

A putative RNA binding protein from *Plasmodium vivax* apicoplast

Sofía M. García-Mauriño¹, Antonio Díaz-Quintana¹, Francisco Rivero-Rodríguez¹, Isabel Cruz-Gallardo², Christian Grüttner¹, Marian Hernández-Vellisca¹ and Irene Díaz-Moreno¹

¹ Instituto de Investigaciones Químicas (IIQ), Centro de Investigaciones Científicas Isla de la Cartuja (cicCartuja), Universidad de Sevilla, Consejo Superior de Investigaciones Científicas (CSIC), Sevilla, Spain

² King's College London, Randall Division of Cell & Molecular Biophysics, London, UK

Keywords

apicoplast RNA binding protein; malarial *Plasmodium* parasite; protein aggregation

Correspondence

I. Díaz-Moreno, Instituto de Investigaciones Químicas (IIQ), Centro de Investigaciones Científicas Isla de la Cartuja (cicCartuja), Universidad de Sevilla, Consejo Superior de Investigaciones Científicas (CSIC), Avenida Américo Vespucio 49, 41092 Sevilla, Spain
Fax: +34 954 460065
Tel: +34 954 489513
E-mail: idiazmoreno@us.es

(Received 3 October 2017, revised 3 November 2017, accepted 14 November 2017)

doi:10.1002/2211-5463.12351

Malaria is caused by Apicomplexa protozoans from the *Plasmodium* genus entering the bloodstream of humans and animals through the bite of the female mosquitoes. The annotation of the *Plasmodium vivax* genome revealed a putative RNA binding protein (apiRBP) that was predicted to be trafficked into the apicoplast, a plastid organelle unique to Apicomplexa protozoans. Although a 3D structural model of the apiRBP corresponds to a noncanonical RNA recognition motif with an additional C-terminal α -helix (α_3), preliminary protein production trials were nevertheless unsuccessful. Theoretical solvation analysis of the apiRBP model highlighted an exposed hydrophobic region clustering α_3 . Hence, we used a C-terminal GFP-fused chimera to stabilize the highly insoluble apiRBP and determined its ability to bind U-rich stretches of RNA. The affinity of apiRBP toward such RNAs is highly dependent on ionic strength, suggesting that the apiRBP–RNA complex is driven by electrostatic interactions. Altogether, apiRBP represents an attractive tool for apicoplast transcriptional studies and for antimalarial drug design.

Malaria is one of the most devastating parasitic diseases in the world, causing death to 1–2 million people per year, mostly children. This disease is caused by Apicomplexa protozoans from the *Plasmodium* genus passing onto humans and animals through the bite of the female mosquitoes from the *Anopheles* genus [1]. The discovery of an essential plastid organelle, the so-called apicoplast, has rekindled current search for new drugs to fight malaria [2]. Widely found in Apicomplexa except for *Cryptosporidium* species [3], the apicoplast is a vestigial nonphotosynthetic plastid surrounded by four membranes due to its secondary

endosymbiosis origin. Indeed, an ancient eukaryotic cell engulfed a cyanobacterium to become a photosynthetic eukaryotic alga. Then, a *Plasmodium* predecessor ingested the eukaryotic alga to establish a new symbiosis and preserved it as a plastid [4]. Hence, the prokaryote-derived metabolic pathways within the apicoplast are substantially different to those from the human host, bringing new opportunities to design drugs against malaria [5].

The apicoplast genome of *Plasmodium* spp. consists of a highly conserved ~35-kb circular, double-stranded DNA lacking the genes encoding proteins

Abbreviations

apiRBP, apicoplast RNA binding protein; ARE, AU-rich element; CD, circular dichroism; DDM, n-dodecyl- β -D-maltoside; DTT, dithiothreitol; ER, endoplasmic reticulum; IPTG, isopropyl- β -D-thiogalactopyranoside; ITC, isothermal titration calorimetry; LB, Luria–Bertani; MD, molecular dynamics; NLS, nuclear localization signal; PATS, prediction of apicoplast-targeted sequences; PlasmAP, *Plasmodium* apicoplast prediction; PMSF, phenylmethylsulfonyl fluoride; RBD, RNA binding domain; RBP, RNA binding protein; RNP1, ribonucleoprotein consensus sequence 1; RNP2, ribonucleoprotein consensus sequence 2; RRM, RNA recognition motif; SP, signal peptide; TP, transit peptide; U-rich, uridine-rich.

involved in photosynthesis [6]. Unlike mRNAs produced in the parasite nucleus, those synthesized in the apicoplast are polycistronic, and they mainly provide the essential machinery for transcription and translation needed for organelle housekeeping functions [7]. An important exception is the presence of the SufB protein (*ycf24*), which is a constituent of the [Fe-S] biogenesis pathway [8]. To our knowledge, a small number of proteins involved in transcription, control of mRNA stability, and translation within the organelle have been studied [7,9,10].

Many genes initially encoded in the apicoplast have been transferred to the parasite nucleus to avoid deleterious mutations of nonrecombinant genomes. Consequently, apicoplast biogenesis and function critically relies on targeting nuclear encoded proteins back to the organelle, by distinctive apicoplast targeting motifs. Apicoplast proteins present a bipartite leader sequence at its N terminus consisting on a hydrophobic signal peptide (SP) and a chloroplast-like transit peptide (TP). While the SP allows the entry into the secretory pathway, the TP is a relatively simple and flexible trafficking signal for post-translational targeting and translocation to the apicoplast [2,11] (Fig. 1).

Up to 466 apicoplast predicted proteins were identified in the *P. falciparum* genome using the PlasmAP and PATS algorithms that extract amino acid features from TP that target proteins to the organelle [12,13]. A similar analysis in the *P. vivax* genome revealed the presence of 316 proteins predicted to be targeted to

the apicoplast [14,15]. One of them (PVX_084415 or apiRBP in this manuscript) is a putative RNA binding protein (RBP) displaying a predicted N-terminal signal peptide and a single RNA recognition motif (RRM). In *P. falciparum*, 189 genes have been annotated as putative RBPs: 179 of them possess an ortholog in *P. vivax*, including the one coding apiRBP [16]. Unfortunately, most of them lack definitive functional annotations and only a few have been structurally characterized (e.g., PDB: 2N7C and 2MYF).

To unveil fundamental aspects of RNA metabolism in *Plasmodium* that may hint to new targets for anti-malarial therapy, we took the challenge of characterizing apiRBP, which is extremely prone to aggregation. In this work, we resorted to theoretical solvation analysis of a 3D model of the protein to design a soluble GFP-fused chimera. Finally, we proved by calorimetry assays that apiRBP is indeed able to recognize target RNA stretches, presumably driven by electrostatic interactions.

Materials and methods

apiRBP plasmid constructs

Three different apiRBP plasmid constructs were used in this study. The first one, His-apiRBP* (residues 76–182), was obtained by amplification of the PVX_084415 gene from *P. vivax* SalI cDNA and was cloned into the NcoI–NotI cloning sites of the pETM-11 vector (Invitrogen) for protein expression in *Escherichia coli*. The second one,

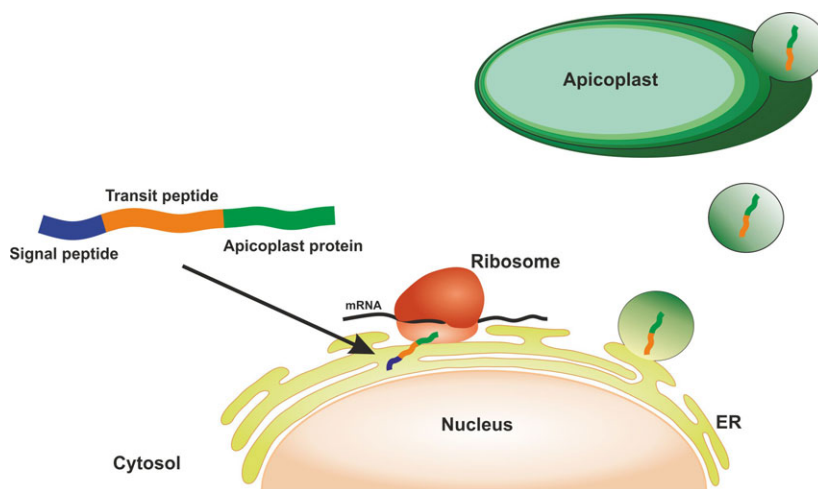


Fig. 1. Protein targeting to the *Plasmodium* spp. apicoplast. Most apicoplast proteins traffic to the organelle due to N-terminal bipartite leaders (SP and TP). The nascent apicoplast proteins are targeted to the endoplasmic reticulum (ER) membrane, where its SP is removed during cotranslation by a signal peptidase. The exact mechanisms that lead the transport from the ER lumen to the apicoplast are not fully understood but may involve ER–apicoplast communication by vesicular transport and TP recognition [2].

