

# RRP6 from *Trypanosoma brucei*: Crystal Structure of the Catalytic Domain, Association with EAP3 and Activity towards Structured and Non-Structured RNA Substrates

Rosicler L. Barbosa, Pierre Legrand, Frank Wien, Blandine Pineau, Andrew Thompson, Beatriz G. Guimarães\*

Synchrotron SOLEIL, Gif-sur Yvette, France

## Abstract

RRP6 is a 3′–5′ exoribonuclease associated to the eukaryotic exosome, a multiprotein complex essential for various RNA processing and degradation pathways. In *Trypanosoma brucei*, RRP6 associates with the exosome in stoichiometric amounts and was localized in both cytoplasm and nucleus, in contrast to yeast Rrp6 which is exclusively nuclear. Here we report the biochemical and structural characterization of *T. brucei* RRP6 (*TbRRP6*) and its interaction with the so-called *T. brucei* Exosome Associated Protein 3 (*TbEAP3*), a potential orthologue of the yeast Rrp6 interacting protein, Rrp47. Recombinant *TbEAP3* is a thermo stable homodimer in solution, however it forms a heterodimeric complex with *TbRRP6* with 1:1 stoichiometry. The crystallographic structure of the *TbRRP6* catalytic core exposes for the first time the native catalytic site of this RNase and also reveals a disulfide bond linking two helices of the HRDC domain. RNA degradation assays show the distributive exoribonuclease activity of *TbRRP6* and novel findings regarding the structural range of its RNA substrates. *TbRRP6* was able to degrade single and double-stranded RNAs and also RNA substrates containing stem-loops including those with 3′ stem-loop lacking single-stranded extensions. Finally, association with *TbEAP3* did not significantly interfere with the *TbRRP6* catalytic activity *in vitro*.

**Citation:** Barbosa RL, Legrand P, Wien F, Pineau B, Thompson A, et al. (2014) RRP6 from *Trypanosoma brucei*: Crystal Structure of the Catalytic Domain, Association with EAP3 and Activity towards Structured and Non-Structured RNA Substrates. PLoS ONE 9(2): e89138. doi:10.1371/journal.pone.0089138

**Editor:** Bostjan Kobe, University of Queensland, Australia

**Received:** November 20, 2013; **Accepted:** January 20, 2014; **Published:** February 18, 2014

**Copyright:** © 2014 Barbosa et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by the Synchrotron Soleil, France. RLB received a fellowship from CNPq, Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: beatriz.guimaraes@synchrotron-soleil.fr

## Introduction

Kinetoplastids are flagellated protists many of them pathogenic to men and domestic animals. The most studied members of the class are the trypanosomes and leishmanias, which cause a number of serious diseases. For example, *Leishmania* species cause different types of infections regarded as cutaneous, mucocutaneous or visceral leishmaniasis in tropical countries around the world. *Trypanosoma brucei* is the causative agent of sleeping sickness in humans and nagana disease in cattle in sub-Saharan Africa while a related species, *T. cruzi*, causes Chagas disease in Latin America. These protozoa are unique in the sense that Kinetoplastidae genomes are deficient in regulatory transcription factors with control of gene expression, relying almost exclusively on post-transcriptional mechanisms (reviewed in [1–4]). Thus, a major mechanism available for kinetoplastids to regulate gene expression involves control of RNA processing and degradation rates.

Most of the enzymes involved in RNA metabolism in yeast and mammals have orthologues in trypanosomes and leishmanias, such as the exosome complex, although significant regulatory and biochemical differences are expected to be found given the central role of RNA stability control for gene expression in Trypanosomatids. The exosome is a 3′–5′ exoribonuclease complex that plays a central role in numerous pathways related to RNA processing and degradation, both in the nucleus and in the

cytoplasm. Initially described in yeast [5], exosomes are found in archaea and eukaryotes from protozoa to mammals. The eukaryotic exosome core is constituted by nine subunits (Exo-9), structurally organized in a ring of three heterodimers formed by RNase PH-related proteins. This hexameric ring is capped by three subunits with homology to KH and S1 RNA-binding domains [6,7]. In yeast and human cells the Exo-9 core is devoid of catalytic activity, the exosome ribonuclease activity is provided by the association of two nucleases, Rrp44 (also known as Dis3) and Rrp6 [6,8]. Rrp44, the exosome tenth subunit, interacts with Exo-9 in the nucleus and cytoplasm and is essential for cell viability [5,9]. It presents endonuclease and processive hydrolytic 3′–5′ exonuclease activities [6,10,11], which are modulated by the interaction with the Exo-9 core [12,13]. In yeast, crystallographic and biochemical studies have shown that the Exo-9 central channel directs the RNA substrate to degradation by Rrp44 which is located on the opposite side of the S1/HK subunits [7,12].

The eleventh exosome subunit, RRP6, is a member of the DEDD superfamily (DEDD-Y subgroup) of divalent metal dependent exonucleases [14,15]. In yeast, Rrp6 is found exclusively in the nuclear exosome, whereas in human, RRP6 is more concentrated in the nucleoli but is also found, at lower concentration, associated to the nucleoplasmic and the cytoplasmic exosome [16]. Rrp6 is not essential for cell viability but deletion of the gene in yeast causes slow growth and high

temperature sensitive phenotypes and accumulation of extended forms of 5.8S and snoRNAs [17,18]. Interaction of Rrp6 with the exosome stimulates both Rrp44 exo and endoribonuclease activities. On the other hand, the Rrp6 catalytic activity is inhibited by a mutation in the Rrp44 exoribonuclease active site of the Exo-11 complex [13]. The crystal structure of the yeast Exo-10 bound to an RNA substrate and the C-terminal region of Rrp6 evidenced that Rrp6 indirectly stabilizes the complex exosome-RNA without a direct contact with the RNA [7].

Rrp6 is composed of an N-terminal PMCN2NT domain which was shown to interact with a cofactor Rrp47 [19], the EXO domain containing the catalytic active site, an HRDC (helicase and RNaseD C-terminal) domain, and a C-terminal region responsible for the interaction with the exosome [7,18]. The crystal structures of the catalytic core of yeast and human RRP6 have shown the structural organization of the EXO and HRDC domains and revealed the conformation of the RRP6 active site residues in the presence of metals and/or nucleotides [20,21]. RNA degradation assays have shown that human RRP6 is more efficient to degrade structured RNA substrates in comparison with yeast Rrp6, and the accessibility to the active site was proposed to play a role in this substrate selectivity [21].

The Rrp6 cofactor in yeast, Rrp47, and its human orthologue C1D are nuclear exosome-associated proteins that bind both RNA and DNA molecules with an apparent specificity for double-stranded DNA and structured RNA substrates [19,22–24]. Rrp47/C1D proteins are conserved throughout eukaryotes and they are composed of a Sas10/C1D domain (*Pfam*: protein families data base domain PF04000) in the N-terminal region, and a more variable C-terminal. Genetic complementation assays identified the Sas10/C1D domain of Rrp47 as critical for yeast normal growth, sufficient for Rrp47 function *in vivo* and responsible for the interaction with the N-terminal domain of Rrp6. However, stable binding of Rrp47 to RNA *in vitro* requires both N-terminal and C-terminal regions. The C-terminal of Rrp47 was also shown to be involved in snoRNA maturation [24].

