6868–6882 *Nucleic Acids Research, 2016, Vol. 44, No. 14* *Published online 28 June 2016*
doi: 10.1093/nar/gkw575

SLIRP stabilizes LRPPRC via an RRM–PPR protein interface

Henrik Spåhr[1,*], Agata Rozanska[2], Xinping Li[3], Ilian Atanassov[3], Robert N. Lightowlers[2], Zofia M. A. Chrzanowska-Lightowlers[2], Oliver Rackham[4,5] and Nils-Göran Larsson[1,6,*]

[1]Department of Mitochondrial Biology, Max Planck Institute for Biology of Ageing, D-50931 Cologne, Germany, [2]Wellcome Trust Centre for Mitochondrial Research, Newcastle University, Newcastle upon Tyne, NE2 4HH, UK, [3]Proteomics Core Facility, Max Planck Institute for Biology of Ageing, D-50931 Cologne, Germany, [4]Harry Perkins Institute of Medical Research and Centre for Medical Research, The University of Western Australia, Nedlands, Western Australia 6009, Australia, [5]School of Chemistry and Biochemistry, The University of Western Australia, Crawley, Western Australia 6009, Australia and [6]Department of Medical Biochemistry and Biophysics, Karolinska Institutet, 171 77 Stockholm, Sweden

Received May 20, 2016; Revised June 14, 2016; Accepted June 16, 2016

ABSTRACT

LRPPRC is a protein that has attracted interest both for its role in post-transcriptional regulation of mitochondrial gene expression and more recently because numerous mutated variants have been characterized as causing severe infantile mitochondrial neurodegeneration. LRPPRC belongs to the pentatricopeptide repeat (PPR) protein family, originally defined by their RNA binding capacity, and forms a complex with SLIRP that harbours an RNA recognition motif (RRM) domain. We show here that LRPPRC displays a broad and strong RNA binding capacity *in vitro* in contrast to SLIRP that associates only weakly with RNA. The LRPPRC–SLIRP complex comprises a hetero-dimer via interactions by polar amino acids in the single RRM domain of SLIRP and three neighbouring PPR motifs in the second quarter of LRPPRC, which critically contribute to the LRPPRC–SLIRP binding interface to enhance its stability. Unexpectedly, specific amino acids at this interface are located within the PPRs of LRPPRC at positions predicted to interact with RNA and within the RNP1 motif of SLIRP's RRM domain. Our findings thus unexpectedly establish that despite the prediction that these residues in LRPPRC and SLIRP should bind RNA, they are instead used to facilitate protein–protein interactions, enabling the formation of a stable complex between these two proteins.

INTRODUCTION

Human mitochondria contain a small, circular genome that encodes 22 tRNAs, 2 rRNAs and 13 proteins, all of which are necessary for adenosine triphosphate (ATP) production by oxidative phosphorylation (OXPHOS). However, the vast majority of OXPHOS components are encoded by the nuclear genome, which necessitates a coordinated regulation of gene expression in response to the physiological demands of the cell. Mitochondrial gene expression is predominantly regulated at the post-transcriptional level by factors controlling mRNA maturation, mRNA stability, mitoribosomal biogenesis and coordinated translation (1). Relatively little structural information is available on factors regulating these processes, however, recently structures of the complete mitochondrial ribosome (2,3) and of members of the MTERF family critical for mitoribosomal biogenesis have emerged (4–6). In addition to the MTERF proteins, there are other examples of α-helical proteins with RNA binding capacity including Pumilio, FBF homology proteins (PUF) and the pentatricopeptide repeat (PPR) proteins. The PUF proteins recognize single-stranded RNA (ssRNA) in a sequence specific modular manner defined by a universal code (7,8) whereas the RNA binding by MTERF proteins appears less specific (9,10). The PPR protein family is constantly expanding with certain higher plant genomes encoding over 400 members with roles in mitochondria or chloroplasts in RNA processing, editing, splicing, stability and translation by recognizing ssRNA according to a predicted PPR code (11,12). These proteins are, like the PUFs and MTERFs, all α-helical proteins composed of repeats, but instead of triangular three-helix motifs, PPRs harbour characteristic 2–30 PPR motifs, each composed of about 35 amino acids that fold into two anti-parallel α-helices. Struc-

*To whom correspondence should be addressed. Tel: +49 221 37970 700; Fax: +49 221 37970 88700; Email: Larsson@age.mpg.de
Correspondence may also be addressed to Henrik Spåhr. Tel: +49 221 37970 700; Fax: +49 221 37970 88700; Email: Spahr@age.mpg.de

tural work on the well characterized maize chloroplast protein PPR10 has shown that 6 nt of its specific RNA substrate, *PSAJ,* stack between six corresponding PPR motifs in a modular fashion (13). Although harbouring similarities in how RNA selectivity is determined, PPR proteins can differ in their mode of RNA association. The PPR10 homodimer, for example, monomerizes upon RNA binding (14), whereas the plant splicing factor THA8 dimerizes (15). Deciphering aspects of the PPR code has recently enabled construction of artificial PPR scaffolds for programmable RNA binding and revealed how they interact with RNA in a sequence-specific manner (16,17). In mammalian mitochondria, merely seven PPR proteins have been identified (18) and structural information for them is only beginning to emerge (19,20). The PPR protein LRPPRC was originally discovered when a recessive mutation causing an A354V amino acid substitution led to the French-Canadian variant of Leigh syndrome, characterized by cytochrome *c* oxidase deficiency and severe neurodegeneration (21). Whereas plant PPR proteins have developed very specialized roles and recognize RNA in a sequence specific manner, LRPPRC stabilizes nearly all mitochondrial (mt) mRNAs, promotes their poly(A) tail length extension and coordinates mitochondrial translation (22–26). It forms a complex with the much smaller SLIRP protein (23,24), which harbours an RNA recognition motif (RRM) and is required for correct association of mt-mRNAs with the mitochondrial ribosome to enable efficient translation (27). The RRM domain is one of the most abundant protein domains in eukaryotes and although most RRM proteins bind RNA, the domain has now been shown to exhibit versatile molecular recognition capacities, including protein–protein interactions (28). The association of LRPPRC and SLIRP is conserved through evolution, and in *Drosophila melanogaster* two orthologues of both LRPPRC and SLIRP exist (29,30). These proteins are interdependent as reduced levels of either component leads to decreased levels of the other (23,24,27).

In a first step towards a structural understanding of the LRPPRC–SLIRP complex, we show that SLIRP forms a hetero-dimer with LRPPRC to enhance its stability. By using a combination of crosslinking-mass spectrometry (MS) and truncation/mutational analysis, we report that SLIRP and LRPPRC associate by a novel RRM–PPR binding mode. SLIRP, despite being composed almost entirely of an RRM domain, only weakly associates with RNA, in sharp contrast to LRPPRC that has a strong and broad RNA binding capacity, and specific residues predicted to bind RNA are unexpectedly the ones that facilitate the protein–protein bridges of the hetero-dimeric complex.

MATERIALS AND METHODS

Cloning and purification

Codon-optimized (DNA 2.0) DNA constructs corresponding to the mature form of human LRPPRC (amino acid 60–1394) or SLIRP (18–109) were cloned in a pJexpress 401 (DNA 2.0) vector and the LRPPRC–SLIRP complex in pCDFDuet-1 (Novagen). LRPPRC harbours a 6×His fusion tag at the N-terminus in both vectors whereas SLIRP was His-tagged in the pJexpress 401 vector and untagged in pCDFDuet-1. The proteins were expressed in Rosetta 2 cells (EMD chemicals) by induction with 0.5 mM isopropyl-1-thio-β-D-galactopyranoside (IPTG) at 30°C for 16 h in Enpresso B media (Biosilta). After lysis, the proteins were purified over a His-Select Ni^{2+} (Sigma-Aldrich) resin and dialysed against H-0.2 (25 mM Tris–HCl [pH 7.8], 0.5 mM ethylenediaminetetraacetic acid (EDTA), 10% glycerol, 1 mM dithiothreitol, 200 mM NaCl) after the addition of TEV protease at a 1:50 protease:protein ratio. Further purification was conducted over a heparin column equilibrated in H-0.2. After washing with H-0.2 the proteins were eluted with H-0.6 and purified to homogeneity over a HiLoad 16/60 Superdex 200 pg gel filtration column (GE Healthcare) in buffer H-0.2 lacking glycerol.

Truncation and mutational analysis of LRPPRC–SLIRP

His-tagged LRPPRC and untagged SLIRP truncations or mutants (Mutagenex) as specified in the figures were co-expressed and purified over Ni^{2+} as described above.

Gel filtration and stoichiometry analysis

LRPPRC was incubated with five times molar excess SLIRP, RNA or SLIRP and RNA in H-0.15 (25 mM Tris–HCl [pH 7.8], 0.5 mM EDTA, 1 mM dithiothreitol, 5 mM $MgCl_2$ 150 mM NaCl) at room temperature (RT) for 30 min. The formed complexes were separated on a Superose 6 or Superose 12 column in H-0.15. The stoichiometry of LRPPRC–SLIRP was determined by absolute quantification as described previously (31).

Blue-Native PAGE

Samples were mixed with 4× Native PAGE sample buffer and 5% [w/v] Coomassie Brilliant Blue G-250 (Thermo Fischer). BN-PAGE samples were resolved on 3–12% NativePAGE Novex Bis-Tris gels and subjected to Coomassie staining.

RNA electrophoretic mobility shift assays (RNA EMSA)

Purified LRPPRC, LRPPRC–SLIRP or SLIRP (0, 0.02, 0.04, 0.08, 0.16, 0.36, 0.64, 1.28, 2.56 μM) was incubated at room temperature for 15 min with 40 ng fluorescein labelled RNAs in 10 mM HEPES (pH 8.0), 1 mM EDTA, 60 mM NaCl, 3 mM $MgCl_2$, 1 mM DTT, 0.1 mg/ml fatty acid-free BSA and 5% glycerol. Reactions were separated by electrophoresis in a 6% DNA retardation gel (Invitrogen) in 0.5× TBE and fluorescence was detected using a Typhoon FLA 9500 biomolecular imager (GE).

