RNA Biology 10:9, 1426–1432; September 2013; © 2013 Landes Bioscience

Pentatricopeptide repeats
Modular blocks for building RNA-binding proteins

Aleksandra Filipovska[1,2] and Oliver Rackham[1,2,*]

[1]Western Australian Institute for Medical Research and Centre for Medical Research; The University of Western Australia; Perth, WA Australia; [2]School of Chemistry and Biochemistry; The University of Western Australia; Crawley, WA Australia

Keywords: RNA-binding, pentatricopeptide repeat, repeat proteins, synthetic biology, RNA-protein interactions

Pentatricopeptide repeat (PPR) proteins control diverse aspects of RNA metabolism across the eukaryotic domain. Recent computational and structural studies have provided new insights into how they recognize RNA, and show that the recognition is sequence-specific and modular. The modular code for RNA-binding by PPR proteins holds great promise for the engineering of new tools to target RNA and identifying RNAs bound by natural PPR proteins.

Introduction: Modularity in Molecular Recognition

Throughout evolution, simple structural motifs have been reused, specialized and combined to generate the diverse families of proteins in modern organisms. The assembly of proteins from simpler individual motifs or domains has imbued many proteins with different degrees of modularity. Modularity can be defined as a characteristic of complex systems that consist of multiple units with discrete functions.[1] Modules have the following features: they have identifiable interfaces with other modules; they can be modified and evolved with some degree of independence and they often maintain their functions when isolated and rearranged. Modularity has been a particularly powerful concept that has facilitated the understanding of biological systems, as in many cases, the functions of proteins can be inferred from the sum of their parts, and in biotechnology, where assembly of heterologous domains has enabled the engineering and manipulation of biological systems.[1,2] These manipulations range from using well-characterized antibody epitopes or fluorescent protein domains to tag proteins of interest, to more complex applications such as the widely used yeast two-hybrid system, where splitting the DNA-binding and transcription activation domains from a transcription factor and fusing them to potential interacting proteins enables the study of protein–protein interactions with high-throughput in living cells.[3,4] The principles of protein modularity have been particularly useful for synthetic biology, with a goal of engineering biological systems with either improved properties or new functions.[1,5-11]

Repeat proteins represent an interesting subset of proteins characterized by small, structural motifs of 20–50 amino acids that are repeated a variable number of times in tandem within each protein.[12,13] The repeated motifs stack on each other to form elongated structures, which provide a large surface area that is particularly advantageous in forming macromolecular interactions when compared with typical globular proteins. Often the entire array is required to form the binding surface; however, in a subset of repeat proteins that bind nucleic acids, the repeats act in a modular fashion, with each repeat interacting with a single nucleotide base. This was first observed in proteins of the Pumilio and FBF homology (PUF) family of RNA-binding proteins,[14] and recently it has been seen in transcription activator-like effector (TALE) proteins that bind DNA.[15-18] These properties have enabled PUFs and TALEs to be used in many research applications involving the manipulation of RNA and DNA, respectively.[19-32] The pentatricopeptide repeat (PPR) proteins are a large family of eukaryotic RNA-binding proteins that have been widely researched because of their diverse and important roles in organelle RNA metabolism, including RNA editing of functional important RNA bases by deamination.[33,34] However, despite intensive study, determining their modes of RNA binding and whether they are modular has proved particularly elusive. The exquisite specificity of PPR proteins in locating individual editing sites within organelle transcriptomes provided indirect evidence that RNA-binding by PPRs might be sequence-dependent[35,36] and modular akin to PUFs and TALEs,[37] while observations that the closely related tetratricopeptide repeat (TPR) protein family can recognize RNA in a sequence-independent manner,[38,39] potentially conflicted with this idea. Now recent studies have shown that PPRs likely recognize their RNA targets in a modular and sequence-specific manner, similarly to PUF and TALE proteins.

The Modes of RNA-Recognition by PPR Proteins

PPR proteins contain a repeated motif that is typically 35 amino acids in length (hence their name) and forms two anti-parallel α helices (**Fig. 1A**).[40,41] Proteins have been observed to contain between two and 30 individual PPRs.[33] Some PPR proteins appear to consist almost entirely of tandem PPRs, while others contain other domains, such as endonuclease or editing domains.[33,42-44] There are two classes of PPR proteins that have been described. Proteins with arrays of PPRs that are all approximately 35 amino acids long are designated as "P class" proteins. The second type of PPR proteins are designated as "PLS class,"

*Correspondence to: Oliver Rackham; Email: rackham@waimr.uwa.edu.au
Submitted: 03/18/2013; Revised: 04/19/2013; Accepted: 04/22/2013
http://dx.doi.org/10.4161/rna.24769

