



Review

RNA-Binding Proteins in Dinoflagellates

Mariia Berdieva ^{*,†} , Pavel Safonov [†] and Sergei Skarlato

Laboratory of Cytology of Unicellular Organisms, Institute of Cytology of the Russian Academy of Sciences, 194064 Saint Petersburg, Russia; safonov.incras@gmail.com (P.S.); sergei.skarlato@mail.ru (S.S.)

* Correspondence: maria.berd4@yandex.ru or maberdieva4@gmail.com

[†] These authors contributed equally to this work.

Abstract

The described features of dinoflagellate gene expression indicate the predominance of post-transcriptional and translational regulation over transcriptional control. These microorganisms also exhibit extensive RNA editing and distinctive splicing characteristics. This regulatory landscape underscores the central role of RNA-binding proteins in dinoflagellate biology. In this review, we summarize current knowledge on major RNA-binding protein groups identified or bioinformatically annotated in dinoflagellates, including RNA recognition motif domain-containing proteins, Sm and Sm-like family, KH domain-containing proteins, zinc-finger proteins, and Pumilio family proteins, S1 domain-containing and cold shock domain-containing proteins, DEAD/DEAH-box RNA helicases, and pentatricopeptide repeat proteins. We focus on the features of their conserved domains, their functions in eukaryotes, and available data on their presence and putative roles in dinoflagellate cells. Integrating genomic, transcriptomic, and proteomic evidence, and where possible experimental data, we highlight both their overall conservation and potential lineage-specific traits. Our aim is to provide a concise synthesis of current knowledge, identify key uncertainties, and outline promising directions for future research into the evolution and cellular roles of RNA-binding proteins in this ecologically and biologically remarkable group.

Keywords: conserved domains; dinoflagellate; gene expression; regulation; RNA-binding proteins

1. Introduction

Dinoflagellates are a phylogenetically distinct and ecologically important group of unicellular eukaryotes. They contribute significantly to marine primary production, form endosymbioses essential for coral reef ecosystems, and include species responsible for harmful algal blooms that impact marine food webs and human health [1–3]. Dinoflagellates possess complex and variable life cycle routes and stages, which contribute to their ubiquitous distribution and adaptive potential [4–6].

It is well established that dinoflagellates exhibit a suite of unusual molecular and cellular features, including extremely large genomes, permanently condensed chromosomes with distinct chromatin architecture, extensive gene duplication, gene organization patterns such as tandem arrays, specific features of splicing, secondary metabolite biosynthesis and other processes [7–13].

A defining characteristic of dinoflagellate gene expression is the predominance of post-transcriptional and translational regulation over transcriptional control (reviewed in [13]). Although dinoflagellates do demonstrate transcriptional responses to external or internal cues, numerous studies show that mRNA levels often remain stable even



Academic Editor: Hunseung Kang

Received: 23 November 2025

Revised: 25 December 2025

Accepted: 26 December 2025

Published: 1 January 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

when corresponding protein levels oscillate across the diel cycle or under stress. This indicates that post-transcriptional and translational regulation serves as the primary control mechanisms. Such patterns have been documented, for example, for components of the bioluminescence system, metabolic pathways, and cell cycle control machinery [14–18]. Moreover, certain dinoflagellate proteins have been shown to be regulated at the post-translational level via phosphorylation [19].

Several features of dinoflagellate cellular and molecular biology are thought to influence their transcription dynamics, likely contributing to the predominance of post-transcriptional regulation of gene expression. Dinoflagellate chromosomes remain condensed throughout the cell cycle and exhibit a 1:10 protein-to-DNA ratio (in contrast with the more typical equimolar ratio observed in other eukaryotes). Moreover, dinoflagellates encode fewer components of RNA polymerase II compared to classical eukaryotic models, and sequence-specific transcriptional factors in dinoflagellates are known to be scarce [13,20,21].

The trans-splicing phenomenon also occupies a special place in dinoflagellate gene expression. In this process, a short RNA fragment from a small noncoding RNA (the splice leader RNA) is transplanted to the 5' end of independently transcribed pre-mRNAs to yield mature mRNAs [22]. This mechanism is ubiquitous in dinoflagellates, and most, if not all, of their transcripts appear to undergo post-transcriptional 5' end modification through trans-splicing (see for review [9,13]).

Dinoflagellates also show widespread RNA editing. This is true for organellar genomes: extensive RNA editing occurs in chloroplasts, where it is essential for the proper folding and function of photosystem proteins [23], and nucleotide modifications are likewise actively employed in mitochondria, potentially contributing to mitochondrial translation efficiency [24,25]. mRNA editing is also observed in nuclear-encoded genes [26,27]. Moreover, most genes appear to be regulated predominantly through one of the post-transcriptional pathways—either RNA editing or alternative splicing—whereas a subset of genes associated with photosynthesis and stress response employs both modes [26]. These observations raise questions regarding the diversity and functioning of the molecular machinery that enables such processes in dinoflagellates. In addition, chromatin-enriched and nuclear proteome analyses reveal an overrepresentation of RNA-binding proteins relative to DNA-binding ones, suggesting an RNA-centric nuclear regulatory architecture [28].

This regulatory landscape places RNA metabolism, and therefore RNA-binding proteins (RBPs), at the core of gene expression control pathways in dinoflagellates. RBPs are typically thought of as proteins that bind RNA targets (single- or double-stranded) through one or multiple globular RNA-binding domains (RBDs) and modulate the fate or function of the bound RNAs [29]. Many RBDs possess modular structure, consisting of multiple repeats of a few basic domains or of several cooperating units [30]. Such an architecture provides advantages in versatility, increased specificity and affinity of binding, and effective regulation of domains or proteins activity [30].

Comparative genomic and transcriptomic studies specify that some RBP families in dinoflagellates exist as large sets of paralogues, in some cases more numerous than in other eukaryotic lineages (e.g., [6,31]). Some RBP-coding genes are arranged in tandem arrays, consistent with the well-documented tendency of dinoflagellate genomes toward extensive duplication and tandem expansion [26,32,33]. The functional significance of these expansions remains unclear but may relate to lineage-specific adaptation and functional diversification—possibilities that are particularly intriguing given the important role of RBPs in post-transcriptional control.

Taken together, these observations highlight RBPs as a functionally crucial protein group in dinoflagellates and suggest that RNA metabolism may represent an influential and even dominant regulatory system in their cells. Large-scale sequencing initiatives, as well as multiple high-quality dinoflagellate genome assemblies, have substantially expanded the catalog of predicted RBPs in this lineage. However, functional studies of specific proteins remain rare. For many families, only domain-level predictions are available, with little biochemical, cellular, or genetic characterization, underscoring a major gap between rapidly accumulating sequence data and mechanistic understanding.

In this review, we provide a concise overview of major RBP families, including features of their conserved domains and canonical roles in eukaryotic cells, and summarize all available evidence for their presence and function in dinoflagellates, drawing on transcriptomic and genomic surveys, nuclear proteomics, and, where possible, experimental studies. In some cases, we also conduct additional sequence-level analyses to supplement or clarify existing data. For several RBP families, sufficient evidence allows meaningful conclusions, whereas for others only basic information about the presence of homologues is available. Our aim is to present an accurate and restrained summary of current knowledge, highlight key uncertainties, and outline promising directions for future research to bridge the gap between sequence-based predictions and cellular functions of RBPs in these biologically and ecologically remarkable organisms.

2. RNA-Binding Proteins Review

In the following sections, we examine the principal classes of RNA-binding proteins, encompassing RNA recognition motif-containing proteins, Sm and Sm-like proteins, KH domain-containing and zinc-finger proteins, Pumilio family proteins, oligosaccharide-/oligonucleotide-binding fold protein group including S1 domain-containing and cold-shock domain-containing proteins, DEAD/DEAH-box RNA helicases, and pentatricopeptide repeat proteins. For each group, we first outline the defining features of their conserved domains—covering amino acid composition, fold architecture, and principles of RNA binding and activity—and summarize their canonical functions in eukaryotic cells. We then compile the evidence currently available for dinoflagellates, including their presence and diversity, structural characteristics, and potential cellular roles. Table 1 provides a concise synthesis of the major groups considered here, highlighting their characteristic domain architectures and representative folds, broad cellular functions, and the present state of knowledge for dinoflagellates (ranging from evidence of presence based on genomic, transcriptomic, and/or proteomic data, to differential gene expression analyses, sequence and structural characterization, and experimental studies).

Table 1. Major RNA-binding protein groups analyzed in this review.

RBP Group	Canonical Domain Composition	Fold Architecture	General Functions in Eukaryotic Cells	Level of Characterization in Dinoflagellates	Relevant Studies in Dinoflagellates
RRM domain-containing proteins (MEI2 and MEI2-like)	Three RRM domains *	$\beta_1\alpha_1\beta_2\beta_3\alpha_2\beta_4$ (the third RRM domain has a $\beta_1\alpha_1\beta_2\beta_3\alpha_2\beta_4\alpha_3\beta_5\beta_6$ fold)	post-transcriptional regulation of developmental gene expression, incl. via binding and control of non-coding RNA and mRNA stability, mRNA sequestration	Presence in genomic/transcriptomic data, DGE **, sequence analysis	[6,31,32,34–40]
RRM domain-containing proteins (PABP)	RRM domains + C-terminal PABP-specific domain	$\beta_1\alpha_1\beta_2\beta_3\alpha_2\beta_4$	mRNA circularization, polyadenylation, nuclear export, degradation	Presence in genomic/transcriptomic data	[26]
Sm/Lsm proteins	Sm domain with motifs Sm1 and Sm2 + linker	N-terminal α -helix + 5 antiparallel β -strands	mRNA splicing, chaperone activity, protection from exonucleases, promotion of mRNA decay	Presence in genomic/transcriptomic data, immunocytochemical labeling	[13,22,26,33,41–46] Sm-binding site: [22,47–49]
KH domain-containing proteins	Number of KH-domains per protein varies	Type I: $\beta_1\alpha_1\alpha_2\beta_2\beta'\alpha'$ Type II: $\beta'\alpha'\beta_1\alpha_1\alpha_2\beta_2$	transcription, mRNA processing, translation	Presence in genomic/transcriptomic/ proteomic data	[13,26,44,46,50,51]
ZF proteins	Number of ZF-domains per protein varies	C2H2: $\beta\beta\alpha$ CCCH: short α -helices + 3_{10} -helices CCHC: 2 short β -strands + zinc knuckle + short α -helix	RNA folding, processing, turnover	Presence in genomic/transcriptomic/ proteomic data	[13,26,28,43,44,50,51]
PUF proteins	Pumilio homology domain (eight tandem motifs)	3 α -helices in a triangular configuration	promotion or repression of translation, mRNA stability regulation	Presence in genomic/transcriptomic/ proteomic data, DGE	[13,26,28,43,44,46,50,52–55]
OB-fold (S1 domain-containing proteins)	S1 domain	Greek key β -barrel (5 anti-parallel β -strands)	facilitation of RNA binding and degradation, ribosome biogenesis, control of translation initiation	Presence in genomic/transcriptomic/ proteomic data	[13,26,28,43,44,56]
OB-fold (CSD-containing proteins)	Cold shock domain with RNP1 and RNP2 motifs		enhancing RNA stability, facilitating translation initiation	Presence in genomic/transcriptomic/ proteomic data, DGE, sequence analysis, experimental data (binding tests)	[26,28,32,33,45,50,57–64]

Table 1. Cont.

RBP Group	Canonical Domain Composition	Fold Architecture	General Functions in Eukaryotic Cells	Level of Characterization in Dinoflagellates	Relevant Studies in Dinoflagellates
DEAD/DEAH-box helicases	helicase core, 8 or 9 conserved motifs (Q—in DEAD-box only, I, Ia, Ib, II-VI)	domain 1 + domain 2, each of 5 β -strands + 5 α -helices	translation initiation, splicing, RNA export, degradation, ribosome biogenesis, organelle gene expression	Presence in genomic/transcriptomic/proteomic data, DGE	[26,28,32,33,43–46,50,52,56,65,66]
YTH domain-containing proteins	C-terminal YTH domain	mixed α -helix/ β -sheet fold (4 α -helices + 6 β -strands or 3 α -helices, 8 β -strands, and 2 3_{10} -helices)	regulation of mRNA stability and degradation, splicing, translation, RNA nuclear export	Presence in genomic/transcriptomic data	[13,26,40,44,45]
PPR proteins	Pentatricopeptide repeats (~35 a.a.)	right-handed α -solenoid of helix-turn-helix hairpin structures	organelle gene expression (RNA stabilization, splicing, processing, translation, editing)	Presence in genomic/transcriptomic data, DGE, sequence analysis	[11,23,26,32,33,36,44,45,51,52,67–78]

* Only the third RRM domain in identified MEI2 homologues in myzozoans, including dinoflagellates. ** DGE—differential gene expression data.

3. RNA Recognition Motif Domain-Containing Proteins

RNA recognition motif (RRM) domain-containing proteins are found in all major branches of life, including viruses. However, they are especially abundant in eukaryotes, where the RRM is the most common RBD [79–82]. In eukaryotes, RRM-containing proteins participate in all major post-transcriptional processes, including pre-mRNA processing, splicing, RNA stabilization, and editing [79,83,84].

A single RRM domain is about 90 amino acids long and contains two conserved ribonucleoprotein (RNP) motifs essential for binding ssRNA: RNP1 ([KR]G[FY][GA][FY][VIL]X[FY]) and RNP2 ([IVL][FY][IVL]XNL) [85,86] (Figure 1a). Structurally, RRM comprises both β -strands and α -helices, arranged in a characteristic $\beta_1\alpha_1\beta_2\beta_3\alpha_2\beta_4$ fold, with the central β -strands bearing the two RNP motifs [87,88] (Figure 1a). RNA recognition therefore occurs on the surface of the β -sheet. Three conserved residues are the most crucial for binding: Arg/Lys binds the RNA backbone, while two aromatic residues interact with bases. Additional secondary structures outside the two conserved motifs can also interact with RNA, increasing the specificity of recognition (reviewed in [30]).

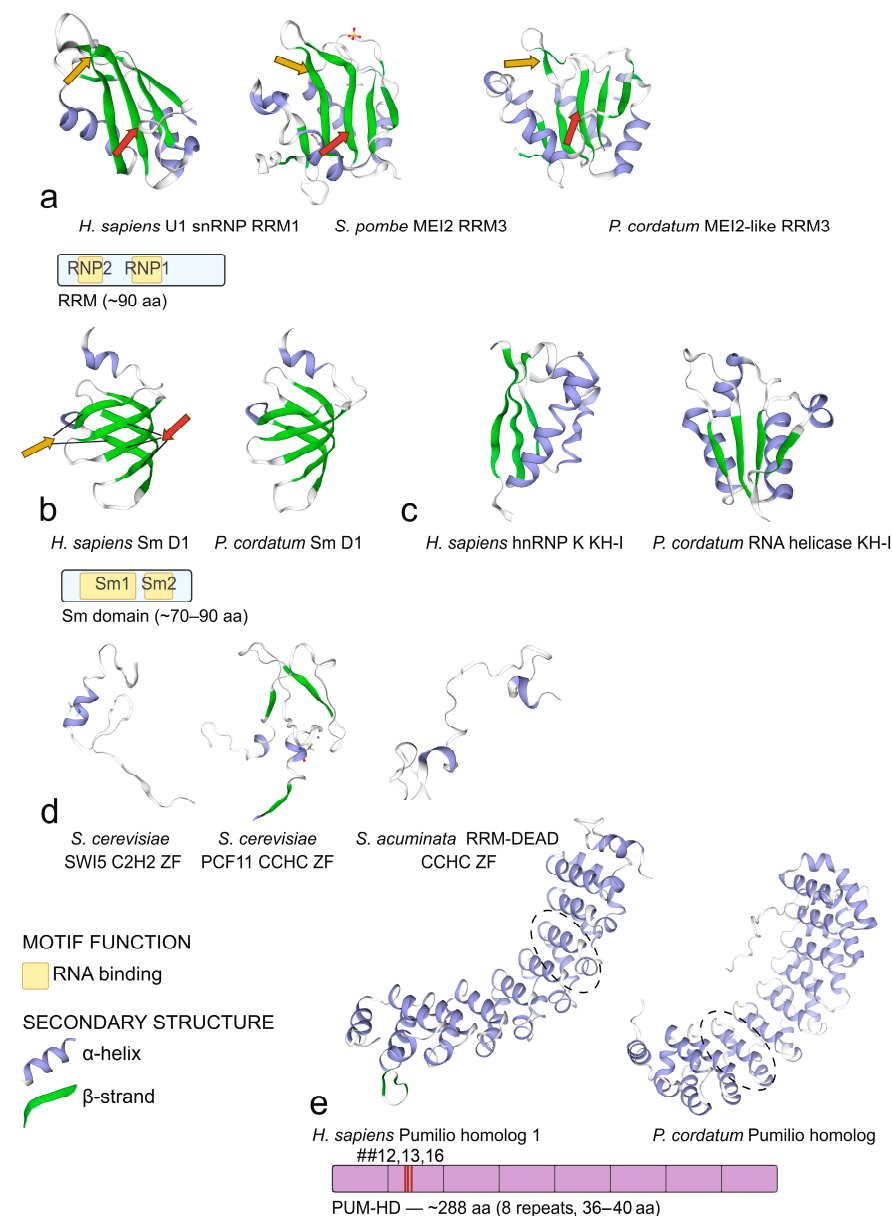


Figure 1. Cont.

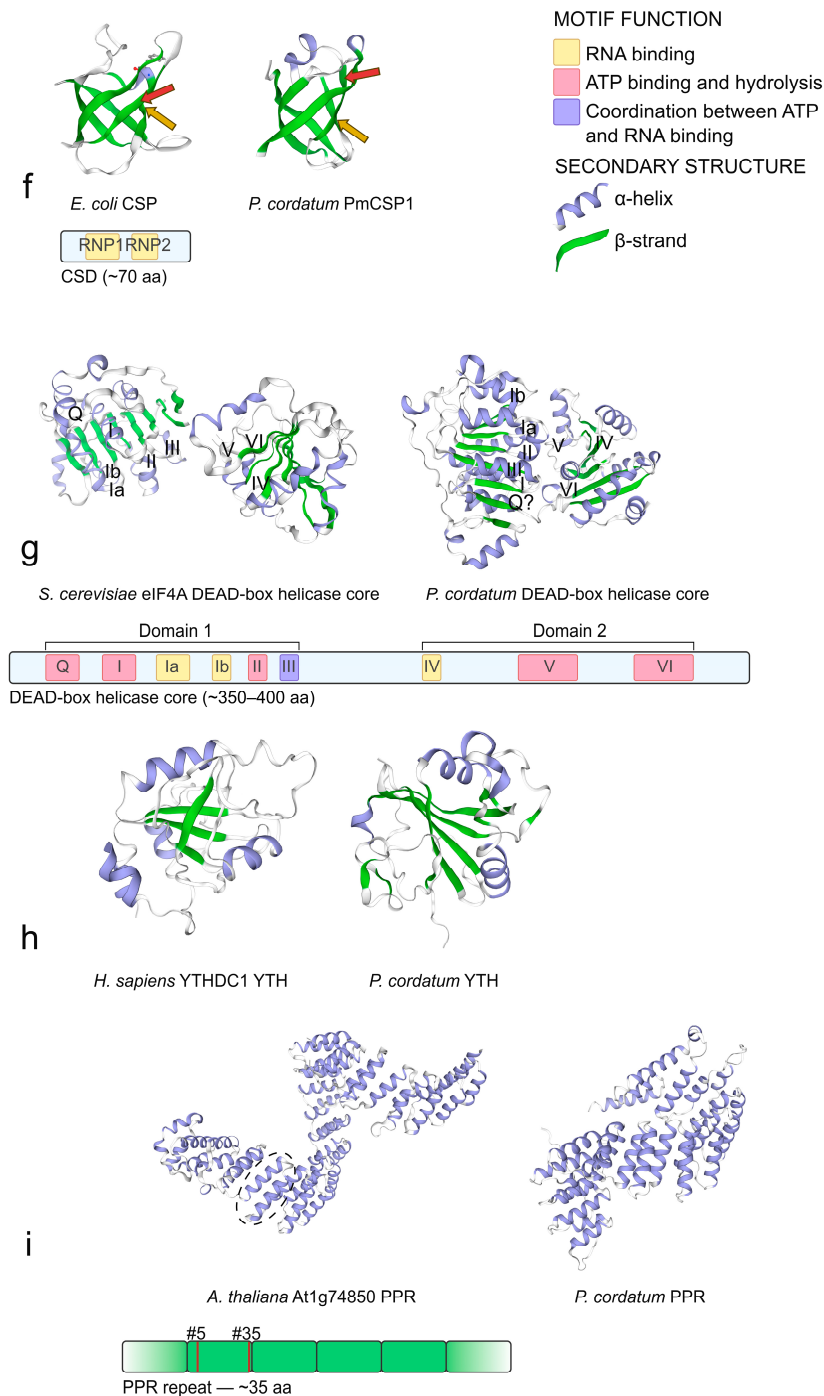


Figure 1. Tertiary structure models of conserved domains and general domain organization of RNA-binding proteins considered in this review. (a) RNA recognition motif (RRM) domain (RRM1) from *Homo sapiens* U1 small nuclear ribonucleoprotein 70 kDa (from model AF-P08621-2-F1-v6), RRM3 from *Schizosaccharomyces pombe* MEI2 protein (PDB code 7eio), and RRM3 from *Prorocentrum cordatum* MEI2-like homolog CAK0832983.1 (AlphaFold Server <https://alphafoldserver.com/> (accessed on 12 December 2025) prediction, pTM = 0.85). (b) Sm domain from *H. sapiens* small nuclear ribonucleoprotein Sm D1 (from model AF-P62314-F1-v6) and *P. cordatum* putative small nuclear ribonucleoprotein Sm D1 CAK0795168.1 (AlphaFold Server prediction, pTM = 0.62). (c) type I KH domain from *Homo sapiens* heterogeneous nuclear ribonucleoprotein K (PDB code 1zzk) and *P. cordatum* CAK0809374.1 putative KH-containing RNA helicase (AlphaFold Server prediction, pTM = 0.73). (d) Zinc-finger domains: C2H2 type from *Saccharomyces cerevisiae* factor SWI5 (PDB code 1ncs), CCHC type from *S. cerevisiae* component of the cleavage factor PCF11 (PDB code 2nax), and CCHC type in modification CX₂CX₃HX₄C from *Scrippsiella acuminata* RRM-DEAD helicase 20130930|942_1 (AlphaFold Server

prediction, pTM = 0.37). (e) Pumilio domain from *H. sapiens* Pumilio homolog 1 (PDB code 1ib2) and *P. cordatum* Pumilio homolog CAK0819378.1 (AlphaFold Server prediction, pTM = 0.89). (f) Cold shock domain from *Escherichia coli* CSP (PDB code 8rxs) and *P. cordatum* PmCSP1 CSD-containing protein QYF11783.1 (AlphaFold Server prediction, pTM = 0.64). (g) DEAD-box helicase core from *S. cerevisiae* translation initiation factor 4A eIF4A (PDB code 1fuu) and from *P. cordatum* DEAD-box helicase CAK0847872.1 (AlphaFold Server prediction, pTM = 0.86). In the *P. cordatum* protein model, the region predicted to correspond to the Q domain is indicated by “Q?”. (h) YTH domain from *H. sapiens* (PDB code 6we8) and *P. cordatum* YTH-containing protein CAK0851154.1 (AlphaFold Server prediction, pTM = 0.75). (i) pentatricopeptide repeat proteins—*Zea mays* PPR10 (PDB code 4m57) and *P. cordatum* PPR protein CAK0882574.1 (AlphaFold Server prediction, pTM = 0.59). In Pumilio and PPR proteins, a single structural unit composed of one repeat is indicated by a dashed oval. For RRM (a), Sm domain (b), CSD (f), and DEAD-box helicase core (g), generalized domain organization schemes are shown. The block color reflects the functional role of key motifs; red arrow indicates motif 1 in the domain (RNP1 in RRM and CSD, Sm1 in Sm domain), sand arrow indicates motif 2 in the domain (RNP2 in RRM and CSD, Sm2 in Sm domain). The generalized organization of Pumilio and PPR proteins, which consist of multiple tandem repeats, is also shown (e,i), with the positions of RNA-binding residues indicated by red lines.