Thermofluor screen

The thermofluor screen for LRPPRC–SLIRP or LRPPRC was performed using SYPRO Orange with a temperature ramping from 5 to 95°C (1°C/min) using a MyIQ RT-PCR instrument (BioRad) as described previously (32).

Crosslinking and sample preparation for LC/MSMS

Purified LRPPRC or LRPPRC–SLIRP proteins were dialysed against 25 mM Hepes pH 8.0 and 200 mM NaCl

for 4 h at 4°C. The dialysed samples were crosslinked by the addition of 50 molar excess BS^3 H12/D12 (Creative Molecules), incubated for 15 min at room temperature and the reaction was stopped by the addition of Tris–HCl, pH 7.4 to a final concentration of 50 mM. Following sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and Coomassie Brilliant Blue staining, the band corresponding to the crosslinked product was excised and digested with 7.5 ng/μl trypsin at 37°C overnight. Before being subjected to MS analysis, the digest was cleaned with reverse-phase C18 StageTip and eluted using 40 μl of 60% acetonitrile in 0.1% formic acid. The peptides were dried in a SpeedVac and re-suspended in 5 μl of 0.1% formic acid.

LC/MSMS analysis

The peptides were analysed using a Q Exactive Plus mass spectrometer (ThermoFisher Scientific) with a nano-electrospray ion source, coupled with an EASY-nLC 1000 (Thermo Fisher Scientific) UHPLC. A 25 cm long reverse-phase C18 column with 75 μm inner diameter (PicoFrit, LC Packings) was used for separating peptides. The LC run lasted 50 min with a concentration of 5% solvent B (0.1% formic acid in acetonitrile) increasing to 25% over 35 min, to 50% over 10 min and to 90% over 5 min. The column was subsequently washed and re-equilibrated. MS spectra were acquired in a data-dependent manner with a top 10 method. MS spectra were acquired with a mass range of 300−1800 m/z and 70 000 resolution at 200m/z. The AGC target of MS was set to 3e6, and the maximum injection time was 20 ms. Peptides were fragmented with higher collision decomposition (HCD) with collision energy of 25. The AGC target of MSMS was set to 2e5 and the maximum injection time was 80 ms.

Crosslink search

The Raw data were converted to MGF file with MSconvert (ProteoWizard). The light and heavy form of isotopic peptide pairs were screened based on the characteristic isotopic mass shift with a Python script. The filtered spectra were searched by xQuest with MS2 spectra of light and heavy pairs (http://proteomics.ethz.ch/orinner/public/htdocs/xquest/xquest_review.html). The MS1 tolerance was set to 20 ppm and the MS2 to 0.1 m/z. The Xquest data was further analysed with xVis (33). After visualization, the data was exported and the protein–protein crosslinks were filtered for uniqueness by keeping the highest scoring crosslinks.

Identification of oligoribonucleotides bound *invivo* by crosslinking immunoprecipitation (CLIP)

Assays were performed essentially as described in (34). Briefly, HEK293-FlpIn TRex cells capable of expressing FLAG-tagged SLIRP were induced for 3 days with 1 μg/ml tetracycline, harvested at 80% confluency (4 × 15 cm^2 plates), washed twice in PBS and UV-irradiated at 400 mJ/cm2 in a Stratalinker (Stratagene). Cells were lysed, the bound ribonucleoprotein (RNP) was treated with RNase T1 to remove unprotected RNA species, and SLIRP bound RNP was immunoprecipitated via the FLAG moiety using the Sigma FLAG-IP kit (Sigma-Aldrich). Bound RNA was dephosphorylated and ligated to the 3′ linker as described in (34). The complex was visualized by labelling the 5′ termini with [γ -^{32}P]ATP (3000 Ci/mol, Perkin Elmer) and PNK (T4 polynucleotide kinase; New England Biolabs), separated by SDS-PAGE (10% Novex precast gels), transferred on to nitrocellulose (BA-85 Whatman) and subjected to autoradiography. RNA was isolated from bound RNP by excision from the nitrocellulose, which was treated with proteinase K and then the RNA precipitated following phenol/chloroform extraction. Ligation of the 5′ terminus and reverse transcription were as described (34). The final PCR (polymerase chain reaction) products were prepared for IonTorrent sequencing following manufacturer's instructions. Sequence data for 100 000–190 000 reads was collected, aligned to mitochondrial DNA (mtDNA) as a reference sequence using the Torrent Suite software on the IonTorrent server. Alignments were then viewed using IGV (integrative genomics viewer) and presented against a linear depiction of human mtDNA.

Structure prediction and docking of LRPPRC–SLIRP

The sequence of human SLIRP (aa 18–94) and LRPPRC (aa 404–503) was used for structure prediction by I-Tasser (35), which calculated models with confidence scores of 0.73 and 0.56, respectively. The models were used for docking with the Haddock2.2 web server (36) using interacting amino acids identified in the mutational analysis as active residues. The figure was prepared with Pymol (http://www.pymol.org/).

RESULTS

SLIRP stabilizes LRPPRC

To examine the interdependence of LRPPRC and SLIRP *in vitro* we developed protocols for expression and purification of the two human proteins. SLIRP expressed well in *Escherichia coli* in contrast to LRPPRC, which only expressed at low levels. In addition, a large fraction of LRPPRC migrated in the void volume on a gel filtration column (data not shown), indicating aggregation. Nevertheless, enough soluble protein could be obtained and was purified to homogeneity (Figure 1A). In an attempt to overcome the problem of LRPPRC aggregation, we generated constructs to co-express it with SLIRP. The presence of SLIRP significantly stabilized LRPPRC and enhanced the expression levels by more than an order of magnitude.

To examine the stability effect of SLIRP on LRPPRC further, a thermofluor assay was performed both for the LRPPRC–SLIRP complex and LRPPRC alone. A pH screen revealed that the LRPPRC–SLIRP complex was most stable at a neutral pH with a T_m value of 43.7°C (Figure 1B, left panel). At extreme pH values, both high and low, the complex was partly unfolded even before the temperature was raised from the starting temperature 5°C indicating that core hydrophobic residues were accessible to the solvent. This effect was more pronounced for LRPPRC alone, also at a neutral pH (Figure 1B, right panel). The estimated T_m value (39.0°C) was also lower than the value for the

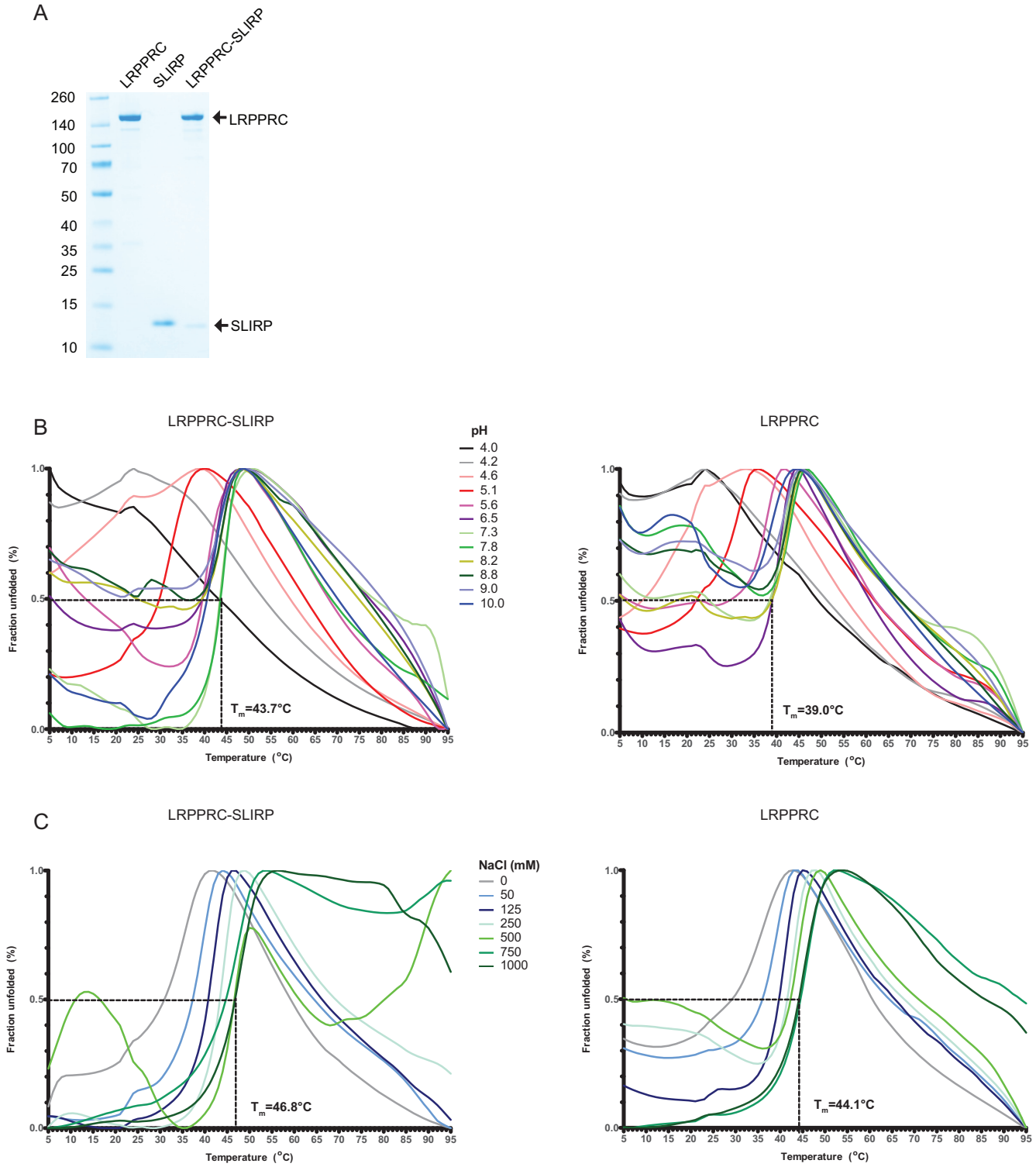

Figure 1. SLIRP stabilizes LRPPRC. (**A**) Purification of LRPPRC, SLIRP and the LRPPRC–SLIRP complex. The gel was stained with Coomassie brilliant blue, purified proteins are indicated on the right and molecular masses according to a standard in kDa on the left. (**B** and **C**) Thermofluor profiles of LRPPRC–SLIRP or LRPPRC. The melting temperature (T_m) corresponds to the temperature where 50% of the protein is unfolded. (B) Effect of pH on LRPPRC–SLIRP or LRPPRC thermal stability. The buffer composition was citric Acid/CHES/HEPES [2:4:3] with the pH indicated. (C) Effect of NaCl concentration on LRPPRC–SLIRP or LRPPRC thermal stability. The buffer composition was 50 mM HEPES pH 7.5 and 0–1000 mM NaCl.