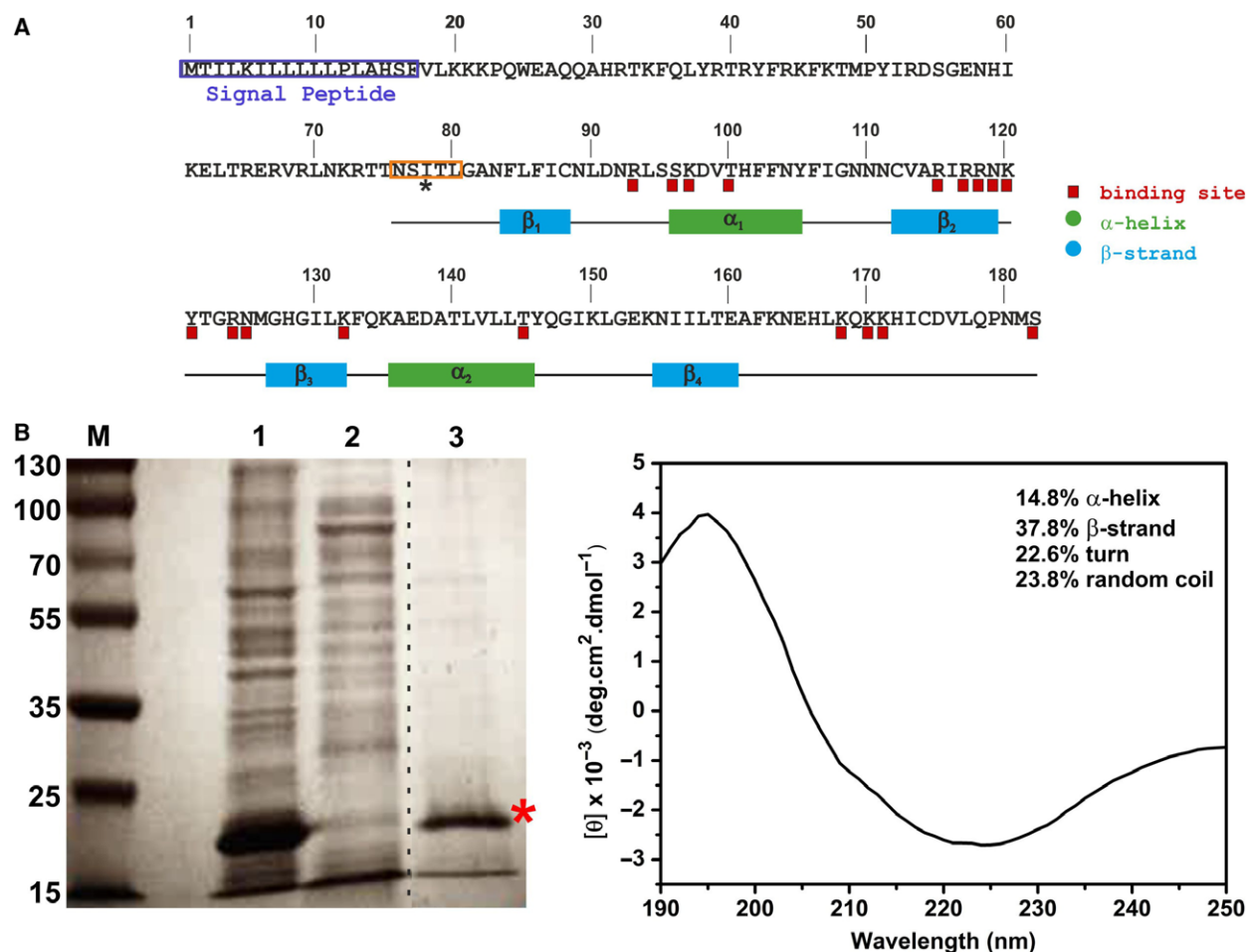


Fig. 2. apiRBP amino acid sequence, solubilization from inclusion bodies and secondary structure characterization. (A) apiRBP full-length sequence along with the constructs used in this study: His-apiRBP* (residues 76–182) and His-apiRBP (residues 81–182). The five N-terminal amino acids that are included in the first construct, but not in the second, are marked by an orange square and an asterisk. The apiRBP SP is colored in purple. Predicted secondary structure elements by JPred V.4 [30] (α -strand and β -sheet) and predicted RNA binding sites by BindN software [31] are also indicated. (B) (Left) SDS/PAGE Coomassie gel (12%) of His-apiRBP* (residues 76–182) expression in *E. coli*. M stands for molecular mass marker in kDa. Lanes 1 and 2 are the pellet and the supernatant from cell sonication, while lane 3 stands for the supernatant of inclusion bodies refolding. The protein was found in fractions from lanes 1 and 3 (marked with a red asterisk) with a molecular weight of 19 kDa. All lanes come from the same gel, but it has been spliced and put together (dashed line between lanes 2 and 3). (Right) Far-UV (190–250 nm) CD spectrum of the refolded protein indicating the contents of secondary structure elements.

His-apiRBP (residues 81–182), is a derivative construct from His-apiRBP* but lacking the first five amino acids (NSITL) (Fig. 2A). It was cloned into the pIVEX2.4d cell-free expression vector by the In-Fusion cloning kit (Clontech) following the manufacturer's instructions. The third construct, apiRBP-GFP-His, was cloned in a modified pET21a(+) vector. Such vector was kindly provided by Prof. Frank Bernhard (Frankfurt, Germany) and contains a C-terminal GFP-His-tag (superfolderGFP) as a quantitative reporter of gene expression. In this variant, the apiRBP gene also comprehends residues 81–182. All primer sequences are available upon request.

Protein expression

His-apiRBP* (residues 76–182) recombinant protein was expressed in Luria-Bertani (LB) medium in *E. coli* BL21 (DE3) cells. Cultures were grown at 37 °C with continuous agitation of 180 r.p.m. Protein expression was induced by the addition of isopropyl- β -D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM once an OD_{600} of 0.8 was reached. Cells were collected by centrifugation after 5 h of continuous agitation at 30 °C. To isolate the inclusion bodies, cells obtained from 1 L of culture were resuspended in 20 mL of 100 mM Tris/HCl buffer (pH 8) with

100 mM NaCl, 1 mM dithiothreitol (DTT), and 1 mM phenylmethylsulfonyl fluoride (PMSF). Lysozyme was added to the cell suspension ($0.35 \text{ mg}\cdot\text{mL}^{-1}$), which was then sonicated. After that, the suspension was treated with DNase I ($20 \text{ mg}\cdot\text{mL}^{-1}$) for 1 h at 37°C and centrifuged at $30\,000 \text{ g}$ for 15 min at 4°C . The pellet, containing the inclusion bodies, was washed twice with ice-cold PBS buffer enriched in 1% of Triton X-100. The inclusion bodies were then resuspended in 2 mL of 100 mM Tris/HCl ice-cold buffer (pH 8) containing 100 mM NaCl, 1 mM DTT, and 6 M guanidinium chloride. The suspension was incubated at 20°C for 2 h and centrifuged at $100\,000 \text{ g}$ for 20 min at 4°C . The inclusion bodies were soluble in the supernatant fraction, which mainly contained the recombinant His-*apiRBP**, so protein refolding was performed without further purification. Two milliliters of inclusion bodies was added to 200 mL of ice-cold refolding buffer containing 100 mM Tris/HCl (pH 8), 5 mM EDTA, and 0.5 M L-arginine with low agitation at 15°C [17]. After 24-h incubation at 15°C with continuous agitation, the protein solution was centrifuged and most of the expressed His-*apiRBP** was soluble in the supernatant which was dialyzed against 100 mM Tris/HCl (pH 8) buffer.