RRM-containing proteins are abundant in dinoflagellates. On the proteomic level, a study of the chromatin-associated proteins in *Lingulaulax polyedra* (syn. *Lingulodinium polyedra*; in this review, we follow currently accepted species names according to the AlgaeBase <https://www.algaebase.org/> (accessed on 19 November 2025) and WoRMS <https://www.marinespecies.org/> (accessed on 19 November 2025) databases, providing synonyms in parentheses upon first mention if they were used in the original publications) detected 56 proteins harboring RRMs. Of these, 44 contained only RRM domains—either one, two, or three—while the remainder also carried other domains [28]. Simultaneously, several groups of RRM-containing proteins have been bioinformatically annotated in dinoflagellates at the genomic and transcriptomic levels. Considering the diversity of RRM-containing proteins, below we briefly characterize each of those evolutionary and/or functional groups separately.

3.1. MEI2-like Proteins

MEI2 is an RNA-binding protein originally described in fission yeast *Schizosaccharomyces pombe*, where it was shown to be a key regulator of transition from mitosis to meiosis [89]. In general, it binds and stabilizes long non-coding meiRNA that serves as a sponge for Mmi1 (meiotic mRNA interceptor 1) protein during meiotic initiation; besides, it sequesters Mmi1 mRNA [90]. Subsequently, homologues of MEI2 were predicted and identified in higher plants [91–93] and diverse protists (e.g., [93–95]), including dinoflagellates [6,32,34,35,37,39,40]. Structurally, MEI2 contains three RRM domains [94], with the third RRM domain (RRM3) being the most conserved [94]. Moreover, RRM3 possesses a unique architecture as it is extended by an additional C-terminal α -helix followed by two β -strands [94] (Figure 1a). Remarkably, in myzozoans, i.e., all alveolates except ciliates [96], the annotated MEI2-like homologues contain only RRM3, except for two homologues of *Kryptoperidinium triquetrum* (syn. *Kryptoperidinium foliaceum*) that apparently originated from the diatom endosymbiont, and two proteins of *Perkinsus chesapeaki* and *P. olseni* [31,94]. The architecture of dinoflagellate RRM3 appears to be similar to the canonical one (Figure 1a). The study by Palii and co-authors [31] also revealed varying numbers of homologues with the same structure in other eukaryotic groups, indicating that the loss of RRM1 and RRM2 occurred independently in different lineages (Table S1 in [31]; here and in following cases, the sources of the mentioned supplementary data from other works are listed in Table S1). In dinoflagellates, genomic and transcriptomic analyses indicate that the repertoire of MEI2-like proteins expanded during the evolution of the group [6,31]. Early-

branching taxa encode relatively few homologues (four in *Amoebophrya* species and about 24 in *Oxyrrhis marina*), whereas the core dinoflagellate *Prorocentrum cordatum* possesses as many as 89 homologues [31].

MEI2 homologues are widely present in dinoflagellate transcriptomes and genomes. However, the absence of direct studies of proteins or loss-of-function phenotypes limits our ability to functionally characterize dinoflagellate MEI2 homologues. Nevertheless, differential gene expression analyses, supplemented with phylogenetic reconstructions, provide important insights into their possible roles in these organisms.

Accumulating transcriptomic evidence links changes in MEI2-like gene expression to distinct life-cycle stages and responses to external stressors in dinoflagellates. As discussed previously in [31], two broad transcriptional patterns are emerging. In the present review, we further refined and expanded this distinction and summarized it in Table 2. In the first pattern, certain MEI2-like genes show altered expression during vegetative life-cycle stages, particularly during cyst formation. In cold-stress-induced temporary cysts of *P. cordatum*, five differentially expressed genes (DEGs) showed reduced expression (protein product sequences CAK0884715.1, CAK0887363.1, CAK0909770.1, CAK0844294.1, and CAK0818609.1), with two of them (CAK0884715.1, CAK0818609.1) exhibiting elevated expression during preparation for excystment [36]. Three DEGs (protein products CAK0836878.1, CAK0828848.1, and CAK0884998.1) also differed in expression relative to motile cells [36]. In dark- and cold-induced temporary (pellicular) cysts of *Scrippsiella acuminata* (syn. *Scrippsiella trochoidea*), both up- and down-regulated DEGs encoding MEI2 homologues were annotated in the same dataset [34,38]. Stress-induced coccoid cells in *Effrenium voratum* demonstrated down-regulated MEI2-like DEGs [37]. Additionally, phosphate deficiency suppressed MEI2-coding genes [71].

Unlike the expression pattern described for *P. cordatum* in [36], *S. acuminata* cysts showed increased expression of gene coding for SMC3 (structural maintenance of chromosomes protein 3), which may be associated with meiosis [38]. SMC3 is a component of the cohesin complex that is essential for sister chromatid pairing prior to division [97]; nevertheless, this complex mediates cohesion in both mitosis and meiosis [98].

Given the rapid return of *S. acuminata* cysts to motility under favorable conditions [38], together with the *P. cordatum* data [36], some MEI2 homologues may play a role in excystment. Up- and down-regulated MEI2-like genes have also been reported in *S. acuminata* resting cysts [34]. Notably, *SPO11* and *DMC1* genes, both listed in the meiosis detection toolkit [99], were also up-regulated. This may indicate recombination events in *S. acuminata* cysts coupled with sexual process. However, because mature resting cysts were sampled for RNA sequencing [34], preparation for excystment could also have occurred in this case.

In the second pattern, regulation of MEI2-like genes coincides with the regulation of a larger number of meiosis-specific genes, which more confidently indicates the preparation for meiosis. In *Alexandrium minutum* under nitrogen deficiency, 47 GO terms related to meiosis were enriched, strongly suggesting the induction of sexuality. Among these, a MEI2 homologue was up-regulated [40]. Moreover, active meiotic processes during bloom events have been documented in several dinoflagellate species [6,39]. The concomitant up-regulation of at least some genes encoding MEI2 homologues is consistent with the proposed role of MEI2 as a regulator of meiosis. Nevertheless, given the large number of MEI2-like genes showing altered expression (see Figure S2 in [6]), we cannot exclude that a subset of these genes may belong to another functional group of MEI2-like proteins and may instead be involved, for example, in rapid post-transcriptional regulation of gene expression during bloom development under natural conditions.

Table 2. Two putative functional groups of genes encoding MEI2-like proteins in dinoflagellates: genes demonstrating transcriptional responses linked to vegetative life-cycle stages and stress, and those potentially associated with the sexual process. Based on Table 1 from [31], with modifications. Changes in gene expression: ↑—increase, ↓—decrease.

Dinoflagellate Species *	Conditions	Life Cycle Stage	MEI2-Like DEGs Number/Expression Changes	Simultaneously Up-Regulated Meiosis-Associated Genes	References
Vegetative life-cycle stage- and stress-related response					
<i>Effrenium voratum</i>	Photobleaching (heat stress, 32 °C)	Coccoid cells	5/↓	-	[37]
<i>Fugacium kawagutii</i>	Phosphorous deficiency (1 week)	n/a	2/↓	-	[71]
<i>Prorocentrum cordatum</i>	Cold stress (2 °C)	Temporary cyst	5/↓; (2/↑ during excystation)	-	[36]
<i>Scrippsiella acuminata</i>	50–60-day cultures	Mature resting cysts	10 (4/↑; 6/↓)	<i>SPO11, DMC1</i>	[34]
	Dark + cold stress (8 °C)	Temporary (pellicle) cysts	10 (9/↑; 1/↓)	<i>SMC3</i>	[38]
Potential sexual life-cycle phase-related response					
<i>Alexandrium minutum</i>	Nitrogen deficiency (72 h)	n/a	n/a/↑	<i>TOP2</i> , 47 enriched GO terms related to meiosis	[40]
<i>Karenia mikimotoi</i>	Bloom, natural conditions	n/a	n/a/↑	<i>RAD54, MSH2, RAD21, SPO11, DMC1, MND1, SMC3</i>	[6]
<i>Noctiluca scintillans</i>	Bloom, natural conditions	n/a	6/↑	<i>SMC1, SMC3, RAD24, HOP2, MND1, REC8, DMC1, MSH5, EXO1, MRE11, MER3/HFM1</i>	[39]
<i>Prorocentrum shikokuense</i>	Bloom, natural conditions	n/a	n/a/↑ (some—↓?)	<i>DMC1, SPO11, MND1, MSH6, RAD21, RAD51, EXO1, MNS1, SMC3, SMC2, SMC1</i>	[6]

* Here and in the following tables, we used currently accepted species names according to the AlgaeBase and WoRMS databases.

Thus, dinoflagellates potentially possess at least two functional sets of MEI2 homologues corresponding to the above activity patterns. Importantly, it is not surprising that some MEI2-like proteins could have functions beyond meiosis regulation. For example, in plants, MEI2 homologues were shown to be expressed in vegetative tissues [100], where some of them are involved in control of leaf initiation [92,93,100]. Furthermore, phylogenetic analysis of a large dataset of dinoflagellate MEI2-like proteins revealed at least six distinct clades of homologues [31]. Notably, the MEI2 homologues implicated in the formation of *P. cordatum* temporary cysts clustered within two of these clades [36]. These results indicate that possibly even more pools of MEI2 homologues coexist in dinoflagellates.

3.2. hnRNP Proteins, snRNP Proteins and SR Proteins

The hnRNP proteins constitute the protein component of heterogeneous nuclear ribonucleoproteins (hnRNPs). These RNA-binding proteins associate with nascent transcripts and influence splicing, mRNA export, stability, localization, and translation (reviewed in [79]). Most of hnRNP proteins contain one or several RRM domains [79]; for example, members of the hnRNP A/B family possess two N-terminal RRMs (reviewed in [101]).

The small nuclear ribonucleoproteins (snRNPs) are core components of the spliceosome. Each snRNP is composed of a specific small nuclear RNA (snRNA), a common set of RNA-binding Sm proteins (considered separately below), and a set of auxiliary proteins, that are specific to each particle (reviewed in [102,103]). Several of these auxiliary proteins carry RRM domains, which bind the corresponding snRNA molecules (e.g., [104–106]).

SR proteins represent a family of RNA-binding factors involved in the control of both constitutive and alternative splicing [107,108]. Structurally, SR proteins contain at least one RRM domain and an Arg-Ser-rich region known as the RS domain; some family members also contain a second RRM or additional auxiliary domains [107,108]. Strong evidence supports the antiquity of this family: phylogenetic reconstruction of SR protein RRMs from a plethora of diverse organisms indicates that the diversification of their main structural types predates the divergence of the major eukaryotic lineages [80]. Notably, the analysis by Califice et al. [80] did not include dinoflagellate SR proteins, although it did feature diverse homologues from apicomplexans.

In dinoflagellates, cis-splicing (“canonical”, i.e., intramolecular splicing within a single pre-mRNA molecule) is generally thought to be less prominent than in many other eukaryotes given the paucity of introns in their genes. Nonetheless, both snRNAs and some of snRNP proteins were shown to be highly conserved between mammals and dinoflagellates (reviewed in [20,21]). At least some RRM-containing auxiliary snRNP proteins are conserved across metazoans, yeasts and plants, suggesting their evolutionary antiquity [108]. Indeed, a brief blastp search against the NCBI ClusteredNR database using human U1-70K protein (NP_001287998.1) as a query revealed strongly-supported homologues in both basal and core dinoflagellates, e.g., CAD7962409.1 from *Amoebophrya* sp. A25 (E-value = 6×10^{-52}), and CAE7225710.1 from *Symbiodinium pilosum* (E-value = 4×10^{-51} ; here and further, the list of the revealed homologues is non-exhaustive). Subsequent analysis utilizing the CD-search tool (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>, accessed on 12 December 2025) revealed U1-70K-specific N-terminal domain (U1snRNP70_N), followed by an RRM domain in both homologues (Figure S1a, multiple sequence alignment was generated using MAFFT7 [109] implemented in the Unipro UGENE software package [110]).

Furthermore, nearly all nuclear-encoded mRNAs in dinoflagellates undergo trans-splicing, in which an exon derived from a separately transcribed splice leader (SL) RNA is joined to the 5' end of an mRNA during its maturation (reviewed in [13,20]). Alternative splicing, previously thought to be absent in dinoflagellates [20], has since been

confirmed by genomic analyses (reviewed in [10]). Taken together, these points indicate that the repertoire of dinoflagellate hnRNP, snRNP, and SR proteins warrants detailed investigation. A blastp search against the NCBI ClusteredNR database using human SRSF2 protein (NP_001182356.1) as a query retrieved a number of dinoflagellate homologues, e.g., CAE8617636.1 (E-value = 4×10^{-22}), and CAJ1443725.1 (E-value = 7×10^{-21}). For these sequences, the CD-search tool predicted an N-terminal RRM domain (Figure S1b). Although the algorithm did not predict the RS domain in the retrieved homologues, it also failed to detect this domain in the query sequence. Nevertheless, an Arg–Ser-rich region can be clearly distinguished in the multiple sequence alignment (Figure S1b). Notably, some of the homologues are annotated correspondingly as SR proteins (e.g., CAE6927798.1 and CAE7359325.1 of *Symbiodinium* sp. CCMP2592).

3.3. RRM-Containing Translation Initiation Factors

In contrast to translation initiation in bacteria, which involves only three factors (IF1–3) [111], the eukaryotic initiation machinery includes a much larger set of proteins and multisubunit complexes [112], several of which harbour RRM domains. Among them is eIF4B—a protein mediating the association between the small (40S) ribosomal subunit and mRNA [113,114]. Besides an N-terminal RRM domain that binds 18S rRNA, eIF4B also contains a second C-terminal RBD, termed the basic domain (BD), which binds mRNA [115,116].

The eIF3 complex, essential for binding the 40S subunit, mRNA and other initiation factors, comprises five core proteins in yeasts, whereas mammals possess a larger set in which these five core subunits are supplemented with seven additional ones, for a total of twelve [117,118]. Two of the core subunits—eIF3b and eIF3g—contain RRMs [119–122]. eIF3b participates in the nucleation of the entire eIF3 complex, promoting its stabilization on the 40S ribosomal subunit through interaction with 18S rRNA. Notably, the RRM domain of eIF3b also mediates protein–protein contacts with other eIF3 subunits [118,122,123]. eIF3g, by means of its RRM domain, contributes to linear scanning of the mRNA 5' UTR for the start codon and may also mediate reinitiation following translation termination on short upstream ORF [124]. In addition, the RRM of eIF3g participates in 18S rRNA binding [125,126].

Poly(A)-binding proteins (PABPs), although not formally classified as initiation factors, also contain RRM domains and, through interaction with eIF4G, mediate mRNA circularization, thereby promoting recruitment of the 40S ribosomal subunit. Besides translation initiation, PABPs participate in mRNA polyadenylation, nuclear export, and degradation. Canonically, cytoplasmic PABPs bear four RRM domains, and some also contain a C-terminal PABP-specific domain MLLE, whereas nuclear PABPs typically have only a single RRM and lack the MLLE domain (reviewed in [127,128]). Earlier reviews (e.g., [127]) noted that, while some metazoans and plants encode multiple PABP genes, unicellular eukaryotes (except yeasts) typically possess a single homologue. However, subsequent research has shown that multiple PABPs in single-celled organisms may be more common. In *Leishmania major*, three PABP paralogues have been identified, and *Trypanosoma* species possess at least two; all identified trypanosomatid PABPs are characterized by a four-RRM + C-terminal MLLE architecture [129]. More recently, two PABPs were identified in *Plasmodium yoelii*: one with a four-RRM + C-terminal MLLE architecture and another bearing only a single RRM [130].

In general, dinoflagellates possess a largely canonical repertoire of eukaryotic initiation factors [13]. Nevertheless, certain factors, including RRM-containing ones, appear to be missing. A notable example is eIF4B, which has not been identified or predicted in dinoflagellates [13]. In *L. polyedra*, four of the five core subunits of the eIF3 complex—a, b,

d, and e—have been identified at the proteomic level [131], i.e., eIF3g has not been detected. As for the diversity of dinoflagellate PABPs, no direct information is currently available in the literature. Although proteins annotated as PABPs have been listed in genomic datasets (e.g., [26]), such general annotations do not allow the domain composition to be inferred. We performed a blastp search against the NCBI ClusteredNR database using the human PABPC1 protein (NP_002559) as a query, and it yielded several dinoflagellate homologues that we considered *bona fide* PABPs (E-value $< 3 \times 10^{-90}$ for all). According to conserved domain predictions, these sequences contain both RRM domains (RRM1–4, except for CAE8617444.1 from *Polarella glacialis* CCMP1383, which contains only RRM2–4) and a C-terminal PABP-specific domain (Figure S1c). At least three homologues were identified in *P. glacialis* CCMP1383 (CAE8641676.1, CAE8622414.1, CAE8617444.1), two in *P. glacialis* CCMP2088 (CAE8685442.1, CAE8738554.1), and two in *P. cordatum* (CAK0907260.1, CAK0870724.1) proteomic datasets corresponding to genome assemblies GCA_905237085.1, GCA_905237095.1, and GCA_963575745.1, respectively [26]. These results indicate the likely coexistence of multiple paralogues in at least some dinoflagellate species.

4. Sm and Sm-like (Lsm) Proteins

Sm and Sm-like (Lsm) proteins represent an evolutionarily conserved family of small RNA-binding proteins ubiquitously found in bacteria, archaea, and eukaryotes [132,133]. In both eukaryotes and prokaryotes, each (L)Sm monomer folds into a characteristic structure consisting of N-terminal α -helix followed by five strongly bent antiparallel β -strands (β 1– β 5) [134] (Figure 1b). Canonical proteins assemble into multimeric ring structures: eukaryotic Sm proteins form a heptameric ring [135], whereas Lsm proteins typically form heteromers or homomers of six or seven subunits [136]. Sm and Lsm complexes preferentially bind uridine-rich single-stranded 3' ends. RNA oligonucleotides generally bind inside the central lumen of the ring, one nucleotide per subunit (monomer) (see for review [137]).

The heptameric Sm ring consists of the core Sm proteins (subunits) of subtypes B (or alternatively B' or N), E, F, G, D1, D2, and D3, which serve as its structural elements. This heteromeric complex is localized in the nucleus and binds spliceosomal U1, U2, U4, and U5 snRNAs, thereby constituting snRNPs, which in turn serve as building blocks of the spliceosome [136]. Notably, Sm proteins can be involved not only in cis-splicing but also in trans-splicing, a process that occurs in dinoflagellates. This is consistent with the fact that SL RNAs share similarities with spliceosomal snRNAs, including a U-rich Sm-protein-binding site [138].

The nuclear Lsm2–8 complex can act as a chaperone that interacts with U6 snRNA and promotes the reassociation of the U4 and U6 snRNPs after splicing. In contrast, the cytoplasmic Lsm1–7 complex recognizes oligoadenylated mRNAs and protects 3' end of the mRNA from exonucleases, while also recruiting the decapping machinery and promoting mRNA decay (see for review [133]).

Sm proteins share a conserved bipartite Sm domain composed of two sequence motifs, Sm1 and Sm2, separated by a variable linker [139,140]. This domain is followed by a glycine- and proline-rich C-terminal tail [140]. The Sm1 motif typically lies in the first three β -strands, while Sm2 covers strands four and five [133,141]. Analyses of eukaryotic Sm protein sequences reveal several highly conserved residues in Sm1. In particular, these include a Gly residue in β 1, which contributes to inter-subunit contacts, an Asp residue that is important for proper structural orientation, and an Asn, the third of the three key RNA-binding residues located in the loop between β 2 and β 3 [139–141]. In Sm2, the conserved region corresponds to the loop between β 4 and β 5, where adjacent Arg (or Lys, as in Sm-E) and Gly (or Cys, as in Sm-F) residues are typically found—the former participating in RNA

binding and the latter contributing to structural stability. A nonpolar hydrophobic region precedes them in $\beta 4$ [139–141].

Sm proteins appear to be well-conserved in dinoflagellates and largely similar to canonical ones (Figure 1b), despite the distinctive features of dinoflagellate splicing. Random sequencing of SL-based cDNA libraries revealed several sequences encoding Sm proteins highly similar to canonical Sm ring components in *Pfiesteria piscicida* and *Karlodinium veneficum* (syn. *Karlodinium micrum*) [22]. Lsm domain-containing sequences were also listed among components of the splicing machinery inventoried in *Pyrocystis lunula* [46]. Although not explicitly discussed, Lsm domains or Lsm-coding genes are annotated in dinoflagellates from diverse lineages, including basal groups (e.g., [13,26,33,43–45]). Sm ring components can also be found in various datasets: for example, all core Sm proteins B, F, E, G, D1, D2, and D3 are present in the *P. glacialis* reference proteome UP000654075 in UniProtKB database (<https://www.uniprot.org/>, accessed on 12 December 2025), and nearly all components—except G subunit—are annotated in *P. cordatum* genome assembly [26]. Random testing of Sm domains confirmed a high level of conservation in dinoflagellate sequences, reflected in their close similarity to canonical eukaryotic Sm proteins and the strong preservation of residues essential for Sm function.

