Family of pentatricopeptide repeat proteins in non-plant organisms

Authors

Czubak Pawel[1]

Raczynska Katarzyna Dorota[1]*

Affiliation

[1]Department of Gene Expression, Institute of Molecular Biology and Biotechnology, Adam Mickiewicz University in Poznan, Poznan, Poland.

Correspondence

Raczynska Dorota, +48 61 829 5953; Fax: +48 61 829 5949; Email: doracz@amu.edu.pl

Abstract

Pentatricopeptide repeat (PPR) proteins belong to the large family of nucleic acid binding proteins. Their name is defined by the presence of the so-called pentatricopeptide repeat (PPR), a degenerate 35-amino acid motif tandemly repeated 2–26 times per protein. The PPR protein family includes some of the major mediators of organelle post-transcriptional control. These proteins localize mainly to mitochondria and chloroplasts, where they are involved in different steps of RNA metabolism: transcription, splicing, editing, polyadenylation, RNA stability and translation. Some of them can also act as adaptors in protein-protein interactions. PPR protein genes were discovered in many eukaryotic genomes, but they are especially abundant in plants. In humans, some PPR proteins are implicated in human diseases, such as cancer, Alzheimer's disease and Parkinson's disease.

Abbreviations

AD - Alzheimer's Disease; APP - amyloid beta precursor protein; APPsw - APP carrying the so-called Swedish mutation; DmLRPPRC - *Drosophila melanogaster*'s leucine-rich pentatricopeptide repeat-containing homologue; HAT - half-A-TPR; KPAF1 - kinetoplast polyadenylation/uridylation factor 1; KPAF2 - kinetoplast polyadenylation/uridylation factor 2; KPAP1- kinetoplast poly(A) polymerase 1; KRIPP - kinetoplast ribosomal PPR protein; LSFC - Leigh syndrome French Canadian; LRPPRC - leucine-rich PPR cassette; LRRK2 (PARK8) - leucine rich repeat kinase 2; MRPS27 - mitochondrial ribosomal protein of the small subunit 27; MRPP3 - mitochondrial RNase P protein 3; PD-Parkinson's disease; POLRMT - RNA polymerase, mitochondrial; PPR - pentatricopeptide repeat; PTCD - pentatricopeptide repeat domain; PSEN – presenilin; PTEN - phosphatase and tensin; PUF - Pumilio and FBF homology; RBP – RNA binding protein; RET1 TUTase - RNA editing TUTase 1; SEL-1 - suppressor of Lin-12; SLIRP - SRA - stem-loop-interacting RNA binding protein; TALE - transcription activator-like effector; TMG - tRNA guanine-N7 methyltransferase; TPR – tetratricopeptide repeat.

Pentatricopeptide repeat proteins

Pentatricopeptide repeat proteins were first described in 2000 as "PPR cassette proteins" by Small and Peeters (1). The main feature of PPR proteins is the presence of a 35-amino acid (pentatricopeptide) structural motif that is tandemly repeated 2–26 times per protein (1). Initially, PPR proteins were thought to bind proteins. Further investigations revealed that the PPR domain is involved in RNA–protein interactions (1, 2). Therefore, PPR proteins belong to a class of RNA binding proteins (RBPs) referred to as heterogeneous ribonucleoproteins that associate with certain transcripts and influence all steps of post-transcriptional regulation of RNAs, including pre-mRNA splicing, mRNA polyadenylation, localization, stability and translation. Some RBPs can mediate RNA transport from the nucleus to mitochondria and from mitochondria to the nucleus (3-6).

PPR proteins are found in all eukaryotes, though they are most prevalent in plants (7-11). For example, in *Arabidopsis thaliana*, there are 441 PPR proteins. They belong to one of the largest nucleus-encoded helical repeat protein families in plants (1, 12-15). Some PPR-encoding genes have also been found in prokaryotes, including pathogenic and symbiotic members of the genera *Rhodobacter*, *Ralstonia*, *Simkania*, *Erwinia* and *Legionella*. They are thought to have been acquired *via* eukaryote-to-prokaryote horizontal gene transfer events (2, 10, 16-19).

Pentatricopeptide repeat proteins structure

The PPR is a degenerate, 35-amino-acid motif repeated in tandem. The number of motifs ranges from 2 to over 26 within a protein (1, 2). The PPR motif forms two anti-parallel α-helices that interact to produce a helix-turn-helix motif. The series of helix-turn-helix motifs throughout the protein stack on each other to form an elongated structure called a superhelix, with a central groove that binds RNA. Residues at position 4 and 34 of each PPR motif are responsible for modular recognition of sequence motifs within the transcript (Fig.1) (2, 20-23).

The sequence of the PPR motif is similar to the closely related tetratricopeptide repeat (TPR) motif that is prevalent in prokaryotes. Therefore, some have hypothesized that the PPR motif emerged from the TPR motif during the early stages of eukaryotic evolution (10). PPR motifs, together with the TPR motif, the suppressor of Lin-12 (SEL-1)–like motif and the half-A-TPR (HAT) repeats, belong to a large family of solenoid repeat structures formed of α-α repeats (1, 24, 25). Similar to Pumilio and FBF homology (PUFs) and transcription activator-like effector (TALE) proteins, PPR proteins interact with nucleic acids in a sequence-dependent manner. Therefore, they are very specific according to the locations of individual editing sites within organelle transcriptomes (26-28). In contrast, the interaction of TPR proteins with RNA is sequence independent (29, 30).