LRPPRC–SLIRP complex, indicating that LRPPRC appeared less stable in the absence of SLIRP. The effect of salt on LRPPRC–SLIRP or LRPPRC stability was also tested. The addition of NaCl significantly stabilized the proteins with a maximum T_m value of 46.8°C for LRPPRC–SLIRP (Figure 1C, left panel) and 44.1°C for LRPPRC alone (Figure 1C, right panel).

LRPPRC has a broad RNA binding capacity

In silico analysis predicted human LRPPRC to be all α-helical (PSIPRED and I-Tasser, data not shown) and detected 22 PPR motifs (Supplementary Figure S1) by TPRpred (37). However, more motifs could be assigned by manual inspection of the sequence of the α-helices that constitute LRPPRC, and it is probable that the protein contains more than 30 PPR motifs (Supplementary Figure S2). The PPR motifs are flanked by additional sequences including the mitochondrial targeting sequence (MTS) at the N-terminus. Extensive characterization of plant PPR proteins has revealed their specialized roles in RNA maturation or stability facilitated by their sequence specific recognition of RNA (13,15). We attempted to assign the sequence specificity of LRPPRC according to the PPR code by using a position-specific scoring matrix where the RNA recognition residues are typically found at position 4 and 34 in each PPR motif. However, in contrast to PPR proteins in plants, the potential RNA recognition residues were predicted to display a broad specificity (Supplementary Figure S2). To test the specificity experimentally, we performed RNA electrophoretic mobility shift assays (REMSA) for LRPPRC, SLIRP and LRPPRC–SLIRP with RNA homopolymers composed of adenines, guanines, uracils or cytosines (Supplementary Figure S3). LRPPRC and the LRPPRC–SLIRP complex both bound each of the homopolymers well and only produced a slightly less clear shift for poly(A). In contrast, SLIRP bound to none of the homopolymers except poly(G), but with very low affinity. To test the preference of LRPPRC for single or double stranded RNA, we performed REMSA with a single stranded 40mer sequence from *MTND1* in addition to *mt-tRNA^{Pro}* (Supplementary Figure S4). LRPPRC–SLIRP and LRPPRC were able to bind both substrates with a preference for single stranded RNA, consistent with the behaviour of PPR proteins in plants (13,15,38). In contrast, SLIRP produced no shifts with either substrate, suggesting an inability to bind.

SLIRP has a very modest RNA binding capacity

There was no detectable binding of RNA by SLIRP in the *in vitro* RNA-shift assays with any of the RNA substrates tested (Supplementary Figures S3 and 4). To investigate the ability of SLIRP to bind RNA *in vivo*, and to define any potential binding sites on mitochondrial RNA, we performed crosslinking immunoprecipitation (CLIP) using a HEK293T cell line that could inducibly express FLAG-tagged SLIRP. SLIRP protected sequences were found to all regions of mitochondrially encoded mRNAs except *MTND6* (Figure 2A and Supplementary Table S1). These data are in agreement with previous findings where mouse knockouts of either LRPPRC or SLIRP both compromise the stability of these mt-mRNAs (23,27). The binding sites with the highest number of reads were located in five mitochondrial mRNAs: *MTND2*, *MTCO1*, *MTATP8/6* (*RNA14*), *MTND3* and *MTND4*. Surprisingly, one of the regions most frequently bound by SLIRP was positioned near the H-strand replication origin. In addition, a small number of reads were obtained for *mt-tRNA^{Pro}*, 12S rRNA and 16S rRNA. Analysis of the identified SLIRP RNA binding sites did not show any discernible specificity for either a defined sequence or structure that was common to all RNA fragments.

To validate the site specificity, we followed up the CLIP results by performing REMSAs with the RNA sequences that produced the highest level of binding. LRPPRC alone or the LRPPRC–SLIRP complex produced very similar results with clear shifts demonstrating an affinity of ∼150 nM (Figure 2B). The most pronounced difference was observed at high protein concentrations where LRPPRC alone produced supershifts reflecting the formation of higher order complexes or aggregation. Surprisingly, no shifts could be observed by the addition of SLIRP in these conditions (Figure 2B). We therefore repeated the REMSAs with 10-fold higher SLIRP concentrations, however, the only shifts observed were indistinct with an affinity at least one order of magnitude lower than for LRPPRC or the LRPPRC–SLIRP complex (Supplementary Figure S5). Thus, these results indicate that SLIRP plays only a minor role, if at all contributing, to the RNA binding of the LRPPRC–SLIRP complex. When taken together, the CLIP and REMSA data suggest that the SLIRP RNA binding identified by CLIP, which was size selected by excision from an SDS-PAGE gel, reflects RNA-protein interactions that formed in the context of the LRPPRC–SLIRP complex in cells and may represent LRPPRC–SLIRP targets rather than SLIRP targets. This is further supported by the detection of SLIRP bound CLIP sequences to all mtDNA-encoded mRNAs except *MTND6*, which is the only mt-mRNA that is not stabilized nor polyadenylated in the presence of LRPPRC (23).

LRPPRC forms a hetero-dimer with SLIRP

The theoretical molecular mass of the mature forms of LRPPRC and SLIRP are 153 and 11 kDa, respectively, and correspond to the migration pattern of the recombinant proteins on SDS-PAGE (Figure 1A). The LRPPRC–SLIRP *in vivo* complex has been reported to migrate at 250 kDa (23,24), which is larger than the sum of the two proteins. This observation, together with a measured 1:1 stoichiometry between LRPPRC and SLIRP after immunoprecipitating SLIRP from control fibroblast mitochondria, has lead to the suggestion that the complex is a hetero-tetramer (25). We combined purified LRPPRC and SLIRP *in vitro* and subjected the resulting complex to gel filtration (Figure 3A) and blue-native (BN) PAGE (Figure 3B). In agreement with the *in vivo* work, our *in vitro* assembled complex migrated at 250 kDa and was identified with MS analysis to be a 1:1 complex (Supplementary Table S2). In addition, we investigated the stoichiometry in the presence of an RNA fragment corresponding to a sequence present in *MTATP6*, which had one of the highest numbers of reads in the CLIP assay and displayed a clear shift in the REMSA. The stoichiometry between LRPPRC and SLIRP, how-

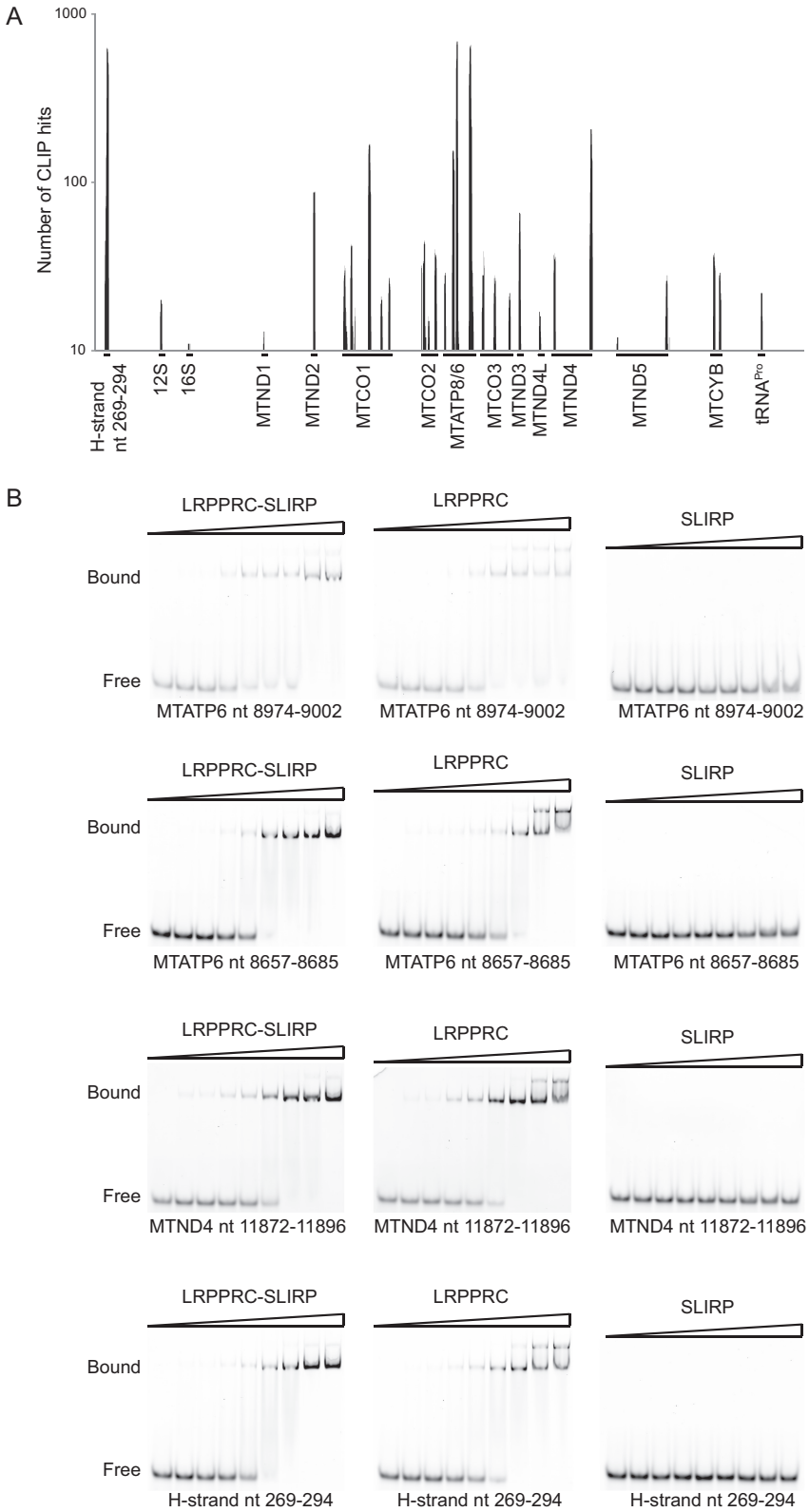

Figure 2. RNA binding by LRPPRC, SLIRP and the LRPPRC–SLIRP complex. (**A**) CLIP analysis of SLIRP. The number of CLIP hits is shown on the y-axis and identified targets on the x-axis. (**B**) REMSA of LRPPRC–SLIRP, LRPPRC or SLIRP. RNA template sequences are indicated. Protein concentrations used were 0, 0.02, 0.04, 0.08, 0.16, 0.36, 0.64, 1.28, 2.56 μM, respectively.