The characterization of the exosome complex in *Trypanosoma brucei* showed that its composition is similar to the yeast and human counterparts, being composed of six RNase PH-related proteins (*TbRRP41A*, *TbRRP41B*, *TbRRP45*, *TbEAP1*, *TbEAP2* and *TbEAP4*) and three subunits related to the S1 domain proteins (*TbRRP4*, *TbRRP40* and *TbCSL4*). The RRP6 subunit, which is specific to the nuclear exosome in yeast, is localized in both nucleus and cytoplasm in *T. brucei* [25–27]. In contrast, Rrp44-like protein was not detected in *T. brucei* purified exosome fractions [26]. An additional subunit, *TbEAP3* (for Exosome Associated Protein 3), which resembles yeast Rrp47, was identified and its interaction with *TbRRP6* was detected by two-hybrid analysis [26]. The structural characterization of the native exosome from *Leishmania tarentolae* by electron microscopy revealed a molecular envelope which is consistent with the Exo-9 structure, and evidenced an additional density at the top of the Exo-9 core that could account for the *TbRRP6* and *TbEAP3* associated proteins [28]. Surprisingly, the purified *L. tarentolae* exosome, containing the RRP6 protein, showed no RNase D-like activity *in vitro* [28].

Despite the amount of data presently available on the exosome complex, structural and functional information about the subunit RRP6 and its cofactor Rrp47 are still lacking. We have been especially interested in the study of their orthologues in kinetoplastids, motivated by the differences in the composition and activity of the trypanosome exosome regarding the subunits associated to the Exo-9 core. To better investigate the catalytic activity of *T. brucei* RRP6 and its interaction with EAP3 we have produced recombinant constructs of both proteins and performed

structural studies and degradation assays against different RNA substrates.

## Materials and Methods

### Cloning and Site-direct Mutagenesis

The genes encoding *TbRRP6* (Tb927.4.1630) and *TbEAP3* (Tb927.7.5460) were synthesized by GeneArt Gene Synthesis (Life Technologies). The corresponding *TbRRP6* DNA sequences were subcloned into the pET28a vector (Novagen) to express *TbRRP6* $\Delta$ sig (residues 20–736), *TbRRP6* $\Delta$ C (residues 20 to 540) and *TbRRP6*CAT (residues 176 to 540) constructs in fusion with a C-terminal His-tag. *TbEAP3* coding sequence was subcloned into the pET21a vector (Novagen) to express the full-length *TbEAP3* protein and the truncated forms *TbEAP3* $\Delta$ C1 (residues 1 to 183) and *TbEAP3* $\Delta$ C2 (residues 1 to 144) (Figures 1A and 2A).

Site-direct mutagenesis was performed in *TbRRP6*CAT and *TbRRP6* $\Delta$ C constructs using the GeneArt Site-Directed Mutagenesis System (Life Technologies) in order to introduce two single-point mutations in the catalytic site, converting D271 to N and Y393 to A; and two single-point mutations in the HRDC domain, converting both C496 and C515 to S. Gene amplification and bacterial transformation followed the protocol described by the manufacturer and the mutated DNA sequences were screened by PCR. The presence of mutations was confirmed by nucleotide sequencing. Table S1 shows the list of the primers used to generate the mutations.

### Protein Expression and Purification

The same general protocol was used for expression and purification of all proteins/complexes studied in this work. Specific differences are described when appropriated. *Escherichia coli* cells C3022 (New England Biolabs) were transformed or co-transformed with the pET expression vectors. The cultures were grown at 37°C in LB media containing ampicillin 100  $\mu$ g/mL or kanamycin 50  $\mu$ g/mL, for the clones on pET21a and pET28a, respectively; whereas both antibiotics were used for co-expression of the proteins in the complexes *TbRRP6* $\Delta$ C-EAP3 $\Delta$ C1 and *TbRRP6* $\Delta$ C-EAP3 $\Delta$ C2. When the culture reached the OD<sub>600</sub> ~ 0.6, the temperature was reduced to 18°C and protein expression was induced with 0.3 mM of isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG), overnight. Cells were harvested by centrifugation at 4000 g for 20 minutes and the pellet was stored at –80°C until protein purification.

The frozen cells from 1 L culture were suspended and lysed in 30 mL of buffer A (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 15 mM Imidazole, 5% glycerol) containing protease inhibitor cocktail (Roche Applied Science) and 1 mg/mL of lysozyme. After incubation of 30 minutes at 4°C, benzonase (25 units/mL) was added to the suspension. Cell extracts were isolated by sonication and centrifugation at 40000 g for 30 minutes at 4°C. The extracts were loaded onto a 5 mL His-Trap FF column (GE Healthcare Life Sciences) equilibrated in buffer A. The proteins were eluted with a 10 column volumes linear gradient from 0 to 400 mM of imidazole in the same buffer. Fractions containing the target proteins were pooled, concentrated and loaded onto a Superdex 200 16/60 column (GE Healthcare Life Sciences) equilibrated with 50 mM Tris-HCl pH 8.0, 500 mM NaCl, 5% glycerol. During the size exclusion chromatography step, both *TbEAP3* $\Delta$ C1 and *TbEAP3* $\Delta$ C2 proteins were eluted in two major well resolved peaks, one of them corresponding to aggregates. Only the non-aggregated fractions were used in the next experiments. Fractions containing the target proteins were pooled and dialyzed against 50 mM Tris-HCl pH 8.0, 10 mM NaCl, 5% glycerol for salt





**Table 1.** Crystallographic data and refinement statistics.

	<i>TbRRP6CAT</i>	<i>TbRRP6CAT-C496S</i>
<i>Data statistics</i>		
Source	SOLEIL-PX1	SOLEIL-PX1
Wavelength (Å)	0.9801	0.9801
Resolution (Å)	50–2.4 (2.54–2.40)	50–2.15 (2.28–2.15)
Space group	P2 <sub>1</sub>	P2 <sub>1</sub>
Unit cell (Å)	39.5, 92.7, 49.9/β = 104.8	39.4, 93.4, 49.9/β = 105.2
Number of observations	39831 (4851)	71910 (10744)
Number of unique reflections	13174 (1877)	19025 (3074)
Completeness	95.9 (85.9)	99.5 (98.5)
Redundancy	3.0 (2.6)	3.7 (3.5)
Mean <i>I</i> / <i>σ</i> ( <i>I</i> )	10.21 (2.8)	9.13 (1.6)
R <sub>meas</sub> (%)	10.3 (45.3)	13.5 (89.6)
<i>Refinement statistics</i>		
R <sub>work</sub> /R <sub>free</sub>	0.16/0.22	0.17/0.22
Bond RMSD length (Å)/angle (°)	0.01/1.1	0.01/1.1
Average B-protein/water (Å <sup>2</sup> )	36.1/38.9	38.1/44.0
Non-hydrogen atoms (except waters)	2866	2854
Number of water molecules	185	212
<i>Ramachandran plot (Molprobit)</i>		
Favored (%)	98.3	98.6
Outliers (%)	0	0
<i>PDB code</i>	4NLB	4NLC

Values in parenthesis correspond to the outer resolution shell.  
doi:10.1371/journal.pone.0089138.t001

TBE and 0.5% glycerol on ice. The gels were visualized using a Fuji LAS-3000 imager with the emission and detection wavelengths set at 460 nm and 510 nm, respectively, and exposure time of 20 seconds.