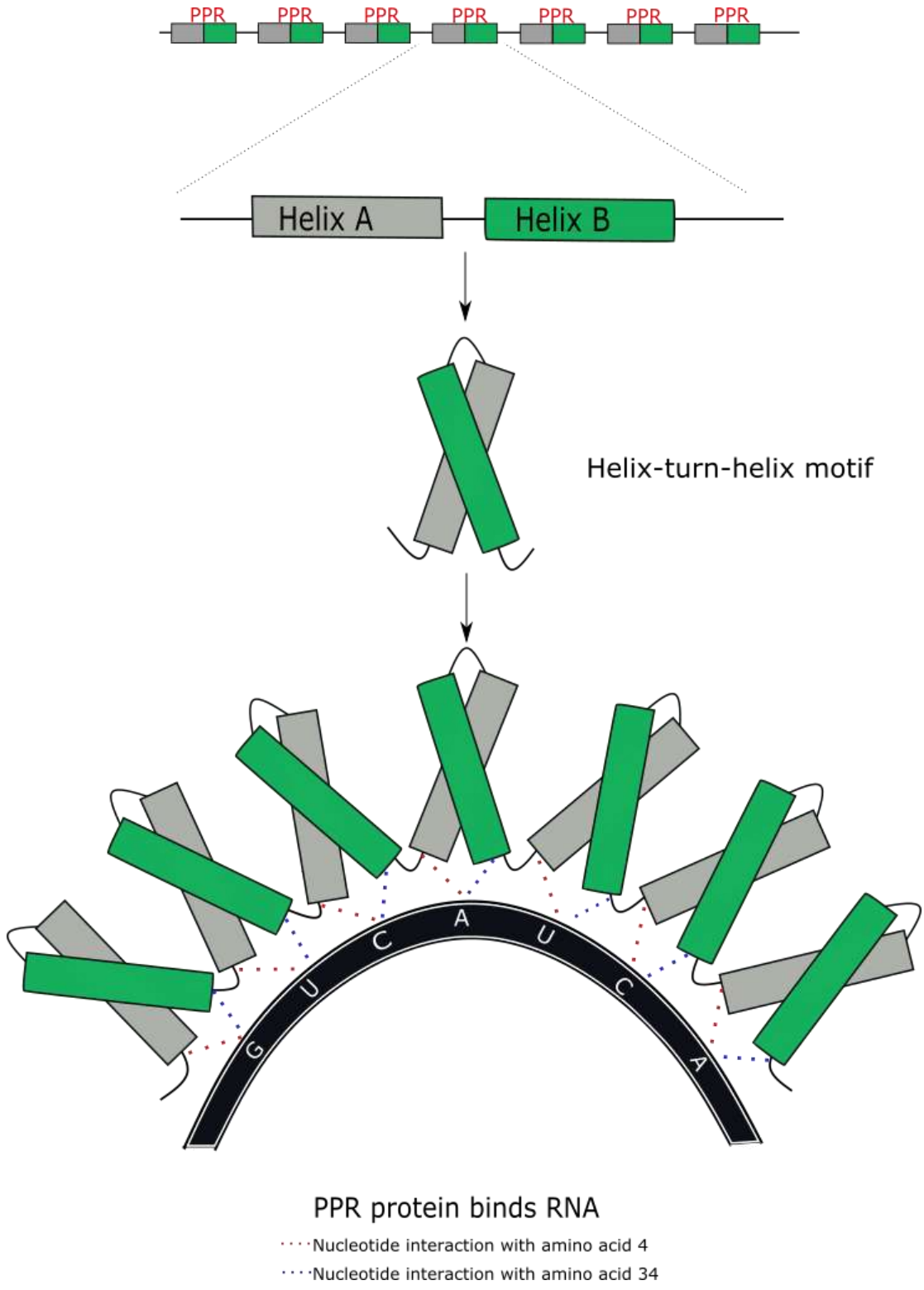

Figure 1. Schematic structure of the PPR protein and the mechanism of RNA recognition. Each helix-turn-helix motif is formed by two α-helices included in the PPR domain. Helix-turn-helix motifs form a superhelix, containing an RNA binding groove. Amino acids at positions 4 and 34 of each PPR motifs are responsible for modular transcripts recognition (adopted from (31)).

Yeast PPR motifs are quite similar to plant PPR motifs, although in yeasts, the PPR motifs are significantly more divergent. Thus, according to the evolutionary distance between yeasts and plants and the relatively sparse knowledge about yeast PPRs, we cannot assume that substrate recognition uses the same residues in the PPR motifs. Similarly, *Trypanosoma* motifs and human motifs also show evolutionary divergence, and their sequences are different from the general motif consensus observed in plants. However, although the PPR motifs in different organisms differ, the key residues are often partially conserved (7, 32).