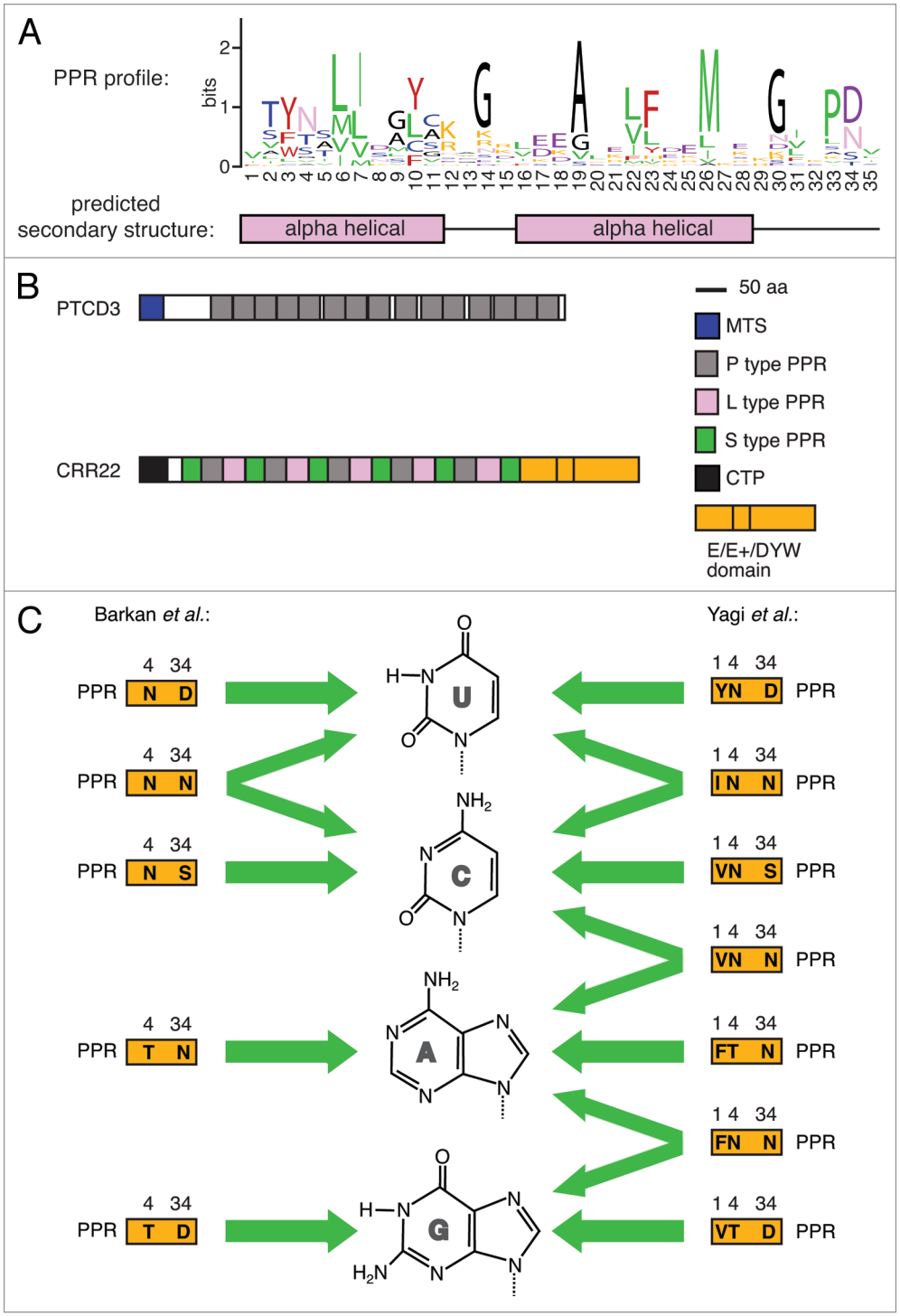

Figure 1. Pentatricopeptide repeats. (**A**) A sequence logo illustrating the characteristic amino acid composition of PPR sequences. The logo was derived from 14,466 PPRs found in the PROSITE PPR entry (PDOC51375) using WebLogo.[83] These sequences are derived from the following taxonomic groups: 86% plants, 5.7% fungi, 4.3% animals, 1.8% algae, 1% trypanosomes, and 1.2% others. Amino acids are color-coded according to the physio-chemical properties of their side chains: small (A, G) in black, nucleophilic (C, S, T) in blue, hydrophobic (I, L, V, M, P) in green, aromatic (F, W, Y) in red, acidic (D, E) in purple, amides (Q, N) in pink and basic (H, K, R) in orange. Regions of α-helical structure are shown below. Amino acids are numbered based on the Pfam model, which functions as a minimal unit.[54] Residue 34 is also defined as ii according to Kobayashi et al.,[54] while the numbering scheme used by Fujii et al.[53] is shifted to the N terminus by two amino acids such that amino acids 1, 4 and 34 in the Pfam model are annotated as 3, 6, and 1, respectively. (**B**) Schematic structures of a typical P class PPR protein, human PTCD3,[84] and a typical PLS class PPR protein, *Arabidopsis* CRR22.[85] PPRs, mitochondrial targeting sequence (MTS), chloroplast targeting peptide (CTP) and the E/E+/DYW domain, often associated with editing PPR proteins, are highlighted. (**C**) The recognition code of PPRs for RNA bases. Only representative predictions by Yagi et al. are shown; for a full list, refer to the original research paper.[57]