### Synchrotron Radiation Circular Dichroism (SRCD)

Synchrotron radiation circular dichroism (SRCD) experiments, covering the UV spectral range from 260–170 nm, were conducted at the DISCO beam line of the Synchrotron SOLEIL [29,30]. The use of SRCD allowed us to extend the spectral range and improve the signal to noise ratio obtained (as compared to conventional CD [31]) in the far UV, with peaks down to 190 nm clearly identifiable. The proteins *TbRRP6CAT* (200 μM), *TbEAP3ΔC1* (300 μM), *TbEAP3ΔC2* (500 μM) and the complexes *TbRRP6ΔC-EAP3ΔC1* (100 μM) and *TbRRP6ΔC-EAP3ΔC2* (100 μM) were dialyzed against 100 mM sodium phosphate buffer pH 8.0. Protein concentrations were calculated from their extinction coefficients [32] and the measured absorption at 280 nm (in triplets) using a nanodrop spectrophotometer (ThermoFisher). Temperature scans were performed between 18°C and 78°C in 3°C steps with 2 minutes settling time. All spectra were acquired using a CaF<sub>2</sub> 12 μm optical path cell [33]. For each measurement, the mean value of three spectra was calculated before subtraction of the baseline and zeroing between 263–270 nm. Spectra were calibrated to a standard solution of (+)-camphor-10-sulphonic acid (CSA), and normalized and converted to Δε using the software CDtool [34]. The secondary structure content was determined with the DichroWeb interface [35] using SELCON3 and CONTIN methods, with the reference set SP175

and a spectral cutoff at half of the total high tension variation of the photomultiplier [31]. The thermal denaturation curves were calculated at 222 nm, and the melting temperatures (*T<sub>m</sub>*) were determined from a sigmoidal fit. When required, the proteins' spectra were collected in the presence of 5 mM of MgCl<sub>2</sub> or 10 mM of DTT to evaluate their effect on the protein stability.

### Crystallization, Diffraction Data Collection and Processing

Extensive crystallization trials were performed using commercial screens with all proteins and complexes produced and purified to homogeneity. Hanging and/or sitting drop vapor-diffusion methods were tested at different protein concentrations. Promising hits were obtained only for the *TbRRP6CAT* constructs. Optimization of the crystallization conditions were performed by varying pH, precipitant and protein concentrations. Sea urchin-like crystals were obtained by hanging drop vapor diffusion at 18°C by mixing the *TbRRP6CAT* protein at 10 mg/mL in 50 mM Tris-HCl pH 8.0, 250 mM NaCl, 5% glycerol with crystallization buffer containing 0.1 M Tris-HCl pH 7.5 and 26% (v/v) PEG 3350. Before data collection at cryogenic temperature, the crystals were cryoprotected with 20% PEG 400 added to the mother liquor prior to flash-cooling in liquid nitrogen. Crystals of the mutant *TbRRP6CAT-C496S* (9 mg/mL in 50 mM Tris-HCl pH 8.0, 50 mM NaCl, 5% glycerol) were obtained by the sitting drop vapor diffusion method after mixing the protein and the well solution containing 25% PEG 2000. The crystals were cryoprotected by adding 25% PEG 400 to the mother liquor. X-ray diffraction data were collected from needle-shaped crystals of approximately 50 μm length and 5 μm width at the PROXIMA 1

beam line of the Synchrotron SOLEIL using a PILATUS 6 M detector. The diffraction data were processed with the XDS package [36]. *TbRRP6CAT* and *TbRRP6CAT-C496S* crystals belong to space group  $P2_1$  and diffracted to 2.40 Å and 2.15 Å resolution, respectively. Data collection statistics are reported in Table 1.

### Structure Resolution and Refinement

The structure of *TbRRP6CAT* was solved by molecular replacement with the program PHASER [37] using the atomic coordinates of the human RRP6 (PDB code 3SAF) [21] as the search model. Rigid body refinement using the model of the native *TbRRP6CAT* was applied in order to obtain the initial electron density map for the mutant *TbRRP6CAT-C496S*. Refinement of the structures were performed alternating cycles of BUSTER [38] with visual inspection and manual rebuilding using COOT [39]. For *TbRRP6CAT*, a total of 359 residues, out of the 365 expected excluding the His-tag fusion, were modeled for one monomer in the asymmetric unit. 185 water molecules were added during the refinement cycles and the  $R_{\text{factor}}/R_{\text{free}}$  values converged to 0.16/0.22. For the mutant *TbRRP6CAT-C496S*, 355 residues and 212 water molecules were modeled in the asymmetric unit and the  $R_{\text{factor}}/R_{\text{free}}$  values converged to 0.17/0.22. The stereochemistry of the models was analyzed with MolProbity [40] and no outliers were observed in the Ramachandran plot. Refinement statistics are summarized in the Table 1. Electrostatic potential were calculated using the Adaptive Poisson-Boltzmann Solver [41] through the PDB2PQR Server [42].

## Results

### Recombinant *TbEAP3* is a Thermo Stable Homodimer in Solution

EAP3 was identified as a component of the *T. brucei* exosome which interacts with the conserved subunit RRP6 [26]. Position-specific iterated BLAST search using the sequence of *TbEAP3* against a non-redundant data base which excludes Trypanosomatidae proteins results only in poor hits and does not detect the yeast orthologue Rrp47, a small nuclear protein known to interact with Rrp6 and involved in RNA maturation [24]. However, the sequence alignment of *TbEAP3* and yeast Rrp47 using the EMBOSS Needle program for pairwise alignment (available through the EMBL-EBI server; <http://www.ebi.ac.uk>) showed that *TbEAP3* primary structure shares 21% identity with Rrp47. Although the major sequence similarity is found in the N-terminal region, the bioinformatically predicted Sas10/C1D domain, present in Rrp47 and proposed to represent a group of nucleic acid binding proteins, is not detected when the *TbEAP3* sequence is analyzed for Pfam matches (protein families' database; <http://pfam.sanger.ac.uk>). Moreover, several residues of the Rrp47 lysine-rich region, which is required for RNA interaction [24] are not conserved in *TbEAP3* which also presents a C-terminal extension of 17 residues relative to Rrp47 (Figure 1A).

To gain insights into the *TbEAP3* function and its ability to interact with RRP6 we expressed and purified the recombinant *TbEAP3* in *E. coli* cells. The recombinant protein is unstable in solution. Degradation products were observed early after cell disruption and throughout the purification process. In order to identify the protease-sensitive regions of *TbEAP3*, limited proteolysis experiments were performed which evidenced two main stable fragments (Figure 1B, left). These fragments were identified by mass spectrometry (data not shown). Based on the mass spectrometry results, two new constructs were designed to express *TbEAP3ΔC1* and *TbEAP3ΔC2* C-terminal truncated proteins

(Figure 1A). Both variants were expressed in *E. coli* cells and purified to homogeneity (Figure 1B, right).