Other lines of evidence also support the conservation of Sm proteins in dinoflagellates. Reddy et al. [41] demonstrated the presence of U1–U6 snRNAs and Sm proteins associated with them in *Cryptothecodinium cohnii* using human anti-Sm antibodies. These antibodies recognize the so-called “Smith antigen”, i.e., the complex of Sm proteins bound to spliceosomal snRNAs [142]. Alverca et al. [42] used serum recognizing symmetrical dimethylarginine to label nuclear proteins, including Sm proteins, in several dinoflagellate species. Symmetrical dimethylarginine is a post-translational modification of an arginine side chain found in C-terminal tail in some core spliceosomal Sm and Lsm proteins [143], though it also occurs in other nuclear proteins such as hnRNPs and coilin [144,145]. The authors attributed labeling of protein bands of the expected molecular mass to Sm proteins, although at the level of immunofluorescence and immunoelectron microscopy this interpretation remained tentative [42]. In the nuclei of *Amphidinium carterae*, *Prorocentrum micans*, *Akashiwo sanguinea*, *Alexandrium fundyense* (syn. *Alexandrium catenella*), and *Alexandrium minutum* (syn. *Alexandrium lusitanicum*), they observed perichromosomal labeling related to perichromosomal granules enriched in splicing factors, as well as labeling in Cajal-like bodies [42].

Considering conservation of the Sm complex, it should be noted that the Sm-binding site in SL RNA appears to exhibit distinctive features in dinoflagellates. According to data obtained by Zhang et al. [22,48], SL RNAs from several species (*Karenia brevis*, *P. piscicida*, *K. veneficum*, *P. cordatum*, *P. glacialis*, *Heterocapsa arctica*), contain a conserved Sm binding motif AUUUUGG (AUGUUGG in one SL RNA type in *P. glacialis*), located within a stem-loop structure in the exon region of the 50–60 bp SL RNA that is ultimately trans-spliced to mRNAs. In other trans-splicing organisms, by contrast, the Sm site is located in the intron. Because the Sm ring functions as part of the nuclear import signal, its retention on the mature mRNA could potentially interfere with cytoplasmic targeting [22]. A different study [47], questioned this exon-based localization in *K. brevis* and identified a U-rich AGGCCUCUUUUG region downstream of the exon in the 150-bp SL RNA. More recently, Song et al. [49] described a U-rich motif GUUUUC in introns of SL genes in *Fugacium kawagutii* (syn. *Symbiodinium kawagutii*) as a candidate unusual Sm-binding site in SL RNAs of 103–292 bp. Further research is needed not only to resolve discrepancies among reported SL RNA sequences and possible diversity in Sm-binding sites, but also to clarify Sm-complex assembly and dynamics during mRNA maturation, ultimately shedding light on the unique features of splicing in dinoflagellates.

5. KH Domain-Containing Proteins

First identified in the human heterogeneous nuclear ribonucleoprotein K (hnRNP K), the K homology (KH) domain is present in proteins across bacteria, archaea, and eukaryotes, often occurring in tandem copies. In general, eukaryotic proteins tend to contain more KH domains than prokaryotic proteins, although some eukaryotic KH proteins bear only a single domain [146].

KH domains are approximately 70 amino acids long and can adopt two topologies that share the same $\beta_1\alpha_1\alpha_2\beta_2$ core (Figure 1c). Eukaryotic proteins usually possess type I domains with a $\beta_1\alpha_1\alpha_2\beta_2\beta'\alpha'$ architecture, while prokaryotic proteins typically bear domains that adopt a $\beta'\alpha'\beta_1\alpha_1\alpha_2\beta_2$ fold (type II KH). In type I KH domains, the β -strands form a $\beta_1\beta'\beta_2$ sheet in which β' is antiparallel to β_1 and β_2 [146]. A short loop connecting the α_1 and α_2 helices contains a GXXG motif that is crucial for RNA binding: proteins with the typical KH fold but modified sequence of this motif do not exhibit RNA-binding activity [147]. The binding of the nucleic acid backbone occurs in a cleft formed by the two α -helices and the β -sheet. Typically, an individual KH domain binds four nucleotides, and the specificity of recognition increases when multiple KH domains are present within a single protein [146,147]. KH domain-containing proteins are involved in the regulation of transcription, mRNA processing, and translation [79]. In both prokaryotes and eukaryotes, some of them have been shown to be membrane-associated [148].

Together with RRM and zinc-finger domains, KH domains are among the most common RBDs [82]. Consequently, it is not surprising that dinoflagellates appear to possess a significant diversity of KH domain-containing proteins (see an example of the predicted tertiary structure in Figure 1c). At the transcriptomic level, such proteins have been annotated in basal parasitic dinoflagellates *Amoebophrya* sp. and *Hematodinium* sp. [13,43]. KH domain-containing proteins or KH domains were also predicted in genomes and/or transcriptomes of several core dinoflagellates, including *Symbiodinium microadriaticum*, *S. natans*, *S. tridacnidorum*, *Breviolum minutum* (syn. *Symbiodinium minutum*), *F. kawagutii*, *L. polyedra*, *C. cohnii*, *Durinskia* sp., *Pyrocystis lunula*, and *P. cordatum* [13,26,44,46,50,51]. At the proteomic level, KH domains were identified among chromatin-associated proteins of *L. polyedra*: the study by Beauchemin and Morse [28] identified 34 proteins harbouring 1–6 KH domains, of which 28 possessed only them.

6. Zinc-Finger Proteins

Zinc-finger proteins are among the most widespread nucleic acid-binding factors found across all three domains of life, with their greatest diversification occurring in eukaryotes [149–151]. They bear compact (about 30 amino acids long), Zn^{2+} -stabilized domains (ZF) that differ in the number and arrangement of zinc-coordinating cysteines and histidines, as well as in their preferred interaction partners [150,152].

Types of ZF domains are designated according to the composition of conserved amino acids that bind zinc ions. The most abundant eukaryotic ZF is characterized by a $\text{CX}_4\text{CX}_{12}\text{HX}_3\text{H}$ motif, hence the designation C2H2. These domains function primarily as DNA-binding units, although numerous C2H2-containing proteins have been shown to interact with RNA [150,153–156]. In contrast to the C2H2 ZFs, the CCCH, CCHC, and CCCC subtypes are known primarily as RBDs found in proteins involved in RNA folding, processing, and turnover [153,157–162].

The specificity of RNA recognition and the mechanism of binding differ among ZF types. C2H2 domains adopt a $\beta\beta\alpha$ fold and recognize RNA mainly through side chains of residues on the α -helix (Figure 1d). At least two modes of recognition are known: in transcription factor IIIA (TFIIIA) of *Xenopus* oocytes, different C2H2 domains bind either the backbone or individual bases of the 5S rRNA molecule [163]. CCCH domains

are compact structures composed of short α -helices and 3_{10} -helices [164] and either bind specific RNA sequences or secondary structures [153,160]. In CCCH-containing protein tristetraprolin (TTP), the recognition of AU-rich elements in 3' UTRs of mRNAs is achieved through interaction of protein main chain with RNA bases [153]. The CCHC domain is composed of two short β -strands, connected by a specific turn called “zinc knuckle”, and a short α -helix [162] (Figure 1d). These domains recognize stem-loop structures in either ssDNA or RNA by interacting with both bases in the loop (facilitating sequence-specific recognition of, for example, G-rich sequences) and the backbone [162,165].

As in the case of the KH domain-containing proteins, zinc-finger proteins bearing C2H2, CCCH and CCHC types of domains were listed both in genomic/transcriptomic and proteomic datasets of dinoflagellates (e.g., [13,26,28,43,44,50,51])

7. Pumilio Family Proteins

The Pumilio RNA-binding protein family, or the PUF (Pumilio and Fem-3 binding factor) family, comprises highly conserved eukaryotic proteins [166] that contain the sequence-specific RNA-binding Pumilio homology domain (PUM-HD), or Puf domain (reviewed in [167,168]). The PUM-HD consists of eight tandem copies of repeated motif of 36–40 amino acids and N- and C-terminal flanking regions [167,169,170]. Each repeat folds into three α -helices arranged in a triangular configuration; together, the repeats pack into a right-handed superhelix curved in the shape of half a doughnut [170] (Figure 1e). RNA is bound in a sequence-specific manner on the concave surface of the superhelix, with each repeat recognizing a single RNA base through three amino acid side chains at conserved positions 12, 13, and 16 [171]. The type of these residues depends on the position and repeat number. In most cases, PUF proteins bind specific sites in the 3' untranslated regions of target mRNAs and also can recruit other protein machineries; they mediate post-transcriptional control of gene expression, either by promoting or repressing translation and by regulating mRNA stability [168]. They have been shown to regulate a wide range of processes across eukaryotes, including stem cell control, developmental patterning and gametogenesis, organelle biogenesis, and stress response [168].

Dinoflagellates appear to possess repertoire of PUF proteins. The PUF-coding genes have been found in transcriptomic data from early-branching parasitic dinoflagellates *Amoebophrya* sp. [43] and *Hematodinium* sp. [13]. Among the core groups, PUM-HDs (domain pfam00806) were inventoried in genomes and/or transcriptomes of Symbiodiniaceae species from various clades (*S. microadriaticum*, *B. minutum*, *F. kawagutii*, *Cladocopium goreau*, *Symbiodinium aenigmaticum*, *S. pseudominutum*, *S. psymphilum*), as well as in omics data from *K. brevis*, *P. cordatum*, *L. polyedra*, *A. carterae*, *C. cohnii*, *P. lunula*, and *Alexandrium minutum* [13,26,44,46,50,54,55]. Pumilio domain was listed among RNA binding domains detected during the proteomic analysis of the chromatin-enriched samples of *L. polyedra* [28].

We did not find any dedicated studies analyzing the amino acid composition, domain architecture, or folding features of PUM-HD-carrying proteins in dinoflagellates. Nevertheless, to obtain a first approximation of PUM-HD conservation, we selected amino acid sequences containing either a Pumilio-family RNA binding repeat (pfam00806) or Pumilio-family RNA binding domain (cd07920) from the *P. cordatum* genome assembly GCA_963575745.1 proteome dataset [26] (Figure S2). The multiple amino acid sequence alignment confirmed that the regions within the repeats corresponding to the inner surface of the protein are relatively well conserved (Figure S2). Indeed, dinoflagellate Pumilio-containing proteins were among the RNA-binding proteins showing the highest degree of sequence similarity (best E-values in tblastn analysis) to mammalian queries [13]. The

architecture of dinoflagellate PUM-HD-carrying proteins also appears to be similar to the canonical one (Figure 1e).

Data on the potential functioning of PUF proteins in dinoflagellate cells remain extremely limited. In *S. microadriaticum* cells, the expression of a gene annotated as coding for Pumilio homologue 5 increased approximately 2.5-fold under heat shock conditions (36 °C for 4 h) ([50], see supplementary data, Additional file 6). Su et al. [54] reported that one gene encoding a Pumilio domain-containing protein was significantly down-regulated in *C. goreau* following one week of microplastic exposure. Additionally, in *K. brevis*, transcripts encoding Pumilio 1 homologue, along with other regulators of gene expression, showed a significant decrease in abundance throughout the stationary phase of culture growth [52]. Here it should also be noted that Roy and Morse [53] examined potential substrates for casein kinase 2, a kinase that can be implicated in circadian regulation, in *L. polyedra*. Among the proteins bearing confirmed phosphosites, they identified not only PUM-HD-carrying proteins but also the previously discussed RRM-, KH-, and ZF-domain-bearing RBPs, as well as the OB-fold family proteins described below. Thus, RBPs emerged as one of the major groups of potential phosphorylation targets for this kinase [53].

8. OB-Fold Proteins

The oligosaccharide-/oligonucleotide-binding (OB) fold group comprises proteins with structurally similar core framework—a Greek key β -barrel composed of five anti-parallel β -strands, i.e., an OB-fold domain [172,173] (Figure 1f). These proteins are found in all three domains of life, and owing to the plasticity of the loops connecting β -strands, individual OB-fold domains range in length from 70 to 150 amino acids. This structural variability allows them to bind diverse ligands, including nucleic acids, oligosaccharides, phospholipids, and proteins. Ligand binding typically occurs on a side surface of the β -barrel, with strands β 2 and β 3 forming the binding site (reviewed in [174–176]).

The overall level of sequence similarity between different OB-fold subgroups is low. Nevertheless, their ubiquity, the conservation of the ligand-binding face, and the same mode of recognition in nucleic acid-binding OB-fold proteins, coupled with the ancestral architecture reconstruction results, suggest that this and several other groups of β -barrel proteins share the same evolutionary origin, with diversification predating the last universal common ancestor (LUCA) [172,175,177–180].

Among OB-fold proteins, nucleic acid-binding proteins comprise the largest and most diverse subgroup [181,182]. Among the most extensively studied members of this subgroup known to bind RNA are the S1 domain-containing and cold shock domain-containing proteins.

8.1. S1 Domain-Containing Proteins

The S1 domain was originally described in the *Escherichia coli* ribosomal protein S1 [183,184] that binds both rRNA and mRNA (reviewed in [82,175]). Often a single protein bears multiple tandemly repeated S1 domains: results of the taxonomic distribution analysis indicate that while the number of domains in bacterial proteins typically does not exceed six, eukaryotic proteins can contain up to fifteen copies [185]. With that, the level of sequence conservation between S1 domains of a single protein can vary [186]. In bacteria, S1 domains are found in ribonucleases (including degradosome components), transcription factors, and proteins involved in adaptation to various stressors (reviewed in [187]). In eukaryotes they have been described in some of the exosome proteins facilitating RNA binding and degradation, in subunits of RNA polymerase II, as well as in proteins involved in ribosome biogenesis, and in an α -subunit of the eukaryotic initiation factor 2 (eIF2), essential for the translation initiation control (reviewed in [187]).

In the *E. coli* S1 protein, an RNA-binding site includes F19, F22, H34, D64, and R68 residues, located on the same side surface as in other typical OB-folds. The aromatic residues provide side chains for the interaction with the RNA nucleobases [184]. While these five residues are mostly conserved in bacteria, in archaeal and eukaryotic S1 domains only the first F residue and residues D and R remain conserved [187]. These observations highlight a significant level of divergence characteristic to S1 domains.

Despite specific studies of the S1 domain-containing proteins of dinoflagellates are virtually absent, S1 domain-containing proteins and/or individual S1 domains were predicted and/or identified in both the genomic/transcriptomic and proteomic databases of different dinoflagellate species (e.g., [13,26,43,44,56]). In the *L. polyedra* nuclear proteome, RBPs containing from one to four S1 domains were identified; also, at least two proteins contained S1 together with other functional domains [28].

8.2. Cold Shock Domain-Containing Proteins

The cold shock domain (CSD), found in eukaryotic proteins, is homologous to bacterial cold shock proteins (reviewed in [188]), originally described in *E. coli* [189]. Whereas bacterial cold shock proteins typically consist of a single cold shock domain, CSD-containing proteins (CSD proteins) commonly possess multiple CSDs or additional functional modules such as basic/aromatic (BA) islands and zinc-finger domains [188]. Despite differences in structural organization and distribution across the domains of life, both bacterial cold shock proteins and eukaryotic CSD-containing proteins are often referred to as CSPs (e.g., [59,190]). However, to avoid confusion, we use the acronym CSP strictly to denote bacterial cold shock proteins unless it is part of a specific protein name.

In bacteria, CSPs are expressed in response to cold shock and can act as transcription activators of cold-shock induced genes. They also enhance RNA stability and facilitate translation initiation by preventing the formation of RNA secondary structures and stabilizing ribosome binding (reviewed in [191]). In plants, CSD-containing proteins have likewise been shown to participate in cold stress response and acquisition of freezing tolerance, as well as in regulating seed germination and growth [190,192,193].

Within the CSD, strands $\beta 2$ and $\beta 3$ bear the canonical nucleic acid-binding motifs RNP1 [KR]G[FY][GA]F[VI]X[FY] and RNP2 [LI][FY][VI][GK][NG]L—or modified variants [KS]G[FKY]G[FL]IXX and [LIV][FQ][VAL]HX[STR], respectively (reviewed in [191]) (Figure 1f). Single stranded DNA or RNA binding occurs in a groove on the β -barrel surface, primarily through interactions between aromatic amino acid side chains and nucleobases. The nucleic acid backbone is not involved in the interaction, which may explain why CSDs display limited sequence specificity and can bind both ssDNA and RNA (reviewed in [191,194]). Interestingly, at both the level of RNP motif composition and overall domain topology, CSDs show some resemblance to the RRM domains. However, this similarity is generally regarded as a case of convergence rather than homology [195,196].

Dinoflagellates appear to have an extensive repertoire of CSD-containing proteins. In *L. polyedra*, they constitute a significant fraction of nucleic acid-binding proteins in the nuclear proteome [28]. In the transcriptome of *L. polyedra*, 68% of all DNA-binding domains (in this and several other studies CSDs are classified within this category) corresponded to CSDs [58]. In transcriptomes of two *Symbiodinium* species, CSD was listed as the most common domain (>60%) of transcription factors [57]; the same was noted for the *S. microadriaticum* transcriptome [50]. *P. glacialis*, another member of the order Suessiales, also showed a strong over-representation of CSDs (PF00313) in transcripts [33]. In *P. cordatum*, Kalvelage et al. [66] reported 59 genes encoding putative CSD-containing proteins. Consistently, in the supplementary data of the *P. cordatum* genome annotation published by Dougan et al. [26], 57 protein-coding sequences were annotated as “CSP” or “CSP-like”

proteins. Moreover, genes encoding CSD-containing proteins in dinoflagellates frequently occur in tandem arrays (see Supplementary Data in [26,32,33]).

All reported dinoflagellate CSDs carry two specific RNP motifs, with consensus sequences KGFGFI (RNP1) and VFVHXX (RNP2), as inferred from published alignments [59,61–64]. Predicted tertiary structures of *P. cordatum* (syn. *P. minimum*) PmCSP1 and PmCSP2 indicate that dinoflagellate CSD-containing proteins also fold into the canonical Greek key β -barrel [63] (Figure 1f). Similar to bacterial CSPs, which can bind DNA as transcription factors in addition to their RNA-binding function [191], dinoflagellate CSD-containing proteins also exhibit the ability to bind both DNA (dsDNA and ssDNA) and RNA non-specifically, as demonstrated in *L. polyedra* and *F. kawagutii* [59,64]. Notably, they display a preference for ssDNA and RNA over dsDNA binding [59,64]; moreover, RNA appears to compete efficiently with DNA [64].

In dinoflagellates, Beauchemin et al. [59] distinguished several types of CSD-containing proteins based on domain architecture: (1) proteins containing a single CSD, (2) CSD plus a C-terminal glycine-rich (GR) domain, (3) CSD plus a C-terminal GR domain and a zinc-finger domain, and (4) proteins with multiple CSDs and one/several RRM. Most sequences fall into the first two types [59], and CSD proteins predicted in various dinoflagellate species correspond to these categories [61–63] (Table 3).

Table 3. Variants of CSD-containing proteins domain architectures in dinoflagellates and their occurrence in different species.

Domain Architecture	Dinoflagellate Species	References
CSD	<i>Lingulaulax polyedra</i>	[59]
	<i>Pfiesteria piscicida</i>	[62]
CSD + GR	<i>Alexandrium tamarense</i>	[59]
	<i>Karenia brevis</i>	[63] *
	<i>Karlodinium veneficum</i>	[62]
	<i>Lingulaulax polyedra</i>	[59]
	<i>Prorocentrum cordatum</i>	[63]
	<i>Scrippsiella acuminata</i>	[61] **, [63]
	<i>Symbiodinium</i> sp.	[59]
CSD + GR + ZF	<i>Alexandrium pacificum</i>	[62]
	<i>Lingulaulax polyedra</i>	[59]
RRM + (CSD) ⁿ + RRM	<i>Lingulaulax polyedra</i>	[59]

* Wang et al. [63] presented this CSD-containing protein (GenBank accession number FK848095.1) as possessing CSD + ZF + GR domain architecture (see Figure 3 in [63]); however, this sequence appears to lack any recognizable ZF and contain only GR region. ** Deng et al. [61] described this CSD-containing protein as possessing only CSD; however, Wang et al. [63] presented it as also exhibiting a C-terminal GR region (see Figure 3 in [63]) and provided amino acid sequence evidence supporting this domain organization.

To date, CSD-containing proteins have been annotated or identified in various dinoflagellate species with different life histories, although functional data on their activity in dinoflagellate cells remains scarce. The only direct evidence for the involvement of CSD proteins in the cold shock response comes from *P. cordatum*, where two genes encoding proteins with a CSD + GR architecture showed enhanced expression after exposure to low temperatures (shift down from 20 °C to 16 °C, 12 °C, or 4 °C for 3 h or 6 h) [63]. Stephens et al. [45] listed CSD (PF00313) among the over-represented domains in the transcriptomes of psychrophilic dinoflagellates compared with transcriptome datasets from all other studied species. However, no significant changes in CSD-containing proteins expression were observed during cold-induced temporary cysts formation in *P. cordatum* [36] or *L. polyedra* [59,60]. In *S. acuminata*, a gene encoding a CSD protein did not show altered

activity in response to low-temperature treatments [61], but such genes were significantly up-regulated in resting cysts [34,61]. In the early-branching dinoflagellate *Amoebophrya* sp., several genes coding for CSD proteins were most active at the free-swimming infective dinospore stage and during the maturation stage inside the host cell [43]. In *Alexandrium pacificum*, the expression of a gene encoding a protein with a CSD + GR + ZF domain composition increased both during the logarithmic growth phase and induced logarithmic growth phase (under high-phosphate/high-manganese conditions) [62].

9. DEAD/DEAH-Box Helicases

The DEAD/DEAH-box helicase family belongs to the superfamily 2 (SF2) of ATP-dependent, nucleic acid-binding RecA-like helicases, together with related groups such as DEAH-, DEXH-, and DEXD-box proteins [197]. This long-established and large protein family is widely distributed across all eukaryotes, most prokaryotes, and also in viruses [198–200].