His-*apiRBP* (residues 81–182) was expressed by cell-free protein synthesis. For this purpose, 50 μL of reaction in batch mode was produced. Briefly, $10 \mu\text{g}\cdot\text{mL}^{-1}$ of His-*apiRBP* in pIVEX2.4d was added to the reaction mixture. The reaction mixture consisted of 1 mM amino acid mix, 0.8 mM ribonucleotides (guanosine-, uracil-, and cytidine-triphosphate), 1.2 mM adenosine triphosphate, 55 mM HEPES, pH 7.5, 68 μM folinic acid, 0.64 mM cyclic adenosine monophosphate, 3.4 mM dithiothreitol, 27.5 mM ammonium acetate, 2 mM spermidine, 80 mM creatine phosphate, 208 mM potassium glutamate, 16 mM magnesium acetate, 250 $\mu\text{g}\cdot\text{mL}^{-1}$ creatine kinase, 27 $\mu\text{g}\cdot\text{mL}^{-1}$ T7 RNA polymerase, 0.175 $\mu\text{g}\cdot\text{mL}^{-1}$ tRNA, and 67 $\mu\text{L}\cdot\text{mL}^{-1}$ S30 *E. coli* bacterial extract. Incubation was carried out at 23°C with agitation for 16 h.

Finally, the *apiRBP*-GFP-His chimera was produced in LB medium in *E. coli* BL21 (DE3) cells. Cultures were grown at 37°C with continuous agitation of 180 r.p.m. Protein expression was induced by the addition of IPTG to a final concentration of 0.5 mM after reaching OD_{600} of 0.3. Cells were collected by centrifugation after 16 h of continuous agitation at 20°C . Posterior to cell disruption by sonication in the presence of 1 mM PMSF and cOmplete Protease Inhibitors (one tablet per 50 mL extraction solution; Sigma), extracts were centrifuged at $17\,000 \text{ g}$ and the His-tagged protein was then purified by nickel affinity chromatography (Ni Sepharose 6 Fast Flow; GE Healthcare) applying an imidazole gradient (10–300 mM) in 20 mM Tris and 100 mM NaCl, pH 7.4. The purified fractions containing *apiRBP*-GFP-His were then submitted to FPLC (AKTA Prime) to remove protein contaminants. Protein concentrations were determined spectrophotometrically by

the Bradford assay, and the molecular weight of the constructs was verified by MALDI-TOF spectroscopy.

Circular dichroism spectroscopy

Circular dichroism (CD) spectrum was recorded on a Jasco J-815 spectropolarimeter equipped with a Peltier temperature control system. The secondary structure analysis of His-*apiRBP** (residues 76–182) was performed by recording far-UV CD spectra (190–250 nm) with 3 μM samples in H_2O at 25°C . The spectrum was an average of 20 scans. The α -helix and β -sheet content was obtained with CDPRO software [18], which includes the algorithms CONTIN, SELCON and CDSSTR and the CLSTR option to compare the protein folding with a set of similar folded proteins.

Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) experiments were performed using a microcalorimeter (TA Instruments) at 25°C titrating *apiRBP*-GFP-His or GFP-His (10 μM) over 10-mer U-rich RNA oligonucleotides (1 μM ; Sigma). Both the proteins and oligonucleotide samples were in 20 mM Tris (pH 7.4) buffer with/without 50 mM NaCl (indicated in each experiment). Measurements were repeated at least twice. The integrated data of heat per injection normalized per mol of injectant versus molar ratio were analyzed with AFFINIMETER software v2.1608 and were fitted to a 1 : 1 interaction model.

Structural model

The 3D model for *apiRBP* (81–182) was built with Ginzu and Rosetta [19,20] as implemented in the Robetta server (<http://rosetta.bakerlab.org>). The Ginzu prediction assigned the sequence to a cluster of RRM-containing proteins, whose reference parent is the histone-lysine N-methyltransferase SETD1A (Chao *et al.*, unpublished, PDB: 3S8S).

Molecular dynamics computations

Molecular dynamics (MD) computations were carried out in a periodic orthorhombic box using AMBER12 with the AMBER 14SB force field [21] and PME electrostatics with a Ewald summation cutoff of 9 Å. For this purpose, the system was neutralized with seven Cl^- ions and solvated with 6058 OPC [22] water molecules. The protein model side chains were relaxed by energy minimization. Then, solvent and counter-ions were subjected to 5000 energy minimization steps followed by 500-ps NPT-MD computations using isotropic molecule position scaling and a pressure relaxation time of 2 ps at 298 K. Temperature was controlled by a Langevin thermostat [23] with a collision

frequency of 5 ps^{-1} . The density of the system reached a plateau during the first 50 ps. Then, the whole system was energy-minimized and submitted to NVT-MD computations at 298 K. The SHAKE algorithm [24] was used to constrain bonds involving hydrogen atoms. Trajectory analyses, including grid inhomogeneous solvation theory (GIST) [25] analyses, were carried out with CPPTRAJ [26]. Grid spacing was 0.5 Å. UCSF Chimera [27] was the software used for molecular graphics and modeling interface.

Results

apiRBP annotation and secondary structure

A search of the apiRBP gene (transcript ID PVX_084415) in PlasmoDB (the official database of *Plasmodium* sequencing projects [28]) predicted an RNA binding protein (RBP) of 182 amino acids. In its N-terminal end, a 17-residue SP was also annotated by combining the predictions of SignalP 3.0 with orthology information (Fig. 2A). However, the plastid TP of apiRBP has not been bounded yet, probably due to the fact that they are normally variable in length, have no primary consensus sequence, and are only distinguished by positively charged residues and an abundance of hydroxylated residues [29]. The InterPro Domain search revealed that apiRBP contains a single RNA recognition motif (RRM) of about 90 amino acids whose secondary structure analysis by JPred V.4 [30] corresponded to the canonical $\beta_1\alpha_1\beta_2\beta_3\alpha_2\beta_4$ topology, with four-stranded β -sheets packed against two α -helices. Different lengths of the RRM are proposed, depending on the superfamily taken as a reference; for instance, SSF54928 superfamily aligns with residues 75–169, while the PS50102 Prosite domain aligns with residues 83–163. Hence, different construct lengths were tested in this report (see below). BindN software [31] also identified several potential RNA-binding residues within the RRM taking into account the pK_a value, hydrophobicity index, and molecular mass of each amino acid. Of note, among those binding residues, is the prevalence of positively charged residues (Arg and Lys stretches), especially the ones localized in the β_2 -strand and the C-terminal tail (Fig. 2A).