Previous size exclusion chromatography assays have described yeast Rrp47 as a hexamer in solution [19]. Based on analytical gel filtration chromatography, the calculated molecular mass of *TbEAP3* is compatible with the size of a trimer (data not shown). To confirm this result, we submitted both *TbEAP3ΔC1* and *TbEAP3ΔC2* variants to size exclusion chromatography combined with multi-angle light scattering (SEC-MALS) analysis. SEC-MALS results showed unique monodisperse peaks for *TbEAP3ΔC1* and *TbEAP3ΔC2* with molecular masses estimated in 41.0 kDa and 32.5 kDa, respectively, which correspond to homodimers in solution (Figure 1C). In addition, mass spectrometry analysis also revealed a dimer of *TbEAP3ΔC1* (measured mass of 42.5 kDa) under denaturing conditions and indicated the presence of an intermolecular disulfide bond (data not shown). The inaccuracy of the initial results obtained for *TbEAP3* by analytical gel filtration chromatography may be explained by the presence of large flexible/unfolded regions, such as those susceptible to limited proteolysis, that increase the hydrodynamic radius of the protein. The C-terminally truncated constructs, however, show the expected molecular mass for homodimers. These findings are in agreement with recent data reported for yeast Rrp47, showing by analytical ultracentrifugation, that it is also purified as a homodimer [43].

Secondary structure content and thermal stability of *TbEAP3ΔC1* and *TbEAP3ΔC2* were determined using SRCD. CD spectra were measured within a wavelength range of 170 to 260 nm and data analyses show a secondary structure content estimate of 36% and 41% of alpha-helix, 15% of beta-sheet and 49% and 44% of random structures for *TbEAP3ΔC1* and *TbEAP3ΔC2*, respectively (Table 2). The higher content of random structures found in *TbEAP3ΔC1* indicates that the C-terminal region of *TbEAP3* might be unfolded. The mutants were then submitted to thermal denaturation and the melting temperatures were estimated by analysis of the peaks at 222 nm, indicating a structural stability up to around 55°C for both constructs. SRCD results are summarized in Table 2 and the thermal denaturation spectra of the *TbEAP3ΔC1* variant is shown in the supporting information (Figure S1). To investigate the possible role of an intermolecular disulfide bond in the stabilization of the *TbEAP3* dimer, we performed SEC-MALS experiments and thermal denaturation under highly reducing conditions. In the presence of 10 mM DTT, CD analyses revealed no changes in the secondary structure content estimated at 20°C (Table 2), although the melting temperatures of both variants decreased by at least 10°C. Moreover, SEC-MALS analysis showed that *TbEAP3ΔC1* and *TbEAP3ΔC2* keep the dimeric structure in solution under reducing condition (data not shown). These results indicate that the rupture of the intermolecular disulfide bond affects *TbEAP3* thermal stability but the quaternary structure is maintained by non-covalent interactions. Interestingly, yeast Rrp47, which does not contain any cysteine residue in its sequence, was also described as a homodimer in solution [43].

Previous studies performed with yeast Rrp47 showed its ability to bind double-stranded DNA and structured RNA but not single-stranded or poly(A) substrates [19]. The human Rrp47 orthologue was shown to interact with tRNA and poly(G) but not poly(A), poly(C) or poly(U) [23]. We performed electrophoretic mobility shift assays to test the binding of the EAP3 variants *TbEAP3ΔC1* and *TbEAP3ΔC2* to single and double-stranded RNA and DNA substrates and to a stem-loop containing RNA (GNRA20). We were not able to detect protein-RNA or DNA interaction for any of these protein-substrate combinations, despite several attempts to

**Table 2.** Secondary structures content and melting temperatures of the target proteins based on SRCD measurements.

Sample	% $\alpha$ -helix	% $\beta$ -sheet	% turn	% others	RMSD	T <sub>m</sub> 222 nm (°C)
EAP3 $\Delta$ C1	36	15	17	32	0.04	55 $\pm$ 2
EAP3 $\Delta$ C1/DTT	36	16	16	32	0.07	44 $\pm$ 1
EAP3 $\Delta$ C2	41	15	14	30	0.06	55 $\pm$ 1
EAP3 $\Delta$ C2/DTT	39	18	13	30	0.02	42 $\pm$ 2
RRP6CAT	39	19	11	31	0.08	35.2 $\pm$ 0.2
RRP6 $\Delta$ C-EAP3 $\Delta$ C1	45	14	11	30	0.04	50.6 $\pm$ 0.3
RRP6 $\Delta$ C-EAP3 $\Delta$ C2	48	13	11	28	0.06	40.0 $\pm$ 0.6

RMSD is the root-mean-square deviation between the experimental and the calculated spectra. The melting temperatures (T<sub>m</sub>) were estimated based on thermal denaturation curves calculated at 222 nm. When indicated the samples were treated with 10 mM of DTT prior to the measurements.  
doi:10.1371/journal.pone.0089138.t002

optimize the experimental conditions. Regarding a possible conserved function, the experiments with Rrp47 truncated mutants, evidencing that the C-terminal lysine-rich region is essential for its RNA binding activity *in vitro* [24], are consistent with the lack of interaction of the *Tb*EAP3 $\Delta$ C1 and *Tb*EAP3 $\Delta$ C2 truncated proteins with RNA. Unfortunately, we were not able to confirm whether the C-terminal of *T. brucei* EAP3 mediates interactions with RNA since the intact protein could not be purified.

#### *Tb*RRP6 forms a Heterodimer with *Tb*EAP3

To study the *T. brucei* RRP6 activity and its interaction with *Tb*EAP3 we initially worked with three *Tb*RRP6 constructs. *Tb*RRP6 $\Delta$ sig (residues 20–736) is the largest construct and lacks the N-terminal residues predicted to act as a signal peptide, *Tb*RRP6 $\Delta$ C (residues 20–540) lacks the C-terminal region shown to be responsible for the interaction with the exosome complex in yeast [7,18] but keeps the N-terminal PMC2NT domain, and *Tb*RRP6CAT (residues 176–540) that retains the EXO and HRDC domains responsible for the exoribonuclease activity (Figure 2A). Attempts to express the single recombinant *Tb*RRP6 $\Delta$ sig and *Tb*RRP6 $\Delta$ C variants failed and the crystallization and catalytic assays were performed only with the *Tb*RRP6CAT variant. However, we were able to co-express both *Tb*RRP6 $\Delta$ sig and *Tb*RRP6 $\Delta$ C with the *Tb*EAP3 constructions. Recombinant production of the complex *Tb*RRP6 $\Delta$ sig-EAP3 resulted in very low yield of only partially purified products despite our efforts to optimize the expression and purification conditions. In contrast, *Tb*RRP6 $\Delta$ C was co-expressed with *Tb*EAP3, *Tb*EAP3 $\Delta$ C1 or *Tb*EAP3 $\Delta$ C2 and the complexes were successfully purified, giving an average yield of 6 mg of the complexes per liter of culture (Figure 2B). The co-purification of the *Tb*EAP3 constructs with *Tb*RRP6 $\Delta$ C was directed by the introduction of a His-tag fusion in the *Tb*RRP6 $\Delta$ C construct whereas the *Tb*EAP3 variants did not contain any tag. The complexes first isolated on a Ni-affinity chromatography were very stable during the subsequent analyses, indicating a strong interaction between *Tb*EAP3 and *Tb*RRP6 proteins *in vitro*. However, *Tb*EAP3 itself proved to be unstable even in complex with *Tb*RRP6 $\Delta$ C, since we observe a degradation band which corresponds to one of the fragments previously identified by limited proteolysis (Figure 2B).