A hallmark of DEAD-box helicases is the complex organization of the helicase core, which is its key functional region. The auxiliary N- and C-terminal extensions are generally non-conserved. The canonical helicase core contains nine conserved motifs responsible for ATP binding and hydrolysis and for RNA binding: the Q motif and motifs I, Ia, Ib, and II–VI [201] (Figure 1g). Additionally, conserved coupling elements referred to as Ic, IVa and Va are distinguished, giving a total of 12 conserved motifs [200,202]. Motifs I and II—Walker motifs A and B, respectively—are responsible for binding and hydrolysis of ATP. Motif I (AXTG[ST]GKT) forms the phosphate-binding loop, while motif II (DEAD), which gave the family its name, hydrolyzes bound ATP; these two motifs act in coordination with the Mg^{2+} ion [201]. The Q motif, specific to DEAD-box helicases (and absent in DEAH- and DEXH-box helicases), recognizes the adenine base of ATP. It interacts with the ligand-bound motif I and is considered to regulate ATPase activity by ensuring proper substrate binding [197,201]. Motif VI has also been proposed to bind the phosphoryl group of ATP and interact with motif II. Motifs involved in RNA binding include motif V (TDV[AG]ARGID), assisted by motifs Ia–Ic, and motif IV (minimal sequence reported [LV]IF) with IVa [200,201]. RNA binding sites contact the sugar phosphate backbone of RNA. Motifs III (SAT) and Va (when defined, extends into part of motif V and several downstream residues) are considered as linking ATPase and helicase activities. Structurally, the helicase core adopts a two-domain architecture (domains 1 and 2): each domain is composed of five β -strands and five α -helices (Figure 1g). These domains are connected by a flexible linker and form a cleft that constructs the ATP-binding site. The RNA binding site is located on the opposite site, perpendicular to the cleft.

The helicase core provides the ATP binding/hydrolysis cycles that are accompanied with conformational changes and RNA binding followed by unwinding. In addition, RNA clamping and annealing activities were shown for some helicases, as well as the ability to remove other proteins bound to RNA [200–202]. Functionally, DEAD-box helicases participate in a wide range of cellular processes associated with RNA metabolism [202]. A well-known member of this family is the eukaryotic initiation factor 4A (eIF4A), a subunit of the eIF4F complex that promotes translation initiation in eukaryotes [203]. The eIF4A helicase is required for exposure of mRNA to the ribosome. Besides, various DEAD-box helicases are involved in pre-mRNA splicing, ribosome biogenesis, RNA export, and degradation, as well in the processes providing organelle gene expression [199].

Dinoflagellates, like other eukaryotes, are expected to possess a variety of DEAD-box helicases. The motifs within the helicase core described above are highly conserved and well defined in their sequences, consistent with those in other organisms (Figure 1g). A rapid search for dinoflagellate DEAD-box including hits in the UniProtKB

database (filters: “taxonomy_id:2864” AND “ft_domain:DEAD”) confirms this conservation. Literature analysis shows that DEAD-box helicases have been annotated or identified in genomic/transcriptomic/proteomic data from both basal and core dinoflagellate groups [26,43–46,50,52,66]. Besides, some genes coding for DEAD-box helicases are listed among those organized into tandem arrays (see Supplementary Data in [26,32,33]).

Proteomic analysis revealed a high abundance of DEAD-box RNA-helicases, including homologue of the splicing-dependent multiprotein exon junction complex component eukaryotic initiation factor 4A-III, among other nuclear proteins in the nuclear fraction of *P. cordatum* cells [66]. Similarly, Beauchemin and Morse [28] reported a large number of DEAD-box helicases among other RNA-binding proteins in the nuclear proteome of *L. polyedra*. They also presented schematic representations of the revealed RBP domain architecture variants, although without detailed discussion. The predominant compositions included a single DEAD-box domain or a combination of DEAD-box and helicase C, the latter sometimes distinguished as encompassing motifs V and VI. Besides, proteins with DEAD-box + helicase C + HA2 + OB structure were identified [28]. The HA2 appear to be a helicase-associated domain, while the OB represents the aforementioned OB-fold domain; such an architecture resembles that of DEAH-box helicases, for example, the yeast RNA helicase-like Prp2 (precursor RNA processing) factor involved in pre-mRNA splicing [204]. Another identified composition, DEAD-box + helicase C + rRNA + HCT [28], may be interpreted as a DEXD/DEXH-type helicase containing rRNA-processing arch domain essential for proper 5.8S rRNA processing and a helicase C-terminal tail domain. Such a protein may be homologous to the RNA exosome complex cofactor Dob1/Mtr4 helicase [205].

Special mention should be made of proteins reported to possess the “RRM + DEAD” architecture [28]. We could not find any data on such proteins in dinoflagellates; moreover, this configuration appears to be rare among eukaryotes. In the GenBank and RefSeq NCBI databases, we found several DEAD-box sequences with predicted RRM/RRMs from some plants and tardigrades (for example, *Brassica napus* CAF1933040.1 and *Ramazzottius varieornatus* GAU95793.1), as well as a DECD-box chlorophyte sequence (*Bathycoccus prasinos* XP_007510271.1). However, no information about such helicases could be found in the literature. In prokaryotes, DEAD-box helicases with a single C-terminal RRM domain—such as *E. coli* DbpA and *Bacillus subtilis* ortholog YxiN, both involved in large ribosomal subunit biogenesis—have been described [206–208]. These proteins bind 23S rRNA by means of their RRM domain with high affinity and specificity; this interaction defines conformational changes that promote formation of the catalytically important, closed conformation of the protein. In turn, the DEAD-box helicase core acts as the catalytic domain that destabilizes RNA duplexes in an ATP-dependent manner [209,210].

We searched for dinoflagellate proteins characterized by the presence of RRM and DEAD/DEAH-box helicase domains in the UniProtKB database, and then performed blastp analysis in the NCBI databases and MMETSP project transcriptome assemblies [211]. We revealed genomic and/or transcriptomic sequences—one sequence per assembly, with one exception—belonging to representatives of the Symbiodiniaceae and Kareniaceae families. These sequences possess a well-defined DEAD-box helicase core (with the DEAD signature) and upstream (rather, N-terminal in some cases) RRM (e.g., *S. microadriaticum* OLP92746.1, *S. necroappetens* CAE7876399.1, *Symbiodinium* sp. CCMP2456 CAE7718496.1, *K. brevis* CCMP2229 20130916|87251_1) (Figure 2a). The number of RRM varied depending on the prediction tool used and among taxa. In most symbiodiniacean sequences, the CD-search tool recognized four RRM as specific hits (Figure 2a), although variants with one to three RRM were also found. The InterPro [212] predictive models appeared to be more sensitive—or more permissive—identifying five RRM in almost all sequences.

Notably, the tertiary structure model predicted by the AlphaFold Server for *Symbiodinium* sp. CCMP2456 CAE7718496.1 did reveal five RRM-like $\beta\alpha\beta\alpha\beta$ folds (Figure 2b). In kareniacean proteins, four RRM domains were predicted by InterPro (and two to three by CD-search). The regions corresponding to the conserved RRM motifs RNP1 and RNP2 were modified relative to the canonical amino acid composition in the revealed sequences. This variation was especially pronounced in RNP2, while RNP1 was conserved in some domains—e.g., KGFAYVQY motif—or retained a similar pattern. Again, as exemplified by the case above, the characteristic $\beta\alpha\beta\alpha\beta$ fold of RRM domains was preserved.

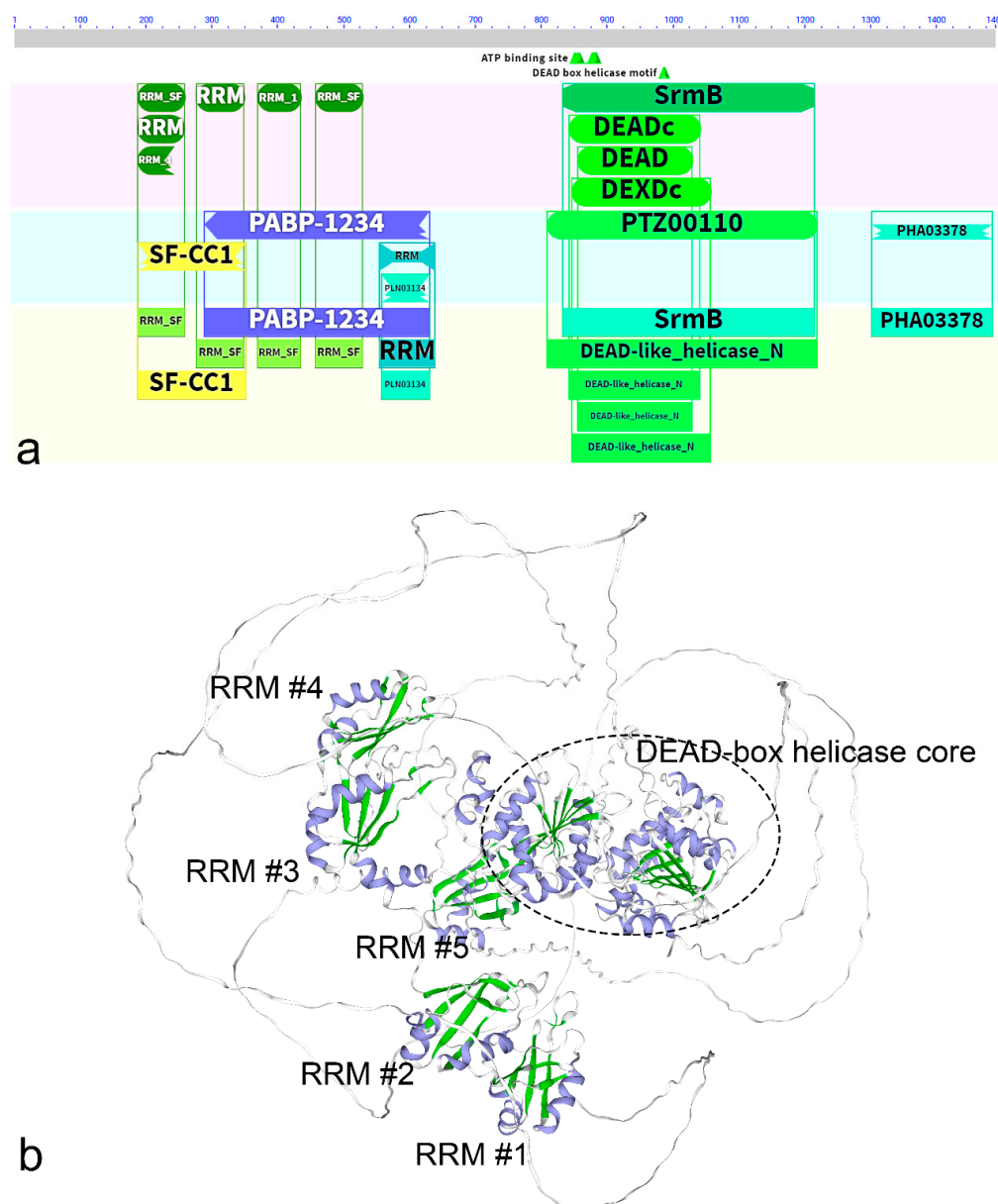


Figure 2. A putative RRM-DEAD protein (*Symbiodinium* sp. CCMP2456 sequence CAE7718496.1 retrieved from the GenBank database): (a) domain composition determined the using NCBI CD-search tool and (b) tertiary structure predicted by the AlphaFold Server. In (b), the colour scheme reflects secondary structure features: α -helices (lilac) and β -sheets (green). Five putative RRM domains and a DEAD-box helicase core (indicated by a dashed oval) are recognizable. Notably, the CD-search tool (a) determined four RRM domains as specific hits (upper pink-shaded panel), i.e., top-ranking hits that meets or exceeds a domain-specific E-value threshold, and a fifth RRM domain as a non-specific hit (middle blue-shaded panel), i.e., hits that meet or exceed the blastp threshold for statistical significance.

Additionally, we revealed similar sequences in transcriptome assemblies from *S. acuminata* and the heterotrophic dinoflagellates *Tripos fusus* (syn. *Ceratium fusus*) and *C. cohnii*. In these proteins, only one RRM domain was recognized. Moreover, the *S. acuminata* and *T. fusus* sequences contained one N-terminal ZF domain of the CCHC type with a motif CX₂CX₃HX₄C (Figure 1d), which is slightly different from the typical CX₂CX₄HX₄C motif, characteristic of the CCHC ZF domains [157,162]. Among DEAD-box helicases, homologues of Vasa identified in invertebrates (except for insects) are known to contain CCHC-type ZFs [213]. Vasa proteins were primarily characterized as functioning during the germ line development and participating in the germ cells specification, proliferation, and maintenance via translation regulation. Besides, in various taxa, genes encoding Vasa homologues are expressed in non-germ cell types such as multipotent stem cells [213]. The ZF domains are believed to contribute to RNA target specificity and may be linked to Vasa functions beyond the germ line [213]. Interestingly, among Vasa homologues we also found DEAD-box helicases with RRM domains. ZF-containing Vasa-2 homologues from hydrozoan cnidarians, such as *Hydra vulgaris* (BAB13308.1), *Nanomia bijuga* (AHI50305.1) [213,214] exhibited N-terminal RRMs, although these domains were not described in the related studies. Their Vasa-1 homologues by contrast, lack RRMs and are phylogenetically distant from Vasa-2 [214].

The conserved composition of the helicase core suggests that DEAD-box helicases in dinoflagellates retain their functional activity. As for their role in cellular processes, available data points, at minimum, to correlations of their activity with cell cycle progression and growth regulation. In *K. brevis*, genes encoding DEAD-box helicases were repressed during the stationary phase of culture growth [52]. Besides, a DEAD box RNA helicase was mentioned as belonging to a cluster of genes, which were up-regulated throughout the light phase of the diel cycle and characterized by a moderate amplitude of expression change during this phase [65]. In *L. polyedra*, two DEXD- and DEXH-containing proteins were overrepresented in the light phase (midday) samples [28]. A DEAD/DEAH box helicase domain-containing protein was highly expressed during the G2/M and, presumably, S phases of the cell cycle in *Prorocentrum donghaiense* [56].

Notably, information on the dinoflagellate homologues of the eukaryotic initiation factor 4A (eIF4A), unlike other translation initiation factors, is very scarce. Mentions of its presence appear in omics datasets [46], and several eIF4A annotated entries exist in the UniProtKB database.

10. YTH Domain-Containing Protein Family

Members of the YT521-B homology (YTH) domain-containing protein family share a conserved domain with the mammalian YT521-B splicing factor and are now known to be broadly distributed across eukaryotes [215,216]. Unlike many other RBPs that occur in both prokaryotes and eukaryotes, YTH domain-containing proteins are exclusive to eukaryotic organisms [216].

Most YTH proteins possess a single C-terminal YTH domain of approximately 145–150 amino acids, typically accompanied by intrinsically disordered N- or C-terminal regions [215,217]. The YTH domain adopts a mixed α -helix/ β -sheet fold (Figure 1h), organized into a four α -helices and six β -strands [215] or three α -helices, eight β -strands, and two 3_{10} -helices [218]. A defining structural feature of YTH domain is an aromatic cage that is composed of conserved Trp and Phe residues located within β -strands; this structure is essential for RNA binding [216,217]. The YTH domain binds single-stranded RNA [216]. The YTH domain-containing proteins have been characterized as so-called m⁶A readers—proteins that specifically recognize N⁶-methyladenosine (modified adenosine

in RNA). The methyl group is accommodated within the aromatic cage and additionally stabilized by residues forming hydrogen bonds with the nucleotide [219,220].

YTH domain-containing proteins are key effectors of the epitranscriptomic m⁶A pathway and participate in multiple aspects of RNA metabolism. They regulate mRNA stability and degradation, splicing, translation, and nuclear export, thereby facilitating rapid remodeling of gene expression programs (see for review [217]). Notably, in fission yeasts, YTH domain-containing protein Mmi1 binds transcripts of meiosis-associated genes to prevent entry into the meiotic program; this repression is later relieved through inactivation of Mmi1 by the RRM-containing protein MEI2 and the long non-coding meiRNA during meiotic initiation [90,221,222].

In dinoflagellates, in contrast to other RBPs, only fragmentary data are available regarding the presence of YTH domain-containing proteins. Two genes encoding proteins with YTH have been annotated in the *P. cordatum* genome (Supplementary Information in [26]), and nine in the *A. minutum* genome [40]. These proteins—for example, *P. cordatum* CAK0851154.1—indeed contain a well-defined C-terminal YTH domain with conserved Trp and Phe residues and exhibit a mixed α -helix/ β -sheet fold similar to canonical structures (Figure 1h). Mentions of YTH domains in dinoflagellates are also found in Roy et al. [13] and Stephens et al. (Supplementary Information in [45]). Additionally, homologs of the YTH-containing DEXH-box RNA helicase YTHDC2 have been predicted in the genomes of *Breviolum minutum* and *F. kawagutii* [44]. In mammalian cells, YTHDC2, similarly to yeast Mmi1, contributes to the transition from mitotic to meiotic divisions [223]. Notably, despite an extensive repertoire of MEI-like proteins, dinoflagellate appear to lack any recognizable homologs of Mmi1 protein.

The activity and cellular functions of YTH-containing proteins in dinoflagellates remain entirely unexplored experimentally. In this context, it is particularly intriguing that in these algae, m¹A mRNA modification, rather than m⁶A, has been reported as the most prevalent internal mRNA modification [224]. This raises the question of whether dinoflagellate YTH proteins interact with RNA in the same manner as their counterparts in other eukaryotes, or whether their ligand specificity and functional roles may differ substantially.

11. Pentatricopeptide Repeat Proteins

The pentatricopeptide repeat (PPR) proteins form a large family of sequence-specific RNA-binding factors, which is widely expanded among eukaryotes. They are structurally similar to tetratricopeptide repeat (TPR) proteins, which are engaged in protein–protein interactions and occur in both prokaryotes and eukaryotes [225]. PPRs are thought to have evolved from TPRs early in eukaryotic history (see for review [226]).

The structural basis of PPR proteins is the PPR motif, a ~35-amino acid (canonical type) repeat arranged in tandem. The repeat is degenerate, i.e., showing diverse amino acid combinations while retaining key structural and functional characteristics [226,227]. The motif conserved patterns include coordination of nucleotides by charged or polar residues and the presence of a hydrophobic pocket containing a proline residue between two adjacent motifs [72,228]. To describe RNA recognition specificity, the concept of a “PPR code” was introduced—specific combinations of amino acids at two key positions within the PPR motif that determine binding to a particular nucleotide [226]. According to the current plant “code”, the fifth and last residues of the 35-amino-acid repeat are primarily responsible for nucleotide recognition [72]. The most common combinations at these positions include Asn and Asp, though Thr and Ser also occur frequently [72].

Each PPR motif adopts a helix–turn–helix hairpin structure consisting of two antiparallel α -helices, helix a and helix b [72,228,229]. Tandem motifs stack into a right-handed α -solenoid, forming a central groove that binds RNA [72,228] (Figure 1i). The RNA-binding

fifth residue corresponds to the helix α , while the last residue lies in the loop interconnecting adjacent motifs [72]. PPRs bind single-stranded RNA in a one-repeat: one-nucleotide manner. Experimental and computational analyses have defined a correspondence between residue pairs and target nucleotides (e.g., [230]). Compared to other RNA-binding proteins, PPRs recognize target sequences of a greater length and diversity (see for review [226]).

Based on their repeat architecture, the PPR family is conventionally divided into two major subfamilies (classes). P-class proteins consist solely of canonical 35-residue P motifs and may lack additional domains; however, several subgroups can be distinguished in which the PPR array is followed by C-terminal domains, such as PPR-SMRs carrying small MutS-related (SMR) domain, PRORP-PPRs with a metallonuclease domain, PPR-TGMs with tRNA guanine-N7 methyltransferase domain, or PPR-TGMs that also possess CCH-type ZF [231].

The PLS-class PPRs contain a characteristic triplet composed of P-, L- (long; ~35–36 amino acids) and S- (short; ~31–32 amino acids) repeats. They are arranged as (P1–L1–S1)ⁿ followed by a P2–L2–S2 triplet and, in some cases, can alternate with or be preceded by SS-type repeats. All PLS-associated repeat variants differ in length and in some positions; for example, P1 and P2 repeats differ from the canonical P repeat in the amino acid composition of the first helix [72]. Subgroups of PLS proteins may also harbor downstream E1/E2-repeats [72]. DYW-type PPRs carry an extended C-terminal DYW region, which includes a zinc-dependent cytidine deaminase signature implicated in cytidine-to-uridine RNA editing [232,233].

Although encoded in the nucleus, most PPR proteins are targeted to mitochondria or plastids, where they act as key post-transcriptional regulators of organellar gene expression by participating in RNA stabilization, splicing, processing, translation, and editing [226,231,234–236].

PPR proteins appear to be well-defined (see as an example Figure 1i) and are reported to be highly abundant in dinoflagellates. Moreover, genomic evidence indicates they are definitely among the PPR-coding genes arranged into blocks of tandemly duplicated genes (see Supplementary Data in [26,32,33]). The abundance of PPRs has been shown for species from various taxa: PPR domains were among the most numerous in the genomes of *Amphidinium gibbosum* [75], as well in symbiodiniacean and suessiacean species [11,33,44,45,51,70]. In thecate photosynthetic species, several hundred putative P-class PPRs can be distinguished: 676 and 536 genomic-based protein sequences were reported for *P. glacialis* and *P. cordatum*, respectively [78]; more than 200 transcriptome shotgun assembly entries were reported for *Alexandrium tamarense*, and about 800 for *L. polyedra* [69] (Table 4). Symbiodiniaceans appear to demonstrate comparable numbers of PPRs: 556 and 265 PPR-containing gene models were annotated for *S. microadriaticum* and *F. kawagutii*, respectively [44] (Table 4). For *B. minutum*, the number of annotated P-class PPR-coding genes varies from 493 [44] to 620 [23]. Cheng et al. [72] listed 687 (per the main text)/784 (per the supplementary data) P-class PPRs in *B. minutum* (Table 4).

Notably, the athecate dinoflagellate *K. brevis* (Gymnodiniales) possesses about 100 [67] or 120 [69] PPRs (Table 4). If confirmed by broader datasets, this may indicate that the major expansion of PPR-coding genes occurred after the divergence of Gymnodiniales. In addition, non-photosynthetic dinoflagellates—which are thought to retain a cryptic plastid [237–239]—also possess fewer but still substantial numbers of PPRs. Transcriptome data from *C. cohnii* revealed 57 P-class PPRs, and even the early-branching basal species *O. marina* has 46 homologues [78] (Table 4).

Table 4. Diversity, classification, and abundance of PPR proteins in dinoflagellates based on available genomic and transcriptomic data. The table summarizes the presence of PPR proteins belonging to different subgroups of the P and PLS classes in dinoflagellates. For each taxon, the number of predicted PPR proteins or PPR-coding genes is indicated together with the data source of—genome (G) or transcriptome (T); proteome datasets derived from genome assemblies (e.g., in *P. cordatum* and *P. glacialis* cases) are also denoted as “G”.