His-apiRBP* (residues 76–182, more alike to the SSF54928 superfamily) was produced in *E. coli* cells and isolated from inclusion bodies due to its low solubility in the supernatant fraction (Fig. 2B, *left*). Correct folding and secondary structure content were assessed by CD (Fig. 2B, *right*). The secondary structure content was 14.8% α -helix, 37.8% β -strand, 22.6% turn, and 23.8% random coil, in agreement with the expected values for an RRM domain, although the α -helix content

was slightly higher than expected [32,33]. Unfortunately, protein stability was too low to perform further functional assays. To test whether such instability was caused by the presence of the end of the TP in the N terminus of the apiRBP construct, we designed a shorter construction excluding amino acids 76–80, His-apiRBP (residues 81–182, more alike to the PS50102 Prosite domain) (Fig. 2A); and cloned it in a vector compatible with cell-free protein expression. An increase in protein solubility was achieved, still insufficient (with or without detergents) to allow the structural and functional characterization of the protein under any of the conditions tested (Fig. S1).

apiRBP structural modeling

A 3D structural model for apiRBP was obtained by the comparative methods implemented in the Robetta server (<http://rosetta.bakerlab.org>) (Fig. 3A). The structural model displayed the expected $\beta_1\alpha_1\beta_2\beta_3\alpha_2\beta_4$ topology of canonical RRM followed by an extra α -helix (herein named as α_3) using the structure of the histone-lysine N-methyltransferase SETD1A (PDB: 3S8S) and other RRM-containing proteins as templates. Moreover, a structural homology search using DALI database [34] returned a Z-score of 14.4, with an RMSD value of 1.8 Å, between apiRBP and SETD1A, including the novel C-terminal α_3 -helix found in apiRBP. Comparison of apiRBP with another RBPs, such as ELAV-like protein 1 (PDB: 4FXV), returned a Z-score of 12.3, although DALI consistently excluded the α_3 -helix of apiRBP from such structural alignments. The 10-residue α_3 -helix is oriented parallel with respect to the β -sheet, mainly through contacts between Leu167 at α_3 and the β -sheet face (Phe86 in β_1 and Ile130 in β_3), along with the H-bond involving Asn164 (α_3 -O_{δ1}) and Arg117 (β_2 -guanidinium groups). Additionally, the sequence of apiRBP shows a two-residue insertion at β_2 (Ala114 and Arg115) with respect to homologous RRM domains such as those from ELAV-like protein 1 (31.2% identity; PDB: 4FXV), the RNA recognition motif 1 from HuR (31.2% identity; PDB: 3HI9), or RNA binding domain 1 from HuC (31.9% identity; PDB: 1D8Z). However, the secondary structure of β_2 seems to be unaffected in the model. Interestingly, the ribonucleoprotein consensus sequences (RNP2: [ILV]-[FY]-[ILV]-X-N-L and RNP1: [RK]-G-[FY]-[GA]-[FY]-[ILV]-X-[FY]) of RRM domains at the β_1 - and β_3 -strands [33] are not conserved in apiRBP showing only two aromatic residues (Phe84 and Phe86) at the RNP2 of the β_1 -strand (Fig. 3A). Further, the loop joining β_4 and α_3 covers the aromatic ring of Phe86, probably preventing it to form π -stacking interactions with RNA.

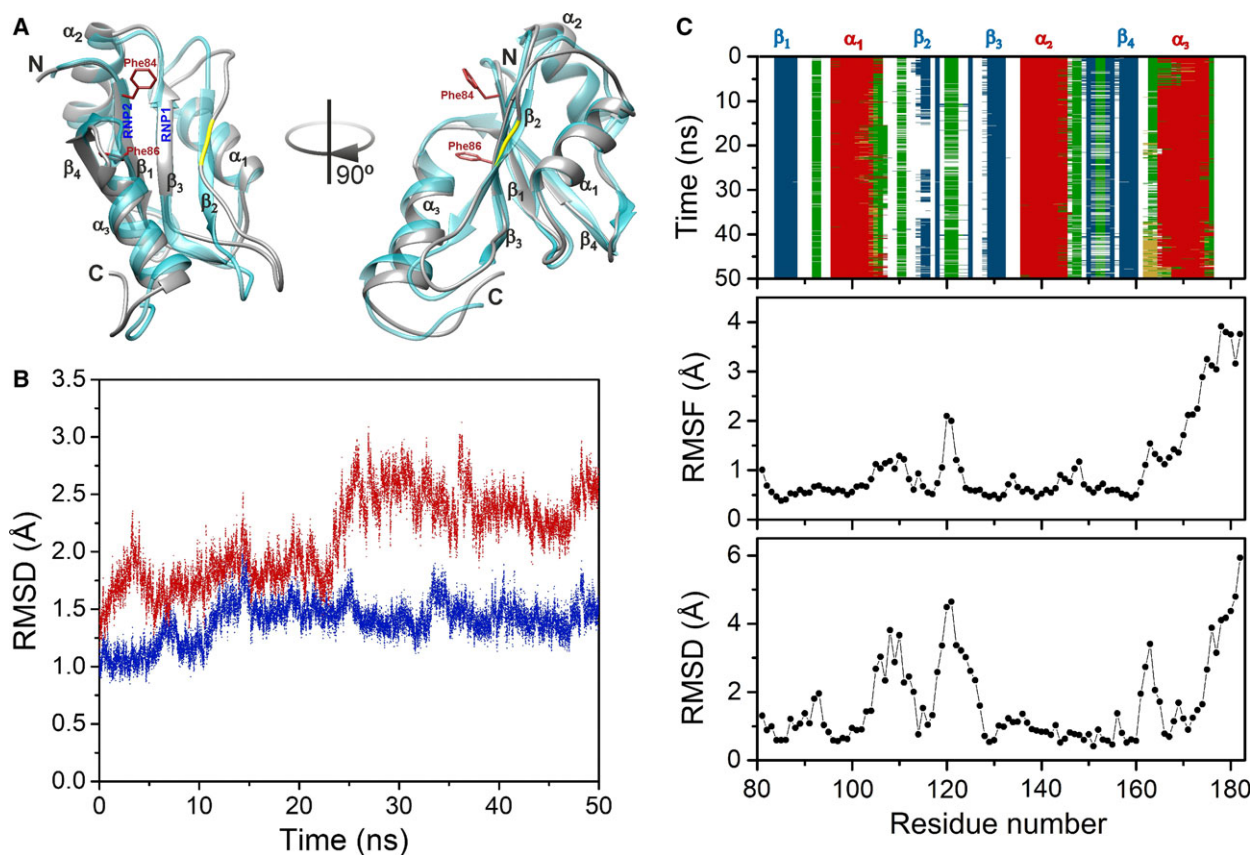


Fig. 3. apiRBP structural modeling. (A) Overlay of the apiRBP structural model obtained with Robetta (cyan) (<http://robetta.bakerlab.org>) and the average one (gray) from MD computations in ribbon representation (residues 81–182). The side chains of aromatic residues of the β_1 -sheet (Phe84 and Phe86) are displayed in red sticks. The residues in β_2 that are absent in the sequences of homologous RRM s are highlighted in yellow. The location where the RNP1 and RNP2 consensus sequences should appear is also marked. Each view is rotated 90° around the vertical axis of the molecule. (B) Time course of RMSD values for the MD trajectory. Data corresponding to the full model (residues 81 to 182) are depicted in red and those computed for the canonical RRM domain (residues 81 to 161) in blue. (C) (*Upper panel*) Secondary structure time course along the MD trajectory. β -Strands are in blue, bends in gray, turns in green, 3_{10} -helices in ocher, and α -helices in red. (*Center panel*) Per-residue atomic fluctuations along the last 10 ns of the trajectory. (*Bottom*) Per-residue backbone RMSD computed between the model and the average structure from MD.

To check the behavior of β_2 and α_3 , and to understand the experimentally observed instability of the protein, a 50-ns molecular dynamics computation was reckoned (Fig. 3). The RMSD evolution depended on whether residues 81–161 or 81–182 were used for the structural alignments (Fig. 3B). This indicates that the C-terminal end of apiRBP—including α_3 —is more mobile than the canonical RRM core. The secondary structure timeline (Fig. 3C) shows that the structure of α_3 is stable along the trajectory. Thus, the RMSD differences in Figure 3B can be attributed to a slight rigid-body motion of α_3 and the high mobility of the last C-terminal residues (stretch 176–182). Interestingly, the three-amino acid sequence (from 162 to 164) neighboring the N terminus of α_3 -helix is folded in a short 3_{10} -helix along MD (highlighted in ocher in

Fig. 3C). Consistently, residues 162–182 show the highest fluctuations and RMSD values when comparing the average structure to the initial, energy-minimized model (Fig. 3C). In addition, β_2 resulted in being unstable probably because of the above-mentioned two-residue insertion (highlighted in yellow in Fig. 3A) with respect to other RRM s. Notably, the sequence stretches flanking the insertion showed high fluctuations and RMSD values with respect to the initial, energy-minimized model (Fig. 3A and 3C).