Recent data using co-purification combined with electrophoretic profile quantification and western-blot analyses showed that yeast Rrp47 and the N-terminal region Rrp6 form a heterodimer *in vitro* [43]. To better characterize the interaction between *T. brucei* RRP6 and EAP3 the complexes *Tb*RRP6 $\Delta$ C-EAP3 $\Delta$ C1 and

*Tb*RRP6 $\Delta$ C-EAP3 $\Delta$ C2 were submitted to SEC-MALS analyses. The molecular masses were estimated in 75.2 kDa and 71.1 kDa for the complexes *Tb*RRP6 $\Delta$ C-EAP3 $\Delta$ C1 and *Tb*RRP6 $\Delta$ C-EAP3 $\Delta$ C2, respectively, which correspond to heterodimers with stoichiometry of 1:1 (Figure 2C). The absence of additional peaks corresponding to the individual proteins indicates the high stability of the complex. This result was confirmed by native mass spectrometry (data not shown) and is in accordance with the description of the complex in yeast. The oligomeric state of the *Tb*RRP6CAT variant was also analyzed by SEC-MALS. Its estimated molecular mass was 40.2 kDa which is consistent with the theoretical mass of a monomer (Figure 2C).

Secondary structure content and thermal stability of *Tb*RRP6CAT and the complexes *Tb*RRP6 $\Delta$ C-EAP3 $\Delta$ C1 and *Tb*RRP6 $\Delta$ C-EAP3 $\Delta$ C2 were investigated using SRCD. The melting temperatures, calculated at 222 nm, and the estimated secondary structure contents are listed in Table 2. The *Tb*RRP6CAT variant was revealed to be very sensitive to temperature increase, showing a calculated T<sub>m</sub> of approximately 35°C. The melting temperatures of the complexes *Tb*RRP6 $\Delta$ C-EAP3 $\Delta$ C1 and *Tb*RRP6 $\Delta$ C-EAP3 $\Delta$ C2 were 50°C and 40°C, respectively (Table 2). The difference observed between the complexes suggests that the C-terminal region of *Tb*EAP3 $\Delta$ C1, that is lacking in *Tb*RRP6 $\Delta$ C-EAP3 $\Delta$ C2, may have a role in the complex stability. Because of the importance of divalent metals for RRP6 activity, we tested the effect of magnesium on its folding and stability. No significant differences in melting temperatures or in secondary structure contents were observed for *Tb*RRP6CAT and for the *Tb*RRP6 $\Delta$ C-EAP3 $\Delta$ C1 complex upon addition of 5 mM of MgCl<sub>2</sub> (data not shown).

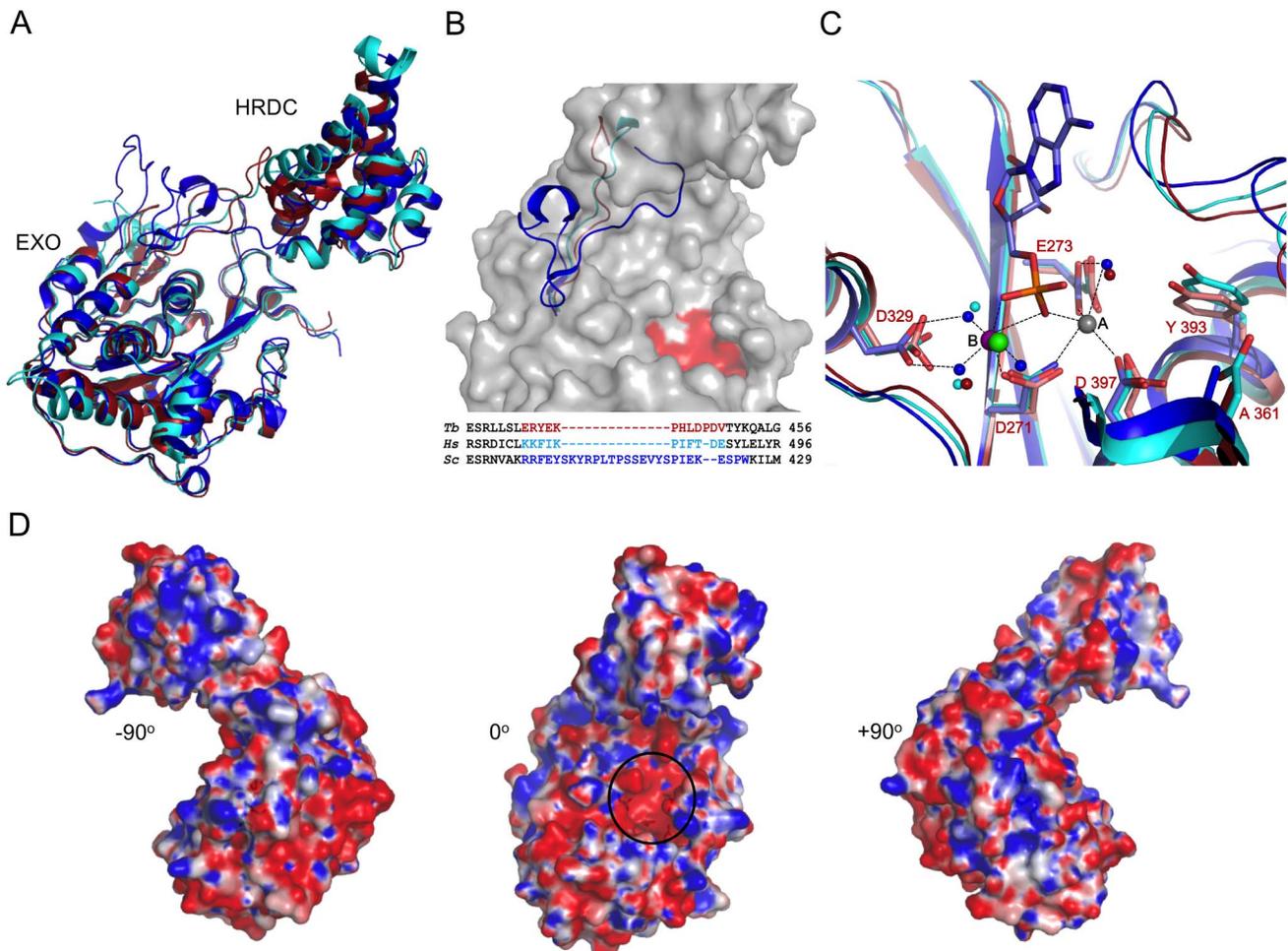
#### The Crystal Structure of the Catalytic Core of *Tb*RRP6