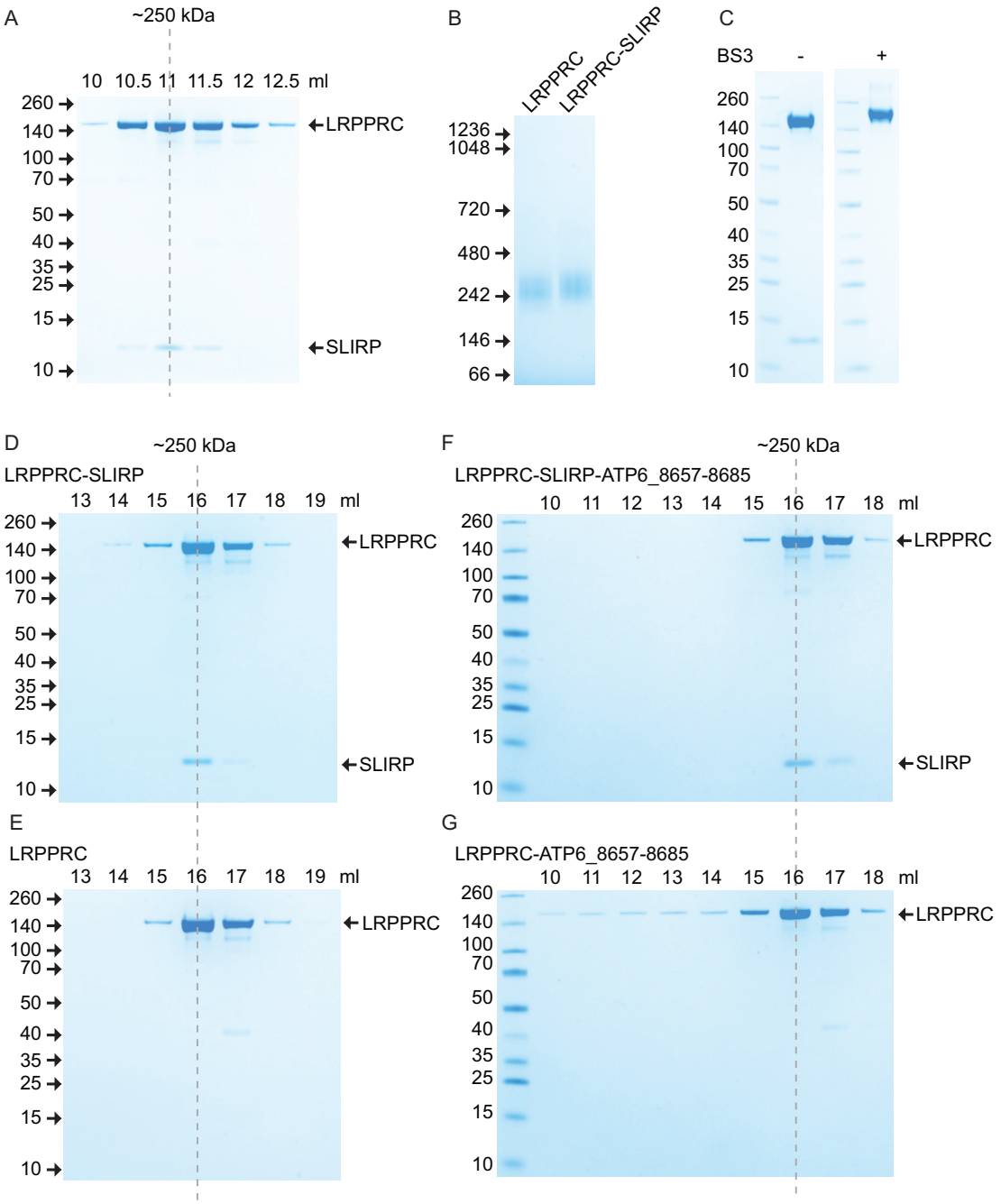

Figure 3. LRPPRC–SLIRP is a hetero-dimer. (**A**) Formation of the LRPPRC–SLIRP complex *in vitro*. The *in vitro* formed LRPPRC–SLIRP complex was subjected to gel filtration on a Superose 12 column and peak fractions were separated by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE). (**B**) BN-PAGE of LRPPRC and the LRPPRC–SLIRP complex. (**C**) Crosslinking of the LRPPRC–SLIRP complex with BS^3. Non-crosslinked and crosslinked LRPPRC–SLIRP complexes were analysed on a SDS-PAGE gel. (**D** and **E**) Gelfiltration analysis of LRPPRC–SLIRP or LRPPRC. Proteins were loaded on a Superose 6 column and analysed on SDS-PAGE. (**F** and **G**) Gelfiltration analysis of LRPPRC–SLIRP or LRPPRC when bound to a RNA fragment corresponding to nucleotides 8657–8685 of *MTATP6*. All gels were stained with Coomassie brilliant blue with purified proteins indicated on the right, elution volume on top and molecular masses according to a standard in kDa on the left.

ever, remained unaffected by the presence of RNA (Supplementary Table S2). To investigate the oligomeric state of LRPPRC–SLIRP, we analysed the complex by SDS-PAGE after the addition of the lysine crosslinker BS^3. SLIRP was effectively crosslinked to LRPPRC and the complex migrated with an apparent mass of 170 kDa on a SDS-PAGE (Figure 3C), which corresponds well to the sum of the two proteins. We conclude that the LRPPRC–SLIRP complex is a hetero-dimer, which due to an elongated shape adopts a modified migration suggestive of a larger molecular mass when analysed by gel filtration or BN-PAGE.

The PPR proteins THA8 and PPR10 in plants have been shown to change their oligomeric states upon RNA binding (14,15). To determine if any such change occurs with LRPPRC, we examined the effect of SLIRP or RNA on LRPPRC by generating LRPPRC–SLIRP, LRPPRC-RNA or LRPPRC–SLIRP–RNA complexes for gel filtration analysis. We used the same RNA fragment as in the stoichiometry analysis above. SLIRP did not affect the oligomeric state of LRPPRC, which migrated as a monomer (Figure 3D and E). Similarly, the addition of RNA did not change the migration pattern for the majority of the complexes (Figure 3F and G). However, a sub-population of LRPPRC formed larger complexes upon addition of RNA, which therefore eluted in earlier fractions over gel filtration (Figure 3G) consistent with the REMSA results (Figure 2B). This effect was completely absent in the presence of SLIRP where only one distinct peak was observed (Figure 3F).

Crosslinking/MS analysis of LRPPRC–SLIRP

As there is currently no structural data for either LRPPRC or SLIRP, crosslinking/MS analysis of the human LRPPRC–SLIRP complex was carried out to gain insight into the architecture of the complex. We used isotope-labelled BS^3 crosslinker that reacts with primary amines in lysine side chains and protein N-termini. We estimated that for crosslinking to occur, the distance between α-carbon pairs must be $\leq$30 Å, corresponding to the length of the BS^3 spacer (11.4 Å), two times the length of a lysine side chain (6.5 Å), and an estimated error of 3 Å for flexible lysine side chain ends in mobile parts of the protein. MS analysis identified a complex crosslink pattern with 95 intra-molecular crosslinks within LRPPRC and 8 inter-molecular crosslinks between LRPPRC and SLIRP (Figure 4A, Supplementary Tables S3 and 4). Evidently, bonds were formed between lysines located in PPR motifs that were both distant and close to each other in the linear sequence of LRPPRC, displaying a star shaped intra-molecular crosslink pattern in a circular plot (Figure 4A). Interestingly, crosslinks between N- and C-terminus proximal lysine residues indicated that the two termini of LRPPRC are not far from each other in the folded structure. This is in contrast to the previously reported structure of PPR10, which has its termini located at each end of a solenoid shape (13). The lysines in SLIRP that crosslinked to LRPPRC were all located in the C-terminal half of the protein and the 8 LRPPRC–SLIRP crosslink pairs were 424–106, 453–54, 853–88, 991–97, 991–106, 991–107 and 1297–88 respectively. Interestingly, the pattern suggests that SLIRP sits at a five-way junction between PPR motifs of LRPPRC that are relatively distant from each other in the sequence (Figure 4A and Supplementary Figure S2). To investigate if SLIRP had any influence on the structure of LRPPRC, we performed crosslinking-MS for LRPPRC alone in order to compare the crosslink pattern with the one generated by the LRPPRC–SLIRP complex. Overall the crosslink patterns appeared similar and displayed the same prominent crosslinks, suggesting a similar structure (Figure 4A and B). However, at least some degree of a conformational change likely occurs in LRPPRC in the presence of SLIRP, as only 30 of the 95 crosslinks observed for the complex were present also for LRPPRC alone (Figure 4B and C; Supplementary Tables S3 and 4). In addition, only 70 total crosslinks were observed for LRPPRC compared to 95 for the complex, which may reflect a more compact LRPPRC structure in the presence of SLIRP, thereby allowing additional lysines to be close enough for crosslinking to occur.