There are two PPR classes that differ by their domain architecture: P-class and PLS-class. P-class PPR proteins possess the canonical 35-amino-acid motifs without additional domains (P means canonical PPR). Members of this class play roles in most aspects of organelle gene expression and commonly regulate RNA processing, splicing, stability and translation. PLS-class PPR proteins have three different types of PPR motifs, which vary in length: P (PPR, 35 amino acids), L (long, 35-36 amino acids) and S (short, ~31 amino acids). Members of this class mainly function in RNA editing (2, 31, 33). Subtypes of the PLS class are further categorized based on additional C-terminal domains: motif E (91 amino acids), motif E+ (33 amino acids) and motif DYW (106 aminoacids). The first PLS subclass contains motif E, which is not catalytic but is predicted to be a protein-protein interaction motif that recruits the editing enzyme. The second PLS subclass contains E, E+ and DYW domains. The DYW domain was named for its aspartate-tyrosine-tryptophan (DYW) tripeptide located at the C-terminus. DYW domains contain the same conserved catalytic residues as cytidine deaminase enzymes, which mediate C-to-U conversion. Thus, the DYW domain has been proposed to be an editing domain. Therefore, the presence of PLS class PPR proteins correlates with the occurrence of organelle RNA editing, and they are absent in organisms where organelle RNA editing does not occur. The PLS class was originally thought to be specific to plants, but the DYW-containing PPR proteins were subsequently identified in the heterolobosean protist *Naegleria gruberi* and later in several non-plant organisms. Therefore, DYW-type PPR proteins are thought to have been acquired from plants *via* horizontal gene transfer events.

It is worth mentioning that PPR proteins differ according to the content of domains. Some of them are built almost entirely of tandem PPRs, whereas others contain different domains, such as endonuclease or protein interaction domains (2, 34, 35). For example, the PPR protein in *Trypanosoma brucei*, PtcE, contains a C-terminal methyltransferase domain. As the domain is thought to be active, PtcE, like PPR proteins in *D. discoideum* mitochondria, is a potential candidate that can modify the stability of mitochondrial transcripts by introducing posttranscriptional modification (9).

In 2008, the crystal structure of human mitochondrial RNA polymerase II, also called POLRMT (RNA polymerase mitochondrial), was revealed (36). It contains a T7-like catalytic C-terminal domain, an N-terminal domain (with a T7 promoter binding domain) and a flexible N-terminal extension. The N-terminal domain consists of nine α-helices, four of which comprise two PPR motifs found in

plant mitochondrial proteins. Each PPR motif forms the hydrophobic core of a helix–turn–helix fold and the interface between parallel and antiparallel α-helices (36).

To better understand the mechanism of PPR-RNA binding, Rackham's group designed synthetic PPR domains based on the conservation of residues within PPRs throughout evolution (37). Interestingly, they were able to generate different PPR variants with altered affinities or degenerate nucleotide recognition properties by introducing some changes in PPR sequence. For example, the presence of phenylalanine at position 1 and asparagine at positions 4 and 34 favoured binding to adenine and guanine, while valine at position 1 favoured binding to adenine and cytosine. Therefore, the synthetic domains could recognize RNA targets of interest in a predictable sequence-specific manner according to the appropriate amino acid at positions 1, 4 and 34 (37). Interestingly, it was also suggested that the overall shape of the PPR scaffold partially depends on the amino acids that mediate RNA association. Moreover, mutual interactions of PPR repeats resulted in some plasticity of the PPR scaffold that enabled it to modulate the RNA binding properties of PPRs (37).

Pentatricopeptide repeat protein functions

PPR proteins are mainly localized to mitochondria, where they play various roles at different steps of RNA metabolism: transcription, transcript processing, splicing, editing, polyadenylation, RNA stability and translation. Regulation of organelle gene expression at the post-transcriptional level depends on RNA binding proteins because of the low number of promoters in organelle genomes and the long half-life of organelle RNAs. In this way, members of the PPR protein family are some of the major mediators of expression of organelle genomes and organelle biogenesis (1, 2, 12, 31, 38).

PPR proteins can also be non-catalytic, in which case they function as molecular adaptors in the recruitment of catalytic enzymes or effector proteins to target transcripts (39, 40). However, they can also catalyse RNA processing and editing themselves (2, 26). Several PPR proteins act as stability and/or translation factors, and their absence leads to both the destabilization of a given mitochondrial mRNA and a defect in the accumulation of the corresponding protein. However, it is still not clear whether these factors have a dual role in stability and translation, as was shown, for example, for PPR10 in *Arabidopsis* chloroplasts. An alternative hypothesis assume that i) a primary RNA stability defect causes a secondary translation deficiency or ii) a primary defect in translation destabilizes the mitochondrial mRNA (32).

The most common function of plant PPR proteins is RNA editing (8, 13, 38). In contrast, in *Trypanosoma brucei*, approximately half of the 39 PPR proteins identified have been detected in affinity-purified ribosomal particles from the kinetoplast, a dense nucleoprotein structure containing the mitochondrial genome. They were termed kinetoplast ribosomal PPR proteins (KRIPPs), and they are considered translational activators for a subset of mitochondrial mRNAs (41, 42). Some PPR proteins, such as kinetoplast polyadenylation/uridylation factors (KPAF) 1 and 2, are involved in stabilization and polyadenylation of

kinetoplast transcripts. They have been identified in the polyadenylation complex and facilitate the A/U-tailing of mRNA 3'ends by kinetoplast poly(A) polymerase 1 (KPAP1) and RNA editing TUTase 1 (RET1 TUTase). Others, such as KRIPP1 and KRIPP8, have been identified both in the polyadenylation complex and in the small ribosomal subunit. Therefore, PPR proteins can serve as protein adaptors in regulatory networks that channel mRNAs for 3′ A/U-tailing and translation (41-43).