PPR Class	PPR Subgroup	Dinoflagellate Species	Number	Data Source	References
P	P-class PPR (without additional domains)	<i>Alexandrium tamarense</i>	>200	T	[69]
			493	G	[44]
		<i>Breviolum minutum</i>	620	G	[23]
			611 *	G	[72]
		<i>Cryptothecodinium cohnii</i>	57	T	[78]
		<i>Fugacium kawagutii</i>	265	G	[44]
		<i>Karenia brevis</i>	100	T	[67]
			120	T	[69]
		<i>Lingulaulax polyedra</i>	800	T	[69]
		<i>Oxyrrhis marina</i>	46	T	[78]
		<i>Polarella glacialis</i>	676	G	[78]
		<i>Prorocentrum cordatum</i>	536	G	[78]
		<i>Symbiodinium microadriaticum</i>	556	G	[44]
	PPR-SMR	<i>Alexandrium tamarense</i>	2	T	[78]
		<i>Cladocopium</i> sp.	1	T	
		<i>Cryptothecodinium cohnii</i>	2	T	
		<i>Karenia brevis</i>	1	T	
		<i>Karlodinium veneficum</i>	1	T	
		<i>Lingulaulax polyedra</i>	5	T	
		<i>Prorocentrum cordatum</i>	2	G	
		<i>Polarella glacialis</i>	4	G	
		<i>Tripos fusus</i>	1	T	
		<i>Scrippsiella acuminata</i>	2	T	
PLS?	PPRs with L1, L2, S1, S2, or SS repeats (no P-L-S triplet)	<i>Breviolum minutum</i>	173 *	G	[72]
PLS	DYW-PPR	<i>Breviolum minutum</i> <i>Symbiodinium microadriaticum</i>	2 **	G, T	[23,72] [73]

* Cheng et al. [72] reported 687 P-class PPR genes in *B. minutum* (per the main text). In Supplementary data Table S4 [72], 784 gene-based variants of predicted motif structure excluding DYW-type genes are listed. Among these, 611 variants contain neither L- nor S-repeat types and consist only of P repeats (including P1 and P2 repeats, which we group together here, although P1 and P2 repeats were originally described for PLS-class PPRs [72]). The remaining 173 variants contain L1, L2, S1, S2, or SS repeats in various combinations. Nevertheless, these variants cannot be strictly classified as PLS-class PPRs due to the absence of canonical P-L-S triplets. ** Only one sequence (symbB1.v1.2.035968.t2) in the relevant databases contains the amino acid residues Asn-Glu-Trp (NEW) at the position corresponding to the DYW motif.

The diversity of dinoflagellate PPRs appear to be largely formed by P-class proteins, and, moreover, the main pool likely consists of “sensu stricto” P-class PPRs that lack additional domains. Nevertheless, dinoflagellate PPR repertoires may also include proteins belonging to other PPR subgroups. Core dinoflagellates possess PPR-SMRs [78]—P-class PPR proteins with a C-terminal SMR domain, presumably an endonuclease [240]. According to plant studies, PPR-SMRs can participate in organellar RNA stability, translation, or editing [240–243]. Dinoflagellate PPR-SMRs fold into the typical superhelical PPR structure,

and their SMR domain exhibits the characteristic $\beta\alpha\beta\alpha\beta\beta$ topology [78]. In general, two conserved motifs typical of SMR domains—the N-terminal X_aDXH (with X for any non-polar amino acid residue and X_a for only aliphatic ones) and the putative catalytic $TGXG$ site—are retained, although some individual substitutions may occur [78]. Importantly, dinoflagellates have seized the SMR domain fused to the PPR backbone, whereas dinoflagellate homologues of MutS2—the other protein canonically bearing the SMR domain—exist as MutS domain III-V-containing MutS2-like proteins and lacking an SMR domain [78,244]. Across different species, usually 1–2, and in some cases 4–5, PPR-SMR homologues are present (Table 4). Phylogenetic analyses based on alignments of PPR-motif regions show that they do not separate from the broader diversity of dinoflagellate P-class PPRs without additional domains [78]. Notably, dinoflagellates with tertiary plastids have one PPR-SMR each: sequences from kryptoperidiniacean species with diatom-derived plastids cluster with diatom whereas those from kareniaceans with haptophyte-derived plastids branch together with dinoflagellate PPR-SMRs [78].

The presence of PPRs from other P-class subfamilies in dinoflagellates has yet to be proven. Nevertheless, for example, brief screening of the published *P. cordatum* genomic data already indicates the existence of PPRs with a putative methyltransferase domain [26]. PPR-TGMs with tRNA guanine-N7 methyltransferase domain have been predicted in various protists, including the alveolate *Perkinsus marinus* [235,245]. These proteins are predominantly targeted to mitochondria—plastid targeting remains unclear—and are predicted to function in tRNA metabolism [245].

It should also be noted that among P-class PPRs (both “sensu stricto” and PPR-SMRs), sequences containing a specific and previously undescribed Ser-Ala-Cys-Glu-Lys (SACEK) motif were reported in core dinoflagellates [78]. It occupies the pre-turn region of helix a of the 35-amino-acid PPR hairpin (positions 9–13). Outside dinoflagellates, such sequences were found only in heterokontophytes and in some cryptophyte and haptophyte species—i.e., among lineages of the CASH assemblage. Despite some occasional substitutions (most frequently at position 12), the SACEK motif is highly conserved; in addition, a downstream Trp and several nonpolar residues are generally well conserved. In some sequences, multiple consecutive SACEK-containing PPR repeats were observed. These SACEK-PPRs are abundant, for example, in *P. cordatum*, *P. glacialis*, and *Alexandrium tamarense*. In other taxa, only one to three SACEK repeats (or modified variants) were found. The species listed above also contain such homologues, whereas others (*A. carterae*, *S. acuminata*, *C. cohnii*) possess only these SACEK-PPRs. Phylogenetically, proteins with multiple SACEK repeats form a well-supported group; most SACEK-containing sequences, regardless of the composition, tend to cluster on closely related branches. It is noteworthy that dinoflagellate P-class PPRs lacking SACEK signature group separately; in one such cluster, they group together with other alveolates [78]. Given the location of the motif opposite to the putative nucleotide-recognition site and its amino acid properties, the SACEK motifs may contribute to stabilizing α -helical stacking within the PPR superhelix or modulating its flexibility [78].

The presence of PLS-class PPRs in dinoflagellates remains questionable. As noted above, Cheng et al. [72] reported 784 P-class PPRs in *B. minutum* within their analysis of putative PPR motif structures across 109 eukaryotic species. According to their results, many *B. minutum* PPRs contain not only canonical P repeats but also arrays of P1 or P2 repeats (611 gene-based variants of predicted motif structure in total) (Table 4). Although P1 and P2 repeats were originally described in the context of PLS-class PPRs, these variants are more appropriately classified as P-class PPRs. Besides, 173 genes encoding combinations that also include L1, L2, S1, S2, or SS repeats were listed (Table 4). However, the strict repetitive P-L-S triplets organization required for assigning proteins to the PLS-class was not observed in *B. minutum* sequences. Furthermore, E-type variants were absent [72].

Interestingly, in their study of RNA editing under different stress conditions in *S. microadriaticum*, Liew et al. [73] mentioned two candidate PPR proteins with presumable DYW domains as candidates for components of the RNA-editing machinery. Two PPR genes containing DYW domain were also mentioned in the *B. minutum* genome [23,72] (Table 4). On the one hand, the published alignment of the corresponding region of *B. minutum* sequences with the DYW domain from *Arabidopsis thaliana* proteins [73] shows that the examined dinoflagellate sequences contain the highly conserved cytidine deaminase signature HXEX_nCXXC, a key Zn²⁺-binding motif required for RNA editing [246]. On the other hand, they lack the C-terminal Asp-Tyr-Trp (DYW), or homologous, signature essential for domain activity [73]; only the sequence symbB1.v1.2.035968.t2 contains residues Asn-Glu-Trp (NEW) at the corresponding position (Figure S3). Moreover, our analysis of the full corresponding *S. microadriaticum* and *B. minutum* sequences retrieved from GenBank, ReefGenomics (<http://reefgenomics.org/>, accessed on 29 October 2025), and the OIST Marine Genomics Unit (<https://marinegenomics.oist.jp>, accessed on 29 October 2025) databases revealed, in some cases, an extension downstream of the putative DYW domain containing ankyrin repeats (Figure S3). Only a few similar dinoflagellate sequences can be found in currently available datasets. Therefore, the presence and functional significance of putative DYW-type PPR proteins in dinoflagellates remains an open question.

Another curious finding concerning PPR domains in dinoflagellates comes from the study of polyketide synthases (PKS) and non-ribosomal peptide synthases (NRPS) diversity in representatives of three symbiodiniacean clades—B1, A3, and C (*B. minutum*, *Symbiodinium tridacnidorum*, and *Cladocopium* sp., respectively) [74]. PKS and NRPS are modular enzymes responsible for the biosynthesis of secondary metabolites [247]. The authors noted that PPR repeats, together with ankyrin and HEAT repeats, may occur as accessory units fused with certain monofunctional PKS or NRPS domains, presumably *trans*-acting ones, i.e., encoded separately from the other domains in the module in the biosynthetic gene clusters [248]. We were not able to identify such PPR-containing sequences on the basis of the published materials (in contrast to HEAT- or ankyrin-containing ones). However, a search for PKS/NRPS sequences with PPR domains in public databases revealed two *S. microadriaticum* PKS proteins (OLP78851.1 and CAE7327639.1). As an example, Figure 3 shows the domain organization of the protein OLP78851.1. They consist of putative modules—three modules in OLP78851.1 and one in CAE7327639.1—containing PKS-associated domains in different combinations. Notably, OLP78851.1 possesses acyl transferase domains in these modules, suggesting that it represents a *cis*-PKS rather than a *trans*-PKS (Figure 3). At the C-terminus, both sequences contain an accessory region with a ketoreductase (KR) unit that includes several PPR motifs (Figure 3). These PPRs turned out to be SACEK-containing PPRs, with multiple consecutive SACEK-type repeats.

The KR module comprises two subdomains, a structural and a catalytic one, the latter containing the YXXXN active-site motif; an additional domain such as enoylreductase (ER) may be inserted between them [249]. Indeed, this arrangement is observed in the upstream module of the revealed *S. microadriaticum* PKSs. Interestingly, the KR with the inserted ER contains the canonical NADPH-binding motif TGGTGXLG, whereas the KR with inserted PPRs lack it and retains only the YXXXN motif, which may indicate it represents a different KR type [249]. We found no literature describing PPR repeats associated with a KR module. However, some plant PPR-containing proteins (mainly DYW-type PPRs) with a C-terminal domain belonging to the short-chain dehydrogenase/reductase (SDR) superfamily, to which KRs also belong [249], and with the YXXXX active-site motif, can be retrieved from databases.

Skav226499, Skav228199, and Skav233181) out of seven differentially expressed PPR-coding genes were up-regulated and two genes (IDs Skav205856 and Skav204446) were down-regulated. In cells grown in the phosphate-depleted conditions or in medium with dissolved inorganic phosphate replaced by organic one, differential expression of these genes was also observed [71]. Notably, direction of changes was the same for the genes common for different stress conditions [71]. Cold stress also altered the expression of PPR-coding genes: 32 genes were significantly down-regulated in *P. cordatum* cold-induced temporary cysts, and, in turn, 29 ones became up-regulated during the preparation of cells for excystment (14 DEGs were common for these conditions) [36]. Exposure of *Cladocodium* sp. cells to nano-plastics in the cultivation medium reduced expression of genes coding for PPR domain-containing proteins [77].

12. Perspectives and Emerging Approaches in Dinoflagellate RBP Research

In general, the dinoflagellate RBPs discussed here demonstrate substantial conservation and high similarity to proteins considered canonical in other eukaryotes, suggesting that their core cellular functions have largely been retained; nevertheless, specific lineage-related features can be identified in several cases. Some RBP groups, such as PPR proteins and MEI2-like homologues, show extensive expansions in dinoflagellate genomes. In the case of PPR proteins, this expansion may be primarily related to the extensive RNA editing that occurs in dinoflagellate chloroplasts [23]. At the same time, the expansion of MEI2-like proteins and the suggested existence of putative functional groups may reflect diversification of post-transcriptional regulatory pathways associated with distinct life-cycle stages and adaptive responses to environmental stressors. Notably, genes encoding some RBPs (including PPRs, CSPs, and DEAD-box helicases) are often arranged in tandem arrays [26,32,33].

Despite the increasing amount of sequence-based information, mechanistic characterization of dinoflagellate RBPs remains limited. For most proteins, functional inference relies almost exclusively on homology-based annotation, with few biochemical or experimental studies addressing RNA targets, binding affinities, or physiological roles. Bridging this gap will require a shift from descriptive genomics toward targeted functional analyses.

Promising research directions encompass both relatively straightforward and more technically demanding approaches. These include detailed structural characterization of RBPs, determination of their intracellular localization, targeted analyses of RBP gene expression dynamics across environmental conditions and life-cycle stages, and comparative studies of free-living, symbiotic, and parasitic dinoflagellate lineages to identify evolutionary drivers underlying RBP family expansion.

Another largely unexplored direction concerns the potential involvement of dinoflagellate RBPs in liquid–liquid phase separation (LLPS) during transitions between life-cycle stages, particularly during resting cyst formation. LLPS is a major mechanism underlying the assembly of various membraneless cellular compartments, including nucleoli, Cajal bodies, stress granules (SGs), and processing (P) bodies (reviewed in [256]). The latter two compartments are composed predominantly of mRNA molecules and RBPs and are involved in mRNA sequestration and regulation of mRNA degradation. While SG formation is strictly stress-induced, P-bodies may be constitutively present in the cell [256].

Resting cyst formation in dinoflagellates is frequently triggered by environmental stressors such as nutrient limitation, changes in illumination, or temperature shifts; in many cases, encystment is also preceded by sexual fusion [4,6]. In this context, the increased expression of genes encoding CSD-containing proteins in dinoflagellate resting cysts [34,61] is particularly noteworthy, as it suggests that at least some of these proteins may partici-

pate in the intracellular processes that mediate rearrangements accompanying entry into dormancy. Notably, many proteins implicated in LLPS in other eukaryotes possess intrinsically disordered regions (IDRs) and/or multiple modular interaction domains [257], and some dinoflagellate CSD proteins either possess tandemly repeated domains or glycine-rich IDRs [59]. Together, these observations indicate that elucidating the roles of CSD-containing proteins and possibly other RBP groups during encystment, including their potential contribution to phase separation-mediated organization of RNA-protein assemblies, represents an important avenue for future research.

To understand how RBPs shape the unique biology of dinoflagellates and to uncover the broader principles of RNA metabolism across various eukaryotic lineages, further methodological advances are required. Emerging genetic transformation and functional genomics approaches are gradually making experimental investigation of gene function and regulatory mechanisms in dinoflagellates feasible. Recent studies have established a nuclear gene transformation protocol in the early-branching dinoflagellate *Oxyrrhis marina* [258], and in the coral symbiont *Breviolum minutum* [259], enabling stable transgene expression. In parallel, the dinoflagellate-specific trans-spliced leader, a defining feature of their mRNAs, is exploited to recover and analyze full-length transcripts, particularly from mixed microbial samples [260,261]. In addition, single-cell transcriptomic analysis of early-branching, uncultured dinoflagellates of the genus *Abedinium* has been applied not only to resolve their phylogenetic relationships but also to shed light on the evolutionary history of characteristic cellular features, including proteins involved in chromatin organization [262]. Together with high-throughput proteomics and integrative multi-omics approaches [26,46,263], these methods provide a framework for moving from descriptive genomics toward an understanding of functional role of distinct proteins and protein groups in life-history transitions, stress adaptation, and other cellular processes in dinoflagellates.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms27010462/s1>.

Author Contributions: Conceptualization, M.B. and P.S.; Writing—Original Draft Preparation, M.B. and P.S.; Writing—Review & Editing, M.B., P.S. and S.S., Funding Acquisition, S.S. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the Russian Science Foundation, project # 22-14-00056-II.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

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

References

1. Hackett, J.D.; Anderson, D.M.; Erdner, D.L.; Bhattacharya, D. Dinoflagellates: A Remarkable Evolutionary Experiment. *Am. J. Bot.* **2004**, *91*, 1523–1534. [[CrossRef](#)] [[PubMed](#)]
2. Glibert, P.; Burford, M. Globally Changing Nutrient Loads and Harmful Algal Blooms: Recent Advances, New Paradigms, and Continuing Challenges. *Oceanography* **2017**, *30*, 58–69. [[CrossRef](#)]
3. Saldarriaga, J.F.; Taylor, F.J.R. Dinoflagellata. In *Handbook of the Protists*; Archibald, J.M., Simpson, A.G.B., Slamovits, C.H., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 625–678, ISBN 978-3-319-28147-6.
4. Bravo, I.; Figueroa, R.I. Towards an Ecological Understanding of Dinoflagellate Cyst Functions. *Microorganisms* **2014**, *2*, 11–32. [[CrossRef](#)] [[PubMed](#)]
5. Figueroa, R.I.; Estrada, M.; Garcés, E. Life Histories of Microalgal Species Causing Harmful Blooms: Haploids, Diploids and the Relevance of Benthic Stages. *Harmful Algae* **2018**, *73*, 44–57. [[CrossRef](#)]

6. Lin, S.; Yu, L.; Wu, X.; Li, M.; Zhang, Y.; Luo, H.; Li, H.; Li, T.; Li, L. Active Meiosis during Dinoflagellate Blooms: A ‘Sex for Proliferation’ Hypothesis. *Harmful Algae* **2022**, *118*, 102307. [\[CrossRef\]](#)
7. Rizzo, P.J. Those Amazing Dinoflagellate Chromosomes. *Cell Res.* **2003**, *13*, 215–217. [\[CrossRef\]](#)
8. 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. [\[CrossRef\]](#)
9. Lin, S. Genomic Understanding of Dinoflagellates. *Res. Microbiol.* **2011**, *162*, 551–569. [\[CrossRef\]](#)
10. Lin, S. A Decade of Dinoflagellate Genomics Illuminating an Enigmatic Eukaryote Cell. *BMC Genom.* **2024**, *25*, 932. [\[CrossRef\]](#)
11. 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. [\[CrossRef\]](#)
12. Fukuda, Y.; Suzaki, T. Unusual Features of Dinokaryon, the Enigmatic Nucleus of Dinoflagellates. In *Marine Protists: Diversity and Dynamics*; Ohtsuka, S., Suzaki, T., Horiguchi, T., Suzuki, N., Not, F., Eds.; Springer: Tokyo, Japan, 2015; pp. 23–45, ISBN 978-4-431-55130-0.
13. Roy, S.; Jagus, R.; Morse, D. Translation and Translational Control in Dinoflagellates. *Microorganisms* **2018**, *6*, 30. [\[CrossRef\]](#)
14. 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. [\[CrossRef\]](#)
15. Milos, P.; Morse, D.; Hastings, J.W. Circadian Control over Synthesis of Many *Gonyaulax* Proteins Is at a Translational Level. *Naturwissenschaften* **1990**, *77*, 87–89. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Mittag, M. The Function of Circadian RNA-binding Proteins and Their *Cis*-acting Elements in Microalgae. *Chronobiol. Int.* **2003**, *20*, 529–541. [\[CrossRef\]](#) [\[PubMed\]](#)
17. 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. [\[CrossRef\]](#)
18. Bowazolo, C.; Tse, S.P.K.; Beauchemin, M.; Lo, S.C.-L.; Rivoal, J.; Morse, D. Label-Free MS/MS Analyses of the Dinoflagellate *Lingulodinium* Identifies Rhythmic Proteins Facilitating Adaptation to a Diurnal LD Cycle. *Sci. Total Environ.* **2020**, *704*, 135430. [\[CrossRef\]](#)
19. Liu, B.; Lo, S.C.-L.; Matton, D.P.; Lang, B.F.; Morse, D. Daily Changes in the Phosphoproteome of the Dinoflagellate *Lingulodinium*. *Protist* **2012**, *163*, 746–754. [\[CrossRef\]](#)
20. Roy, S.; Morse, D. Transcription and Maturation of mRNA in Dinoflagellates. *Microorganisms* **2013**, *1*, 71–99. [\[CrossRef\]](#)
21. Zaheri, B.; Morse, D. An Overview of Transcription in Dinoflagellates. *Gene* **2022**, *829*, 146505. [\[CrossRef\]](#) [\[PubMed\]](#)
22. 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. [\[CrossRef\]](#)
23. Mungpakdee, S.; Shinzato, C.; Takeuchi, T.; Kawashima, T.; Koyanagi, R.; Hisata, K.; Tanaka, M.; Goto, H.; Fujie, M.; Lin, S.; et al. Massive Gene Transfer and Extensive RNA Editing of a Symbiotic Dinoflagellate Plastid Genome. *Genome Biol. Evol.* **2014**, *6*, 1408–1422. [\[CrossRef\]](#)
24. Lin, S.; Zhang, H.; Spencer, D.F.; Norman, J.E.; Gray, M.W. Widespread and Extensive Editing of Mitochondrial mRNAs in Dinoflagellates. *J. Mol. Biol.* **2002**, *320*, 727–739. [\[CrossRef\]](#)
25. Shoguchi, E.; Yoshioka, Y.; Shinzato, C.; Arimoto, A.; Bhattacharya, D.; Satoh, N. Correlation between Organelle Genetic Variation and RNA Editing in Dinoflagellates Associated with the Coral *Acropora digitifera*. *Genome Biol. Evol.* **2020**, *12*, 203–209. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Dougan, K.E.; Deng, Z.-L.; Wöhlbrand, L.; Reuse, C.; Bunk, B.; Chen, Y.; Hartlich, J.; Hiller, K.; John, U.; Kalvelage, J.; et al. Multi-Omics Analysis Reveals the Molecular Response to Heat Stress in a “Red Tide” Dinoflagellate. *Genome Biol.* **2023**, *24*, 265. [\[CrossRef\]](#)
27. Chen, Y.; Dougan, K.E.; Bhattacharya, D.; Chan, C.X. Nuclear Genomes of Dinoflagellates Reveal Evolutionarily Conserved Pattern of RNA Editing Relative to Stress Response. *Front. Protistol.* **2024**, *2*, 1320917. [\[CrossRef\]](#)
28. Beauchemin, M.; Morse, D. A Proteomic Portrait of Dinoflagellate Chromatin Reveals Abundant RNA-Binding Proteins. *Chromosoma* **2018**, *127*, 29–43. [\[CrossRef\]](#)
29. Hentze, M.W.; Castello, A.; Schwarzl, T.; Preiss, T. A Brave New World of RNA-Binding Proteins. *Nat. Rev. Mol. Cell Biol.* **2018**, *19*, 327–341. [\[CrossRef\]](#)
30. Lunde, B.M.; Moore, C.; Varani, G. RNA-Binding Proteins: Modular Design for Efficient Function. *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 479–490. [\[CrossRef\]](#) [\[PubMed\]](#)
31. Pali, O.; Safonov, P.; Berdieva, M. MEI2-like Proteins in Dinoflagellates: Diversity and Possible Functions in Life Cycle Transitions. *Protistology* **2024**, *18*, 249–262. [\[CrossRef\]](#)
32. 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. [\[CrossRef\]](#) [\[PubMed\]](#)