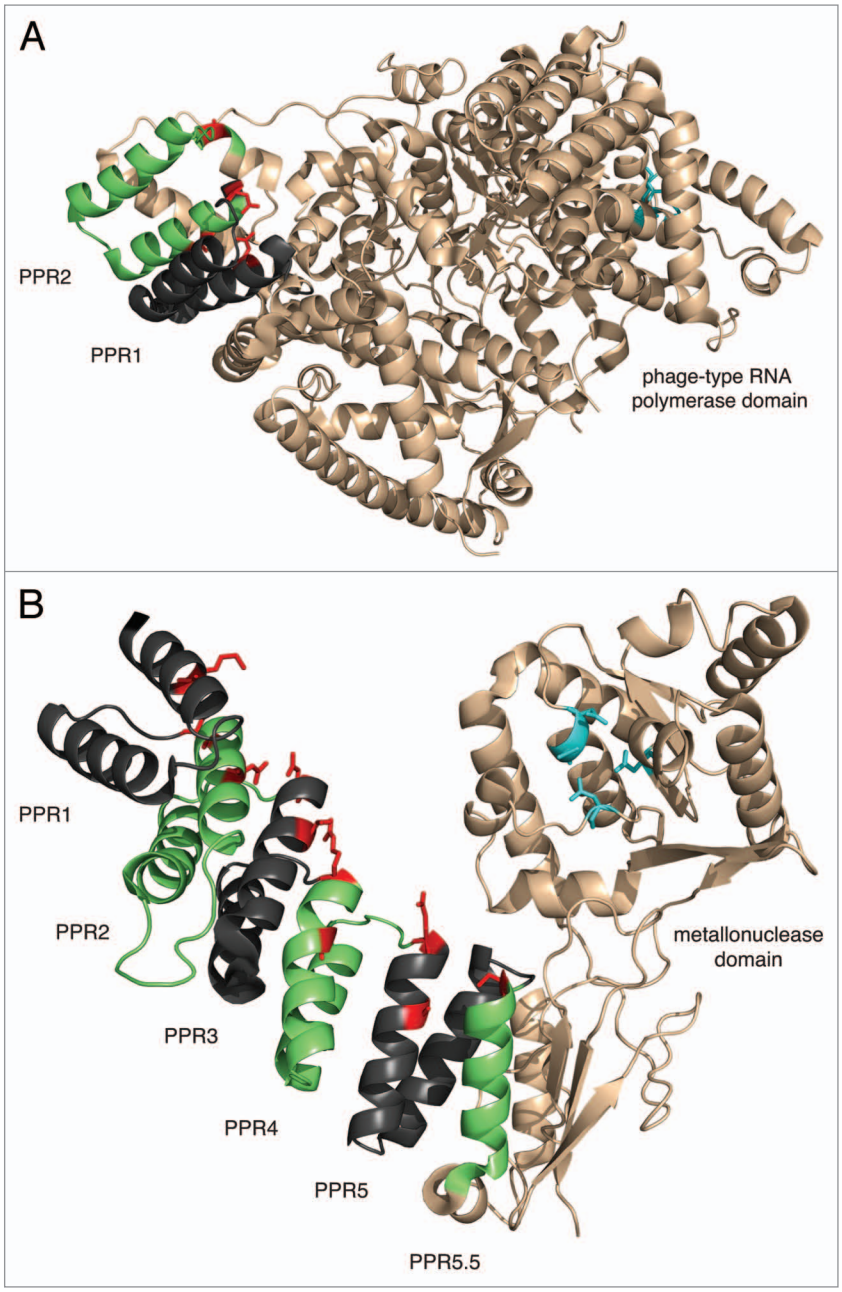

Figure 2. Structures of PPR-containing proteins. (**A**) Crystal structure of two tandem PPRs within the human POLRMT protein (PDB accession code 3SPA).[60] The N-terminal PPR is colored in gray and the C-terminal PPR is colored in green. Residues 4 and 34 of each PPR are highlighted in red. The active site residues predicted to be in close proximity to the phosphate groups in incoming nucleoside triphosphates are highlighted in blue. (**B**) Crystal structure of five-and-a-half tandem PPRs within the *Arabidopsis thaliana* PRORP1 protein (PDB accession code 4G24).[61] PPRs are colored in alternating gray and green. Residues 4 and 34 of each PPR are highlighted in red and active site residues of the metallonuclease domain are shown in blue.

and this class includes the PPR RNA editing proteins within chloroplasts and mitochondria of land plants.[45-47] The PLS class of PPR proteins contain C-terminal domains required for RNA editing and a distinctive PPR architecture. These editing proteins contain triplet repeats alternating between a typical PPR, a longer PPR of 35 or 36 amino acids and a short PPR of only 31 amino acids (**Fig. 1B**).[45] Although the significance of this observation is not clear, it may relate to the binding properties of the PPR array since editing PPR proteins associate transiently with their target RNAs while P class proteins often form very stable complexes with their targets.[48-52]

The first insight about how PPRs bound their RNA targets came from studies on the evolutionary conflict between nuclear and mitochondrial genomes in plants.[53] The coevolution of PPR proteins and their binding sites results in accelerated evolution that would be expected to be most evident in amino acids required for RNA binding. The amino acids where these effects were most evident were at positions 1, 4, and 34 within PPRs. Furthermore,

a computational model derived from analysis of co-varying residues predicted that these amino acids might be co-located on the internal face of one of the PPR helices.[53] A detailed study by Kobayashi et al. used truncations of the *Arabidopsis* HCF152 protein, consisting of two adjacent PPRs, to perform extensive mutagenesis to identify amino acids that are important for the affinity or specificity of RNA-binding by these proteins.[54] This study identified residues 1, 4, 8, 12, and 34 as the most important for maintaining high-affinity RNA-binding. The close overlap between the positions identified by co-variation analysis[53] and mutagenesis[54] suggested that these residues play a key role in the recognition of RNA.[55]

Barkan et al. recently used computational analysis of correlations between specific PPR residues and RNA bases within their binding sites to elucidate the code for RNA recognition by PPR proteins.[56] This was facilitated by the elucidation of the footprints of a few well-characterized PPR proteins on their target RNAs.[48,50,51] Alignment of all possible combinations of amino acids of each PPR to each base showed a significant correlation between the identities of amino acids at positions 4 and 34 and particular bases within the RNA footprint (**Fig. 1C**).[56] This code for RNA recognition was confirmed experimentally by mutating residues 4 and 34 in two PPRs within the well-characterized PPR10 protein to show that the recoded mutant PPRs had the new, predicted specificities according to the code. When the code was applied to the PLS class of editing PPR proteins, they all aligned exactly four bases upstream of the edited base, indicating that the PPR array provided a very accurate "molecular ruler" for modification of specific bases. Interestingly, although the code accurately predicted the binding preferences of the canonical and short PPRs within editing PPR proteins, every third, long PPR could not be accommodated by the predicted code. This led to the proposal that every third, long PPR might not directly bind RNA bases, potentially explaining the ability of editing PPR proteins to dissociate from their RNA targets more easily.[56]

Yagi et al. recently used an alternative computational approach to determine the code for RNA recognition by PPRs.[57] Analysis of amino acids with significantly low variability when found opposite a particular nucleotide base confirmed the importance of residues 4 and 34, as observed by Barkan et al.,[56] but also this group predicted that the amino acid at position 1 in the PPR motif plays a role in PPR specificity (**Fig. 1C**). They found that the amino acid at position 4 played the most significant role in base specificity and mainly discriminated between purines and pyrimidines, while the amino acid at position 34 distinguished amino and keto groups. Amino acids at position 1 appeared to be responsible for fine-tuning the recognition of particular bases. The codes described by the two groups are highly overlapping and generally harmonious, although in the code predicted by Yagi et al., the specificity of amino acids at positions 4 and 34 can sometimes be significantly modulated by particular amino acids at position 1 (**Fig. 1C**). Like Barkan et al., they found that editing PPR proteins bound four bases upstream of their target base; however, they found that the identity of some amino acids in the long PPRs correlated with certain bases, implying that some of these repeats in PLS class proteins might influence RNA-binding specificity.[57] Another interesting difference between P and PLS class proteins, observed by both groups, was that P class proteins typically had internal sequences in their RNA targets that did not appear to be recognized by PPRs. These sections of RNA are probably looped out from the RNA-protein interface and might occur because of a mismatch between the curvature of the PPR proteins and the distance between RNA bases, as has been observed for PUF proteins.[14,26,56-59] A precise understanding of how amino acids within each PPR recognize individual bases and determine their binding specificities will clearly require detailed structural studies.