Despite extensive trials, we were not able to crystallize any of the *Tb*EAP3 variants or the complexes *Tb*RRP6 $\Delta$ C-EAP3 $\Delta$ C1 and *Tb*RRP6 $\Delta$ C-EAP3 $\Delta$ C2. The *Tb*RRP6CAT construct crystallized in a sea urchin-like form composed of very thin needles. The crystals showed to be very hard to reproduce and to optimize and a single data set could be collected from an eventual protuberant needle. *Tb*RRP6CAT crystal structure was refined at 2.4 Å resolution to final R<sub>factor</sub>/R<sub>free</sub> of 16%/22%, respectively (Table 1). The model covers residues 176 to 541 and includes 185 solvent molecules. The polypeptide chain was clearly defined by the electron density, except the residues 416 to 423 that could not be modeled. The 3D structure of the RRP6 catalytic domain was previously described for the yeast and human counterparts [20,21]. *Tb*RRP6CAT shares 41% and 40% of sequence identity

with the corresponding catalytic core of the yeast and human proteins respectively and, as expected, conserves their overall architecture. The EXO domain consists of a classical  $\alpha/\beta$  fold composed by a six-stranded  $\beta$ -sheet flanked by  $\alpha$ -helices and the HRDC domain is constituted of six  $\alpha$ -helices (Figure 3A). Superposition of the *T. brucei* RRP6 structure with the human and yeast orthologues results in RMSD of 1.38 Å for 343 C-alpha atoms aligned and 1.42 Å for 333 C-alpha aligned, respectively. As previously described [20,21] the EXO and HRDC domains are connected by a linker. Comparison of the structures of the yeast and human orthologues showed that the longer linker of yeast Rrp6 narrows the active site entrance which was proposed to affect the ability of the yeast enzyme to degrade structured RNA substrates [21]. The *T. brucei* RRP6 structure shows a closer match

to the human RRP6, both proteins presenting a shorter linker and a more accessible active site (Figure 3B).

*TbRRP6CAT* is the first RRP6 described up to now that was crystallized with the native catalytic site DEDD-Y comprised by the residues D271, E273, D329, D397 and Y393. In the *TbRRP6CAT* crystal structure, these residues exhibit a configuration similar to the mutated active sites of the human and the yeast structures, even in the absence of metal ions or nucleotides (Figure 3C). The apo *TbRRP6CAT* catalytic site shows only moderate movements of the side chains comparing with Mn/AMP bound Rrp6 structure, and a water molecule that is coordinated by Y393 and E273 residues could represent the hydrolytic water (Figure 3C). The aspartate D404 of the human RRP6, proposed to have a role in modulating the enzyme activity [21], is replaced by alanine in both yeast and *T. brucei*.



**Figure 3. Structure of *T. brucei* RRP6 catalytic core.** A) Overall structural comparison of the catalytic core of apo *T. brucei* RRP6 (red), *H. sapiens* RRP6 (cyan) (PDB code 3SAF) and *S. cerevisiae* Rrp6 (dark blue) (PDB code 2HBL). EXO and HRDC domains are indicated. B) Structural comparison of the linker region between the EXO and HRDC domains of *T. brucei*, *H. sapiens* and *S. cerevisiae* RRP6 proteins. The linkers are colored as in (A). The molecular surface of *TbRRP6* is shown in gray with the active site residues highlighted in red. A structure-based sequence alignment is shown at the bottom of the picture. C) DEDD-Y active site of apo *TbRRP6* (red) superposed to the Mg-bound *HsRRP6*-D313N mutant (cyan) and *ScRrp6*-Y361A mutant (dark blue) bound to one AMP, a zinc and a manganese ion. Water molecules are represented in the same color as the protein. Manganese and zinc (*ScRrp6* structure) are represented in purple (metal B) and gray (metal A) and magnesium (*HsRRP6* structure) is represented in green. Residues numbers correspond to the *T. brucei* structure. *ScRrp6* active site interactions are indicated by dotted lines. We observe that *TbRRP6* conserves a water molecule in the position of the hydrolytic water that interacts with Y393. Alanine residues (A361 in *TbRRP6*) replace the aspartate D404 of *HsRRP6* in *T. brucei* and yeast. D) Electrostatic surface of the *TbRRP6* catalytic core. The bounds for potential contour map visualization are  $\pm 5$  kT/e. The active site cavity is indicated with a black circle.  
doi:10.1371/journal.pone.0089138.g003

The electron density map revealed the presence of a disulfide bond between the residues C496 and C515, linking  $\alpha$ -helix 3 and the end of  $\alpha$ -helix 4 of the HRDC domain (Figure S2). C515 is conserved in kinetoplastids and in human RRP6 but it is not present in the yeast orthologue. In contrast, C496 is not conserved among the kinetoplastid homologues. In order to study the effect of the disulfide bond disruption on the protein structure and activity, site directed mutations were introduced in the cysteine residues generating the variants *TbRRP6CAT-C496S* and *TbRRP6CAT-C595S*. Crystallization trials were performed and crystals suitable for X-ray diffraction experiments were obtained for the mutant *TbRRP6CAT-C496S*. A complete data set was collected and the *TbRRP6CAT-C496S* structure was refined at 2.15 Å resolution to final  $R_{\text{factor}}/R_{\text{free}}$  of 17%/22% (Table 1). The model covers residues 179 to 540 and includes 212 waters and 2 PEG molecules. Similarly to *TbRRP6CAT* the residues 416 to 423 could not be modeled. Superposition of *TbRRP6CAT* and the mutant *TbRRP6CAT-C496S* structures resulted in an RMSD of 0.51 Å for 352 C-alpha atoms aligned. The structural changes resulted from the disulfide bond disruption are restricted to one of the helices involved in the SS bond. In the oxidized protein, the short turn that follows helix 4 moves away from helix 3 for the SS bond formation. An investigation of a putative role of the disulfide bond in the *TbRRP6* activity is presented in the next paragraph.

### ***TbRRP6* is more Efficient in the Presence of Manganese Ion and SS Bond Disruption does not Affect RNA Degradation Activity in vitro**

The exoribonucleolytic activity of *TbRRP6CAT* was initially tested in single-point RNA degradation assays. Reactions were performed at different pH and temperatures, using either magnesium or manganese as cofactor. A 30-mer single-stranded 5'-fluorescein labeled RNA (ssRNA, see material and methods) was used as substrate. As expected, no degradation activity was observed when the reaction mixtures do not contain any divalent metal. On the other hand, a significant increase in RNA degradation efficiency is observed when manganese ion is used instead of magnesium (Figure 4A). In addition, we observe that *TbRRP6CAT* conserves the catalytic activity in the range of temperature (20–37°C) and pH (6.5–8.0) tested (Figure 4A). In the presence of manganese, *TbRRP6CAT* degraded the RNA substrate completely under all the conditions assayed. However, in the presence of magnesium *TbRRP6CAT* is more efficient at 37°C and pH 8.0 (Figure 4A). To confirm the *TbRRP6CAT* preference for the manganese ion, time course assays were performed which indicated that *TbRRP6CAT* is at least five times more efficient in the presence of manganese as compared to the same reaction in the presence of magnesium (Figure 4B).