Truncation analysis of LRPPRC and SLIRP

To complement the crosslinking-MS data and to elucidate a more detailed view of the interaction between human LRPPRC and SLIRP, we performed truncation analysis of both proteins. Initially, four constructs were designed to independently encode regions spanning the entire LRPPRC sequence, ensuring that predicted α-helices were not disrupted (Figure 5A). These four fragments incorporated an N-terminal His-tag and were co-expressed with untagged SLIRP. In this way, the interaction of individual LRPPRC domains with SLIRP could be identified by co-purification over a Ni^{2+} column. Only the second domain (385–724) of LRPPRC was absolutely required for the interaction with SLIRP (Figure 5B). To fine map the SLIRP binding site, we designed four new overlapping fragments of LRPPRC (residues 367–571, 404–615, 504–675 and 535–724) (Figure 5A). By this approach, we identified that only fragments 367–571 and 504–675 both interacted with SLIRP (Figure 5C). We, therefore, conclude that the amino acids of LRPPRC responsible for interacting with SLIRP were located between residues 404–503.

SLIRP is predominantly composed of an RRM domain. To investigate if the RRM domain is responsible for the interaction with LRPPRC, we co-expressed SLIRP lacking the flanking residues at both the N- and C-terminus, together with the LRPPRC domain comprising residues 404–675. This extended LRPPRC fragment was selected and used hereafter, as the expression was significantly improved by extending the C-terminus from amino acid 615 to 675. Interestingly this N- and C- terminal truncated SLIRP was still able to bind LRPPRC, confirming that the RRM domain is responsible for the interaction with LRPPRC (Figure 5D). A parallel experiment using a shorter LRPPRC fragment, comprising residues 443–675, was also expressed with SLIRP. This fragment was prone to degradation and was unable to fully capture SLIRP (Figure 5D), suggesting that residues within the 404–442 window critically contribute to the binding, however it doesn't preclude that residues within the 443–503 window contribute to the binding as well.

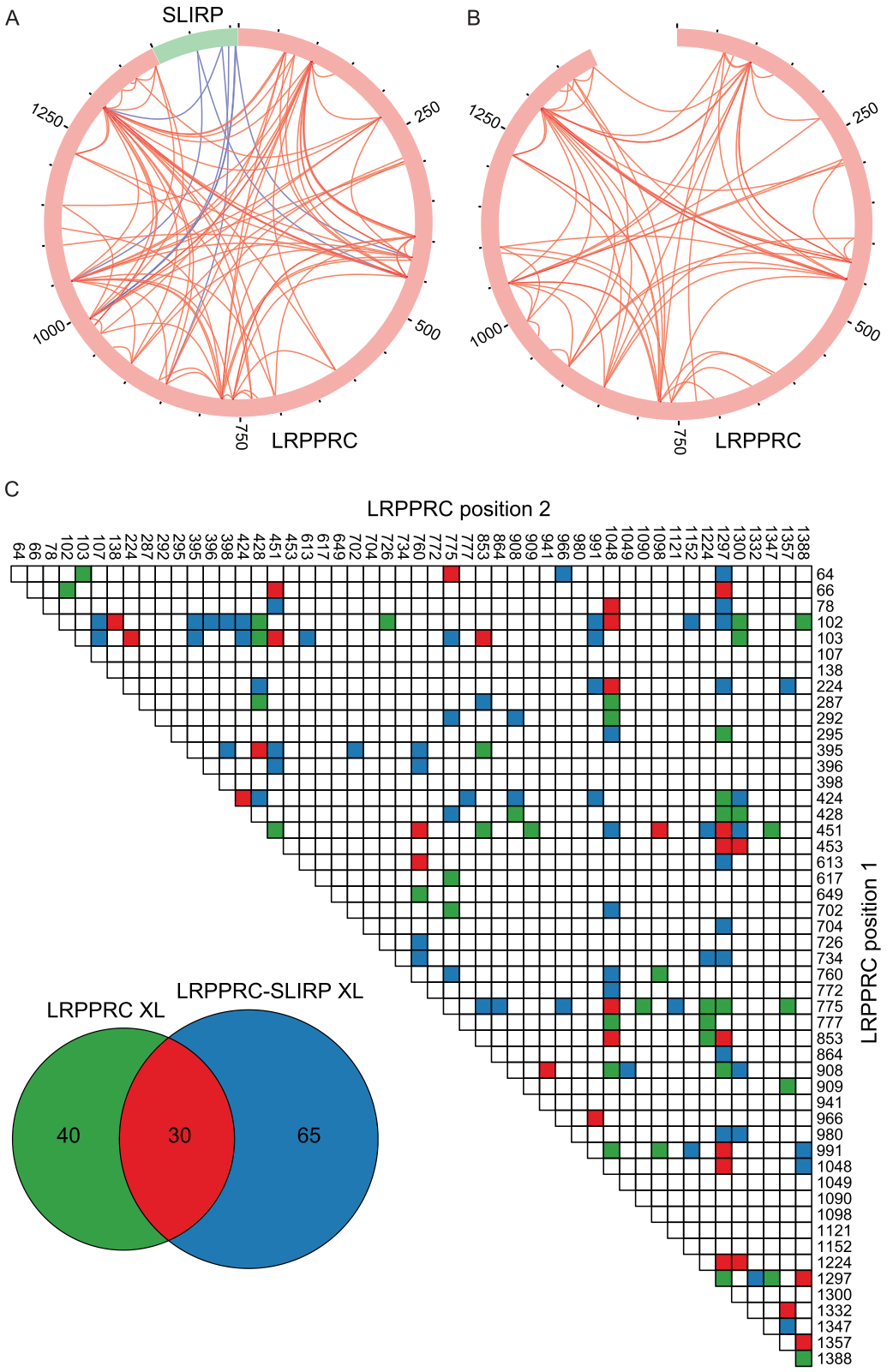

Figure 4. Crosslinking pattern of LRPPRC in complex with SLIRP (**A**) and LRPPRC alone (**B**). Circular plots were prepared with xVis (33) showing intra-molecular crosslinks in red and inter-molecular crosslinks in blue. Sequence numbers in 250 aa increments are indicated. (**C**) Numbers and pattern of unique LRPPRC intra-molecular crosslinks identified from LRPPRC alone (green and red), LRPPRC in complex with SLIRP (blue and red), in both states (red).

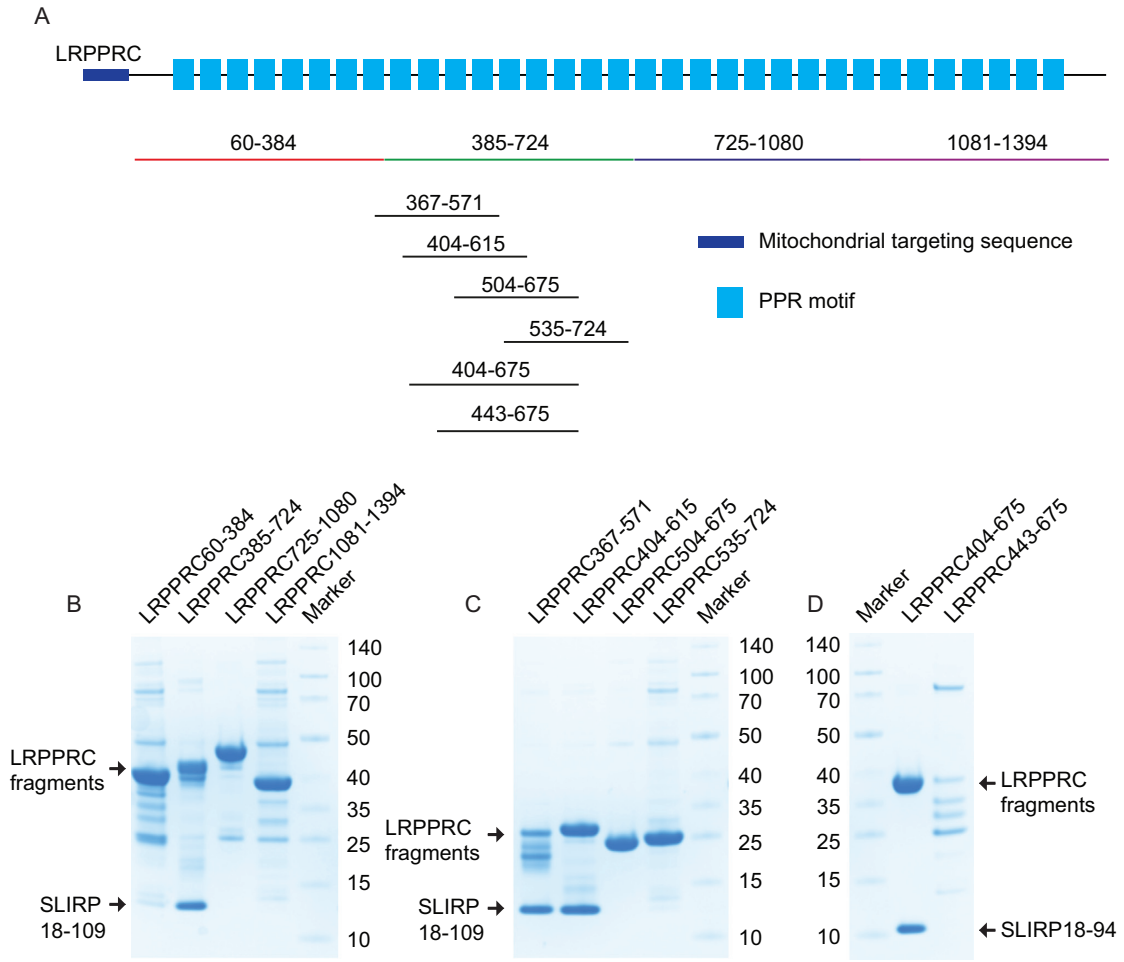

Figure 5. Identification of the LRPPRC–SLIRP binding interface. (**A**) LRPPRC truncations. A schematic representation of LRPPRC is shown, where the mitochondrial targeting sequence (MTS) is depicted in purple and the predicted PPR domains as blue rectangles. PPR motifs are annotated according to Supplementary Figure S2. Sequence numbers of LRPPRC fragments co-expressed with SLIRP are indicated. (**B**–**D**) Binding ability of LRPPRC fragments to SLIRP. Co-expressed LRPPRC fragments and SLIRP after Ni^{2+} purification. Proteins were separated by 4–12% SDS-PAGE and stained with Coomassie brilliant blue. Expressed LRPPRC fragments are indicated on the top and molecular masses according to standard in kDa on the left or right.