Although not as numerous as in plants and trypanosomes, PPR proteins from other organisms also participate in translational control of mitochondrial mRNAs. In *S. cerevisiae* and closely related species, PPR proteins facilitate initiation of translation by recognizing and binding to the 5′ UTRs of specific mRNAs. Thus, they act as mitochondrial mRNA-specific translational activators and are a part of the feedback control loop that adjusts the translation rate to the assembly of nascent polypeptides into respective respiratory complexes (32, 44). In humans, leucine-rich pentatricopeptide repeat-containing (LRPPRC) protein stimulates poly(A) polymerase for mRNA polyadenylation to protect protein-encoding transcripts from the mitochondrial degradosome activity. Similarly, *Drosophila melanogaster*'s homologue DmLRPPRC1 is necessary for mitochondrial mRNA polyadenylation. Loss of the LRPPRC or DmLRPPRC protein leads to decreased stability of mitochondrial transcripts and reduction of their abundance (45-49).

LRPPRC associates with SRA stem-loop-interacting RNA binding protein (SLIRP), a mitochondrial RBP. Both proteins are part of a high-molecular-mass ribonucleoprotein complex that regulates the metabolism of mitochondrial transcripts (46). In addition to mitochondria, LRPPRC is present in the nucleus, where it is involved in the regulation of mitochondrial biogenesis and energy homeostasis by controlling nuclear gene expression. In this way, LRPPRC might play a role in coordinating mitochondrial transcription with nuclear gene expression, which is critical for proper organelle function (50).

Recently, a whole new family of PPR proteins - PPR-TGM (PPR-tRNA guanine-N7 methyltransferase) proteins, were discovered. What is interesting, PPR-TGM proteins are missing in plants and animals. Instead, they are found in single-celled eukaryotic microbes organisms, including cellular slime moulds, entamoebae, algae and diatoms. In addition to the PPR tract, these proteins contain a C-terminal tRNA guanine-N7 methyltransferase domain originated from a chlamydial TGM-encoding gene, acquired *via* horizontal gene transfer from bacteria. The domain architecture of PPR-TGM proteins suggests they function in tRNA metabolism both in mitochondria, chloroplasts and in the cytoplasm (51).

In mammals, seven PPR proteins have been identified to date: POLRMT, pentatricopeptide repeat domain 1 (PTCD1), PTCD2, PTCD3, LRPPRC, mitochondrial ribosomal protein S27 (MRPS27) and mitochondrial RNase P protein 3 (MRPP3). They are mainly localized in mitochondria and regulate transcription, transcript processing, RNA stability and translation, not RNA editing (Table 1) (31, 52-58). However, the physiological targets of PPR proteins are often unknown, which limits the understanding of their mechanistic roles in organelle gene expression and energy

metabolism.

Table 1. PPR proteins and their function in human cells.

Protein	Short name	Function	Gene location
Mitochondrial RNA polymerase II	POLRMT	mitochondrial gene expression, initiation of the mitochondrial genome replication	19p13.3
Pentatricopeptide repeat domain 1	PTCD1	RNA binding, regulation of translation	7q22.1
Pentatricopeptide repeat domain 2	PTCD2	mitochondrial RNA maturation, mitochondrial respiratory chain function	5q13.2
Pentatricopeptide repeat domain 3	PTCD3	mitochondrial translation, organelle biogenesis and maintenance	2p11.2
Leucine-rich pentatricopeptide-repeat containing protein	LRPPRC	cytoskeletal organization, vesicular transport, transcriptional regulation of both nuclear and mitochondrial genes	2p21
Mitochondrial ribosomal protein S27	MRPS27	mitochondrial protein synthesis, stimulation of mitochondrial mRNA translation of subunit components of the mitochondrial electron transport chain	5q13.2
Mitochondrial RNase P protein 3	KIAA0391 (MRPP3)	mitochondrial tRNA maturation	14q13.2

Pentatricopeptide repeat proteins in medicine

LRPPRC is one of the best-studied PPR proteins, as it is implicated in human diseases, such as cancer. The connection between cancer and mitochondria is based on the accelerated energy consumption of tumour cells. As LRPPRC plays an important role in mitochondrial function, it is also necessary for tumour development. Indeed, LRPPRC is abundantly expressed in various types of tumours, such as lung adenocarcinoma cell lines, oesophageal squamous cell carcinoma, lung adenocarcinoma, stomach adenocarcinoma, colon adenocarcinoma, mammary adenocarcinoma, endometrial adenocarcinoma and lymphoma (59, 60). Moreover, the overexpression of LRPPRC confers resistance to apoptosis of tumour cells (60). In agreement with this, knock-down of LRPPRC reduced the anti-apoptosis, invasion, and *in vitro* colony formation abilities of lung adenocarcinoma and Hodgkin lymphoma cells. This suggests that the knock-down of LRPPRC could be one strategy to prevent carcinogenesis (60).