Solvation analysis of apiRBP

To understand the high propensity of apiRBP to aggregate in solution, MD computations were performed with explicit OPC water molecules, which are

the best to represent the bulk properties of water [22]. Then, the trajectories were analyzed using the grid inhomogeneous solvation theory to assess the number density of water molecules around the domain (Fig. 4A and 4B). Notably, the density of water molecules decreased near the surface patch of the RRM, and the translational entropy of water molecules in this region increased with respect to the bulk. This revealed a hydrophobic cluster near the C terminus involving side chains from Ile173 at α_3 and Val176, Leu177, and Pro179.

apiRBP binding to U-rich RNA stretches

Taking into account the results from the solvation analysis, a new version of the apiRBP construct (residues 81–182) was designed in-frame with a C-terminal GFP-His-tag (namely apiRBP-GFP-His). Presumably, the high stability and solubility of the GFP-tag would serve as a stabilizing factor for the insoluble apiRBP, as previously seen for other protein targets [35]. As anticipated, when recombinant apiRBP-GFP-His expression was performed in *E. coli*, the chimera was stable enough to perform further binding analysis (Fig. S2). As no RNA binding activity has been reported for the structurally apiRBP-related SETD1A RRM domain [36,37], nor for its yeast homologous SET1 RRM1 [38], in this work we tested the ability of apiRBP to bind uridine-rich (U-rich) 10-mer RNA. This is a consensus AU-rich element (ARE) at the

3'-UTR of mRNAs recognized by ELAV family of proteins that share ~31% of sequence identity with apiRBP. Isothermal titration calorimetry (ITC) performed at low ionic strength showed that apiRBP-GFP-His recognized 10-mer U-rich RNA stretches (Fig. 5A, *left*) with an affinity within the μM range when fitting the stoichiometry to $n = 1$ (Table 1). This indicates that one molecule of apiRBP-GFP-His binds to one molecule of RNA (1 : 1). When 50 mM of NaCl was added to the ITC buffer, the recognition of 10-mer U-rich RNA was negligible (Fig. 5B, *left*). Beyond the high isoelectric point (pI) of apiRBP ($pI = 9.72$), its surface electrostatic potential at low ionic strength (Fig. 5A, *center* and *right*) showed a high positively charged face (surrounding RRM β_2 and α_3) and a slightly negative patch in the opposite side of the protein. This evidences a dipole, which is partly attenuated at higher ionic strength (Fig. 5B, *center* and *right*), although no substantial differences in His-apiRBP solubility have been observed upon salt addition (data not shown). The strong dependence on ionic strength of the binding indicated that the apiRBP-RNA complex is mainly driven by electrostatic interactions rather than π -stacking. This is consistent with the fact that aromatic residues placed at RNP2 are occluded by the extra α_3 -helix. However, we cannot exclude that binding to other RNA/DNA stretches would lead to higher protein-target affinities and/or sequence specificity. Finally, a negative control carried out with free GFP-His showed no interaction by ITC

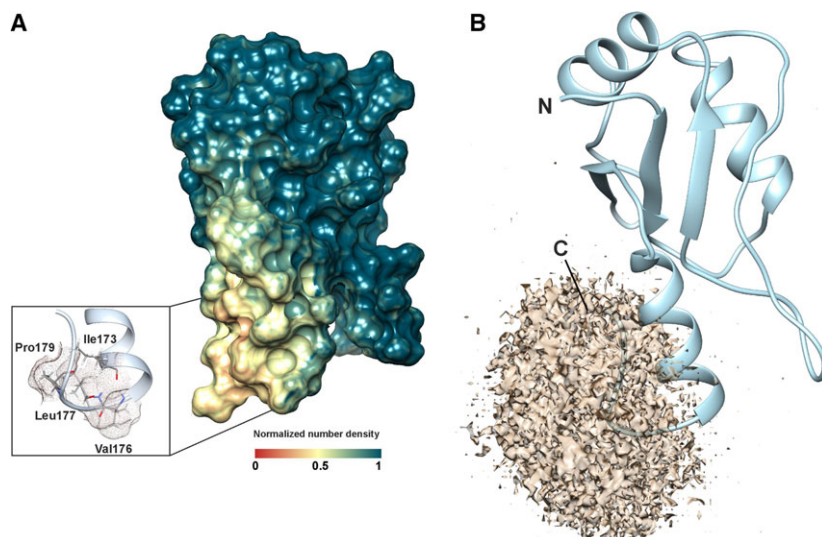


Fig. 4. apiRBP solvation analysis. Number density of water oxygen centers across simulations, normalized with respect to bulk values and calculated by GIST implemented in CPPTRAJ. (A) Map of normalized number density values at the surface of the molecular model. The inset shows the hydrophobic cluster at C-end α_3 . (B) Limit surface for the region showing normalized number densities lower than 0.35 voxel^{-1} .

Table 1. Equilibrium and thermodynamic parameters for the interaction of apiRBP-GFP-His affinity with 10-mer U-rich RNA. Relative errors corresponding to the standard deviation of the fitting: K_D , 20%; ΔH and $-\Delta S$, 5%; ΔG , 2%.

K_D (μM)	3.36
n	1
ΔH ($\text{kcal}\cdot\text{mol}^{-1}$)	-2.04
ΔG ($\text{kcal}\cdot\text{mol}^{-1}$)	-7.5
$-\Delta S$ ($\text{kcal}\cdot\text{mol}^{-1}$)	-5.42

with 10-mer U-rich RNA, proving that apiRBP is leading RNA recognition. Additionally, the apiRBP-GFP-His dilution experiment in 20 mM Tris (pH 7.4) produced no substantial heats (Fig. S3).

Discussion

The cyanobacterial origin of the apicoplast makes this organelle an excellent target for pharmaceutical research [39]. Here, we have characterized a putative

P. vivax RBP that is predicted to be targeted to the apicoplast [14,15]. The elevated protein conservation of its orthologous in the genus *Plasmodium* (72% of amino acid similarity) indicates that this protein may play a key role for the malaria parasite life cycle. Most of RBPs are composed of small RNA binding domains (RBDs) that are needed for their recruitment to specific RNA targets. Within RBDs, there are four prominent families: RNA recognition motifs (RRMs), zinc fingers, KH domains, and double-stranded RNA binding motifs [33].

The identification of secondary structure elements in apiRBP corresponded with a noncanonical RRM domain (four-stranded β -sheet packed against two α -helices), which is extended by a C-terminal α_3 -helix. This is consistent with the increased α -helix content detected by CD (14.8%) in comparison with other RRM domains [32]. Other noncanonical RRM domains that include additional secondary elements have been extensively reported in the literature. Such is the case of the extra