33. Stephens, T.G.; González-Pech, R.A.; Cheng, Y.; Mohamed, A.R.; Burt, D.W.; Bhattacharya, D.; Ragan, M.A.; Chan, C.X. Genomes of the Dinoflagellate *Polarella glacialis* Encode Tandemly Repeated Single-Exon Genes with Adaptive Functions. *BMC Biol.* **2020**, *18*, 56. [\[CrossRef\]](#)
34. Deng, Y.; Hu, Z.; Shang, L.; Peng, Q.; Tang, Y.Z. Transcriptomic Analyses of *Scrippsiella trochoidea* Reveals Processes Regulating Encystment and Dormancy in the Life Cycle of a Dinoflagellate, with a Particular Attention to the Role of Absciscic Acid. *Front. Microbiol.* **2017**, *8*, 2450. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Berdieva, M.A.; Pozdnyakov, I.A.; Kalinina, V.O.; Skarlato, S.O. Putative Meiotic Toolkit in the Dinoflagellate *Prorocentrum cordatum*: Additional Evidence for Sexual Process from Transcriptome. *J. Eukaryot. Microbiol.* **2021**, *68*, e12845. [\[CrossRef\]](#)
36. Berdieva, M.; Safonov, P.; Palii, O.; Prilutsky, M.; Matantseva, O.; Skarlato, S. Transcriptomic Analysis of Cold-Induced Temporary Cysts in Marine Dinoflagellate *Prorocentrum cordatum*. *Int. J. Mol. Sci.* **2025**, *26*, 5432. [\[CrossRef\]](#)
37. Gong, S.; Li, G.; Jin, X.; Qiu, D.; Liang, J.; Yu, K.; Tan, Y.; Ma, X.; Xia, X. Photobleaching and Recovery of Symbiodiniaceae *Effrenium voratum* SCS01 Reveals Life Form Transformation under Thermal Stress. *Front. Mar. Sci.* **2021**, *8*, 740416. [\[CrossRef\]](#)
38. Guo, X.; Wang, Z.; Liu, L.; Li, Y. Transcriptome and Metabolome Analyses of Cold and Darkness-Induced Pellicle Cysts of *Scrippsiella trochoidea*. *BMC Genom.* **2021**, *22*, 526. [\[CrossRef\]](#)
39. Luo, H.; Wang, J.; Goes, J.I.; Gomes, H.d.R.; Al-Hashmi, K.; Tobias, C.; Koerting, C.; Lin, S. A Grazing-Driven Positive Nutrient Feedback Loop and Active Sexual Reproduction Underpin Widespread *Noctiluca* Green Tides. *ISME Commun.* **2022**, *2*, 103. [\[CrossRef\]](#) [\[PubMed\]](#)
40. Akbar, M.A.; Mohd Yusof, N.Y.; Usup, G.; Ahmad, A.; Baharum, S.N.; Bunawan, H. Nutrient Deficiencies Impact on the Cellular and Metabolic Responses of Saxitoxin Producing *Alexandrium minutum*: A Transcriptomic Perspective. *Mar. Drugs* **2023**, *21*, 497. [\[CrossRef\]](#)
41. Reddy, R.; Spector, D.; Henning, D.; Liu, M.H.; Busch, H. Isolation and Partial Characterization of Dinoflagellate U1-U6 Small RNAs Homologous to Rat U Small Nuclear RNAs. *J. Biol. Chem.* **1983**, *258*, 13965–13969. [\[CrossRef\]](#) [\[PubMed\]](#)
42. Alverca, E.; Franca, S.; de la Espina, S.M.D. Topology of Splicing and snRNP Biogenesis in Dinoflagellate Nuclei. *Biol. Cell* **2006**, *98*, 709–720. [\[CrossRef\]](#)
43. Lu, Y.; Wohlrab, S.; Glöckner, G.; Guillou, L.; John, U. Genomic Insights into Processes Driving the Infection of *Alexandrium tamarense* by the Parasitoid *Amoebophrya* sp. *Eukaryot. Cell* **2014**, *13*, 1439–1449. [\[CrossRef\]](#)
44. Aranda, M.; Li, Y.; Liew, Y.J.; Baumgarten, S.; Simakov, O.; Wilson, M.C.; Piel, J.; Ashoor, H.; Bougouffa, S.; Bajic, V.B.; et al. Genomes of Coral Dinoflagellate Symbionts Highlight Evolutionary Adaptations Conducive to a Symbiotic Lifestyle. *Sci. Rep.* **2016**, *6*, 39734. [\[CrossRef\]](#)
45. Stephens, T.G.; Ragan, M.A.; Bhattacharya, D.; Chan, C.X. Core Genes in Diverse Dinoflagellate Lineages Include a Wealth of Conserved Dark Genes with Unknown Functions. *Sci. Rep.* **2018**, *8*, 17175. [\[CrossRef\]](#) [\[PubMed\]](#)
46. Fajardo, C.; Amil-Ruiz, F.; Fuentes-Almagro, C.; De Donato, M.; Martinez-Rodriguez, G.; Escobar-Niño, A.; Carrasco, R.; Mancera, J.M.; Fernandez-Acero, F.J. An “Omic” Approach to *Pyrocystis lunula*: New Insights Related with This Bioluminescent Dinoflagellate. *J. Proteom.* **2019**, *209*, 103502. [\[CrossRef\]](#) [\[PubMed\]](#)
47. 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. [\[CrossRef\]](#)
48. Zhang, H.; Campbell, D.A.; Sturm, N.R.; Lin, S. Dinoflagellate Spliced Leader RNA Genes Display a Variety of Sequences and Genomic Arrangements. *Mol. Biol. Evol.* **2009**, *26*, 1757–1771. [\[CrossRef\]](#)
49. Song, Y.; Zaheri, B.; Liu, M.; Sahu, S.K.; Liu, H.; Chen, W.; Song, B.; Morse, D. *Fugacium* Spliced Leader Genes Identified from Stranded RNA-Seq Datasets. *Microorganisms* **2019**, *7*, 171. [\[CrossRef\]](#) [\[PubMed\]](#)
50. 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. [\[CrossRef\]](#)
51. González-Pech, R.A.; Stephens, T.G.; Chen, Y.; Mohamed, A.R.; Cheng, Y.; Shah, S.; Dougan, K.E.; Fortuin, M.D.A.; Lagorce, R.; Burt, D.W.; et al. Comparison of 15 Dinoflagellate Genomes Reveals Extensive Sequence and Structural Divergence in Family Symbiodiniaceae and Genus *Symbiodinium*. *BMC Biol.* **2021**, *19*, 73. [\[CrossRef\]](#)
52. Johnson, J.G.; Morey, J.S.; Neely, M.G.; Ryan, J.C.; Van Dolah, F.M. Transcriptome Remodeling Associated with Chronological Aging in the Dinoflagellate, *Karenia brevis*. *Mar. Genom.* **2012**, *5*, 15–25. [\[CrossRef\]](#)
53. Roy, S.; Morse, D. The Dinoflagellate *Lingulodinium* Has Predicted Casein Kinase 2 Sites in Many RNA Binding Proteins. *Protist* **2014**, *165*, 330–342. [\[CrossRef\]](#)
54. Su, Y.; Zhang, K.; Zhou, Z.; Wang, J.; Yang, X.; Tang, J.; Li, H.; Lin, S. Microplastic Exposure Represses the Growth of Endosymbiotic Dinoflagellate *Cladocopium goreau* in Culture through Affecting Its Apoptosis and Metabolism. *Chemosphere* **2020**, *244*, 125485. [\[CrossRef\]](#)

55. Le Gac, M.; Mary, L.; Metegnier, G.; Quéré, J.; Siano, R.; Rodríguez, F.; Destombe, C.; Sourisseau, M. Strong Population Genomic Structure of the Toxic Dinoflagellate *Alexandrium minutum* Inferred from Meta-Transcriptome Samples. *Environ. Microbiol.* **2022**, *24*, 5966–5983. [\[CrossRef\]](#) [\[PubMed\]](#)
56. Wang, D.-Z.; Zhang, Y.-J.; Zhang, S.-F.; Lin, L.; Hong, H.-S. Quantitative Proteomic Analysis of Cell Cycle of the Dinoflagellate *Prorocentrum donghaiense* (Dinophyceae). *PLoS ONE* **2013**, *8*, e63659. [\[CrossRef\]](#) [\[PubMed\]](#)
57. 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. [\[CrossRef\]](#) [\[PubMed\]](#)
58. 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. [\[CrossRef\]](#)
59. Beauchemin, M.; Roy, S.; Pelletier, S.; Averback, A.; Lanthier, F.; Morse, D. Characterization of Two Dinoflagellate Cold Shock Domain Proteins. *mSphere* **2016**, *1*, e00034–e15. [\[CrossRef\]](#)
60. Roy, S.; Letourneau, L.; Morse, D. Cold-Induced Cysts of the Photosynthetic Dinoflagellate *Lingulodinium polyedrum* Have an Arrested Circadian Bioluminescence Rhythm and Lower Levels of Protein Phosphorylation. *Plant Physiol.* **2014**, *164*, 966–977. [\[CrossRef\]](#)
61. Deng, Y.; Hu, Z.; Chai, Z.; Tang, Y.Z. Cloning and Partial Characterization of a Cold Shock Domain-Containing Protein Gene from the Dinoflagellate *Scrippsiella trochoidea*. *J. Eukaryot. Microbiol.* **2019**, *66*, 393–403. [\[CrossRef\]](#)
62. Liu, Y.; Sui, Z.; Zhang, S.; Ma, J.; Zhong, J. Isolation and Expression Analysis of Growth-Related Genes at Different Growth Stages of Dinoflagellate *Alexandrium pacificum*. *J. Ocean Univ. China* **2019**, *18*, 710–718. [\[CrossRef\]](#)
63. Wang, H.; Kim, H.; Ki, J.-S. Transcriptomic Identification and Expression Analysis of Cold Shock Domain Protein (CSP) Genes in the Marine Dinoflagellate *Prorocentrum minimum*. *J. Appl. Phycol.* **2021**, *33*, 843–854. [\[CrossRef\]](#)
64. Zaheri, B.; Morse, D. Assessing Nucleic Acid Binding Activity of Four Dinoflagellate Cold Shock Domain Proteins from *Symbiodinium kawagutii* and *Lingulodinium polyedra*. *BMC Mol. Cell Biol.* **2021**, *22*, 27. [\[CrossRef\]](#)
65. 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. [\[CrossRef\]](#)
66. Kalvelage, J.; Wöhlbrand, L.; Schoon, R.-A.; Zink, F.-M.; Correll, C.; Senkler, J.; Eubel, H.; Hoppenrath, M.; Erhard, R.; Braun, H.-P.; et al. The Enigmatic Nucleus of the Marine Dinoflagellate *Prorocentrum cordatum*. *mSphere* **2023**, *8*, e00038–e23. [\[CrossRef\]](#)
67. 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. [\[CrossRef\]](#)
68. Morey, J.S.; Van Dolah, F.M. Global Analysis of mRNA Half-Lives and *de Novo* Transcription in a Dinoflagellate, *Karenia brevis*. *PLoS ONE* **2013**, *8*, e66347. [\[CrossRef\]](#)
69. Schallenberg-Rüdinger, M.; Lenz, H.; Polsakiewicz, M.; Gott, J.M.; Knoop, V. A Survey of PPR Proteins Identifies DYW Domains like Those of Land Plant RNA Editing Factors in Diverse Eukaryotes. *RNA Biol.* **2013**, *10*, 1549–1556. [\[CrossRef\]](#)
70. Shoguchi, E.; Beedessee, G.; Tada, I.; Hisata, K.; Kawashima, T.; Takeuchi, T.; Arakaki, N.; Fujie, M.; Koyanagi, R.; Roy, M.C.; et al. Two Divergent *Symbiodinium* Genomes Reveal Conservation of a Gene Cluster for Sunscreen Biosynthesis and Recently Lost Genes. *BMC Genom.* **2018**, *19*, 458. [\[CrossRef\]](#) [\[PubMed\]](#)
71. Lin, S.; Yu, L.; Zhang, H. Transcriptomic Responses to Thermal Stress and Varied Phosphorus Conditions in *Fugacium kawagutii*. *Microorganisms* **2019**, *7*, 96. [\[CrossRef\]](#)
72. Cheng, S.; Gutmann, B.; Zhong, X.; Ye, Y.; Fisher, M.F.; Bai, F.; Castleden, I.; Song, Y.; Song, B.; Huang, J.; et al. Redefining the Structural Motifs That Determine RNA Binding and RNA Editing by Pentatricopeptide Repeat Proteins in Land Plants. *Plant J.* **2016**, *85*, 532–547. [\[CrossRef\]](#)
73. Liew, Y.J.; Li, Y.; Baumgarten, S.; Voolstra, C.R.; Aranda, M. Condition-Specific RNA Editing in the Coral Symbiont *Symbiodinium microadriaticum*. *PLoS Genet.* **2017**, *13*, e1006619. [\[CrossRef\]](#) [\[PubMed\]](#)
74. Beedessee, G.; Hisata, K.; Roy, M.C.; Van Dolah, F.M.; Satoh, N.; Shoguchi, E. Diversified Secondary Metabolite Biosynthesis Gene Repertoire Revealed in Symbiotic Dinoflagellates. *Sci. Rep.* **2019**, *9*, 1204. [\[CrossRef\]](#) [\[PubMed\]](#)
75. Beedessee, G.; Kubota, T.; Arimoto, A.; Nishitsuji, K.; Waller, R.F.; Hisata, K.; Yamasaki, S.; Satoh, N.; Kobayashi, J.; Shoguchi, E. Integrated Omics Unveil the Secondary Metabolic Landscape of a Basal Dinoflagellate. *BMC Biol.* **2020**, *18*, 139. [\[CrossRef\]](#) [\[PubMed\]](#)
76. Maor-Landaw, K.; van Oppen, M.J.H.; McFadden, G.I. Symbiotic Lifestyle Triggers Drastic Changes in the Gene Expression of the Algal Endosymbiont *Breviolum minutum* (Symbiodiniaceae). *Ecol. Evol.* **2020**, *10*, 451–466. [\[CrossRef\]](#)
77. Ripken, C.; Khalturin, K.; Shoguchi, E. Response of Coral Reef Dinoflagellates to Nanoplastics under Experimental Conditions Suggests Downregulation of Cellular Metabolism. *Microorganisms* **2020**, *8*, 1759. [\[CrossRef\]](#)

78. Safonov, P.; Berdieva, M. Diversity and Evolution of P-Class Pentatricopeptide Repeat Proteins in Dinoflagellates. *Protistology* **2025**, *19*, 212–228.
79. Dreyfuss, G.; Kim, V.N.; Kataoka, N. Messenger-RNA-Binding Proteins and the Messages They Carry. *Nat. Rev. Mol. Cell Biol.* **2002**, *3*, 195–205. [[CrossRef](#)]
80. Califice, S.; Baurain, D.; Hanikenne, M.; Motte, P. A Single Ancient Origin for Prototypical Serine/Arginine-Rich Splicing Factors. *Plant Physiol.* **2012**, *158*, 546–560. [[CrossRef](#)]
81. Prall, W.; Sharma, B.; Gregory, B.D. Transcription Is Just the Beginning of Gene Expression Regulation: The Functional Significance of RNA-Binding Proteins to Post-Transcriptional Processes in Plants. *Plant Cell Physiol.* **2019**, *60*, 1939–1952. [[CrossRef](#)]
82. Corley, M.; Burns, M.C.; Yeo, G.W. How RNA-Binding Proteins Interact with RNA: Molecules and Mechanisms. *Mol. Cell* **2020**, *78*, 9–29. [[CrossRef](#)]
83. Maris, C.; Dominguez, C.; Allain, F.H.-T. The RNA Recognition Motif, a Plastic RNA-binding Platform to Regulate Post-transcriptional Gene Expression. *FEBS J.* **2005**, *272*, 2118–2131. [[CrossRef](#)] [[PubMed](#)]
84. Cléry, A.; Blatter, M.; Allain, F.H.-T. RNA Recognition Motifs: Boring? Not Quite. *Curr. Opin. Struct. Biol.* **2008**, *18*, 290–298. [[CrossRef](#)] [[PubMed](#)]
85. Adam, S.A.; Nakagawa, T.; Swanson, M.S.; Woodruff, T.K.; Dreyfuss, G. mRNA Polyadenylate-Binding Protein: Gene Isolation and Sequencing and Identification of a Ribonucleoprotein Consensus Sequence. *Mol. Cell. Biol.* **1986**, *6*, 2932–2943. [[CrossRef](#)] [[PubMed](#)]
86. Dreyfuss, G.; Swanson, M.S.; Piñol-Roma, S. Heterogeneous Nuclear Ribonucleoprotein Particles and the Pathway of mRNA Formation. *Trends Biochem. Sci.* **1988**, *13*, 86–91. [[CrossRef](#)]
87. Nagai, K.; Oubridge, C.; Jessen, T.H.; Li, J.; Evans, P.R. Crystal Structure of the RNA-Binding Domain of the U1 Small Nuclear Ribonucleoprotein A. *Nature* **1990**, *348*, 515–520. [[CrossRef](#)]
88. Oubridge, C.; Ito, N.; Evans, P.R.; Teo, C.-H.; Nagai, K. Crystal Structure at 1.92 Å Resolution of the RNA-Binding Domain of the U1A Spliceosomal Protein Complexed with an RNA Hairpin. *Nature* **1994**, *372*, 432–438. [[CrossRef](#)]
89. Watanabe, Y.; Yamamoto, M.S. *pombe* Mei2+ Encodes an RNA-Binding Protein Essential for Premeiotic DNA Synthesis and Meiosis I, Which Cooperates with a Novel RNA Species meiRNA. *Cell* **1994**, *78*, 487–498. [[CrossRef](#)]
90. Mukherjee, K.; Futcher, B.; Leatherwood, J. *Mmi1* and *Rep2* mRNAs Are Novel RNA Targets of the Mei2 RNA-Binding Protein during Early Meiosis in *Schizosaccharomyces pombe*. *Open Biol.* **2018**, *8*, 180110. [[CrossRef](#)]
91. Hirayama, T.; Ishida, C.; Kuromori, T.; Obata, S.; Shimoda, C.; Yamamoto, M.; Shinozaki, K.; Ohto, C. Functional Cloning of a cDNA Encoding Mei2-like Protein from *Arabidopsis thaliana* Using a Fission Yeast Pheromone Receptor Deficient Mutant. *FEBS Lett.* **1997**, *413*, 16–20. [[CrossRef](#)]
92. Veit, B.; Briggs, S.P.; Schmidt, R.J.; Yanofsky, M.F.; Hake, S. Regulation of Leaf Initiation by the *Terminal Ear 1* Gene of Maize. *Nature* **1998**, *393*, 166–168. [[CrossRef](#)]
93. Anderson, G.H.; Alvarez, N.D.G.; Gilman, C.; Jeffares, D.C.; Trainor, V.C.W.; Hanson, M.R.; Veit, B. Diversification of Genes Encoding Mei2-like RNA Binding Proteins in Plants. *Plant Mol. Biol.* **2004**, *54*, 653–670. [[CrossRef](#)]
94. Jeffares, D.C.; Phillips, M.J.; Moore, S.; Veit, B. A Description of the Mei2-like Protein Family; Structure, Phylogenetic Distribution and Biological Context. *Dev. Genes Evol.* **2004**, *214*, 149–158. [[CrossRef](#)] [[PubMed](#)]
95. Ji, N.; Huang, J.; Zhang, Z.; Zhou, L.; Shen, X.; Lin, S. Identification and Expression Analysis of Meiosis-Related Genes in the Harmful Alga *Heterosigma akashiwo* (Raphidophyceae). *Harmful Algae* **2020**, *92*, 101736. [[CrossRef](#)]
96. Cavalier-Smith, T.; Chao, E.E. Protalveolate Phylogeny and Systematics and the Origins of Sporozoa and Dinoflagellates (Phylum Myzozoa Nom. Nov.). *Eur. J. Protistol.* **2004**, *40*, 185–212. [[CrossRef](#)]
97. Wang, Z.; Christman, M.F. Replication-Related Activities Establish Cohesion between Sister Chromatids. *Cell Biochem. Biophys.* **2001**, *35*, 289–301. [[CrossRef](#)]
98. Lee, J.Y.; Orr-Weaver, T.L. The Molecular Basis of Sister-Chromatid Cohesion. *Annu. Rev. Cell Dev. Biol.* **2001**, *17*, 753–777. [[CrossRef](#)] [[PubMed](#)]
99. Schurko, A.M.; Logsdon, J.M. Using a Meiosis Detection Toolkit to Investigate Ancient Asexual “Scandals” and the Evolution of Sex. *BioEssays* **2008**, *30*, 579–589. [[CrossRef](#)]
100. Kaur, J.; Sebastian, J.; Siddiqi, I. The *Arabidopsis*-Mei2-like Genes Play a Role in Meiosis and Vegetative Growth in *Arabidopsis*. *Plant Cell* **2006**, *18*, 545–559. [[CrossRef](#)] [[PubMed](#)]
101. Thibault, P.A.; Ganesan, A.; Kalyaanamoorthy, S.; Clarke, J.-P.W.E.; Salapa, H.E.; Levin, M.C. hnRNP A/B Proteins: An Encyclopedic Assessment of Their Roles in Homeostasis and Disease. *Biology* **2021**, *10*, 712. [[CrossRef](#)]
102. Valadkhan, S.; Jaladat, Y. The Spliceosomal Proteome: At the Heart of the Largest Cellular Ribonucleoprotein Machine. *Proteomics* **2010**, *10*, 4128–4141. [[CrossRef](#)]
103. Will, C.L.; Lührmann, R. Spliceosome Structure and Function. *Cold Spring Harb. Perspect. Biol.* **2011**, *3*, a003707. [[CrossRef](#)] [[PubMed](#)]