LRPPR proteins can also be used as a specific biomarkers, for instance, in prostate cancer. Prostate cancer is the most common non-cutaneous malignancy and the second leading cause of cancer death among men in the United States (61). By immunochemistry analysis Jiang et al. (62) showed that prostate tissues at late stage of prostate adenocarcinomas express higher level of LRPPRC. Moreover, LRPPRC levels were significantly lower in patients surviving longer than five or ten years than in patients surviving shorter than five or ten years. From this reason the level of LRPPRC can be a biomarker for late-stage prostate adenocarcinomas patients with poor prognosis (62). They also confirmed

the association of high LRPPRC levels with hormone therapy insensitivity in prostate adenocarcinomas tissue samples collected from prostate-specific phosphatase and tensin homolog (PTEN) deficient mice or hormone-dependent and independent prostate cancer cell lines (62).

Mitochondrial pathophysiology due to mutation in the LRPPRC gene underlies a human neurodegenerative disorder, the French-Canadian variant of Leigh syndrome (Leigh syndrome French Canadian; LSFC). Individuals with LSFC often appear normal at birth but begin to lose basic skills such as head control, sucking, walking, and talking in infancy or early childhood. They may also present with intellectual disabilities, dysmorphic features, irritability, vomiting, and seizures. At the molecular level, LSFC is caused by a point mutation in the gene encoding LRPPRC that affects the stability of most mitochondrial mRNAs. However, LSFC's pathophysiological effect has mainly been attributed to a reduced level of cytochrome *c* oxidase mRNA. This in turn results in defective cytochrome oxidase assembly and deficiencies of mitochondrial complex I, complex IV and ATP synthase. As a final consequence, mitochondrial respiration and ATP production are impaired (63-66). At the molecular level, the LRPPRC protein shows some homology to the yeast protein PET309, which is also required for the efficient expression of the*COX* gene (63, 66).

LRPPRC protein is also thought to be involved in the occurrence of sporadic Alzheimer's disease (AD) cases. The Alzheimer's Disease is characterized by inflammatory changes, accumulation of misfolded proteins and oxidative damage, which results in region-specific loss of synaptic contacts and neuronal cell death. The etiology of AD is mostly sporadic (95% of the cases) and have a late onset at about the age of 65. The risk of AD occurrence can be altered by various environmental exposures, such as nutrition, smoking, head trauma, infections or systemic inflammation, as well as psychosocial factors, such as education or physical activity (67). AD pathogenesis may also be due to epigenetic mechanisms. Studies on human postmortem brain samples, peripheral leukocytes and transgenic animal models have revealed that AD is strongly linked to aging and epigenetic deregulations, including abnormal DNA methylation and histone modifications (67, 68).

In contrast, in case of the familial form of AD (5% of cases) some early onset cases were reported. Familiar form of AD is connected with the mutations in the genes encoding amyloid precursor protein (APP), presenilin-1 (PSEN1) and presenilin-2 (PSEN2) (69). LRPPRC preferentially interacts with amyloid beta precursor protein carrying the so-called Swedish mutation (APPsw). Such interaction is responsible for some cases of early-onset Alzheimer's disease (70).

Examination of the brains from Alzheimer patients compared with controls shows an elevated levels of another human PPR protein, PTCD2, in the cerebral cortex of AD patients (71). Therefore, PTCD2 can also be used as a biomarker for Alzheimer disease (AD).

Parkinson's disease (PD) is another neurodegenerative disorder linked to mutation in PPR protein gene. Parkinson's disease is characterized by tremor, bradykinesia, rigidity and postural instability. First symptom of about 70% cases of disease is motoric instabilities with resting tremor at a typical frequency

of 3–5 Hz (72). PD neurodegeneration is thought to be caused by genetic and environmental risk factors and interactions between them, that leads to the neuronal death. As a risk factor of PD some environmental toxins (such as rotenone and paraquat) and other factors causing mitochondrial dysfunction have been reported. 90% of PD cases are sporadic, the remaining 10% of cases show familiar inheritance. In the latter cases, mutations in the leucine rich repeat kinase 2 (LRRK2, PARK8) are the most frequent (73).

Parkinson's disease (PD) is also associated with mutations in the pentatricopeptide repeat protein gene - *LRPPRC* gene. At the molecular level, four intronic substitutions influence the splicing of the *LRPPRC* pre-mRNA, leading to skipping of some exons (e.g., exon 9 or exon 33). This leads to the creation of a premature stop codon that targets the mature mRNA for further degradation by the nonsense mediated decay (NMD) pathway. LRPPRC protein depletion causes mitochondrial dysfunction. Interestingly, due to the tissue-specific availability of some splicing factors, a given mutation can have different effects in different tissues. LRPPRC mutations linked to the origin and progression of disease are observed mainly in organs where the LRPPRC steady-state level is quite low. Moreover, the clinical impact of the LRPPRC mutation depends on whether the mutation is homozygous or heterozygous. Finally, some intronic LRPPRC variants could lead to complete or partial LRPPRC protein depletion or to its functional changes, which can influence the risk of Parkinson's disease and other neurodegenerative disorders (e.g., LSFC) (66, 74, 75).