Structural Insights into PPR Protein Function

Although a PPR protein in complex with its RNA target has not been observed at an atomic level to date, the structures of two PPR proteins have been solved recently, providing insights into their potential binding mechanisms. The first structures of PPRs were observed in the human mitochondrial RNA polymerase (POLRMT).[60] POLRMT is the sole RNA polymerase in mammalian mitochondria and closely resembles bacteriophage RNA polymerases such as that from T7. In addition, it contains a distinctive N-terminal extension that includes two PPRs (**Fig. 2A**). This region is essential for transcription of double-stranded promoter DNA but not transcription from a melted promoter; however, the exact mechanisms are not known.[60] The tandem PPRs within POLRMT form a pair of anti-parallel α helices, as predicted from their similarity to TPRs; however, they are nestled within and stabilized by surrounding structural elements, which is likely to be atypical for PPR proteins with long PPR arrays.[60] The positioning of the PPRs and their importance within the N-terminal domain, implies that they might interact with promoter DNA bases or nascent RNA transcripts, although this has not been examined experimentally.

An array of PPRs was recently observed in the structure of a protein-only RNase P enzyme.[61] Mature tRNAs are produced from larger precursor transcripts by the action of RNase P and RNase Z enzymes that cleave at their 5' and 3' ends, respectively.[62-64] Although RNase P has been known for decades to be a protein-bound ribozyme,[65] new breakthroughs have shown that convergent evolution has also produced a protein-only RNase P enzyme.[43,44,61,66-68] The protein-only RNase P was first identified in human mitochondria as a complex of three distinct proteins that are all required for RNase P activity: a tRNA methyltransferase (MRPP1/TRMT10C), a short chain dehydrogenase/reductase (MRPP2/SDR5C1) and a PPR protein (MRPP3).[44] MRPP3 contains a C-terminal metallonuclease domain and was predicted to perform the RNA cleavage reaction.[44] Studies of the plant and Trypanosome homologs of MRPP3, the proteinaceous RNase P protein (PRORP) family, revealed that unlike the human protein, they all function efficiently as isolated proteins and that different paralogs are responsible for mitochondrial and nuclear tRNA processing.[43,66,68] Remarkably, despite their distinct compositions and evolutionary origins, PRORP proteins can functionally substitute for RNA-based RNase P enzymes in both *E. coli* and yeast nuclei.[43,68]

The crystal structure of the *Arabidopsis* protein-only RNase P, PRORP1, has revealed the most insightful picture of PPR structure to date.[61] PRORP1 forms an L-shaped structure, where the array of five and a half PPRs forms one arm and the metallonuclease domain forms the other (**Fig. 2B**). The 11 α-helices of the PPR array stack to form a right handed super helix, reminiscent of the well-characterized TPR proteins, with which they share sequence similarity. The electrostatic potential of the array's surface is overall neutral, implying that it interacts predominantly with the RNA bases of its target tRNAs. Furthermore, it is compelling to note that the amino acids at positions 4 and 34 of each PPR form a path to the active site of the metallonuclease domain (**Fig. 2B**). Detailed tRNA-binding studies have shown that single-stranded bases within the D and TψC loops are required for PRORP1 binding.[67] These loops are also bound by the RNA-based RNase P from *E. coli* and the two enzymes might bind tRNAs in a similar mode, providing a functionally similar outcome, despite different types of catalysis.[61] Importantly, studies of PRORP1 provide insights into the likely structures of P class PPR arrays.

Future Directions: Understanding and Engineering PPR Proteins

A potentially complicating factor in using the concept of modularity to manipulate and understand biological systems is that, over the course of evolution, motifs and domains that may once have functioned independently can evolve new functions and interactions that depend on neighboring domains or other macromolecules or small molecules within their cellular network. This means that motifs and domains are seldom truly modular in biological systems. For PPRs, this is evident due to the inherent structures of repeat arrays, where stacking between individual repeats stabilizes the entire array. Furthermore, it has become clear that many PPR proteins require other interacting proteins to function. For example, the human mitochondrial RNase P protein MRPP3 has catalytic activity only when it forms a complex with MRPP1/TRMT10C and MRPP2/SDR5C1.[44] The mammalian mitochondrial PPR protein, LRPPRC, which is important for polyadenylation and translation, forms an intimate complex with another RNA-binding protein, SLIRP, such that each protein requires the other for their stability.[69-71] Importantly, the editing PPR proteins of *Arabidopsis* must associate with multiple organelle RNA editing factor (MORF) family proteins or, in some cases, an autonomous DYW domain protein, DYW1, to deaminate their target RNA bases.[72-75] To date, a complete description of the proteins required for RNA editing has not been elucidated. Curiously, the MORF family proteins interact directly with the PLS PPR array,[74] further complicating the picture of how the editing PPR proteins recognize RNA. However, the recently elucidated code for PPR-RNA recognition provides evidence that each PPR can recognize a specific base independently of its neighboring repeat's specificity. This has enabled the prediction of the RNA targets of plant PPR proteins;[57] however, divergence from the typical code by yeast and mammalian PPR proteins might make prediction of their targets challenging.[57,76-79] Nevertheless, the fact that individual PPRs within proteins can be altered to modify their specificities provides further evidence of their modularity and indicates that they might be engineered to target RNAs of interest, as has been the case for PUFs.[19-21,25-32]

Before the code for PPR-RNA binding can be widely applied to understand natural PPR proteins or engineer artificial RNA-binding proteins, a number of outstanding issues must be resolved. These include the structural details of PPR-RNA binding, the deviation from the RNA recognition code of long PPRs in PLS class proteins and PPRs of diverse organisms outside of the plant kingdom, molecular details determining the balance between affinity and specificity and an understanding of the additional factors that are required for the activity of PPR proteins. Furthermore, the development of robust methods to produce recombinant PPR proteins, which are predominantly insoluble, would help the analyses of their affinity and specificity for RNA targets in vitro. Nevertheless, PPR proteins have essential roles in many cellular processes,[33,37,44,45,49,55,62,66,68,76-78,80,81] which makes resolving these gaps in our knowledge of fundamental importance. Furthermore, PPR proteins have a number of desirable features that make the development of their applications in biotechnology and synthetic biology quite appealing. Natural PPR proteins have been observed to contain between two and 30 individual PPRs,[33] providing considerable flexibility in the complexity of the RNAs they might bind. The PPR proteins characterized to date operate in mitochondria, chloroplasts and nuclei; locations where the most common RNA-directed tool, RNA interference, cannot function or functions poorly to target RNA.[82] Furthermore, proteins that contain PPRs have often been observed to contain many other domains with diverse roles in RNA metabolism, such as RNA cleavage, modification and control of translation;[33] illustrating their potential structural compatibility as a fusion partner. These qualities will likely be very useful for making new research tools to manipulate aspects of RNA biology that have been neglected due to a lack of appropriate reagents and for controlling gene networks to build cells with new properties in synthetic biology. Whatever comes next in PPR biology and biotechnology, it is clear that we are at a defining juncture in this research field.