104. Hall, K.B.; Stump, W.T. Interaction of N-Terminal Domain of U1A Protein with an RNA Stem/Loop. *Nucleic Acids Res.* **1992**, *20*, 4283–4290. [[CrossRef](#)] [[PubMed](#)]
105. Conrad, H.; Lührmann, R.; Will, C. The U11/U12 snRNP 65K Protein Acts as a Molecular Bridge, Binding the U12 snRNA and U11-59K Protein. *EMBO J.* **2005**, *24*, 3057–3069. [[CrossRef](#)]
106. Williams, S.G.; Hall, K.B. Human U2B'' Protein Binding to snRNA Stemloops. *Biophys. Chem.* **2011**, *159*, 82–89. [[CrossRef](#)]
107. Ma, X.; He, F. Advances in the Study of SR Protein Family. *Genom. Proteom. Bioinform.* **2003**, *1*, 2–8. [[CrossRef](#)]
108. Busch, A.; Hertel, K.J. Evolution of SR Protein and hnRNP Splicing Regulatory Factors. *Wiley Interdiscip. Rev. RNA* **2012**, *3*, 1–12. [[CrossRef](#)]
109. Katoh, K.; Standley, D.M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [[CrossRef](#)]
110. Okonechnikov, K.; Golosova, O.; Fursov, M. Unipro UGENE: A Unified Bioinformatics Toolkit. *Bioinformatics* **2012**, *28*, 1166–1167. [[CrossRef](#)]
111. Laursen, B.S.; Sørensen, H.P.; Mortensen, K.K.; Sperling-Petersen, H.U. Initiation of Protein Synthesis in Bacteria. *Microbiol. Mol. Biol. Rev.* **2005**, *69*, 101–123. [[CrossRef](#)]
112. Jackson, R.J.; Hellen, C.U.T.; Pestova, T.V. The Mechanism of Eukaryotic Translation Initiation and Principles of Its Regulation. *Nat. Rev. Mol. Cell Biol.* **2010**, *11*, 113–127. [[CrossRef](#)]
113. Trachsel, H.; Erni, B.; Schreier, M.H.; Staehelin, T. Initiation of Mammalian Protein Synthesis: II. The Assembly of the Initiation Complex with Purified Initiation Factors. *J. Mol. Biol.* **1977**, *116*, 755–767. [[CrossRef](#)]
114. Benne, R.; Hershey, J. Mechanism of Action of Protein-Synthesis Initiation-Factors from Rabbit Reticulocytes. *J. Biol. Chem.* **1978**, *253*, 3078–3087. [[CrossRef](#)]
115. Méthot, N.; Pickett, G.; Keene, J.D.; Sonenberg, N. In Vitro RNA Selection Identifies RNA Ligands That Specifically Bind to Eukaryotic Translation Initiation Factor 4B: The Role of the RNA Remotif. *RNA* **1996**, *2*, 38–50.
116. Fleming, K.; Ghuman, J.; Yuan, X.; Simpson, P.; Szendrői, A.; Matthews, S.; Curry, S. Solution Structure and RNA Interactions of the RNA Recognition Motif from Eukaryotic Translation Initiation Factor 4B. *Biochemistry* **2003**, *42*, 8966–8975. [[CrossRef](#)]
117. Valášek, 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](#)]
118. Valášek, L.S.; Zeman, J.; Wagner, S.; Beznosková, P.; Pavlíková, Z.; Mohammad, M.P.; Hronová, V.; Herrmannová, A.; Hashem, Y.; Gunišová, S. Embraced by eIF3: Structural and Functional Insights into the Roles of eIF3 across the Translation Cycle. *Nucleic Acids Res.* **2017**, *45*, 10948–10968. [[CrossRef](#)] [[PubMed](#)]
119. Méthot, N.; Rom, E.; Olsen, H.; Sonenberg, N. The Human Homologue of the Yeast Pti1 Protein Is an Integral Part of the Eukaryotic Initiation Factor 3 Complex and Interacts with P170. *J. Biol. Chem.* **1997**, *272*, 1110–1116. [[CrossRef](#)] [[PubMed](#)]
120. Block, K.L.; Vornlocher, H.-P.; Hershey, J.W.B. Characterization of cDNAs Encoding the P44 and P35 Subunits of Human Translation Initiation Factor eIF3. *J. Biol. Chem.* **1998**, *273*, 31901–31908. [[CrossRef](#)]
121. Valášek, L.; Phan, L.; Schoenfeld, L.W.; Valášková, V.; Hinnebusch, A.G. Related eIF3 Subunits TIF32 and HCR1 Interact with an RNA Recognition Motif in PRT1 Required for eIF3 Integrity and Ribosome Binding. *EMBO J.* **2001**, *20*, 891–904. [[CrossRef](#)]
122. ElAntak, L.; Tzakos, A.G.; Locker, N.; Lukavsky, P.J. Structure of eIF3b RNA Recognition Motif and Its Interaction with eIF3j: Structural Insights into the Recruitment of eIF3b to the 40 S Ribosomal Subunit. *J. Biol. Chem.* **2007**, *282*, 8165–8174. [[CrossRef](#)]
123. ElAntak, L.; Wagner, S.; Herrmannová, A.; Karásková, M.; Rutkai, E.; Lukavsky, P.J.; Valášek, L. The Indispensable N-Terminal Half of eIF3j/HCR1 Cooperates with Its Structurally Conserved Binding Partner eIF3b/PRT1-RRM and with eIF1A in Stringent AUG Selection. *J. Mol. Biol.* **2010**, *396*, 1097–1116. [[CrossRef](#)]
124. Cuchalová, L.; Kouba, T.; Herrmannová, A.; Dányi, I.; Chiu, W.; Valášek, L. The RNA Recognition Motif of Eukaryotic Translation Initiation Factor 3g (eIF3g) Is Required for Resumption of Scanning of Posttermination Ribosomes for Reinitiation on GCN4 and Together with eIF3i Stimulates Linear Scanning. *Mol. Cell. Biol.* **2010**, *30*, 4671–4686. [[CrossRef](#)]
125. Simonetti, A.; Brito Querido, J.; Myasnikov, A.G.; Mancera-Martinez, E.; Renaud, A.; Kuhn, L.; Hashem, Y. eIF3 Peripheral Subunits Rearrangement after mRNA Binding and Start-Codon Recognition. *Mol. Cell* **2016**, *63*, 206–217. [[CrossRef](#)]
126. Brito Querido, J.; Sokabe, M.; Kraatz, S.; Gordiyenko, Y.; Skehel, J.M.; Fraser, C.S.; Ramakrishnan, V. Structure of a Human 48S Translational Initiation Complex. *Science* **2020**, *369*, 1220–1227. [[CrossRef](#)]
127. Mangus, D.A.; Evans, M.C.; Jacobson, A. Poly(A)-Binding Proteins: Multifunctional Scaffolds for the Post-Transcriptional Control of Gene Expression. *Genome Biol.* **2003**, *4*, 223. [[CrossRef](#)] [[PubMed](#)]
128. Eliseeva, I.A.; Lyabin, D.N.; Ovchinnikov, L.P. Poly(A)-Binding Proteins: Structure, Domain Organization, and Activity Regulation. *Biochemistry* **2013**, *78*, 1377–1391. [[CrossRef](#)]
129. da Costa Lima, T.; Moura, D.; Reis, C.; Vasconcelos, J.; Ellis, L.; Carrington, M.; Figueiredo, R.; de Melo Neto, O. Functional Characterization of Three Leishmania Poly(A) Binding Protein Homologues with Distinct Binding Properties to RNA and Protein Partners. *Eukaryot. Cell* **2010**, *9*, 1484–1494. [[CrossRef](#)] [[PubMed](#)]

130. Minns, A.M.; Hart, K.J.; Subramanian, S.; Hafenstein, S.; Lindner, S.E. Nuclear, Cytosolic, and Surface-Localized Poly(A)-Binding Proteins of *Plasmodium yoelii*. *mSphere* **2018**, *3*, e00435-17. [[CrossRef](#)] [[PubMed](#)]
131. Bowazolo, C.; Morse, D. Insights into Daily Metabolic Changes of the Dinoflagellate *Lingulodinium* from Ribosome Profiling. *Cell Cycle* **2023**, *22*, 1343–1352. [[CrossRef](#)]
132. Sauer, E. Structure and RNA-Binding Properties of the Bacterial LSm Protein Hfq. *RNA Biol.* **2013**, *10*, 610–618. [[CrossRef](#)]
133. Wilusz, C.J.; Wilusz, J. Lsm Proteins and Hfq. *RNA Biol.* **2013**, *10*, 592–601. [[CrossRef](#)]
134. Kambach, C.; Walke, S.; Young, R.; Avis, J.M.; de la Fortelle, E.; Raker, V.A.; Lührmann, R.; Li, J.; Nagai, K. Crystal Structures of Two SM Protein Complexes and Their Implications for the Assembly of the Spliceosomal snRNPs. *Cell* **1999**, *96*, 375–387. [[CrossRef](#)]
135. Walke, S.; Bragado-Nilsson, E.; Séraphin, B.; Nagai, K. Stoichiometry of the Sm Proteins in Yeast Spliceosomal snRNPs Supports the Heptamer Ring Model of the Core Domain. *J. Mol. Biol.* **2001**, *308*, 49–58. [[CrossRef](#)]
136. Khusial, P.; Plaag, R.; Zieve, G.W. LSm Proteins Form Heptameric Rings That Bind to RNA via Repeating Motifs. *Trends Biochem. Sci.* **2005**, *30*, 522–528. [[CrossRef](#)]
137. Weichenrieder, O. RNA Binding by Hfq and Ring-Forming (L)Sm Proteins. *RNA Biol.* **2014**, *11*, 537–549. [[CrossRef](#)] [[PubMed](#)]
138. Hastings, K.E.M. SL Trans-Splicing: Easy Come or Easy Go? *Trends Genet.* **2005**, *21*, 240–247. [[CrossRef](#)]
139. Hermann, H.; Fabrizio, P.; Raker, V.A.; Foulaki, K.; Hornig, H.; Brahms, H.; Lührmann, R. snRNP Sm Proteins Share Two Evolutionarily Conserved Sequence Motifs Which Are Involved in Sm Protein-protein Interactions. *EMBO J.* **1995**, *14*, 2076–2088. [[CrossRef](#)] [[PubMed](#)]
140. Séraphin, B. Sm and Sm-like Proteins Belong to a Large Family: Identification of Proteins of the U6 as Well as the U1, U2, U4 and U5 snRNPs. *EMBO J.* **1995**, *14*, 2089–2098. [[CrossRef](#)]
141. Tharun, S.; Muhlrads, D.; Chowdhury, A.; Parker, R. Mutations in the *Saccharomyces cerevisiae* LSM1 Gene That Affect mRNA Decapping and 3' End Protection. *Genetics* **2005**, *170*, 33–46. [[CrossRef](#)] [[PubMed](#)]
142. Lerner, M.R.; Steitz, J.A. Antibodies to Small Nuclear RNAs Complexed with Proteins Are Produced by Patients with Systemic Lupus Erythematosus. *Proc. Natl. Acad. Sci. USA* **1979**, *76*, 5495–5499. [[CrossRef](#)]
143. Brahms, H.; Meheus, L.; De Brabandere, V.; Fischer, U.; Lührmann, R. Symmetrical Dimethylation of Arginine Residues in Spliceosomal Sm Protein B/B' and the Sm-like Protein LSM4, and Their Interaction with the SMN Protein. *RNA* **2001**, *7*, 1531–1542. [[CrossRef](#)]
144. Boisvert, F.-M.; Côté, J.; Boulanger, M.-C.; Cléroux, P.; Bachand, F.; Autexier, C.; Richard, S. Symmetrical Dimethylarginine Methylation Is Required for the Localization of SMN in Cajal Bodies and Pre-mRNA Splicing. *J. Cell Biol.* **2002**, *159*, 957–969. [[CrossRef](#)]
145. Gao, G.; Dhar, S.; Bedford, M.T. PRMT5 Regulates IRES-Dependent Translation via Methylation of hnRNP A1. *Nucleic Acids Res.* **2017**, *45*, 4359–4369. [[CrossRef](#)]
146. Valverde, R.; Edwards, L.; Regan, L. Structure and Function of KH Domains. *FEBS J.* **2008**, *275*, 2712–2726. [[CrossRef](#)]
147. Nicastro, G.; Taylor, I.A.; Ramos, A. KH–RNA Interactions: Back in the Groove. *Curr. Opin. Struct. Biol.* **2015**, *30*, 63–70. [[CrossRef](#)] [[PubMed](#)]
148. Hasan, M.K.; Brady, J.L. Nucleic Acid-Binding KH Domain Proteins Influence a Spectrum of Biological Pathways Including as Part of Membrane-Localized Complexes. *J. Struct. Biol. X* **2024**, *10*, 100106. [[CrossRef](#)] [[PubMed](#)]
149. Laity, J.H.; Lee, B.M.; Wright, P.E. Zinc Finger Proteins: New Insights into Structural and Functional Diversity. *Curr. Opin. Struct. Biol.* **2001**, *11*, 39–46. [[CrossRef](#)] [[PubMed](#)]
150. Isernia, C.; Malgieri, G.; Russo, L.; D'Abrosca, G.; Baglivo, I.; Pedone, P.V.; Fattorusso, R. Zinc Fingers. In *Transition Metals and Sulfur—A Strong Relationship for Life*; Sosa Torres, M., Kroneck, P., Eds.; Metal Ions in Life Sciences; De Gruyter: Berlin, Germany, 2020; Volume 20, pp. 415–436, ISBN 978-3-11-058975-7.
151. Janczarek, M. The Ros/MucR Zinc-Finger Protein Family in Bacteria: Structure and Functions. *Int. J. Mol. Sci.* **2022**, *23*, 15536. [[CrossRef](#)]
152. Cléry, A.; Allain, F. From Structure to Function of RNA Binding Domains. In *RNA Binding Proteins*; Lorković, Z., Ed.; CRC Press: Boca Raton, FL, USA, 2012; pp. 137–158, ISBN 1-4987-1336-X.
153. Hall, T.M.T. Multiple Modes of RNA Recognition by Zinc Finger Proteins. *Curr. Opin. Struct. Biol.* **2005**, *15*, 367–373. [[CrossRef](#)]
154. Fedotova, A.A.; Bonchuk, A.N.; Mogila, V.A.; Georgiev, P.G. C2H2 Zinc Finger Proteins: The Largest but Poorly Explored Family of Higher Eukaryotic Transcription Factors. *Acta Naturae* **2017**, *9*, 47–58. [[CrossRef](#)]
155. Song, J.; Nabeel-Shah, S.; Pu, S.; Lee, H.; Braunschweig, U.; Ni, Z.; Ahmed, N.; Marcon, E.; Zhong, G.; Ray, D.; et al. Regulation of Alternative Polyadenylation by the C2H2-Zinc-Finger Protein Sp1. *Mol. Cell* **2022**, *82*, 3135–3150.e9. [[CrossRef](#)]
156. Nabeel-Shah, S.; Pu, S.; Burns, J.D.; Braunschweig, U.; Ahmed, N.; Burke, G.L.; Lee, H.; Radovani, E.; Zhong, G.; Tang, H.; et al. C2H2-Zinc-Finger Transcription Factors Bind RNA and Function in Diverse Post-Transcriptional Regulatory Processes. *Mol. Cell* **2024**, *84*, 3810–3825.e10. [[CrossRef](#)]

157. Summers, M.F. Zinc Finger Motif for Single-Stranded Nucleic Acids? Investigations by Nuclear Magnetic Resonance. *J. Cell. Biochem.* **1991**, *45*, 41–48. [[CrossRef](#)] [[PubMed](#)]
158. Loughlin, F.E.; Mansfield, R.E.; Vaz, P.M.; McGrath, A.P.; Setiyaputra, S.; Gamsjaeger, R.; Chen, E.S.; Morris, B.J.; Guss, J.M.; Mackay, J.P. The Zinc Fingers of the SR-like Protein ZRANB2 Are Single-Stranded RNA-Binding Domains That Recognize 5' Splice Site-like Sequences. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 5581–5586. [[CrossRef](#)]
159. Nguyen, C.D.; Mansfield, R.E.; Leung, W.; Vaz, P.M.; Loughlin, F.E.; Grant, R.P.; Mackay, J.P. Characterization of a Family of RanBP2-Type Zinc Fingers That Can Recognize Single-Stranded RNA. *J. Mol. Biol.* **2011**, *407*, 273–283. [[CrossRef](#)]
160. Fu, M.; Blackshear, P.J. RNA-Binding Proteins in Immune Regulation: A Focus on CCCH Zinc Finger Proteins. *Nat. Rev. Immunol.* **2017**, *17*, 130–143. [[CrossRef](#)]
161. Maeda, K.; Akira, S. Regulation of mRNA Stability by CCCH-Type Zinc-Finger Proteins in Immune Cells. *Int. Immunol.* **2017**, *29*, 149–155. [[CrossRef](#)] [[PubMed](#)]
162. Wang, Y.; Yu, Y.; Pang, Y.; Yu, H.; Zhang, W.; Zhao, X.; Yu, J. The Distinct Roles of Zinc Finger CCHC-Type (ZCCHC) Superfamily Proteins in the Regulation of RNA Metabolism. *RNA Biol.* **2021**, *18*, 2107–2126. [[CrossRef](#)] [[PubMed](#)]
163. Lu, D.; Alexandra Searles, M.; Klug, A. Crystal Structure of a Zinc-Finger–RNA Complex Reveals Two Modes of Molecular Recognition. *Nature* **2003**, *426*, 96–100. [[CrossRef](#)]
164. Hudson, B.P.; Martinez-Yamout, M.A.; Dyson, H.J.; Wright, P.E. Recognition of the mRNA AU-Rich Element by the Zinc Finger Domain of TIS11d. *Nat. Struct. Mol. Biol.* **2004**, *11*, 257–264. [[CrossRef](#)]
165. Font, J.; Mackay, J.P. Beyond DNA: Zinc Finger Domains as RNA-Binding Modules. In *Engineered Zinc Finger Proteins: Methods and Protocols*; Methods in Molecular Biology; Mackay, J.P., Segal, D.J., Eds.; Methods in Molecular Biology; Humana Press: Totowa, NJ, USA, 2010; Volume 649, pp. 479–491, ISBN 978-1-60761-753-2.
166. Najdová, V.; Stairs, C.W.; Vinopalová, M.; Voleman, L.; Doležal, P. The Evolution of the Puf Superfamily of Proteins across the Tree of Eukaryotes. *BMC Biol.* **2020**, *18*, 77. [[CrossRef](#)]
167. Wickens, M.; Bernstein, D.S.; Kimble, J.; Parker, R. A PUF Family Portrait: 3'UTR Regulation as a Way of Life. *Trends Genet.* **2002**, *18*, 150–157. [[CrossRef](#)]
168. Wang, M.; Ogé, L.; Perez-Garcia, M.D.; Hamama, L.; Sakr, S. The PUF Protein Family: Overview on PUF RNA Targets, Biological Functions, and Post Transcriptional Regulation. *Int. J. Mol. Sci.* **2018**, *19*, 410. [[CrossRef](#)]
169. Zhang, B.; Gallegos, M.; Puoti, A.; Durkin, E.; Fields, S.; Kimble, J.; Wickens, M.P. A Conserved RNA-Binding Protein That Regulates Sexual Fates in the *C. Elegans* Hermaphrodite Germ Line. *Nature* **1997**, *390*, 477–484. [[CrossRef](#)]
170. Wang, X.; Zamore, P.D.; Tanaka Hall, T.M. Crystal Structure of a Pumilio Homology Domain. *Mol. Cell* **2001**, *7*, 855–865. [[CrossRef](#)] [[PubMed](#)]
171. Wang, X.; McLachlan, J.; Zamore, P.D.; Hall, T.M.T. Modular Recognition of RNA by a Human Pumilio-Homology Domain. *Cell* **2002**, *110*, 501–512. [[CrossRef](#)]
172. Murzin, A.G. OB(Oligonucleotide/Oligosaccharide Binding)-fold: Common Structural and Functional Solution for Non-homologous Sequences. *EMBO J.* **1993**, *12*, 861–867. [[CrossRef](#)]
173. Agrawal, V.; Kishan, K.V.R. OB-Fold: Growing Bigger with Functional Consistency. *Curr. Protein Pept. Sci.* **2003**, *4*, 195–206. [[CrossRef](#)]
174. Arcus, V. OB-Fold Domains: A Snapshot of the Evolution of Sequence, Structure and Function. *Curr. Opin. Struct. Biol.* **2002**, *12*, 794–801. [[CrossRef](#)] [[PubMed](#)]
175. Mihailovich, M.; Militti, C.; Gabaldón, T.; Gebauer, F. Eukaryotic Cold Shock Domain Proteins: Highly Versatile Regulators of Gene Expression. *BioEssays* **2010**, *32*, 109–118. [[CrossRef](#)] [[PubMed](#)]
176. Bianco, P.R. OB-Fold Families of Genome Guardians: A Universal Theme Constructed from the Small β -Barrel Building Block. *Front. Mol. Biosci.* **2022**, *9*, 784451. [[CrossRef](#)]
177. Theobald, D.L.; Mitton-Fry, R.M.; Wuttke, D.S. Nucleic Acid Recognition by OB-Fold Proteins. *Annu. Rev. Biophys.* **2003**, *32*, 115–133. [[CrossRef](#)]
178. Theobald, D.L.; Wuttke, D.S. Divergent Evolution within Protein Superfolds Inferred from Profile-Based Phylogenetics. *J. Mol. Biol.* **2005**, *354*, 722–737. [[CrossRef](#)]
179. Guardino, K.M.; Sheftic, S.R.; Slattery, R.E.; Alexandrescu, A.T. Relative Stabilities of Conserved and Non-Conserved Structures in the OB-Fold Superfamily. *Int. J. Mol. Sci.* **2009**, *10*, 2412–2430. [[CrossRef](#)]
180. Alvarez-Carreño, C.; Penev, P.I.; Petrov, A.S.; Williams, L.D. Fold Evolution before LUCA: Common Ancestry of SH3 Domains and OB Domains. *Mol. Biol. Evol.* **2021**, *38*, 5134–5143. [[CrossRef](#)] [[PubMed](#)]
181. Andreeva, A.; Kulesha, E.; Gough, J.; Murzin, A.G. The SCOP Database in 2020: Expanded Classification of Representative Family and Superfamily Domains of Known Protein Structures. *Nucleic Acids Res.* **2020**, *48*, D376–D382. [[CrossRef](#)]