Nowadays, when we are able to manipulate or design PPR proteins that recognize any RNA sequence we want, the possible application in medicine are opened. By introducing mutations into the coding sequence of known PPR protein we can change the crucial amino acids and direct the binding of the PPR protein to an RNA sequence of interest (22). Moreover, synthetic PPR protein (37) can be used to replace a dysfunctional PPR protein, to regulate endogenous gene expression or even to prevent viral infections by blocking the synthesis of viral RNA. One of the most interesting are PPR-TGM proteins, that are important to the clinically-relevant pathogens but are absent in humans. Therefore, PPR-TGM proteins represent new putative targets for the development of drugs that can specifically inhibit their activity in pathogenic eukaryotes. Drugs, developed to block PPR-TGM function in parasites, can effectively treat parasitic infections without any effects on the human body (51).

Conclusions

PPR proteins are important for expression of organelle genomes and organelle biogenesis. Some of them are implicated in human diseases, such as cancer, Alzheimer's disease and Parkinson's disease. Further analysis of their functions and targets will help to describe the molecular pathways underlying these human disorders. The precise knowledge of PPR proteins' actions and the molecular consequences of their disruption can also be useful in designing therapeutic strategies for patients. A promising avenue for future research is synthetic domains based on PPR motifs

that can recognize RNA targets of interest in a predictable, sequence-specific manner. Therefore, PPR proteins could be used to regulate particular gene transcripts' expression.

Acknowledgments

Funding
This work was supported by: Polish Science Centre under grant UMO-2016/21/B/NZ1/00232 (to KDR); KNOW RNA Research Centre in Poznan under grant 01/KNOW2/2014 (to PC, KDR).

References

1. Small ID, Peeters N. The PPR motif - a TPR-related motif prevalent in plant organellar proteins. Trends Biochem Sci. 2000;25(2):46-7.
2. Schmitz-Linneweber C, Small I. Pentatricopeptide repeat proteins: a socket set for organelle gene expression. Trends Plant Sci. 2008;13(12):663-70.
3. Dong Y, Yoshitomi T, Hu JF, Cui J. Long noncoding RNAs coordinate functions between mitochondria and the nucleus. Epigenetics Chromatin. 2017;10(1):41.
4. Moore MJ. From birth to death: the complex lives of eukaryotic mRNAs. Science. 2005;309(5740):1514-8.
5. Dreyfuss G, Kim VN, Kataoka N. Messenger-RNA-binding proteins and the messages they carry. Nat Rev Mol Cell Biol. 2002;3(3):195-205.
6. Keene JD. RNA regulons: coordination of post-transcriptional events. Nat Rev Genet. 2007;8(7):533-43.
7. 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(10):2935-48.
8. Shikanai T. RNA editing in plant organelles: machinery, physiological function and evolution. Cell Mol Life Sci. 2006;63(6):698-708.
9. Manna S, Brewster J, Barth C. Identification of Pentatricopeptide Repeat Proteins in the Model Organism Dictyostelium discoideum. Int J Genomics. 2013;2013:586498.
10. Barkan A, Small I. Pentatricopeptide repeat proteins in plants. Annu Rev Plant Biol. 2014;65:415-42.
11. O'Toole N, Hattori M, Andres C, Iida K, Lurin C, Schmitz-Linneweber C, et al. On the expansion of the pentatricopeptide repeat gene family in plants. Mol Biol Evol. 2008;25(6):1120-8.
12. Raczynska KD, Augustyniak H. [Family of pentatricopeptide repeat proteins]. Postepy Biochem. 2005;51(4):440-6.
13. Hassani D, Khalid M, Bilal M, Zhang YD, Huang D. Pentatricopeptide repeat-directed RNA editing and their biomedical applications. International Journal of Pharmacology. 2017;13(7):762-72.
14. Filipovska A, Rackham O. Pentatricopeptide repeats: modular blocks for building RNA-binding proteins. RNA Biol. 2013;10(9):1426-32.
15. Lurin C, Andres C, Aubourg S, Bellaoui M, Bitton F, Bruyere C, et al. Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. Plant Cell. 2004;16(8):2089-103.
16. Cazalet C, Gomez-Valero L, Rusniok C, Lomma M, Dervins-Ravault D, Newton HJ, et al. Analysis of the Legionella longbeachae genome and transcriptome uncovers unique strategies to cause Legionnaires' disease. PLoS Genet. 2010;6(2):e1000851.
17. Choudhary M, Zanhua X, Fu YX, Kaplan S. Genome analyses of three strains of Rhodobacter sphaeroides: evidence of rapid evolution of chromosome II. J Bacteriol. 2007;189(5):1914-21.
18. Hallam SJ, Putnam N, Preston CM, Detter JC, Rokhsar D, Richardson PM, et al. Reverse methanogenesis: testing the hypothesis with environmental genomics. Science. 2004;305(5689):1457-62.
19. Schallenberg-Rudinger M, Lenz H,

Polsakiewicz M, Gott JM, 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(9):1549-56.

20. Ke J, Chen RZ, Ban T, Zhou XE, Gu X, Tan MH, et al. Structural basis for RNA recognition by a dimeric PPR-protein complex. Nat Struct Mol Biol. 2013;20(12):1377-82.

21. Yin P, Li Q, Yan C, Liu Y, Liu J, Yu F, et al. Structural basis for the modular recognition of single-stranded RNA by PPR proteins. Nature. 2013;504(7478):168-71.

22. 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(8):e1002910.

23. 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(3):e57286.