Mutational and sequence conservation analysis of the LRPPRC–SLIRP interface

To pinpoint the critical amino acids that are at the human LRPPRC–SLIRP interface, an extensive mutational analysis of both proteins was performed. We mutated charged or polar residues to alanines that were identified by the truncation analysis, to cover the entire interacting window of LRPPRC and SLIRP. The mutations were introduced in doublets or triplets, or in one instance as a quintuplet, before purification over Ni^{2+} as performed for the previous truncations. Two of the triple LRPPRC mutations, R438A-H440A-Y441A and Y477A-D479A-Y480A, were unable to interact with SLIRP, in contrast to the other mutants where the interaction remained intact (Figure 6A). Two interaction sites were also identified on SLIRP, where the H59A-R60A and E80A-N81A-H82A mutations lost the ability to interact with LRPPRC (Figure 6B).

We performed a multiple sequence alignment of LRPPRC (Figure 6C) and SLIRP (Figure 6D) to display the sequence conservation of the critical LRPPRC–SLIRP interface-forming amino acids identified in the mutational analysis. The residues of the R438A-H440A-Y441A triple mutation are located in two neighbouring PPR motifs of LRPPRC and all three residues are highly conserved (Figure 6C and Supplementary Figure S2). Indeed, they are located in one of the most conserved regions of LRPPRC. The residues within the other identified interaction triplet are located on the N-terminal edge of the next PPR motif and display more varied sequence conservation; Y477 is conserved between human, mouse and frogs, but not in the two isoforms of LRPPRC in flies, D479 is less conserved and Y480 is highly conserved. Interestingly, both LRPPRC amino acid triplets critical for LRPPRC–SLIRP interface integrity are located at positions that are predicted to participate in RNA binding according to the PPR code (Supplementary Figure S2).

The multiple sequence alignment of SLIRP displayed pronounced sequence conservation of the RRM domain, with flanking non-conserved sequence elements at the termini (Figure 6D). RRM domains typically contain two consensus sequences, RNP2 (six residues ([I/L/V]-

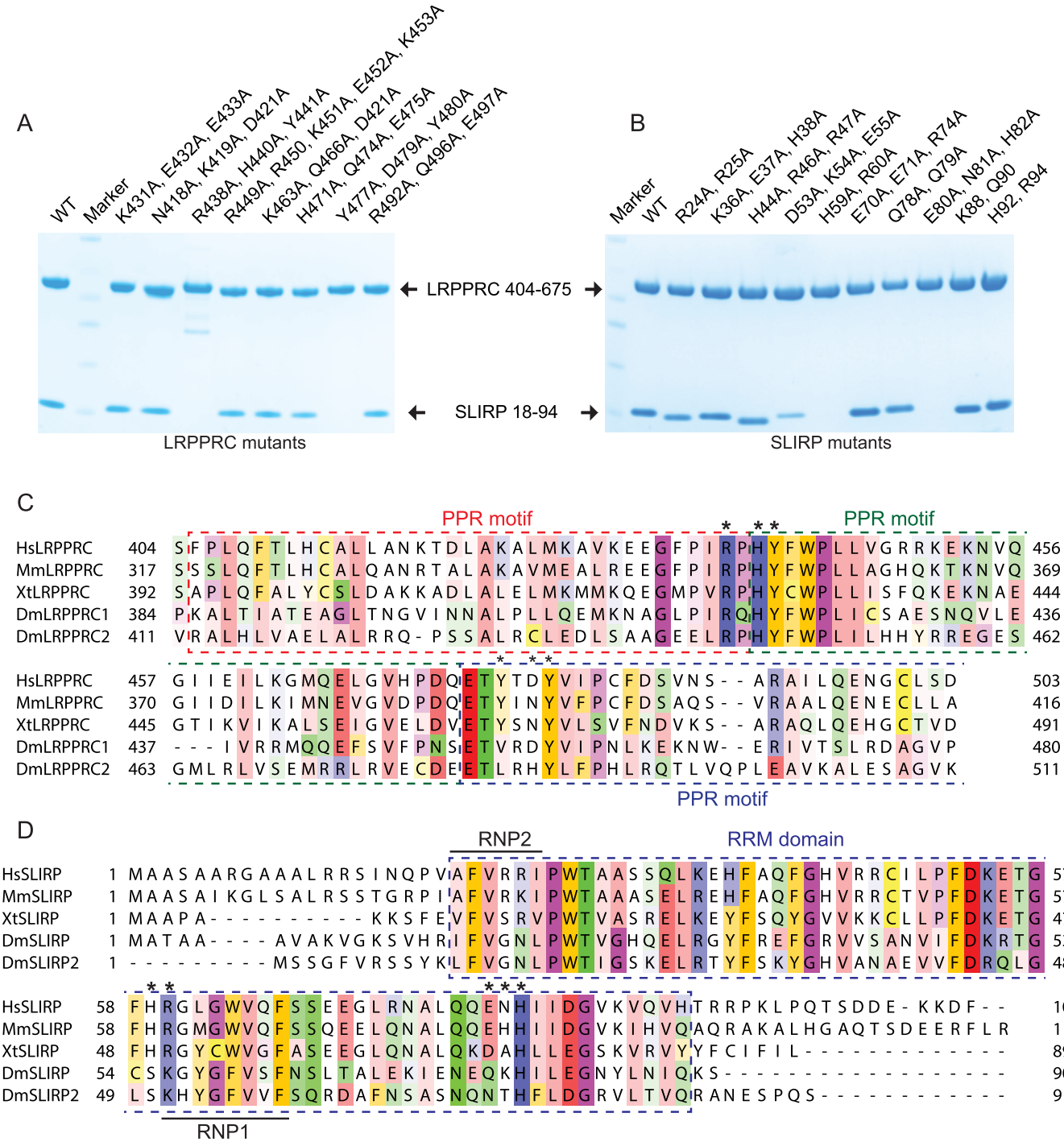

Figure 6. Amino acids crucial for LRPPRC–SLIRP binding interface. (**A** and **B**) Mutagenesis of the LRPPRC–SLIRP interface. Co-expressed LRPPRC (aa 404–675) and SLIRP (aa 18–94) mutants after Ni^{2+} purification. Proteins were separated on a 4–12% SDS-PAGE gel and stained with Coomassie brilliant blue. Mutations are indicated on the top. (**C** and **D**) Sequence alignment of LRPPRC and SLIRP. Sequence conservation is shown with Zappo color code (i.e. aliphatic/ hydrophobic, pink; aromatic, orange; positive, blue; negative, red; hydrophilic, green; conformationally special, magenta; and cysteine, yellow). Species abbreviations are: Hs, *Homo sapiens*; Mm, *Mus musculus*; Xt, *Xenopus tropicalis*; Dm, *Drosophila melanogaster*. Crucial amino acids for LRPPRC–SLIRP interface integrity are indicated with asterisks. The predicted PPR motifs of LRPPRC and the RRM domain of SLIRP are highlighted with coloured dashed windows. RNP1 and RNP2 motifs of SLIRP are indicated.

[F/Y]-[I/L/V]-X-N-L)) and RNP1 (eight residues ([K/R]-G-[F/Y]-[G/A]-[F/Y]-[I/L/V]-X-[F/Y]), which are important for their interactions with RNA. However, the RNP consensus sequences are not strictly adhered to for several amino acids of human SLIRP. At RNP1 position 3, which is in close conjunction with the LRPPRC interacting residue R60 at position 1, SLIRP harbours a Leu residue instead of Phe/Tyr and at position 5 a Trp residue is present rather than Phe or Tyr (Figure 6D). In addition, at RNP2 position 1, the canonical I/L/V is replaced by an alanine and at position 5 an Asp residue is replaced by an arginine. This is interesting as aromatic Phe/Tyr residues in RNP1 and 2, which typically stack with RNA bases or sugars in other RRM proteins, are replaced with aliphatic residues that are not predicted to interact favourably with RNA and instead are used for protein–protein interactions. In contrast to human SLIRP, the two isoforms of SLIRP in flies strictly conform to the consensus sequence.

Both the double H59A-R60A and the triple E80A-N81A-H82A non-binding mutants, have changes spanning regions located in the C-terminal portion of SLIRP, which corresponds well with the crosslinking data (Figure 4A). R60 is universally conserved whereas H59 is conserved only between human, mouse and frogs, but not between the two isoforms of SLIRP in flies. Interestingly, R60 is the first residue of the RNP1 motif, indicating that specific predicted RNA binding residues of both LRPPRC and SLIRP are being used to facilitate protein–protein interactions at the interface between these proteins rather than for RNA interactions. The residues of the triple mutation E80A-N81A-H82A display a varied sequence conservation pattern; E80 is conserved between human, mouse and frogs, but not in the two isoforms of SLIRP in flies, N81 is not conserved and H82 is highly conserved.

Model of the LRPPRC–SLIRP interface

Currently, no structural data exists of either LRPPRC or SLIRP. However, high-resolution structures have been solved for other proteins that are predominantly composed of RRM motifs (28,39) or PPR motifs (13,15). We reasoned that predicting the 3D structures of LRPPRC and SLIRP would supplement the information on the critical interface-forming amino acids that we obtained in the mutational analysis, and docking the models together would shed light on how the interface is organized. SLIRP is composed of an RRM domain with a βαββαββ topology that forms a five-stranded β-sheet packed against two α-helices. The domain of LRPPRC that contacts SLIRP harbours three PPR motifs, of which each is composed of two α-helices. The interaction that generates the LRPPRC–SLIRP complex, therefore, reveals a novel RRM-PPR binding mode (Figure 7). The LRPPRC–SLIRP interface presented in the model buries 1830 $Å^2$ of solvent-accessible surface area. The β-sheet of SLIRP harbouring the two RNP motifs, which typically interacts with ssRNA in other RRM-proteins, here partly faces LRPPRC. On opposite sides of the β-sheet, the amino acids of SLIRP, which were identified to be important for binding LRPPRC in the mutational analysis, are located in loops between β2–β3 and α2–β4. On the LRPPRC side of the interface, three neighbouring PPR motifs participate in binding SLIRP. The model suggests that hydrogen bonds may be formed between Glu80 of SLIRP and Arg438 of LRPPRC, and Arg60 of SLIRP and Asp479 of LRPPRC.