182. Schaeffer, R.D.; Medvedev, K.E.; Andreeva, A.; Chuguransky, S.R.; Pinto, B.L.; Zhang, J.; Cong, Q.; Bateman, A.; Grishin, N.V. ECOD: Integrating Classifications of Protein Domains from Experimental and Predicted Structures. *Nucleic Acids Res.* **2025**, *53*, D411–D418. [[CrossRef](#)] [[PubMed](#)]
183. Subramanian, A.-R. Structure and Functions of Ribosomal Protein S1. In *Progress in Nucleic Acid Research and Molecular Biology*; Cohn, W.E., Ed.; Academic Press: Cambridge, MA, USA, 1983; Volume 28, pp. 101–142, ISBN 0079-6603.
184. Bycroft, M.; Hubbard, T.J.P.; Proctor, M.; Freund, S.M.V.; Murzin, A.G. The Solution Structure of the S1 RNA Binding Domain: A Member of an Ancient Nucleic Acid–Binding Fold. *Cell* **1997**, *88*, 235–242. [[CrossRef](#)]
185. Deryusheva, E.I.; Machulin, A.V.; Selivanova, O.M.; Galzitskaya, O.V. Taxonomic Distribution, Repeats, and Functions of the S1 Domain-Containing Proteins as Members of the OB-Fold Family. *Proteins* **2017**, *85*, 602–613. [[CrossRef](#)]
186. Machulin, A.V.; Deryusheva, E.I.; Selivanova, O.M.; Galzitskaya, O.V. The Number of Domains in the Ribosomal Protein S1 as a Hallmark of the Phylogenetic Grouping of Bacteria. *PLoS ONE* **2019**, *14*, e0221370. [[CrossRef](#)] [[PubMed](#)]
187. Deryusheva, E.I.; Machulin, A.V.; Matyunin, M.A.; Galzitskaya, O.V. Investigation of the Relationship between the S1 Domain and Its Molecular Functions Derived from Studies of the Tertiary Structure. *Molecules* **2019**, *24*, 3681. [[CrossRef](#)]
188. Graumann, P.L.; Marahiel, M.A. A Superfamily of Proteins That Contain the Cold-Shock Domain. *Trends Biochem. Sci.* **1998**, *23*, 286–290. [[CrossRef](#)]
189. Goldstein, J.; Pollitt, N.S.; Inouye, M. Major Cold Shock Protein of *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 283–287. [[CrossRef](#)]
190. Kim, M.-H.; Sasaki, K.; Imai, R. Cold Shock Domain Protein 3 Regulates Freezing Tolerance in *Arabidopsis thaliana*. *J. Biol. Chem.* **2009**, *284*, 23454–23460. [[CrossRef](#)]
191. Horn, G.; Hofweber, R.; Kremer, W.; Kalbitzer, H.R. Structure and Function of Bacterial Cold Shock Proteins. *Cell. Mol. Life Sci.* **2007**, *64*, 1457. [[CrossRef](#)]
192. Park, S.J.; Kwak, K.J.; Oh, T.R.; Kim, Y.O.; Kang, H. Cold Shock Domain Proteins Affect Seed Germination and Growth of *Arabidopsis thaliana* under Abiotic Stress Conditions. *Plant Cell Physiol.* **2009**, *50*, 869–878. [[CrossRef](#)]
193. Manasa, L.S.; Panigrahy, M.; Panigrahi, K.C.S.; Rout, G.R. Overview of Cold Stress Regulation in Plants. *Bot. Rev.* **2022**, *88*, 359–387. [[CrossRef](#)]
194. Heinemann, U.; Roske, Y. Cold-Shock Domains—Abundance, Structure, Properties, and Nucleic-Acid Binding. *Cancers* **2021**, *13*, 190. [[CrossRef](#)] [[PubMed](#)]
195. Graumann, P.; Marahiel, M.A. A Case of Convergent Evolution of Nucleic Acid Binding Modules. *BioEssays* **1996**, *18*, 309–315. [[CrossRef](#)] [[PubMed](#)]
196. Manival, X.; Ghisolfi-Nieto, L.; Joseph, G.; Bouvet, P.; Erard, M. RNA-Binding Strategies Common to Cold-Shock Domain- and RNA Recognition Motif-Containing Proteins. *Nucleic Acids Res.* **2001**, *29*, 2223–2233. [[CrossRef](#)]
197. Fairman-Williams, M.E.; Guenther, U.-P.; Jankowsky, E. SF1 and SF2 Helicases: Family Matters. *Curr. Opin. Struct. Biol.* **2010**, *20*, 313–324. [[CrossRef](#)]
198. Linder, P.; Lasko, P.F.; Ashburner, M.; Leroy, P.; Nielsen, P.J.; Nishi, K.; Schnier, J.; Slonimski, P.P. Birth of the D-E-A-D Box. *Nature* **1989**, *337*, 121–122. [[CrossRef](#)]
199. Rocak, S.; Linder, P. DEAD-Box Proteins: The Driving Forces behind RNA Metabolism. *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 232–241. [[CrossRef](#)] [[PubMed](#)]
200. Linder, P.; Jankowsky, E. From Unwinding to Clamping—the DEAD Box RNA Helicase Family. *Nat. Rev. Mol. Cell Biol.* **2011**, *12*, 505–516. [[CrossRef](#)] [[PubMed](#)]
201. Cordin, O.; Banroques, J.; Tanner, N.K.; Linder, P. The DEAD-Box Protein Family of RNA Helicases. *Gene* **2006**, *367*, 17–37. [[CrossRef](#)] [[PubMed](#)]
202. Putnam, A.; Jankowsky, E. DEAD-Box Helicases as Integrators of RNA, Nucleotide and Protein Binding. *Biochim. Biophys. Acta* **2013**, *1829*, 884–893. [[CrossRef](#)]
203. Rogers, G.W.; Komar, A.A.; Merrick, W.C. eIF4A: The Godfather of the DEAD Box Helicases. *Prog. Nucleic Acid Res. Mol. Biol.* **2002**, *72*, 307–331. [[CrossRef](#)]
204. Edwalds-Gilbert, G.; Kim, D.-H.; Silverman, E.; Lin, R.-J. Definition of a Spliceosome Interaction Domain in Yeast Prp2 ATPase. *RNA* **2004**, *10*, 210–220. [[CrossRef](#)]
205. de la Cruz, J.; Kressler, D.; Tollervey, D.; Linder, P. Dob1p (Mtr4p) Is a Putative ATP-dependent RNA Helicase Required for the 3' End Formation of 5.8S rRNA in *Saccharomyces cerevisiae*. *EMBO J.* **1998**, *17*, 1128–1140. [[CrossRef](#)]
206. Bøddeker, N.; Stade, K.; Franceschi, F. Characterization of DbpA, an *Escherichia coli* DEAD Box Protein with ATP Independent RNA Unwinding Activity. *Nucleic Acids Res.* **1997**, *25*, 537–544. [[CrossRef](#)]
207. Kossen, K.; Uhlenbeck, O.C. Cloning and Biochemical Characterization of *Bacillus subtilis* YxiN, a DEAD Protein Specifically Activated by 23S rRNA: Delineation of a Novel Sub-Family of Bacterial DEAD Proteins. *Nucleic Acids Res.* **1999**, *27*, 3811–3820. [[CrossRef](#)] [[PubMed](#)]

208. Hardin, J.W.; Hu, Y.X.; McKay, D.B. Structure of the RNA Binding Domain of a DEAD-Box Helicase Bound to Its Ribosomal RNA Target Reveals a Novel Mode of Recognition by an RNA Recognition Motif. *J. Mol. Biol.* **2010**, *402*, 412–427. [\[CrossRef\]](#)
209. Rudolph, M.G.; Klostermeier, D. When Core Competence Is Not Enough: Functional Interplay of the DEAD-Box Helicase Core with Ancillary Domains and Auxiliary Factors in RNA Binding and Unwinding. *Biol. Chem.* **2015**, *396*, 849–865. [\[CrossRef\]](#)
210. Wurm, J.P.; Glowacz, K.-A.; Sprangers, R. Structural Basis for the Activation of the DEAD-Box RNA Helicase DbpA by the Nascent Ribosome. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2105961118. [\[CrossRef\]](#)
211. Keeling, P.J.; Burki, F.; Wilcox, H.M.; Allam, B.; Allen, E.E.; Amaral-Zettler, L.A.; Armbrust, E.V.; Archibald, J.M.; Bharti, A.K.; Bell, C.J.; et al. The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): Illuminating the Functional Diversity of Eukaryotic Life in the Oceans through Transcriptome Sequencing. *PLoS Biol.* **2014**, *12*, e1001889. [\[CrossRef\]](#)
212. Blum, M.; Andreeva, A.; Florentino, L.C.; Chuguransky, S.R.; Grego, T.; Hobbs, E.; Pinto, B.L.; Orr, A.; Paysan-Lafosse, T.; Ponamareva, I.; et al. InterPro: The Protein Sequence Classification Resource in 2025. *Nucleic Acids Res.* **2025**, *53*, D444–D456. [\[CrossRef\]](#)
213. Gustafson, E.A.; Wessel, G.M. Vasa Genes: Emerging Roles in the Germ Line and in Multipotent Cells. *BioEssays* **2010**, *32*, 626–637. [\[CrossRef\]](#)
214. Siebert, S.; Goetz, F.E.; Church, S.H.; Bhattacharyya, P.; Zapata, F.; Haddock, S.H.D.; Dunn, C.W. Stem Cells in *Nanomia bijuga* (Siphonophora), a Colonial Animal with Localized Growth Zones. *EvoDevo* **2015**, *6*, 22. [\[CrossRef\]](#) [\[PubMed\]](#)
215. Stoilov, P.; Rafalska, I.; Stamm, S. YTH: A New Domain in Nuclear Proteins. *Trends Biochem. Sci.* **2002**, *27*, 495–497. [\[CrossRef\]](#) [\[PubMed\]](#)
216. Zhang, Z.; Theler, D.; Kaminska, K.H.; Hiller, M.; De La Grange, P.; Pudimat, R.; Rafalska, I.; Heinrich, B.; Bujnicki, J.M.; Allain, F.H.-T.; et al. The YTH Domain Is a Novel RNA Binding Domain. *J. Biol. Chem.* **2010**, *285*, 14701–14710. [\[CrossRef\]](#)
217. Liao, J.; Wei, Y.; Liang, J.; Wen, J.; Chen, X.; Zhang, B.; Chu, L. Insight into the Structure, Physiological Function, and Role in Cancer of m6A Readers—YTH Domain-Containing Proteins. *Cell Death Discov.* **2022**, *8*, 137. [\[CrossRef\]](#)
218. Li, F.; Zhao, D.; Wu, J.; Shi, Y. Structure of the YTH Domain of Human YTHDF2 in Complex with an m6A Mononucleotide Reveals an Aromatic Cage for m6A Recognition. *Cell Res.* **2014**, *24*, 1490–1492. [\[CrossRef\]](#)
219. Luo, S.; Tong, L. Molecular Basis for the Recognition of Methylated Adenines in RNA by the Eukaryotic YTH Domain. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 13834–13839. [\[CrossRef\]](#)
220. Xu, C.; Wang, X.; Liu, K.; Roundtree, I.A.; Tempel, W.; Li, Y.; Lu, Z.; He, C.; Min, J. Structural Basis for Selective Binding of m6A RNA by the YTHDC1 YTH Domain. *Nat. Chem. Biol.* **2014**, *10*, 927–929. Erratum in: *Nat Chem Biol.* **2015**, *11*, 815. [\[CrossRef\]](#)
221. Harigaya, Y.; Tanaka, H.; Yamanaka, S.; Tanaka, K.; Watanabe, Y.; Tsutsumi, C.; Chikashige, Y.; Hiraoka, Y.; Yamashita, A.; Yamamoto, M. Selective Elimination of Messenger RNA Prevents an Incidence of Untimely Meiosis. *Nature* **2006**, *442*, 45–50. [\[CrossRef\]](#) [\[PubMed\]](#)
222. Wang, C.; Zhu, Y.; Jiang, Y.; Xu, C.; Wu, J.; Shi, Y. A Novel RNA-Binding Model of the YTH Domain Reveals the Mechanism for Recognition of Determinant of Selective. *Nucleic Acids Res.* **2016**, *44*, 969–982. [\[CrossRef\]](#) [\[PubMed\]](#)
223. Soh, Y.Q.S.; Mikedis, M.M.; Kojima, M.; Godfrey, A.K.; de Rooij, D.G.; Page, D.C. *Meioc* Maintains an Extended Meiotic Prophase I in Mice. *PLoS Genet.* **2017**, *13*, e1006704. [\[CrossRef\]](#) [\[PubMed\]](#)
224. Li, C.; Li, Y.; Guo, J.; Wang, Y.; Shi, X.; Zhang, Y.; Liang, N.; Ma, H.; Yuan, J.; Xu, J.; et al. Abundant mRNA m1A Modification in Dinoflagellates: A New Layer of Gene Regulation. *EMBO Rep.* **2024**, *25*, 4655–4673. [\[CrossRef\]](#)
225. Blatch, G.L.; Lässle, M. The Tetratricopeptide Repeat: A Structural Motif Mediating Protein-Protein Interactions. *BioEssays* **1999**, *21*, 932–939. [\[CrossRef\]](#)
226. Barkan, A.; Small, I. Pentatricopeptide Repeat Proteins in Plants. *Annu. Rev. Plant Biol.* **2014**, *65*, 415–442. [\[CrossRef\]](#)
227. Small, I.D.; Peeters, N. The PPR Motif—A TPR-Related Motif Prevalent in Plant Organellar Proteins. *Trends Biochem. Sci.* **2000**, *25*, 46–47. [\[CrossRef\]](#) [\[PubMed\]](#)
228. Yin, P.; Li, Q.; Yan, C.; Liu, Y.; Liu, J.; Yu, F.; Wang, Z.; Long, J.; He, J.; Wang, H.-W.; et al. Structural Basis for the Modular Recognition of Single-Stranded RNA by PPR Proteins. *Nature* **2013**, *504*, 168–171. [\[CrossRef\]](#)
229. Ringel, R.; Sologub, M.; Morozov, Y.I.; Litonin, D.; Cramer, P.; Temiakov, D. Structure of Human Mitochondrial RNA Polymerase. *Nature* **2011**, *478*, 269–273. [\[CrossRef\]](#)
230. Barkan, A.; Rojas, M.; Fujii, S.; Yap, A.; Chong, Y.S.; Bond, C.S.; Small, I. A Combinatorial Amino Acid Code for RNA Recognition by Pentatricopeptide Repeat Proteins. *PLoS Genet.* **2012**, *8*, e1002910. [\[CrossRef\]](#)
231. Wang, X.; An, Y.; Xu, P.; Xiao, J. Functioning of PPR Proteins in Organelle RNA Metabolism and Chloroplast Biogenesis. *Front. Plant Sci.* **2021**, *12*, 627501. [\[CrossRef\]](#)
232. Hayes, M.L.; Giang, K.; Berhane, B.; Mulligan, R.M. Identification of Two Pentatricopeptide Repeat Genes Required for RNA Editing and Zinc Binding by C-Terminal Cytidine Deaminase-like Domains. *J. Biol. Chem.* **2013**, *288*, 36519–36529. [\[CrossRef\]](#)
233. Hayes, M.L.; Santibanez, P.I. A Plant Pentatricopeptide Repeat Protein with a DYW-Deaminase Domain Is Sufficient for Catalyzing C-to-U RNA Editing in Vitro. *J. Biol. Chem.* **2020**, *295*, 3497–3505. [\[CrossRef\]](#) [\[PubMed\]](#)

234. Schmitz-Linneweber, C.; Small, I. Pentatricopeptide Repeat Proteins: A Socket Set for Organelle Gene Expression. *Trends Plant Sci.* **2008**, *13*, 663–670. [\[CrossRef\]](#)
235. Manna, S. An Overview of Pentatricopeptide Repeat Proteins and Their Applications. *Biochimie* **2015**, *113*, 93–99. [\[CrossRef\]](#)
236. Ichinose, M.; Sugita, M. RNA Editing and Its Molecular Mechanism in Plant Organelles. *Genes* **2017**, *8*, 5. [\[CrossRef\]](#) [\[PubMed\]](#)
237. Sanchez-Puerta, M.V.; Lippmeier, J.C.; Apt, K.E.; Delwiche, C.F. Plastid Genes in a Non-Photosynthetic Dinoflagellate. *Protist* **2007**, *158*, 105–117. [\[CrossRef\]](#)
238. Slamovits, C.H.; Keeling, P.J. Plastid-Derived Genes in the Nonphotosynthetic Alveolate *Oxyrrhis marina*. *Mol. Biol. Evol.* **2008**, *25*, 1297–1306. [\[CrossRef\]](#)
239. Guo, Z.; Zhang, H.; Liu, S.; Lin, S. Biology of the Marine Heterotrophic Dinoflagellate *Oxyrrhis marina*: Current Status and Future Directions. *Microorganisms* **2013**, *1*, 33–57. [\[CrossRef\]](#)
240. Liu, S.; Melonek, J.; Boykin, L.M.; Small, I.; Howell, K.A. PPR-SMRs. *RNA Biol.* **2013**, *10*, 1501–1510. [\[CrossRef\]](#)
241. Zoschke, R.; Watkins, K.P.; Miranda, R.G.; Barkan, A. The PPR-SMR Protein PPR53 Enhances the Stability and Translation of Specific Chloroplast RNAs in Maize. *Plant J.* **2016**, *85*, 594–606. [\[CrossRef\]](#)
242. Zhou, W.; Lu, Q.; Li, Q.; Wang, L.; Ding, S.; Zhang, A.; Wen, X.; Zhang, L.; Lu, C. PPR-SMR Protein SOT1 Has RNA Endonuclease Activity. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E1554–E1563. [\[CrossRef\]](#)
243. Zhang, J.; Guo, Y.; Fang, Q.; Zhu, Y.; Zhang, Y.; Liu, X.; Lin, Y.; Barkan, A.; Zhou, F. The PPR-SMR Protein ATP4 Is Required for Editing the Chloroplast Rps8 mRNA in Rice and Maize. *Plant Physiol.* **2020**, *184*, 2011–2021. [\[CrossRef\]](#)
244. Berdieva, M.; Kalinina, V.; Palii, O.; Skarlato, S. Putative Muts2 Homologs in Algae: More Goods in Shopping Bag? *J. Mol. Evol.* **2024**, *92*, 815–833. [\[CrossRef\]](#) [\[PubMed\]](#)
245. Manna, S.; Barth, C. Identification of a Novel Pentatricopeptide Repeat Subfamily with a C-Terminal Domain of Bacterial Origin Acquired via Ancient Horizontal Gene Transfer. *BMC Res. Notes* **2013**, *6*, 525. [\[CrossRef\]](#) [\[PubMed\]](#)
246. Boussard, C.; Avon, A.; Kindgren, P.; Bond, C.S.; Challenor, M.; Lurin, C.; Small, I. The Cytidine Deaminase Signature HxE(x)nCxXC of DYW1 Binds Zinc and Is Necessary for RNA Editing of ndhD-1. *New Phytol.* **2014**, *203*, 1090–1095. [\[CrossRef\]](#) [\[PubMed\]](#)
247. Wang, H.; Fewer, D.P.; Holm, L.; Rouhiainen, L.; Sivonen, K. Atlas of Nonribosomal Peptide and Polyketide Biosynthetic Pathways Reveals Common Occurrence of Nonmodular Enzymes. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 9259–9264. [\[CrossRef\]](#)
248. Helfrich, E.J.N.; Piel, J. Biosynthesis of Polyketides by Trans-AT Polyketide Synthases. *Nat. Prod. Rep.* **2016**, *33*, 231–316. [\[CrossRef\]](#)
249. Zheng, J.; Keatinge-Clay, A.T. The Status of Type I Polyketide Synthase Ketoreductases. *Med. Chem. Commun.* **2013**, *4*, 34–40. [\[CrossRef\]](#)
250. D’Andrea, L.D.; Regan, L. TPR Proteins: The Versatile Helix. *Trends Biochem. Sci.* **2003**, *28*, 655–662. [\[CrossRef\]](#)
251. Van Dolah, F.M.; Kohli, G.S.; Morey, J.S.; Murray, S.A. Both Modular and Single-Domain Type I Polyketide Synthases Are Expressed in the Brevetoxin-Producing Dinoflagellate, *Karenia brevis* (Dinophyceae). *J. Phycol.* **2017**, *53*, 1325–1339. [\[CrossRef\]](#) [\[PubMed\]](#)
252. Williams, E.P.; Bachvaroff, T.R.; Place, A.R. A Global Approach to Estimating the Abundance and Duplication of Polyketide Synthase Domains in Dinoflagellates. *Evol. Bioinform.* **2021**, *17*, 11769343211031871. [\[CrossRef\]](#) [\[PubMed\]](#)
253. Fischbach, M.A.; Walsh, C.T.; Clardy, J. The Evolution of Gene Collectives: How Natural Selection Drives Chemical Innovation. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 4601–4608, Erratum in: *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 1679. [\[CrossRef\]](#)
254. Shelest, E.; Heimerl, N.; Fichtner, M.; Sasso, S. Multimodular Type I Polyketide Synthases in Algae Evolve by Module Duplications and Displacement of AT Domains in *Trans*. *BMC Genom.* **2015**, *16*, 1015. [\[CrossRef\]](#) [\[PubMed\]](#)
255. Nivina, A.; Yuet, K.P.; Hsu, J.; Khosla, C. Evolution and Diversity of Assembly-Line Polyketide Synthases. *Chem. Rev.* **2019**, *119*, 12524–12547. [\[CrossRef\]](#)
256. Zhang, L.; Wang, S.; Wang, W.; Shi, J.; Stovall, D.B.; Li, D.; Sui, G. Phase-Separated Subcellular Compartmentation and Related Human Diseases. *Int. J. Mol. Sci.* **2022**, *23*, 5491. [\[CrossRef\]](#) [\[PubMed\]](#)
257. Shen, B.; Chen, Z.; Yu, C.; Chen, T.; Shi, M.; Li, T. Computational Screening of Phase-Separating Proteins. *Genom. Proteom. Bioinform.* **2021**, *19*, 13–24. [\[CrossRef\]](#)
258. Sprecher, B.N.; Zhang, H.; Lin, S. Nuclear Gene Transformation in the Dinoflagellate *Oxyrrhis marina*. *Microorganisms* **2020**, *8*, 126. [\[CrossRef\]](#)
259. Gornik, S.G.; Maegle, I.; Hambleton, E.A.; Voss, P.A.; Waller, R.F.; Guse, A. Nuclear Transformation of a Dinoflagellate Symbiont of Corals. *Front. Mar. Sci.* **2022**, *9*, 1035413. [\[CrossRef\]](#)
260. Lin, S.; Zhang, H.; Zhuang, Y.; Tran, B.; Gill, J. Spliced Leader-Based Metatranscriptomic Analyses Lead to Recognition of Hidden Genomic Features in Dinoflagellates. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 20033–20038. [\[CrossRef\]](#) [\[PubMed\]](#)
261. 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. [\[CrossRef\]](#)

262. Cooney, E.C.; Okamoto, N.; Cho, A.; Hehenberger, E.; Richards, T.A.; Santoro, A.E.; Worden, A.Z.; Leander, B.S.; Keeling, P.J. Single-Cell Transcriptomics of *Abedinium* Reveals a New Early-Branching Dinoflagellate Lineage. *Genome Biol. Evol.* **2020**, *12*, 2417–2428. [[CrossRef](#)] [[PubMed](#)]
263. Morse, D.; Tse, S.P.K.; Lo, S.C.L. Exploring Dinoflagellate Biology with High-Throughput Proteomics. *Harmful Algae* **2018**, *75*, 16–26. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.