24. Main ER, Stott K, Jackson SE, Regan L. Local and long-range stability in tandemly arrayed tetratricopeptide repeats. Proc Natl Acad Sci U S A. 2005;102(16):5721-6.

25. Kobe B, Kajava AV. When protein folding is simplified to protein coiling: the continuum of solenoid protein structures. Trends Biochem Sci. 2000;25(10):509-15.

26. 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(Pt 6):1643-7.

27. 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(9):1930-41.

28. 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 U S A. 2007;104(19):8178-83.

29. Abbas YM, Pichlmair A, Gorna MW, Superti-Furga G, Nagar B. Structural basis for viral 5'-PPP-RNA recognition by human IFIT proteins. Nature. 2013;494(7435):60-4.

30. Pichlmair A, Lassnig C, Eberle CA, Gorna MW, Baumann CL, Burkard TR, et al. IFIT1 is an antiviral protein that recognizes 5'-triphosphate RNA. Nat Immunol. 2011;12(7):624-30.

31. Manna S. An overview of pentatricopeptide repeat proteins and their applications. Biochimie. 2015;113:93-9.

32. Herbert CJ, Golik P, Bonnefoy N. Yeast PPR proteins, watchdogs of mitochondrial gene expression. RNA Biol. 2013;10(9):1477-94.

33. Hall TM. De-coding and re-coding RNA recognition by PUF and PPR repeat proteins. Curr Opin Struct Biol. 2016;36:116-21.

34. Gobert A, Gutmann B, Taschner A, Gossringer M, Holzmann J, Hartmann RK, et al. A single Arabidopsis organellar protein has RNase P activity. Nat Struct Mol Biol. 2010;17(6):740-4.

35. Takenaka M, Zehrmann A, Verbitskiy D, Kugelmann M, Hartel 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 U S A. 2012;109(13):5104-9.

36. Ringel R, Sologub M, Morozov YI, Litonin D, Cramer P, Temiakov D. Structure of human mitochondrial RNA

polymerase. Nature. 2011;478(7368):269-73.
37. Coquille S, Filipovska A, Chia T, Rajappa L, Lingford JP, Razif MF, et al. An artificial PPR scaffold for programmable RNA recognition. Nat Commun. 2014;5:5729.
38. Fujii S, Small I. The evolution of RNA editing and pentatricopeptide repeat genes. New Phytol. 2011;191(1):37-47.
39. Matsushima Y, Hirofuji Y, Aihara M, Yue S, Uchiumi T, Kaguni LS, et al. Drosophila protease ClpXP specifically degrades DmLRPPRC1 controlling mitochondrial mRNA and translation. Sci Rep. 2017;7(1):8315.
40. Barros MH, Tzagoloff A. Aep3p-dependent translation of yeast mitochondrial ATP8. Mol Biol Cell. 2017;28(11):1426-34.
41. Aphasizheva I, Maslov D, Wang X, Huang L, Aphasizhev R. Pentatricopeptide repeat proteins stimulate mRNA adenylation/uridylation to activate mitochondrial translation in trypanosomes. Mol Cell. 2011;42(1):106-17.
42. Aphasizheva I, Maslov DA, Qian Y, Huang L, Wang Q, Costello CE, et al. Ribosome-associated pentatricopeptide repeat proteins function as translational activators in mitochondria of trypanosomes. Mol Microbiol. 2016;99(6):1043-58.
43. Ridlon L, Skodova I, Pan S, Lukes J, Maslov DA. The importance of the 45 S ribosomal small subunit-related complex for mitochondrial translation in Trypanosoma brucei. J Biol Chem. 2013;288(46):32963-78.
44. Herrmann JM, Woellhaf MW, Bonnefoy N. Control of protein synthesis in yeast mitochondria: the concept of translational activators. Biochim Biophys Acta. 2013;1833(2):286-94.
45. Richter-Dennerlein R, Dennerlein S, Rehling P. Integrating mitochondrial translation into the cellular context. Nat Rev Mol Cell Biol. 2015;16(10):586-92.
46. Sasarman F, Brunel-Guitton C, Antonicka H, Wai T, Shoubridge EA, Consortium L. LRPPRC and SLIRP interact in a ribonucleoprotein complex that regulates posttranscriptional gene expression in mitochondria. Mol Biol Cell. 2010;21(8):1315-23.
47. Bratic A, Wredenberg A, Gronke S, Stewart JB, Mourier A, Ruzzenente B, et al. The bicoid stability factor controls polyadenylation and expression of specific mitochondrial mRNAs in Drosophila melanogaster. PLoS Genet. 2011;7(10):e1002324.
48. Ruzzenente B, Metodiev MD, Wredenberg A, Bratic A, Park CB, Camara Y, et al. LRPPRC is necessary for polyadenylation and coordination of translation of mitochondrial mRNAs. EMBO J. 2012;31(2):443-56.
49. 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(16):8033-47.
50. Cooper MP, Qu L, Rohas LM, Lin J, Yang W, Erdjument-Bromage H, et al. Defects in energy homeostasis in Leigh syndrome French Canadian variant through PGC-1alpha/LRP130 complex. Genes Dev. 2006;20(21):2996-3009.
51. 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.
52. Rackham O, Filipovska A. The role of mammalian PPR domain proteins in the regulation of mitochondrial gene

