ubiquitin-like proteins. *Nature* **2009**, *458*, 422–429. [CrossRef] [PubMed]

191. Glickman, M.H.; Ciechanover, A. The ubiquitin-proteasome proteolytic pathway: Destruction for the sake of construction. *Physiol. Rev.* **2002**, *82*, 373–428. [CrossRef] [PubMed]

192. Komander, D.; Rape, M. The ubiquitin code. *Annu. Rev. Biochem.* **2012**, *81*, 203–229. [CrossRef] [PubMed]