Conclusions

Recent breakthroughs using bioinformatics and structural analyses have revolutionized our understanding of the roles of PPRs in RNA binding. These studies have set the scene for future studies to determine the structural basis of PPR-RNA recognition and to test the potential of this code to be applied in a modular manner in synthetic biology and biotechnology.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

Work in our laboratories is supported by fellowships and grants from the Australian Research Council, the National Health and Medical Research Council and the Royal Perth Hospital Medical Research Foundation.

References

1. Filipovska A, Rackham O. Building a parallel metabolism within the cell. ACS Chem Biol 2008; 3:51-63; PMID:18173237; http://dx.doi.org/10.1021/cb700185e.
2. Bhattacharyya RP, Reményi A, Yeh BJ, Lim WA. Domains, motifs, and scaffolds: the role of modular interactions in the evolution and wiring of cell signaling circuits. Annu Rev Biochem 2006; 75:655-80; PMID:16756506; http://dx.doi.org/10.1146/annurev.biochem.75.103004.142710.
3. Fields S, Song O. A novel genetic system to detect protein-protein interactions. Nature 1989; 340:245-6; PMID:2547163; http://dx.doi.org/10.1038/340245a0.
4. Rajagopala SV, Sikorski P, Caufield JH, Tovchigrechko A, Uetz P. Studying protein complexes by the yeast two-hybrid system. Methods 2012; 58:392-9; PMID:22841565; http://dx.doi.org/10.1016/j.ymeth.2012.07.015.
5. Schwille P. Bottom-up synthetic biology: engineering in a tinkerer's world. Science 2011; 333:1252-4; PMID:21885774; http://dx.doi.org/10.1126/science.1211701.
6. Khalil AS, Collins JJ. Synthetic biology: applications come of age. Nat Rev Genet 2010; 11:367-79; PMID:20395970; http://dx.doi.org/10.1038/nrg2775.
7. Davidson EA, Ellington AD. Synthetic RNA circuits. Nat Chem Biol 2007; 3:23-8; PMID:17173026; http://dx.doi.org/10.1038/nchembio846.
8. Benner SA, Sismour AM. Synthetic biology. Nat Rev Genet 2005; 6:533-43; PMID:15995697; http://dx.doi.org/10.1038/nrg1637.
9. Voigt CA. Genetic parts to program bacteria. Curr Opin Biotechnol 2006; 17:548-57; PMID:16978856; http://dx.doi.org/10.1016/j.copbio.2006.09.001.
10. Drubin DA, Way JC, Silver PA. Designing biological systems. Genes Dev 2007; 21:242-54; PMID:17289915; http://dx.doi.org/10.1101/gad.1507207.
11. Endy D. Foundations for engineering biology. Nature 2005; 438:449-53; PMID:16306983; http://dx.doi.org/10.1038/nature04342.
12. Grove TZ, Cortajarena AL, Regan L. Ligand binding by repeat proteins: natural and designed. Curr Opin Struct Biol 2008; 18:507-15; PMID:18602006; http://dx.doi.org/10.1016/j.sbi.2008.05.008.
13. Boersma YL, Plückthun A. DARPins and other repeat protein scaffolds: advances in engineering and applications. Curr Opin Biotechnol 2011; 22:849-57; PMID:21715155; http://dx.doi.org/10.1016/j.copbio.2011.06.004.
14. Wang X, McLachlan J, Zamore PD, Hall TM. Modular recognition of RNA by a human pumilio-homology domain. Cell 2002; 110:501-12; PMID:12202039; http://dx.doi.org/10.1016/S0092-8674(02)00873-5.
15. Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. Science 2009; 326:1501; PMID:19933106; http://dx.doi.org/10.1126/science.1178817.
16. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, et al. Breaking the code of DNA binding specificity of TAL-type III effectors. Science 2009; 326:1509-12; PMID:19933107; http://dx.doi.org/10.1126/science.1178811.
17. Deng D, Yan C, Pan X, Mahfouz M, Wang J, Zhu JK, et al. Structural basis for sequence-specific recognition of DNA by TAL effectors. Science 2012; 335:720-3; PMID:22223738; http://dx.doi.org/10.1126/science.1215670.
18. Mak AN, Bradley P, Cernadas RA, Bogdanove AJ, Stoddard BL. The crystal structure of TAL effector PthXo1 bound to its DNA target. Science 2012; 335:716-9; PMID:22223736; http://dx.doi.org/10.1126/science.1216211.
19. Filipovska A, Rackham O. Modular recognition of nucleic acids by PUF, TALE and PPR proteins. Mol Biosyst 2012; 8:699-708; PMID:22234420; http://dx.doi.org/10.1039/c2mb05392f.
20. Filipovska A, Razif MF, Nygård KK, Rackham O. A universal code for RNA recognition by PUF proteins. Nat Chem Biol 2011; 7:425-7; PMID:21572425; http://dx.doi.org/10.1038/nchembio.577.
21. Filipovska A, Rackham O. Designer RNA-binding proteins: New tools for manipulating the transcriptome. RNA Biol 2011; 8:978-83; PMID:21941129; http://dx.doi.org/10.4161/rna.8.6.17907.
22. Mahfouz MM, Li L. TALE nucleases and next generation GM crops. GM Crop 2011; 2.
23. Zhang F, Cong L, Lodato S, Kosuri S, Church GM, Arlotta P. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. Nat Biotechnol 2011; 29:149-53; PMID:21248753; http://dx.doi.org/10.1038/nbt.1775.
24. Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, et al. A TALE nuclease architecture for efficient genome editing. Nat Biotechnol 2011; 29:143-8; PMID:21179091; http://dx.doi.org/10.1038/nbt.1755.
25. Dong S, Wang Y, Cassidy-Amstutz C, Lu G, Bigler R, Jezyk MR, et al. Specific and modular binding code for cytosine recognition in Pumilio/FBF (PUF) RNA-binding domains. J Biol Chem 2011; 286:26732-42; PMID:21653694; http://dx.doi.org/10.1074/jbc.M111.244889.
26. Lu G, Dolgner SJ, Hall TM. Understanding and engineering RNA sequence specificity of PUF proteins. Curr Opin Struct Biol 2009; 19:110-5; PMID:19186050; http://dx.doi.org/10.1016/j.sbi.2008.12.009.