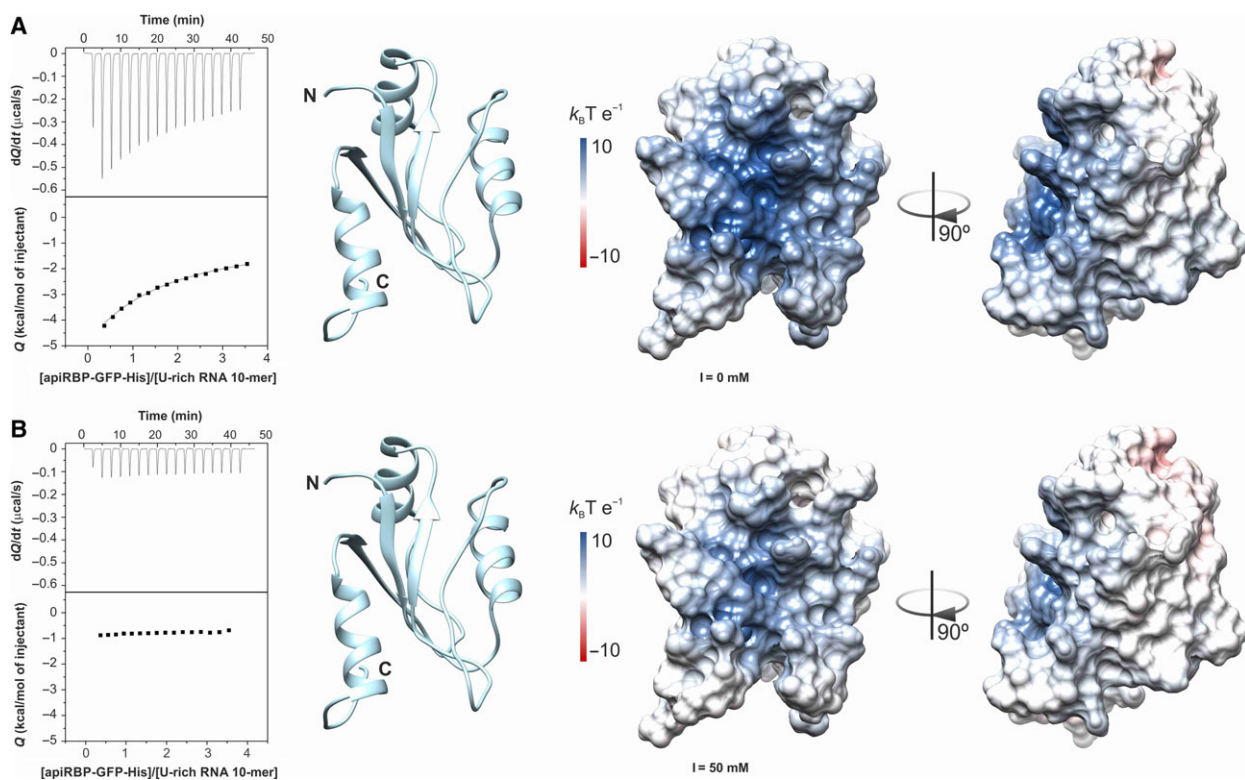


Fig. 5. apiRBP binding to target U-rich RNA oligonucleotides by isothermal titration calorimetry and electrostatic potential surface. (A) (Left) 10-mer U-rich RNA (1 μM , in the calorimetric cell) was titrated with apiRBP-GFP-His (10 μM , in the syringe) in 20 mM Tris (pH 7.4) buffer. (B) (Left) 1 μM 10-mer U-rich RNA was titrated with 10 μM apiRBP-GFP-His in 20 mM Tris (pH 7.4) buffer containing 50 mM NaCl. Each graph shows a thermogram of the heat released versus time in the upper panel and the integrated peak intensities plotted against the molar ratio of the complex in the lower panel. All experiments were conducted at 25 $^{\circ}\text{C}$. (A) and (B) (Center and right) Ribbon representation and electrostatic potential surface with a color ramp for positive (blue) and negative (red) potentials at low (A) and higher (B) ionic strength values. Each view is rotated 90 $^{\circ}$ around the vertical axis of the molecule. I = ionic strength.

N-terminal α -helix (α_0) in the third RRM of TIA-1 (RRM3) [40,41], the β -sheet extended by an additional fifth antiparallel β -strand in PTB RRM2 and RRM3 modules [42], or the RRM with an extra C-terminal helix (xRRM) in La and LARP7 proteins [43,44]. In addition, its RNP1 module is not fully conserved, whereas RNP2 is partly occluded by α_3 -helix.

Eukaryotic RRMs are often found as multiple copies within a protein, and together with other protein domains, they confer different affinity and specificity for the RNA sequences [45–53]. RRMs are also found in prokaryotes, where they tend to occur as single domains in small proteins, typically around 100 amino acids in length (Pfam ID PF0076 and InterPro ID IPR000504 [54,55]). Both the short length and the presence of a single RRM domain in apiRBP point to an endosymbiotic origin of the protein.

In most of the solved RRM–RNA complexes, two of the three aromatic residues in RNP1 make stacking interactions with the nucleobases, while the third can interact hydrophobically with the sugar rings. In addition, an Arg or a Lys residue at RNP1 forms a salt bridge with the phosphodiester backbone [33]. Particularly interesting in apiRBP is the lack of aromatics at the strand β_3 , suggesting that the binding between apiRBP and RNA might occur by electrostatic steering rather than by nucleobase stacking. This is in agreement with the mapping of the surface electrostatic potential showing a highly positively charged face surrounding β_2 and the C-terminal α_3 -helix; and also with the finding that a higher salt concentration in the buffer disrupted the binding of apiRBP to target U-rich RNAs by ITC. Finally, the prediction server BindN [31] localized potential RNA binding sites that were out of the limits of the strands β_1 and β_3 (particularly concentrated in β_2). Then, if RNA recognition was β_2 -mediated, this highly dynamic secondary element could be fixed upon apiRBP–RNA complex formation. Examples of RRMs using unusual RNA binding surfaces are reported in the literature. In such cryptic, atypical RRMs, aromatic residues have been displaced from the β_1 – β_3 motif to the β -connecting loops in qRRMs [56] or to the so-called RNP3 at β_2 in xRRMs [43,44,57]. In the particular case of apiRBP RRM, β_2 -strand could bind to RNA through electrostatic contacts. Additionally, we cannot exclude the participation of the C-terminal end of apiRBP in the binding to RNA as it is also highly positively charged. Both β_2 and α_3 may take part of a wide RNA binding platform. Altogether, these findings suggest an uncommon way of binding for an unusual RRM module and open a new door for further structural and functional characterization to better understand those mechanisms.

The fold of a protein and its stability toward denaturation are governed by the sequence of amino acids and the environment (solvent, salts, pH, temperature, crowding, etc.) [58]. Expression of a soluble recombinant apiRBP in *E. coli* was unsuccessful under the conditions tested in this study. The solvation analysis of apiRBP showed an exposed hydrophobic cluster at the C terminus. We then resolved to fuse a GFP-tag to the C-terminal end of the protein generating a more stable chimera that was useful for functional assays. Indeed, a model of the apiRBP–GFP chimera protein showed that GFP blocked the C-end hydrophobic patch in the RBP, probably preventing aggregation and interfering, somehow, with RNA binding (Fig. S4). A recent thermodynamic and molecular modeling study revealed that reaction in tight nucleic acid–RRM interactions is mostly enthalpy-driven [59]. According to their analysis, all these interactions display a strong enthalpy–entropy compensation effect (see Fig. 4A in [59]). Despite the fact of using a GFP-tagged chimera, plotting our values in Figure 4A of the manuscript would match their fitting at one extreme. Similar compensation effects have been observed in other systems [60,61]. Nevertheless, for a deeper structural approach, it would be of great interest to study the RRM module alone under native conditions.

Additionally, the exact residue where the TP finishes and the RRM domain starts in apiRBP still remains unknown. Interestingly, a study of the TP in *Toxoplasma gondii* showed that positive charges are more influential in the N-terminal portion of the TP, that arginine and lysine are equally suitable, and that the exact position of these charges is not important [11]. The accumulation of positive charges along the β_2 -strand and the extra C-terminal α_3 -helix of the RRM module is particularly interesting in apiRBP. Once proven that apiRBP is targeted to the apicoplast by cellular approaches, the possible role of those positive charges into the transit to the apicoplast organelle could be investigated. This would resemble the functioning of the single RRM from the trypanosome TcUBP1 RBP that behaves as a structural nuclear localization signal (NLS), alternating nuclear import and RNA binding [62]. Truncations and strategic point mutations of apiRBP would help to prove or reject this theory.

Historically, malarial drug design has been focused on compounds that modulate protein function such as doxycycline [63], but recently, RNA has become a new target for pharmaceutical companies [64]. Aminoglycoside antibiotics that target the RNA component of the small ribosomal subunits are being widely used for the treatment of bacterial infections. In addition, novel

approaches to drug discovery are identifying potential target sites in mRNA molecules, in particular the binding sites of proteins that regulate mRNA translation or stability [64]. The future molecular characterizations of proteins with a crucial role in expression of plastid genes, such as apiRBP, are a promising target for RNA-based antimalarial drugs.