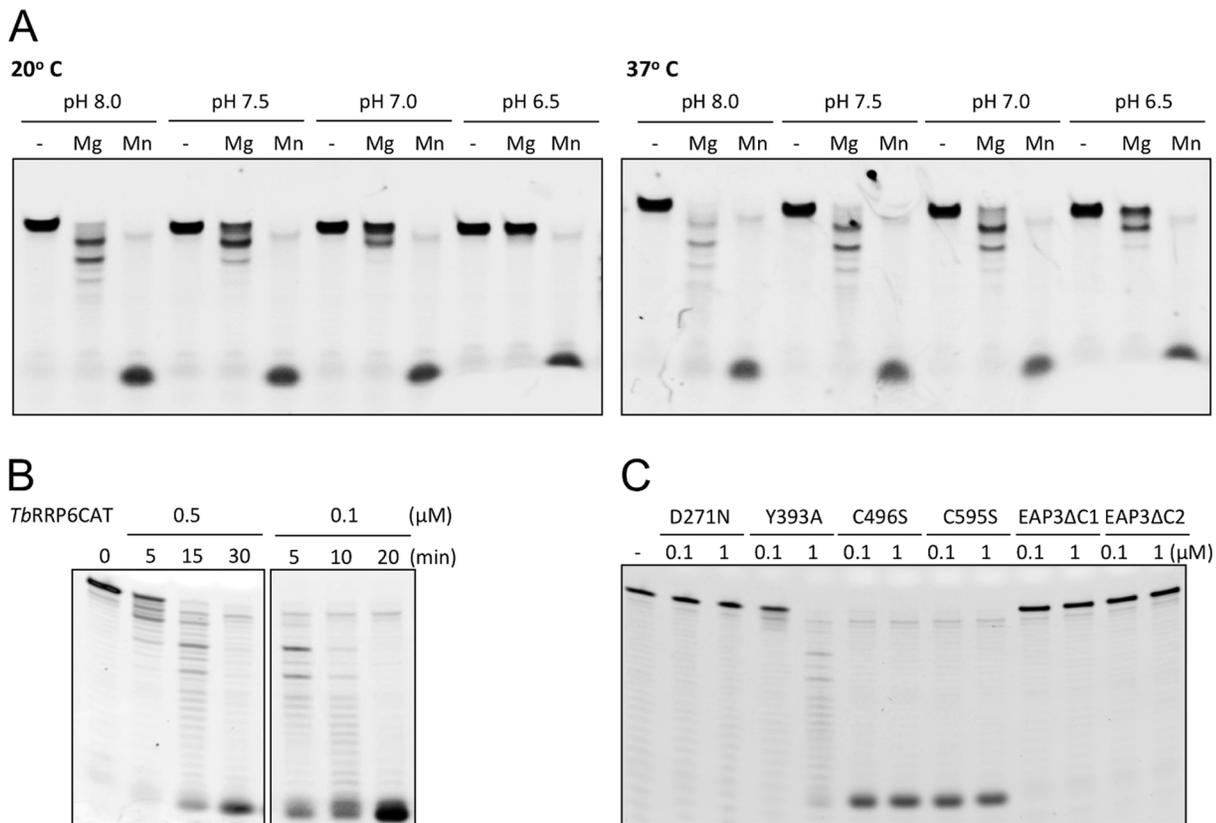
The exoribonucleolytic activity of the point mutants *TbRRP6CAT-C496S*, *TbRRP6CAT-C595S*, *TbRRP6CAT-D271N* and *TbRRP6CAT-Y393A* were also assayed. As previously observed for yeast and human RRP6 proteins [21], the active site mutation D271N abolishes activity, while the mutant Y393A retains activity although the degradation efficiency is highly compromised. On the other hand, the mutants C496S and C595S showed activity comparable with the native protein, indicating that the disruption of the SS bond in the HRDC domain does not affect *TbRRP6* degradation of non-structured substrates *in vitro* (Figure 4C). The variants *TbEAP3ΔC1* and *TbEAP3ΔC2*, which are not expected to present ribonucleolytic activity, were also assayed as negative controls, to verify that our purification protocol is efficient to eliminate any RNase activity from residual bacterial contaminants.

### ***TbRRP6* is able to Degrade Double-stranded and Structured RNA Substrates without a 3' Overhang**

Time-course degradation assays were performed with *TbRRP6CAT* variant and complexes *TbRRP6ΔC-EAP3ΔC1* and *TbRRP6ΔC-EAP3ΔC2* using different synthetic RNA substrates. Initially, protein activity was tested against a 30-mer single-stranded RNA, and the detection of decreasing size intermediates prior to accumulation of the final product indicates the distributive exoribonucleolytic activity of *TbRRP6* (Figure 5, left), similarly to results previously described for yeast and human orthologues [21]. Moreover, our results show that upon association with EAP3, although the activity slows as evidenced by the degradation pattern after 1 and 3 minutes, the complex is still able to degrade the non-structured RNA substrate completely. More surprising is the observation that *TbRRP6CAT* and the *TbRRP6ΔC-EAP3ΔC1* and *TbRRP6ΔC-EAP3ΔC2* complexes also degrade the double-stranded RNA substrate efficiently (Figure 5, right). The mutants *TbRRP6CAT-C496S* and *TbRRP6CAT-C595S*, lacking the SS bond, were also assayed against double-stranded RNA and showed activity comparable with the constructs with native cysteine residues (Figure S3). Previous studies reported that the yeast and human ribonucleases RRP44 and RRP6 require a 3' single-stranded extension to start the substrate degradation [6,10,21]. By contrast, our results show that *T. brucei* RRP6 is able to degrade double-stranded RNA without any 3' overhang. To further evaluate the activity of *T. brucei* RRP6 against structured RNA substrates and to compare with previous results obtained for yeast and human enzymes [21], a set of synthetic RNAs were designed containing a GNRA stem-loop in different positions of an AU-rich chain. The substrates were named GNRA0, GNRA5, GNRA20, GNRA24 and GNRA29, where the ending number indicates the number of nucleotides of the 3' single strand. Time-course degradation assays were performed with *TbRRP6CAT* and the *TbRRP6ΔC-EAP3ΔC1* and *TbRRP6ΔC-EAP3ΔC2* complexes. Again, no significant difference was observed between the activity of *TbRRP6CAT* and the complexes (Figure 6). Intermediates resistant to degradation are present when the secondary structure is positioned close to the 5' end (GNRA24 and GNRA29), indicating that a 5' overhang is needed for efficient activity, as previously described for the yeast and human orthologues [21]. It has been suggested that the HRDC domain would interact with the 5' single strand, stabilizing the binding to the substrate long enough to allow degradation [21]. In contrast, our results show that *TbRRP6* is able to degrade 3' double-stranded RNA and RNA substrates containing stem-loops at the 3'-end without any overhang (GNRA0) (Figure 6). To our knowledge this is the first time that such an activity is reported for an RRP6 orthologue.

### **Discussion**

*T. brucei* RRP6 was previously characterized as an essential structural subunit of the exosome complex found both in the nucleus and in the cytoplasm [26,27], distinguishing the trypanosome exosome from those of humans and yeast. In this work, we aimed to obtain functional information on *T. brucei* RRP6 and data on its substrate preferences and regulation by the putative interacting partner *TbEAP3*. We succeeded in determining the crystal structure of the catalytic domain of *TbRRP6* containing the native catalytic site residues, which is a novel result as compared to the yeast and human RRP6 orthologues whose structures were determined using inactive catalytic site mutants. Another important novel finding was the fact that *TbRRP6* is able to degrade both 3' double-stranded RNA substrates and 3'-end structured



**Figure 4. Exoribonucleolytic activity of *TbRRP6* under different biochemical and temperature conditions.** All assays were performed with 0.1  $\mu\text{M}$  of a 30-mer single-stranded RNA substrate (ssRNA, see materials and methods). A) Single-point RNA degradation assay using *TbRRP6CAT* at 0.5  $\mu\text{M}$  and incubation of 40 minutes. The reactions were performed in absence (–) and in presence of manganese (Mn) or magnesium (Mg) salts and in different pH and temperatures, as indicated at the top of the gels. B) Time course assay in the presence of magnesium (left) or manganese (right) ions. Enzyme concentration and incubation times are indicated at the top of the gel. We observe faster RNA degradation in the presence of manganese even at lower protein concentration. C) Exoribonucleolytic activity tests of the mutants *TbRRP6CAT*-D271N, *TbRRP6CAT*-Y393A, *TbRRP6CAT*-C496S, *TbRRP6CAT*-C595S. Assays were conducted at two protein concentrations, as indicated at the top of the gel. The first lane corresponds to the reaction mixture without protein (–) and *TbEAP3* mutants which were not expected to present ribonucleolytic activity were also used as negative controls.