27. Cheong CG, Hall TM. Engineering RNA sequence specificity of Pumilio repeats. Proc Natl Acad Sci USA 2006; 103:13635-9; PMID:16954190; http://dx.doi.org/10.1073/pnas.0606294103.
28. Wang Y, Cheong CG, Hall TM, Wang Z. Engineering splicing factors with designed specificities. Nat Methods 2009; 6:825-30; PMID:19801992; http://dx.doi.org/10.1038/nmeth.1379.
29. Tilsner J, Linnik O, Christensen NM, Bell K, Roberts IM, Lacomme C, et al. Live-cell imaging of viral RNA genomes using a Pumilio-based reporter. Plant J 2009; 57:758-70; PMID:18980643; http://dx.doi.org/10.1111/j.1365-313X.2008.03720.x.
30. Ozawa T, Natori Y, Sato M, Umezawa Y. Imaging dynamics of endogenous mitochondrial RNA in single living cells. Nat Methods 2007; 4:413-9; PMID:17401370.
31. Cooke A, Prigge A, Opperman L, Wickens M. Targeted translational regulation using the PUF protein family scaffold. Proc Natl Acad Sci USA 2011; 108:15870-5; PMID:21911377; http://dx.doi.org/10.1073/pnas.1105151108.
32. Chen Y, Varani G. Finding the missing code of RNA recognition by PUF proteins. Chem Biol 2011; 18:821-3; PMID:21802002; http://dx.doi.org/10.1016/j.chembiol.2011.07.001.
33. Schmitz-Linneweber C, Small I. Pentatricopeptide repeat proteins: a socket set for organelle gene expression. Trends Plant Sci 2008; 13:663-70; PMID:19004664; http://dx.doi.org/10.1016/j.tplants.2008.10.001.
34. Knoop V. When you can't trust the DNA: RNA editing changes transcript sequences. Cell Mol Life Sci 2011; 68:567-86; PMID:20938709; http://dx.doi.org/10.1007/s00018-010-0538-9.
35. Williams-Carrier R, Kroeger T, Barkan A. Sequence-specific binding of a chloroplast pentatricopeptide repeat protein to its native group II intron ligand. RNA 2008; 14:1930-41; PMID:18669444; http://dx.doi.org/10.1261/rna.1077708.
36. Okuda K, Myouga F, Motohashi R, Shinozaki K, Shikanai T. Conserved domain structure of pentatricopeptide repeat proteins involved in chloroplast RNA editing. Proc Natl Acad Sci USA 2007; 104:8178-83; PMID:17483454; http://dx.doi.org/10.1073/pnas.0700865104.
37. Delannoy E, Stanley WA, Bond CS, Small ID. Pentatricopeptide repeat (PPR) proteins as sequence-specificity factors in post-transcriptional processes in organelles. Biochem Soc Trans 2007; 35:1643-7; PMID:18031283; http://dx.doi.org/10.1042/BST0351643.
38. Abbas YM, Pichlmair A, Górna MW, Superti-Furga G, Nagar B. Structural basis for viral 5'-PPP-RNA recognition by human IFIT proteins. Nature 2013; 494:60-4; PMID:23334420; http://dx.doi.org/10.1038/nature11783.
39. Pichlmair A, Lassnig C, Eberle CA, Górna MW, Baumann CL, Burkard TR, et al. IFIT1 is an antiviral protein that recognizes 5'-triphosphate RNA. Nat Immunol 2011; 12:624-30; PMID:21642987; http://dx.doi.org/10.1038/ni.2048.
40. Small ID, Peeters N. The PPR motif - a TPR-related motif prevalent in plant organellar proteins. Trends Biochem Sci 2000; 25:46-7; PMID:10664580; http://dx.doi.org/10.1016/S0968-0004(99)01520-0.
41. Aubourg S, Boudet N, Kreis M, Lecharny A. In Arabidopsis thaliana, 1% of the genome codes for a novel protein family unique to plants. Plant Mol Biol 2000; 42:603-13; PMID:10809006; http://dx.doi.org/10.1023/A:1006352315928.
42. de Longevialle AF, Hendrickson L, Taylor NL, Delannoy E, Lurin C, Badger M, et al. The pentatricopeptide repeat gene OTP51 with two LAGLIDADG motifs is required for the cis-splicing of plastid ycf3 intron 2 in Arabidopsis thaliana. Plant J 2008; 56:157-68; PMID:18557832; http://dx.doi.org/10.1111/j.1365-313X.2008.03581.x.
43. Gobert A, Gutmann B, Taschner A, Gössringer M, Holzmann J, Hartmann RK, et al. A single Arabidopsis organellar protein has RNase P activity. Nat Struct Mol Biol 2010; 17:740-4; PMID:20473316; http://dx.doi.org/10.1038/nsmb.1812.
44. Holzmann J, Frank P, Löffler E, Bennett KL, Gerner C, Rossmanith W. RNase P without RNA: identification and functional reconstitution of the human mitochondrial tRNA processing enzyme. Cell 2008; 135:462-74; PMID:18984158; http://dx.doi.org/10.1016/j.cell.2008.09.013.
45. Lurin C, Andrés C, Aubourg S, Bellaoui M, Bitton F, Bruyère C, et al. Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. Plant Cell 2004; 16:2089-103; PMID:15269332; http://dx.doi.org/10.1105/tpc.104.022236.
46. Salone V, Rüdinger M, Polsakiewicz M, Hoffmann B, Groth-Malonek M, Szurek B, et al. A hypothesis on the identification of the editing enzyme in plant organelles. FEBS Lett 2007; 581:4132-8; PMID:17707818; http://dx.doi.org/10.1016/j.febslet.2007.07.075.
47. Rüdinger M, Volkmar U, Lenz H, Groth-Malonek M, Knoop V. Nuclear DYW-type PPR gene families diversify with increasing RNA editing frequencies in liverwort and moss mitochondria. J Mol Evol 2012; 74:37-51; PMID:22302222; http://dx.doi.org/10.1007/s00239-012-9486-3.
48. Prikryl J, Rojas M, Schuster G, Barkan A. Mechanism of RNA stabilization and translational activation by a pentatricopeptide repeat protein. Proc Natl Acad Sci USA 2011; 108:415-20; PMID:21173259; http://dx.doi.org/10.1073/pnas.1012076108.
49. Pfalz J, Bayraktar OA, Prikryl J, Barkan A. Site-specific binding of a PPR protein defines and stabilizes 5' and 3' mRNA termini in chloroplasts. EMBO J 2009; 28:2042-52; PMID:19424177; http://dx.doi.org/10.1038/emboj.2009.121.