Acknowledgements

We are very grateful to Prof. Dr. Miguel A. De la Rosa and Carlos A. Elena-Real for helpful discussions and Hilal Cura for technical assistance. We thank Dr. Katuska González-Arzola for GFP-His recombinant protein. SMGM, ICG, FRR, ADQ, and IDM were supported by Andalusian Government (P11-CVI-7216 and BIO-198) and the University of Seville (VI PPIT-US). FRR was also supported by Spanish Ministry of Education, Culture and Sports (FPU013/04373). CG was supported by the Pablo de Olavide Erasmus Program. This work used the Cell Free expression platform (Lionel Imbert) of the Grenoble Instruct Centre (ISBG; UMS 3518 CNRS-CEA-UJF-EMBL) with support from FRISBI (ANR-10-INSB-05-02) and GRAL (ANR-10-LABX-49-01) within the Grenoble Partnership for Structural Biology (PSB). We also thank the BIP-cicCartuja platform and AFFINIMETER with the support from Eva Muñoz to analyze the ITC thermograms. Finally, we would like to acknowledge Dr. Frank Bernhard (GE) for providing the pET21a(+) vector with the C-terminal GFP-His-tag (superfolderGFP) which was used for cloning of apiRBP and Dr. Laura Tomás-Gallardo for the FPLC assistance in the Proteomic service at the CABD, Sevilla.

Author contributions

IDM designed research; CG, ICG, FRR, MH, and SMGM performed research; FRR, SMGM, and ADQ analyzed the data; and SMGM, ADQ, and IDM wrote the manuscript.

References

- López C, Yepes-Pérez Y, Hincapié-Escobar N, Díaz-Arévalo D and Patarroyo MA (2017) What is known about the immune response induced by *Plasmodium vivax* malaria vaccine candidates? *Front Immunol* **8**, 126.
- Kalanon M and McFadden GI (2010) Malaria, *Plasmodium falciparum* and its apicoplast. *Biochem Soc Trans* **38**, 775–782.
- Zhu G, Marchewka MJ and Keithly JS (2000) *Cryptosporidium parvum* appears to lack a plastid genome. *Microbiology* **146** (Pt 2), 315–321.
- Lim L and McFadden GI (2010) The evolution, metabolism and functions of the apicoplast. *Phil Trans R Soc B* **365**, 749–763.
- Goodman CD and McFadden GI (2013) Targeting apicoplasts in malaria parasites. *Expert Opin Ther Targets* **17**, 167–177.
- Marechal E and Cesbron-Delauw MF (2001) The apicoplast: a new member of the plastid family. *Trends Plant Sci* **6**, 200–205.
- Nisbet RE and McKenzie JL (2016) Transcription of the apicoplast genome. *Mol Biochem Parasitol* **210**, 5–9.
- Kumar B, Chaubey S, Shah P, Tanveer A, Charan M, Siddiqi MI and Habib S (2011) Interaction between sulphur mobilisation proteins SufB and SufC: evidence for an iron–sulphur cluster biogenesis pathway in the apicoplast of *Plasmodium falciparum*. *Int. J Parasitol* **41**, 991–999.
- Wilson RJ and Williamson DH (1997) Extrachromosomal DNA in the Apicomplexa. *Microbiol Mol Biol Rev* **61**, 1–16.
- Habib S, Vaishya S and Gupta K (2016) Translation in organelles of apicomplexan parasites. *Trends Parasitol* **32**, 939–952.
- Tonkin CJ, Roos DS and McFadden GI (2006) N-terminal positively charged amino acids, but not their exact position, are important for apicoplast transit peptide fidelity in *Toxoplasma gondii*. *Mol Biochem Parasitol* **150**, 192–200.
- Foth BJ, Ralph SA, Tonkin CJ, Struck NS, Fraunholz M, Roos DS, Cowman AF and McFadden GI (2003) Dissecting apicoplast targeting in the malaria parasite *Plasmodium falciparum*. *Science* **299**, 705–708.
- Zuegge J, Ralph S, Schmuker M, McFadden GI and Schneider G (2001) Deciphering apicoplast targeting signals-feature extraction from nuclear-encoded precursors of *Plasmodium falciparum* apicoplast proteins. *Gene* **280**, 19–26.
- Carlton JM, Adams JH, Silva JC, Bidwell SL, Lorenzi H, Caler E, Crabtree J, Angiuoli SV, Merino EF, Amedeo P *et al.* (2008) Comparative genomics of the neglected human malaria parasite *Plasmodium vivax*. *Nature* **455**, 757–763.
- Neto Ade M, Alvarenga DA, Rezende AM, Resende SS, Ribeiro Rde S, Fontes CJ, Carvalho LH and de Brito CF (2012) Improving N-terminal protein annotation of *Plasmodium* species based on signal peptide prediction of orthologous proteins. *Malar J* **11**, 375.
- Reddy BPN, Shrestha S, Hart KJ, Liang X, Kemirembe K, Cui L and Lindner SE (2015) A

- bioinformatic survey of RNA-binding proteins in *Plasmodium*. *BMC Genom* **16**, 890.
- 17 Rudolph R and Lilie H (1996) *In vitro* folding of inclusion body proteins. *FASEB J* **10**, 49–56.
 - 18 Sreerama N and Woody RW (2000) Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal Biochem* **287**, 252–260.
 - 19 Das R and Baker D (2008) Macromolecular modeling with rosetta. *Annu Rev Biochem* **77**, 363–382.
 - 20 Kim DE, Chivian D, Malmstrom L and Baker D (2005) Automated prediction of domain boundaries in CASP6 targets using GinzU and RosettaDOM. *Proteins* **61** (Suppl 7), 193–200.
 - 21 Hornak V, Abel R, Okur A, Strockbine B, Roitberg A and Simmerling C (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **65**, 712–725.
 - 22 Izadi S, Anandakrishnan R and Onufriev AV (2014) Building water models: a different approach. *J Phys Chem Lett* **5**, 3863–3871.
 - 23 Andersen HC (1980) Molecular dynamics simulations at constant pressure and/or temperature. *J Chem Phys* **72**, 2384–2393.
 - 24 Ryckaert J-P, Ciccotti G and Berendsen HJC (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J Comput Phys* **23**, 327–341.
 - 25 Lazaridis T (1998) Inhomogeneous fluid approach to solvation thermodynamics. 1. Theory. *J Phys Chem B* **102**, 3531–3541.
 - 26 Roe DR and Cheatham TE III (2013) PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J Chem Theory Comput* **9**, 3084–3095.
 - 27 Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC and Ferrin TE (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605–1612.
 - 28 Bahl A, Brunk B, Crabtree J, Fraunholz MJ, Gajria B, Grant GR, Ginsburg H, Gupta D, Kissinger JC, Labo P *et al.* (2003) PlasmoDB: the *Plasmodium* genome resource. A database integrating experimental and computational data. *Nucleic Acids Res* **31**, 212–215.
 - 29 Bruce BD (2000) Chloroplast transit peptides: structure, function and evolution. *Trends Cell Biol* **10**, 440–447.
 - 30 Drozdetskiy A, Cole C, Procter J and Barton GJ (2015) JPred4: a protein secondary structure prediction server. *Nucleic Acids Res* **43**, W389–W394.
 - 31 Wang L and Brown SJ (2006) BindN: a web-based tool for efficient prediction of DNA and RNA binding sites in amino acid sequences. *Nucleic Acids Res* **34**, W243–W248.
 - 32 Scheiba RM, Aroca A and Díaz-Moreno I (2012) HuR thermal stability is dependent on domain binding and upon phosphorylation. *Eur Biophys J* **41**, 597–605.
 - 33 Clery A, Blatter M and Allain FH (2008) RNA recognition motifs: boring? Not quite. *Curr Opin Struct Biol* **18**, 290–298.
 - 34 Holm L and Rosenstrom P (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res* **38**, W545–W549.
 - 35 Drew DE, von Heijne G, Nordlund P and de Gier JW (2001) Green fluorescent protein as an indicator to monitor membrane protein overexpression in *Escherichia coli*. *FEBS Lett* **507**, 220–224.
 - 36 Lee JH and Skalnik DG (2008) Wdr82 is a C-terminal domain-binding protein that recruits the Setd1A Histone H3-Lys4 methyltransferase complex to transcription start sites of transcribed human genes. *Mol Cell Biol* **28**, 609–618.
 - 37 Li Y, Schulz VP, Deng C, Li G, Shen Y, Tusi BK, Ma G, Stees J, Qiu Y, Steiner LA *et al.* (2016) Setd1a and NURF mediate chromatin dynamics and gene regulation during erythroid lineage commitment and differentiation. *Nucleic Acids Res* **44**, 7173–7188.
 - 38 Tresaugues L, Dehé PM, Guérois R, Rodriguez-Gil A, Varlet I, Salah P, Pamblanco M, Luciano P, Quevillon-Cheruel S, Sollier J *et al.* (2006) Structural characterization of Set1 RNA recognition motifs and their role in histone H3 lysine 4 methylation. *J Mol Biol* **359**, 1170–1181.
 - 39 Ralph SA, van Dooren GG, Waller RF, Crawford MJ, Fraunholz MJ, Foth BJ, Tonkin CJ, Roos DS and McFadden GI (2004) Tropical infectious diseases: metabolic maps and functions of the *Plasmodium falciparum* apicoplast. *Nat Rev Microbiol* **2**, 203–216.
 - 40 Aroca A, Díaz-Quintana A and Díaz-Moreno I (2011) A structural insight into the C-terminal RNA recognition motifs of T-cell intracellular antigen-1 protein. *FEBS Lett* **585**, 2958–2964.
 - 41 Wang I, Hennig J, Jagtap PK, Sonntag M, Valcarcel J and Sattler M (2014) Structure, dynamics and RNA binding of the multi-domain splicing factor TIA-1. *Nucleic Acids Res* **42**, 5949–5966.
 - 42 Oberstrass FC, Auweter SD, Erat M, Hargous Y, Henning A, Wenter P, Reymond L, Amir-Ahmady B, Pitsch S, Black DL *et al.* (2005) Structure of PTB bound to RNA: specific binding and implications for splicing regulation. *Science* **309**, 2054–2057.
 - 43 Jacks A, Babon J, Kelly G, Manolaridis I, Cary PD, Curry S and Conte MR (2003) Structure of the C-terminal domain of human La protein reveals a novel RNA recognition motif coupled to a helical nuclear retention element. *Structure* **11**, 833–843.
 - 44 Singh M, Wang Z, Koo BK, Patel A, Cascio D, Collins K and Feigon J (2012) Structural basis for telomerase