expression. Biochim Biophys Acta. 2012;1819(9-10):1008-16.
53. Siira SJ, Spahr H, Shearwood AJ, Ruzzenente B, Larsson NG, Rackham O, et al. LRPPRC-mediated folding of the mitochondrial transcriptome. Nat Commun. 2017;8(1):1532.
54. Perks KL, Ferreira N, Richman TR, Ermer JA, Kuznetsova I, Shearwood AJ, et al. Adult-onset obesity is triggered by impaired mitochondrial gene expression. Sci Adv. 2017;3(8):e1700677.
55. Lightowlers RN, Chrzanowska-Lightowlers ZM. Human pentatricopeptide proteins: only a few and what do they do? RNA Biol. 2013;10(9):1433-8.
56. Rossmanith W. Of P and Z: mitochondrial tRNA processing enzymes. Biochim Biophys Acta. 2012;1819(9-10):1017-26.
57. Xu F, Ackerley C, Maj MC, Addis JB, Levandovskiy V, Lee J, et al. Disruption of a mitochondrial RNA-binding protein gene results in decreased cytochrome b expression and a marked reduction in ubiquinol-cytochrome c reductase activity in mouse heart mitochondria. Biochem J. 2008;416(1):15-26.
58. Rorbach J, Minczuk M. The post-transcriptional life of mammalian mitochondrial RNA. Biochem J. 2012;444(3):357-73.
59. Seo DC, Sung JM, Cho HJ, Yi H, Seo KH, Choi IS, et al. Gene expression profiling of cancer stem cell in human lung adenocarcinoma A549 cells. Mol Cancer. 2007;6:75.
60. Tian T, Ikeda J, Wang Y, Mamat S, Luo W, Aozasa K, et al. Role of leucine-rich pentatricopeptide repeat motif-containing protein (LRPPRC) for anti-apoptosis and tumourigenesis in cancers. Eur J Cancer. 2012;48(15):2462-73.
61. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. CA Cancer J Clin. 2013;63(1):11-30.
62. Jiang X, Li X, Huang H, Jiang F, Lin Z, He H, et al. Elevated levels of mitochondrion-associated autophagy inhibitor LRPPRC are associated with poor prognosis in patients with prostate cancer. Cancer. 2014;120(8):1228-36.
63. Mootha VK, Lepage P, Miller K, Bunkenborg J, Reich M, Hjerrild M, et al. Identification of a gene causing human cytochrome c oxidase deficiency by integrative genomics. Proc Natl Acad Sci U S A. 2003;100(2):605-10.
64. Mourier A, Ruzzenente B, Brandt T, Kuhlbrandt W, Larsson NG. Loss of LRPPRC causes ATP synthase deficiency. Hum Mol Genet. 2014;23(10):2580-92.
65. Sasarman F, Nishimura T, Antonicka H, Weraarpachai W, Shoubridge EA, Consortium L. Tissue-specific responses to the LRPPRC founder mutation in French Canadian Leigh Syndrome. Hum Mol Genet. 2015;24(2):480-91.
66. Xu F, Morin C, Mitchell G, Ackerley C, Robinson BH. The role of the LRPPRC (leucine-rich pentatricopeptide repeat cassette) gene in cytochrome oxidase assembly: mutation causes lowered levels of COX (cytochrome c oxidase) I and COX III mRNA. Biochem J. 2004;382(Pt 1):331-6.
67. Chouliaras L, Sierksma AS, Kenis G, Prickaerts J, Lemmens MA, Brasnjevic I, et al. Erratum: gene-environment interaction research and transgenic mouse models of Alzheimer's disease. Int J Alzheimers Dis. 2011;2010:356862.
68. Day JJ, Sweatt JD. Cognitive neuroepigenetics: a role for epigenetic mechanisms in learning and memory. Neurobiol Learn Mem. 2011;96(1):2-12.

69. Minati L, Edginton T, Bruzzone MG, Giaccone G. Current concepts in Alzheimer's disease: a multidisciplinary review. Am J Alzheimers Dis Other Demen. 2009;24(2):95-121.
70. Hosp F, Vossfeldt H, Heinig M, Vasiljevic D, Arumughan A, Wyler E, et al. Quantitative interaction proteomics of neurodegenerative disease proteins. Cell Rep. 2015;11(7):1134-46.
71. Acharya NK, Nagele EP, Han M, Coretti NJ, DeMarshall C, Kosciuk MC, et al. Neuronal PAD4 expression and protein citrullination: possible role in production of autoantibodies associated with neurodegenerative disease. J Autoimmun. 2012;38(4):369-80.
72. Jankovic J. Parkinson's disease: clinical features and diagnosis. J Neurol Neurosurg Psychiatry. 2008;79(4):368-76.
73. Zimprich A, Biskup S, Leitner P, Lichtner P, Farrer M, Lincoln S, et al. Mutations in LRRK2 cause autosomal-dominant parkinsonism with pleomorphic pathology. Neuron. 2004;44(4):601-7.
74. Gaweda-Walerych K, Mohagheghi F, Zekanowski C, Buratti E. Parkinson's disease-related gene variants influence pre-mRNA splicing processes. Neurobiol Aging. 2016;47:127-38.
75. Samii A, Nutt JG, Ransom BR. Parkinson's disease. Lancet. 2004;363(9423):1783-93.