50. Ruwe H, Schmitz-Linneweber C. Short non-coding RNA fragments accumulating in chloroplasts: footprints of RNA binding proteins? Nucleic Acids Res 2012; 40:3106-16; PMID:22139936; http://dx.doi.org/10.1093/nar/gkr1138.
51. Zhelyazkova P, Hammani K, Rojas M, Voelker R, Vargas-Suárez M, Börner T, et al. Protein-mediated protection as the predominant mechanism for defining processed mRNA termini in land plant chloroplasts. Nucleic Acids Res 2012; 40:3092-105; PMID:22156165; http://dx.doi.org/10.1093/nar/gkr1137.
52. Okuda K, Shikanai T. A pentatricopeptide repeat protein acts as a site-specificity factor at multiple RNA editing sites with unrelated cis-acting elements in plastids. Nucleic Acids Res 2012; 40:5052-64; PMID:22362750; http://dx.doi.org/10.1093/nar/gks164.
53. Fujii S, Bond CS, Small ID. Selection patterns on restorer-like genes reveal a conflict between nuclear and mitochondrial genomes throughout angiosperm evolution. Proc Natl Acad Sci USA 2011; 108:1723-8; PMID:21220331; http://dx.doi.org/10.1073/pnas.1007667108.
54. Kobayashi K, Kawabata M, Hisano K, Kazama T, Matsuoka K, Sugita M, et al. Identification and characterization of the RNA binding surface of the pentatricopeptide repeat protein. Nucleic Acids Res 2012; 40:2712-23; PMID:22127869; http://dx.doi.org/10.1093/nar/gkr1084.
55. Nakamura T, Yagi Y, Kobayashi K. Mechanistic insight into pentatricopeptide repeat proteins as sequence-specific RNA-binding proteins for organellar RNAs in plants. Plant Cell Physiol 2012; 53:1171-9; PMID:22576772; http://dx.doi.org/10.1093/pcp/pcs069.
56. Barkan A, Rojas M, Fujii S, Yap A, Chong YS, Bond CS, et al. A combinatorial amino acid code for RNA recognition by pentatricopeptide repeat proteins. PLoS Genet 2012; 8:e1002910; PMID:22916040; http://dx.doi.org/10.1371/journal.pgen.1002910.
57. Yagi Y, Hayashi S, Kobayashi K, Hirayama T, Nakamura T. Elucidation of the RNA recognition code for pentatricopeptide repeat proteins involved in organelle RNA editing in plants. PLoS One 2013; 8:e57286; PMID:23472078; http://dx.doi.org/10.1371/journal.pone.0057286.
58. Gupta YK, Nair DT, Wharton RP, Aggarwal AK. Structures of human Pumilio with noncognate RNAs reveal molecular mechanisms for binding promiscuity. Structure 2008; 16:549-57; PMID:18328718; http://dx.doi.org/10.1016/j.str.2008.01.006.
59. Wang Y, Opperman L, Wickens M, Hall TM. Structural basis for specific recognition of multiple mRNA targets by a PUF regulatory protein. Proc Natl Acad Sci USA 2009; 106:20186-91; PMID:19901328; http://dx.doi.org/10.1073/pnas.0812076106.
60. Ringel R, Sologub M, Morozov YI, Litonin D, Cramer P, Temiakov D. Structure of human mitochondrial RNA polymerase. Nature 2011; 478:269-73; PMID:21947009; http://dx.doi.org/10.1038/nature10435.
61. Howard MJ, Lim WH, Fierke CA, Koutmos M. Mitochondrial ribonuclease P structure provides insight into the evolution of catalytic strategies for precursor-tRNA 5' processing. Proc Natl Acad Sci USA 2012; 109:16149-54; PMID:22991464; http://dx.doi.org/10.1073/pnas.1209062109.
62. Sanchez MI, Mercer TR, Davies SM, Shearwood AM, Nygård KK, Richman TR, et al. RNA processing in human mitochondria. Cell Cycle 2011; 10:2904-16; PMID:21857155; http://dx.doi.org/10.4161/cc.10.17.17060.
63. Maraia RJ, Lamichhane TN. 3' processing of eukaryotic precursor tRNAs. Wiley Interdiscip Rev RNA 2011; 2:362-75; PMID:21572561; http://dx.doi.org/10.1002/wrna.64.
64. Rossmanith W. Of P and Z: mitochondrial tRNA processing enzymes. Biochim Biophys Acta 2012; 1819:1017-26; PMID:22137969; http://dx.doi.org/10.1016/j.bbagrm.2011.11.003.
65. Harris ME, Christian EL. Recent insights into the structure and function of the ribonucleoprotein enzyme ribonuclease P. Curr Opin Struct Biol 2003; 13:325-33; PMID:12831883; http://dx.doi.org/10.1016/S0959-440X(03)00069-1.
66. Gutmann B, Gobert A, Giegé P. PRORP proteins support RNase P activity in both organelles and the nucleus in Arabidopsis. Genes Dev 2012; 26:1022-7; PMID:22549728; http://dx.doi.org/10.1101/gad.189514.112.
67. Gobert A, Pinker F, Fuchsbauer O, Gutmann B, Boutin R, Roblin P, et al. Structural insights into protein-only RNase P complexed with tRNA. Nat Commun 2013; 4:1353; PMID:23322041; http://dx.doi.org/10.1038/ncomms2358.
68. Taschner A, Weber C, Buzet A, Hartmann RK, Hartig A, Rossmanith W. Nuclear RNase P of Trypanosoma brucei: a single protein in place of the multicomponent RNA-protein complex. Cell Rep 2012; 2:19-25; PMID:22840392; http://dx.doi.org/10.1016/j.celrep.2012.05.021.