DISCUSSION

Previous studies have shown that LRPPRC and SLIRP interact to form a complex and are interdependent on each other (23,24,27). However, little was known about how this interaction is formed. The well characterized PPR10 in plants is thus far the largest all α-helical PPR protein for which a high-resolution structure is available making comparisons with LRPPRC relevant. PPR10 is composed of 19 PPR motifs that form a right-handed two-turn super-helix capped between additional α-helices at either terminal (13). In its apo state, PPR10 is a homo-dimer composed of two intertwined antiparallel solenoids of PPRs, thought to promote stability of the protein, before becoming a monomer upon binding its RNA substrate (14). Here, we show that SLIRP forms a hetero-dimer with LRPPRC to enhance its stability and to prevent formation of larger LRPPRC oligomers or aggregates. Interestingly, the C-terminal half of SLIRP is wedged at a five-way junction between PPR motifs of LRPPRC that are distant in the linear sequence, suggesting that, unlike PPR10, the structure of LRPPRC must fold back on itself. SLIRP may aid this process, as more intra-molecular interactions are observed within LRPPRC in its presence, likely making the overall protein structure more compact. Thus, the significantly larger LRPPRC comprising over 30 PPR motifs instead of 19 of PPR10 may have minimized the instability issues, common for PPR proteins, in this way instead of forming a homo-dimer.

We identify further amino acids from three neighbouring PPR motifs in the second quarter of LRPPRC that critically contribute to the binding of SLIRP. The charged or polar residues located in loops on opposite sides of the β-sheet within the SLIRP RRM motif are required to build the LRPPRC–SLIRP interface via a novel RRM-PPR binding mode. Interestingly, specific amino acid residues important for maintaining the interface integrity are located at positions in the PPRs of LRPPRC that are predicted by the PPR code to interact with RNA, and in the RNP1 motif within the SLIRP RRM domain. Thus, predicted RNA binding residues of both LRPPRC and SLIRP are used for protein–protein interactions, enabling the formation of a stable complex between the proteins. The use of an RNA binding domain to facilitate a protein–protein interaction has been observed before (28,39). Further, direct binding via two RNA binding domains is also known to occur, but typically these interactions are formed between RNA binding domains from the same family i.e. RRM-RRM, PPR-PPR or KH-KH. In plants, the plastid RNA editing protein ORRM1 interacts with the PPR proteins CRR28 or OTP82, although this interaction is mediated through the RIP–RIP domain and not its C-terminal RRM (40). Various structural elements of RRM domains i.e. the two α-helices, the β-sheet and loops in between, have been identified as participating in protein–protein interactions, however, proteins that use these motifs in this way, rather than for RNA binding, typically have non-canonical RNP1 and

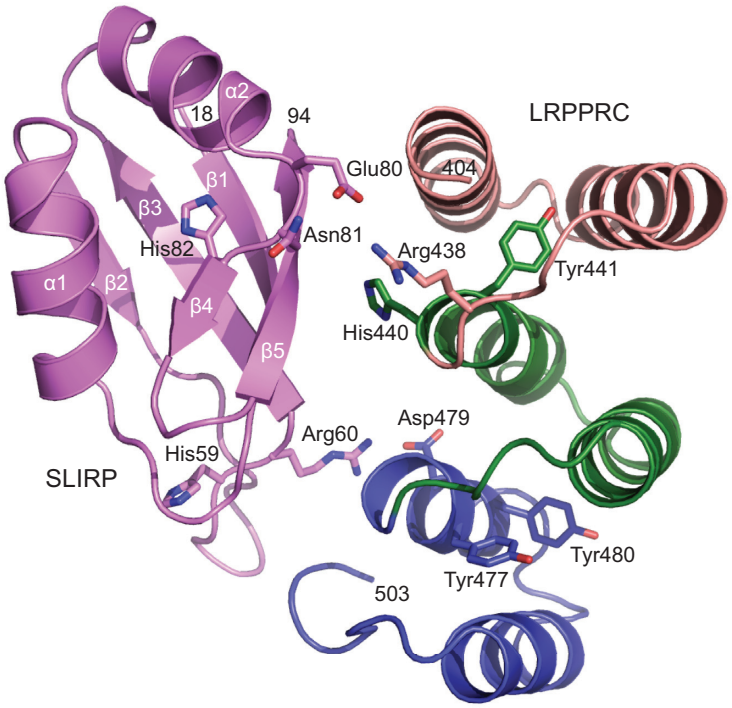

Figure 7. Structure prediction of the LRPPRC–SLIRP interface. SLIRP (amino acids 18–94) is shown in purple and the PPR motifs of LRPPRC (amino acids 404–503) are shown in salmon, green and blue, respectively. Amino acids crucial for LRPPRC–SLIRP interface integrity are shown with sticks.

RNP2 sequences. SLIRP can also be assigned to this category, at least for higher organisms like human and mouse, as the *Drosophila melanogaster* variants follow the canonical RNP consensus much more strictly, and fewer residues that are important for the LRPPRC–SLIRP binding interface are conserved in flies. Whether this means that alternative binding mechanisms between the LRPPRC and SLIRP variants have evolved in flies needs further investigation.

The structures of the heterodimeric splicing factors U2AF35-U2AF65 and U2AF65-SF1 belonging to the UHM family of proteins have revealed a tongue in groove binding mechanism where polar interactions surround a tryptophan from one protein that is inserted into a hydrophobic pocket of a non-canonical RRM motif of the other (41). In a similar way, there is also a highly conserved Trp residue (Trp443) in LRPPRC located in close proximity to R438-H440-Y441, which is required for binding SLIRP, and it is therefore tempting to speculate that this Trp residue may participate in the binding. In addition, UHM proteins have a conserved Arg-X-Phe motif located in a loop of the RRM motif that is of special importance for the binding mechanism. The Arg60 residue of SLIRP, which is involved in LRPPRC binding, is preceded by a phenylalanine and a histidine, constituting a similar motif, but in the reverse order. However, to elucidate if indeed the interaction between LRPPRC and SLIRP resembles the UHM protein binding mechanism and to gain a complete view of the LRPPRC–SLIRP interface, a high-resolution structure is clearly required.

Interestingly, PPR sequences are closely structurally related to tetratricopeptide repeat (TPR) proteins that have well characterized roles in protein–protein interactions. Previous studies have shown that PPRs can be used as sites of homodimerization for PPR proteins (13) and our work demonstrates that PPRs can also be used as protein–protein interaction motifs to engage other proteins. Recent studies have shown that a few proteins use their TPRs to bind RNA, such as the antiviral defence protein IFIT5 (42). Together these findings indicate that there is substantial functional overlap, as well as structural overlap, between these protein families.

We find that LRPPRC binds RNA in a strong and unspecific manner and that SLIRP at most has a very modest RNA binding capacity. The existence of non-canonical RNP sequences of human SLIRP, where even one residue (Arg60) is directly involved in binding LRPPRC, may shed some light on why it has such poor RNA binding capacity. A model of the LRPPRC–SLIRP interface further suggests that the RRM β-sheet of SLIRP partly faces the PPR domains of LRPPRC. An example with a similar binding mode has been observed for the Y14–magoh and UPF3b–UPF2 complexes where the RRM β-sheet of Y14 and UPF3b engages the RNP motifs to bind the α-helical surface of the ligand protein, which prevents RNA binding on the β-sheet surface (43,44).

Based on our findings, we suggest that the primary role of the RRM domain of SLIRP is not to bind RNA, but to bind LRPPRC to enhance its stability, folding and prevent aggregation. We have previously shown that SLIRP protects LRPPRC from degradation and has a role in fine-tuning the rate of mitochondrial protein synthesis by facilitating the ordered association of mature mRNAs with the mitochondrial ribosome (27). If not responsible for binding the mRNAs directly, SLIRP may support the interaction between

LRPPRC and ribosomal subunits or even act as a bridge between them and in this way present the mRNA to the ribosome.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENT

We thank the SPC team at EMBL-Hamburg for technical assistance with thermofluor experiments.

FUNDING

Swedish Research Council (2015-00418 to N.G.L.); Wellcome Trust (096919/Z/11/Z to R.N.L., Z.C.L.); National Health and Medical Research Council (APP1045677 to O.R.); Australian Research Council (DP140104111 to O.R.); Cancer Council of Western Australia (to O.R.); Alexander von Humboldt Foundation (to O.R.). Funding for open access charge: Max Planck Institute for Biology of Aging.
Conflict of interest statement. None declared.