- RNA recognition and RNP assembly by the holoenzyme La family protein p65. *Mol Cell* **47**, 16–26.
- 45 Waris S, García-Mauriño SM, Sivakumaran A, Beckham SA, Loughlin FE, Gorospe M, Díaz-Moreno I, Wilce MCJ and Wilce JA (2017) TIA-1 RRM23 binding and recognition of target oligonucleotides. *Nucleic Acids Res* **45**, 4944–4957.
- 46 Cruz-Gallardo I, Aroca Á, Gunzburg MJ, Sivakumaran A, Yoon JH, Angulo J, Persson C, Gorospe M, Karlsson BG, Wilce JA *et al.* (2014) The binding of TIA-1 to RNA C-rich sequences is driven by its C-terminal RRM domain. *RNA Biol* **11**, 766–776.
- 47 Cruz-Gallardo I, Aroca A, Persson C, Karlsson BG and Díaz-Moreno I (2013) RNA binding of T-cell intracellular antigen-1 (TIA-1) C-terminal RNA recognition motif is modified by pH conditions. *J Biol Chem* **288**, 25986–25994.
- 48 Cukier CD, Hollingworth D, Martin SR, Kelly G, Díaz-Moreno I and Ramos A (2010) Molecular basis of FIR-mediated c-myc transcriptional control. *Nat Struct Mol Biol* **17**, 1058–1064.
- 49 Díaz-Quintana A, García-Mauriño SM and Díaz-Moreno I (2015) Dimerization model of the C-terminal RNA recognition motif of HuR. *FEBS Lett* **589**, 1059–1066.
- 50 Maris C, Domínguez C and Allain FH (2005) The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J* **272**, 2118–2131.
- 51 Cruz-Gallardo I, Del Conte R, Velázquez-Campoy A, García-Mauriño SM and Díaz-Moreno I (2015) A non-Invasive NMR method based on histidine imidazoles to analyze the pH-modulation of protein-nucleic acid interfaces. *Chemistry* **21**, 7588–7595.
- 52 Scheiba RM, de Opakua AI, Diaz-Quintana A, Cruz-Gallardo I, Martinez-Cruz LA, Martinez-Chantar ML, Blanco FJ and Diaz-Moreno I (2014) The C-terminal RNA binding motif of HuR is a multi-functional domain leading to HuR oligomerization and binding to U-rich RNA targets. *RNA Biol* **11**, 1250–1261.
- 53 Díaz-Moreno I, Hollingworth D, Kelly G, Martin S, García-Mayoral M, Briata P, Gherzi R and Ramos A (2010) Orientation of the central domains of KSRP and its implications for the interaction with the RNA targets. *Nucleic Acids Res* **38**, 5193–5205.
- 54 Finn RD, Bateman A, Clements J, Coghill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J *et al.* (2014) Pfam: the protein families database. *Nucleic Acids Res* **42**, D222–D230.
- 55 Mitchell A, Chang HY, Daugherty L, Fraser M, Hunter S, Lopez R, McAnulla C, McMenamin C, Nuka G, Pesseat S *et al.* (2015) The InterPro protein families database: the classification resource after 15 years. *Nucleic Acids Res* **43**, D213–D221.
- 56 Domínguez C, Fisette JF, Chabot B and Allain FH (2010) Structural basis of G-tract recognition and encaging by hnRNP F quasi-RRMs. *Nat Struct Mol Biol* **17**, 853–861.
- 57 Alfano C, Sanfelice D, Babon J, Kelly G, Jacks A, Curry S and Conte MR (2004) Structural analysis of cooperative RNA binding by the La motif and central RRM domain of human La protein. *Nat Struct Mol Biol* **11**, 323–329.
- 58 Lindman S, Hernández-García A, Szczepankiewicz O, Frohm B and Linse S (2010) *In vivo* protein stabilization based on fragment complementation and a split GFP system. *Proc Natl Acad Sci USA* **107**, 19826–19831.
- 59 Samatanga B, Cléry A, Barraud P, Allain FHT and Jelesarov I (2017) Comparative analyses of the thermodynamic RNA binding signatures of different types of RNA recognition motifs. *Nucleic Acids Res* **45**, 6037–6050.
- 60 Díaz A, Hervás M, Navarro JA, De La Rosa MA and Tollín G (1994) A thermodynamic study by laser-flash photolysis of plastocyanin and cytochrome *c*₆ oxidation by photosystem I from the green alga *Monoraphidium braunii*. *Eur J Biochem* **222**, 1001–1007.
- 61 Medina M, Hervás M, Navarro JA, De la Rosa MA, Gómez-Moreno C and Tollín G (1992) A laser flash absorption spectroscopy study of Anabaena sp. PCC 7119 flavodoxin photoreduction by photosystem I particles from spinach. *FEBS Lett* **313**, 239–242.
- 62 Cassola A, Noe G and Frasch AC (2010) RNA recognition motifs involved in nuclear import of RNA-binding proteins. *RNA Biol* **7**, 339–344.
- 63 Dahl EL and Rosenthal PJ (2008) Apicoplast translation, transcription and genome replication: targets for antimalarial antibiotics. *Trends Parasitol* **24**, 279–284.
- 64 Zaman GJ, Michiels PJ and van Boeckel CA (2003) Targeting RNA: new opportunities to address drugless targets. *Drug Discov Today* **8**, 297–306.

Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:

Fig. S1. Solubilization tests of His-apiRBP synthesized in cell-free extracts.

Fig. S2. Protein purification, stability, and identification of apiRBP-GFP-His.

Fig. S3. Control experiments of the apiRBP-GFP-His binding to RNA by isothermal titration calorimetry.

Fig. S4. apiRBP-GFP structural model.