69. Chujo T, Ohira T, Sakaguchi Y, Goshima N, Nomura N, Nagao A, et al. LRPPRC/SLIRP suppresses PNPase-mediated mRNA decay and promotes polyadenylation in human mitochondria. Nucleic Acids Res 2012; 40:8033-47; PMID:22661577; http://dx.doi.org/10.1093/nar/gks506.
70. Ruzzenente B, Metodiev MD, Wredenberg A, Bratic A, Park CB, Cámara Y, et al. LRPPRC is necessary for polyadenylation and coordination of translation of mitochondrial mRNAs. EMBO J 2012; 31:443-56; PMID:22045337; http://dx.doi.org/10.1038/emboj.2011.392.
71. Sasarman F, Brunel-Guitton C, Antonicka H, Wai T, Shoubridge EA; LSFC Consortium. LRPPRC and SLIRP interact in a ribonucleoprotein complex that regulates posttranscriptional gene expression in mitochondria. Mol Biol Cell 2010; 21:1315-23; PMID:20200222; http://dx.doi.org/10.1091/mbc.E10-01-0047.
72. Härtel B, Zehrmann A, Verbitskiy D, van der Merwe JA, Brennicke A, Takenaka M. MEF10 is required for RNA editing at nad2-842 in mitochondria of Arabidopsis thaliana and interacts with MORF8. Plant Mol Biol 2013; 81:337-46; PMID:23288601; http://dx.doi.org/10.1007/s11103-012-0003-2.
73. Takenaka M, Zehrmann A, Verbitskiy D, Kugelmann M, Härtel B, Brennicke A. Multiple organellar RNA editing factor (MORF) family proteins are required for RNA editing in mitochondria and plastids of plants. Proc Natl Acad Sci USA 2012; 109:5104-9; PMID:22411807; http://dx.doi.org/10.1073/pnas.1202452109.
74. Bentolila S, Heller WP, Sun T, Babina AM, Friso G, van Wijk KJ, et al. RIP1, a member of an Arabidopsis protein family, interacts with the protein RARE1 and broadly affects RNA editing. Proc Natl Acad Sci USA 2012; 109:E1453-61; PMID:22566615; http://dx.doi.org/10.1073/pnas.1121465109.
75. Boussardon C, Salone V, Avon A, Berthomé R, Hammani K, Okuda K, et al. Two interacting proteins are necessary for the editing of the NdhD-1 site in Arabidopsis plastids. Plant Cell 2012; 24:3684-94; PMID:23001034; http://dx.doi.org/10.1105/tpc.112.099507.
76. Rackham O, Filipovska A. The role of mammalian PPR domain proteins in the regulation of mitochondrial gene expression. Biochim Biophys Acta 2012; 1819:1008-16; PMID:22051507; http://dx.doi.org/10.1016/j.bbagrm.2011.10.007.
77. Lipinski KA, Puchta O, Surendranath V, Kudla M, Golik P. Revisiting the yeast PPR proteins--application of an Iterative Hidden Markov Model algorithm reveals new members of the rapidly evolving family. Mol Biol Evol 2011; 28:2935-48; PMID:21546354; http://dx.doi.org/10.1093/molbev/msr120.
78. Kühl I, Dujeancourt L, Gaisne M, Herbert CJ, Bonnefoy N. A genome wide study in fission yeast reveals nine PPR proteins that regulate mitochondrial gene expression. Nucleic Acids Res 2011; 39:8029-41; PMID:21727087; http://dx.doi.org/10.1093/nar/gkr511.
79. Lightowlers RN, Chrzanowska-Lightowlers ZM. PPR (pentatricopeptide repeat) proteins in mammals: important aids to mitochondrial gene expression. Biochem J 2008; 416:e5-6; PMID:18939947; http://dx.doi.org/10.1042/BJ20081942.
80. Hammani K, Gobert A, Small I, Giegé P. A PPR protein involved in regulating nuclear genes encoding mitochondrial proteins? Plant Signal Behav 2011; 6:748-50; PMID:21455023; http://dx.doi.org/10.4161/psb.6.5.15148.
81. Mingler MK, Hingst AM, Clement SL, Yu LE, Reifur L, Koslowsky DJ. Identification of pentatricopeptide repeat proteins in Trypanosoma brucei. Mol Biochem Parasitol 2006; 150:37-45; PMID:16837079; http://dx.doi.org/10.1016/j.molbiopara.2006.06.006.
82. Carthew RW, Sontheimer EJ. Origins and Mechanisms of miRNAs and siRNAs. Cell 2009; 136:642-55; PMID:19239886; http://dx.doi.org/10.1016/j.cell.2009.01.035.
83. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. Genome Res 2004; 14:1188-90; PMID:15173120; http://dx.doi.org/10.1101/gr.849004.
84. Davies SM, Rackham O, Shearwood AM, Hamilton KL, Narsai R, Whelan J, et al. Pentatricopeptide repeat domain protein 3 associates with the mitochondrial small ribosomal subunit and regulates translation. FEBS Lett 2009; 583:1853-8; PMID:19427859; http://dx.doi.org/10.1016/j.febslet.2009.04.048.
85. Okuda K, Chateigner-Boutin AL, Nakamura T, Delannoy E, Sugita M, Myouga F, et al. Pentatricopeptide repeat proteins with the DYW motif have distinct molecular functions in RNA editing and RNA cleavage in Arabidopsis chloroplasts. Plant Cell 2009; 21:146-56; PMID:19182104; http://dx.doi.org/10.1105/tpc.108.064667.