REFERENCES

1. Hallberg,B.M. and Larsson,N.G. (2014) Making proteins in the powerhouse. *Cell Metab.*, **20**, 226–240.
2. Amunts,A., Brown,A., Toots,J., Scheres,S.H. and Ramakrishnan,V. (2015) Ribosome. The structure of the human mitochondrial ribosome. *Science*, **348**, 95–98.
3. Greber,B.J., Bieri,P., Leibundgut,M., Leitner,A., Aebersold,R., Boehringer,D. and Ban,N. (2015) Ribosome. The complete structure of the 55S mammalian mitochondrial ribosome. *Science*, **348**, 303–308.
4. Spahr,H., Habermann,B., Gustafsson,C.M., Larsson,N.G. and Hallberg,B.M. (2012) Structure of the human MTERF4-NSUN4 protein complex that regulates mitochondrial ribosome biogenesis. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 15253–15258.
5. Spahr,H., Samuelsson,T., Hallberg,B.M. and Gustafsson,C.M. (2010) Structure of mitochondrial transcription termination factor 3 reveals a novel nucleic acid-binding domain. *Biochem. Biophys. Res. Commun.*, **397**, 386–390.
6. Yakubovskaya,E., Guja,K.E., Mejia,E., Castano,S., Hambardjieva,E., Choi,W.S. and Garcia-Diaz,M. (2012) Structure of the essential MTERF4:NSUN4 protein complex reveals how an MTERF protein collaborates to facilitate rRNA modification. *Structure*, **20**, 1940–1947.
7. Filipovska,A., Razif,M.F., Nygard,K.K. and Rackham,O. (2011) A universal code for RNA recognition by PUF proteins. *Nat. Chem. Biol.*, **7**, 425–427.
8. Wang,X., McLachlan,J., Zamore,P.D. and Hall,T.M. (2002) Modular recognition of RNA by a human pumilio-homology domain. *Cell*, **110**, 501–512.
9. Metodiev,M.D., Spahr,H., Loguercio Polosa,P., Meharg,C., Becker,C., Altmueller,J., Habermann,B., Larsson,N.G. and Ruzzenente,B. (2014) NSUN4 is a dual function mitochondrial protein required for both methylation of 12S rRNA and coordination of mitoribosomal assembly. *PLoS Genet.*, **10**, e1004110.
10. Wredenberg,A., Lagouge,M., Bratic,A., Metodiev,M.D., Spahr,H., Mourier,A., Freyer,C., Ruzzenente,B., Tain,L., Gronke,S. *et al.* (2013) MTERF3 regulates mitochondrial ribosome biogenesis in invertebrates and mammals. *PLoS Genet.*, **9**, e1003178.
11. Barkan,A., Rojas,M., Fujii,S., Yap,A., Chong,Y.S., Bond,C.S. and Small,I. (2012) A combinatorial amino acid code for RNA recognition by pentatricopeptide repeat proteins. *PLoS Genet.*, **8**, e1002910.
12. Yagi,Y., Hayashi,S., Kobayashi,K., Hirayama,T. and Nakamura,T. (2013) Elucidation of the RNA recognition code for pentatricopeptide repeat proteins involved in organelle RNA editing in plants. *PLoS One*, **8**, e57286.
13. Yin,P., Li,Q., Yan,C., Liu,Y., Liu,J., Yu,F., Wang,Z., Long,J., He,J., Wang,H.W. *et al.* (2013) Structural basis for the modular recognition of single-stranded RNA by PPR proteins. *Nature*, **504**, 168–171.
14. Li,Q., Yan,C., Xu,H., Wang,Z., Long,J., Li,W., Wu,J., Yin,P. and Yan,N. (2014) Examination of the dimerization states of the single-stranded RNA recognition protein pentatricopeptide repeat 10 (PPR10). *J. Biol. Chem.*, **289**, 31503–31512.
15. Ke,J., Chen,R.Z., Ban,T., Zhou,X.E., Gu,X., Tan,M.H., Chen,C., Kang,Y., Brunzelle,J.S., Zhu,J.K. *et al.* (2013) Structural basis for RNA recognition by a dimeric PPR-protein complex. *Nat. Struct. Mol. Biol.*, **20**, 1377–1382.
16. Coquille,S., Filipovska,A., Chia,T., Rajappa,L., Lingford,J.P., Razif,M.F., Thore,S. and Rackham,O. (2014) An artificial PPR scaffold for programmable RNA recognition. *Nat. Commun.*, **5**, 5729.
17. Shen,C., Zhang,D., Guan,Z., Liu,Y., Yang,Z., Yang,Y., Wang,X., Wang,Q., Zhang,Q., Fan,S. *et al.* (2016) Structural basis for specific single-stranded RNA recognition by designer pentatricopeptide repeat proteins. *Nat. Commun.*, **7**, 11285.
18. Rackham,O. and Filipovska,A. (2012) The role of mammalian PPR domain proteins in the regulation of mitochondrial gene expression. *Biochim. Biophys. Acta*, **1819**, 1008–1016.
19. Schwinghammer,K., Cheung,A.C., Morozov,Y.I., Agaronyan,K., Temiakov,D. and Cramer,P. (2013) Structure of human mitochondrial RNA polymerase elongation complex. *Nat. Struct. Mol. Biol.*, **20**, 1298–1303.
20. Reinhard,L., Sridhara,S. and Hallberg,B.M. (2015) Structure of the nuclease subunit of human mitochondrial RNase P. *Nucleic Acids Res.*, **43**, 5664–5672.
21. Mootha,V.K., Lepage,P., Miller,K., Bunkenborg,J., Reich,M., Hjerrild,M., Delmonte,T., Villeneuve,A., Sladek,R., Xu,F. *et al.* (2003) Identification of a gene causing human cytochrome c oxidase deficiency by integrative genomics. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 605–610.
22. Chujo,T., Ohira,T., Sakaguchi,Y., Goshima,N., Nomura,N., Nagao,A. and Suzuki,T. (2012) LRPPRC/SLIRP suppresses PNPase-mediated mRNA decay and promotes polyadenylation in human mitochondria. *Nucleic Acids Res.*, **40**, 8033–8047.
23. Ruzzenente,B., Metodiev,M.D., Wredenberg,A., Bratic,A., Park,C.B., Camara,Y., Milenkovic,D., Zickermann,V., Wibom,R., Hultenby,K. *et al.* (2012) LRPPRC is necessary for polyadenylation and coordination of translation of mitochondrial mRNAs. *EMBO J.*, **31**, 443–456.
24. Sasarman,F., Brunel-Guitton,C., Antonicka,H., Wai,T., Shoubridge,E.A. and Consortium,L. (2010) LRPPRC and SLIRP interact in a ribonucleoprotein complex that regulates posttranscriptional gene expression in mitochondria. *Mol. Biol. Cell*, **21**, 1315–1323.
25. Sasarman,F., Nishimura,T., Antonicka,H., Weraarpachai,W., Shoubridge,E.A. and Consortium,L. (2015) Tissue-specific responses to the LRPPRC founder mutation in French Canadian Leigh Syndrome. *Hum. Mol. Genet.*, **24**, 480–491.
26. Wilson,W.C., Hornig-Do,H.T., Bruni,F., Chang,J.H., Jourdain,A.A., Martinou,J.C., Falkenberg,M., Spahr,H., Larsson,N.G., Lewis,R.J. *et al.* (2014) A human mitochondrial poly(A) polymerase mutation reveals the complexities of post-transcriptional mitochondrial gene expression. *Hum. Mol. Genet.*, **23**, 6345–6355.
27. Lagouge,M., Mourier,A., Lee,H.J., Spahr,H., Wai,T., Kukat,C., Silva Ramos,E., Motori,E., Busch,J.D., Siira,S. *et al.* (2015) SLIRP regulates the rate of mitochondrial protein synthesis and protects LRPPRC from degradation. *PLoS Genet.*, **11**, e1005423.
28. Muto,Y. and Yokoyama,S. (2012) Structural insight into RNA recognition motifs: versatile molecular Lego building blocks for biological systems. *Wiley Interdiscip. Rev. RNA*, **3**, 229–246.
29. Baggio,F., Bratic,A., Mourier,A., Kauppila,T.E., Tain,L.S., Kukat,C., Habermann,B., Partridge,L. and Larsson,N.G. (2014) Drosophila melanogaster LRPPRC2 is involved in coordination of mitochondrial translation. *Nucleic Acids Res.*, **42**, 13920–13938.
30. Bratic,A., Wredenberg,A., Gronke,S., Stewart,J.B., Mourier,A., Ruzzenente,B., Kukat,C., Wibom,R., Habermann,B., Partridge,L. *et al.* (2011) The bicoid stability factor controls polyadenylation and

expression of specific mitochondrial mRNAs in Drosophila melanogaster. *PLoS Genet.*, **7**, e1002324.
31. Harmel,J., Ruzzenente,B., Terzioglu,M., Spahr,H., Falkenberg,M. and Larsson,N.G. (2013) The leucine-rich pentatricopeptide repeat-containing protein (LRPPRC) does not activate transcription in mammalian mitochondria. *J. Biol. Chem.*, **288**, 15510–15519.
32. Boivin,S., Kozak,S. and Meijers,R. (2013) Optimization of protein purification and characterization using Thermofluor screens. *Protein Expr. Purif.*, **91**, 192–206.
33. Grimm,M., Zimniak,T., Kahraman,A. and Herzog,F. (2015) xVis: a web server for the schematic visualization and interpretation of crosslink-derived spatial restraints. *Nucleic Acids Res.*, **43**, W362–W369.
34. Ule,J., Jensen,K., Mele,A. and Darnell,R.B. (2005) CLIP: a method for identifying protein-RNA interaction sites in living cells. *Methods*, **37**, 376–386.
35. Roy,A., Kucukural,A. and Zhang,Y. (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.*, **5**, 725–738.
36. van Zundert,G.C., Rodrigues,J.P., Trellet,M., Schmitz,C., Kastritis,P.L., Karaca,E., Melquiond,A.S., van Dijk,M., de Vries,S.J. and Bonvin,A.M. (2016) The HADDOCK2.2 web server: user-friendly integrative modeling of biomolecular complexes. *J. Mol. Biol.*, **428**, 720–725.
37. Karpenahalli,M.R., Lupas,A.N. and Soding,J. (2007) TPRpred: a tool for prediction of TPR-, PPR- and SEL1-like repeats from protein sequences. *BMC Bioinformatics*, **8**, 2.
38. Yagi,Y., Tachikawa,M., Noguchi,H., Satoh,S., Obokata,J. and Nakamura,T. (2013) Pentatricopeptide repeat proteins involved in plant organellar RNA editing. *RNA Biol.*, **10**, 1419–1425.
39. Clery,A., Blatter,M. and Allain,F.H. (2008) RNA recognition motifs: boring? Not quite. *Curr. Opin. Struct. Biol.*, **18**, 290–298.
40. Sun,T., Germain,A., Giloteaux,L., Hammani,K., Barkan,A., Hanson,M.R. and Bentolila,S. (2013) An RNA recognition motif-containing protein is required for plastid RNA editing in Arabidopsis and maize. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, E1169–E1178.
41. Kielkopf,C.L., Lucke,S. and Green,M.R. (2004) U2AF homology motifs: protein recognition in the RRM world. *Genes Dev.*, **18**, 1513–1526.
42. Abbas,Y.M., Pichlmair,A., Gorna,M.W., Superti-Furga,G. and Nagar,B. (2013) Structural basis for viral 5′-PPP-RNA recognition by human IFIT proteins. *Nature*, **494**, 60–64.
43. Kadlec,J., Izaurralde,E. and Cusack,S. (2004) The structural basis for the interaction between nonsense-mediated mRNA decay factors UPF2 and UPF3. *Nat. Struct. Mol. Biol.*, **11**, 330–337.
44. Fribourg,S., Gatfield,D., Izaurralde,E. and Conti,E. (2003) A novel mode of RBD-protein recognition in the Y14-Mago complex. *Nat. Struct. Biol.*, **10**, 433–439.