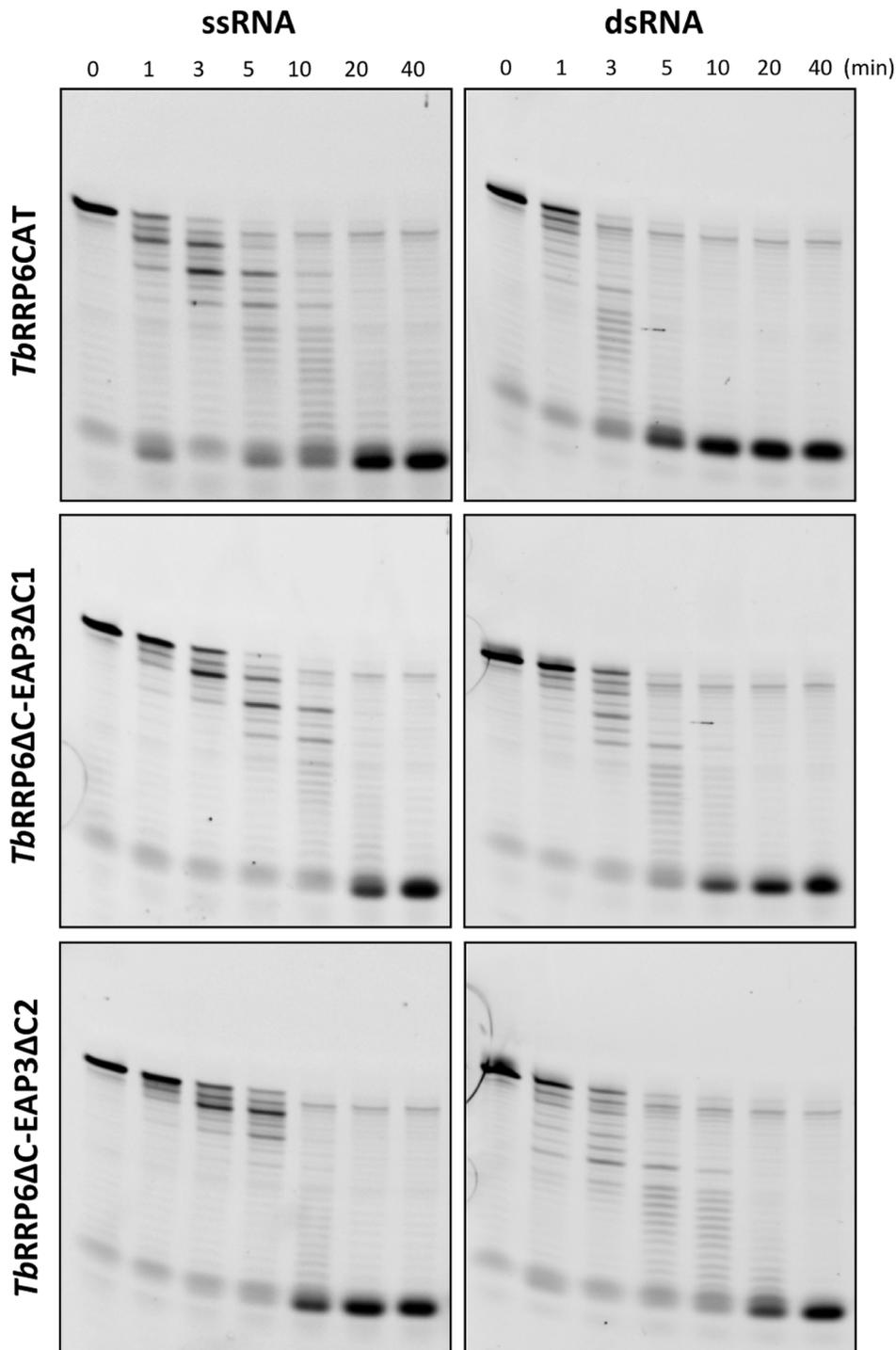
doi:10.1371/journal.pone.0089138.g004

RNA. Furthermore, we reconstituted several *TbRRP6*-*TbEAP3* complex variants and performed RNA degradation assays showing that there is no detectable effect of *TbEAP3* interaction with *TbRRP6* on RNA degradation *in vitro* under the conditions tested in this work. The heterodimer formed by *TbEAP3* and *TbRRP6* *in vitro* is consistent with the interaction detected between these proteins by two-hybrid analysis [26] and suggests that *TbEAP3* belongs to the functional Rrp47/C1D protein family despite their primary structure divergence. We have also shown that recombinant *TbEAP3* is a homodimer in solution. The biological relevance of the oligomer of *TbEAP3* and whether the *TbEAP3* homodimer is dissociated to allow the formation of the heterodimer with *TbRRP6* *in vivo* remain to be determined. Nevertheless, the biochemical behavior of the recombinant *TbEAP3* could be an indication of its role as a platform for protein interaction and binding of nucleic acid substrates given its propensity to form aggregates and its instability in solution.

It was previously reported that the native exosome purified from *Leishmania tarentolae* lacked the hydrolytic RNase-D like activity usually attributed to the RRP6 subunit [28]. The explanation proposed for this finding suggested that kinetoplastid RRP6 could be inactive, the substrates tested were inappropriate or that the association of RRP6 with the exosome or with *EAP3* could

regulate RRP6 activity [28]. More recently, accumulation of mRNA degradation intermediates containing 5' ends was described for *T. brucei* with impaired exosome activity, suggesting that a distributive 3'–5' exoribonuclease activity should be present in trypanosome [44]. In this work, we show that *TbRRP6* does have distributive 3'–5' exoribonuclease activity *in vitro*. Moreover, association with *EAP3* did not significantly affect degradation activity of *TbRRP6* $\Delta\text{C}$  on the substrates tested (Figures 5 and 6). Studies in yeast cells have indicated that Rrp47 functions together with Rrp6 for selection of some RNA substrates to be degraded [22]. Thus, although we have not detected a significant effect of *TbEAP3* association on the *TbRRP6* activity towards the synthetic substrates used in this work we do not exclude that a regulation might occur *in vivo*. We also hypothesize that the C-terminal truncations of the *TbEAP3* variants could have affected the RNA binding properties of *EAP3* and its ability to modulate *TbRRP6* activity.

The crystal structure of the *TbRRP6* EXO-HRDC catalytic core revealed the native catalytic site DEDD-Y which retains the conformation of the ion/nucleotide bound active sites described for the RRP6 mutants from human and yeast orthologues. We have also shown that *TbRRP6* is more efficient in degrading RNA substrates *in vitro* in the presence of manganese instead of

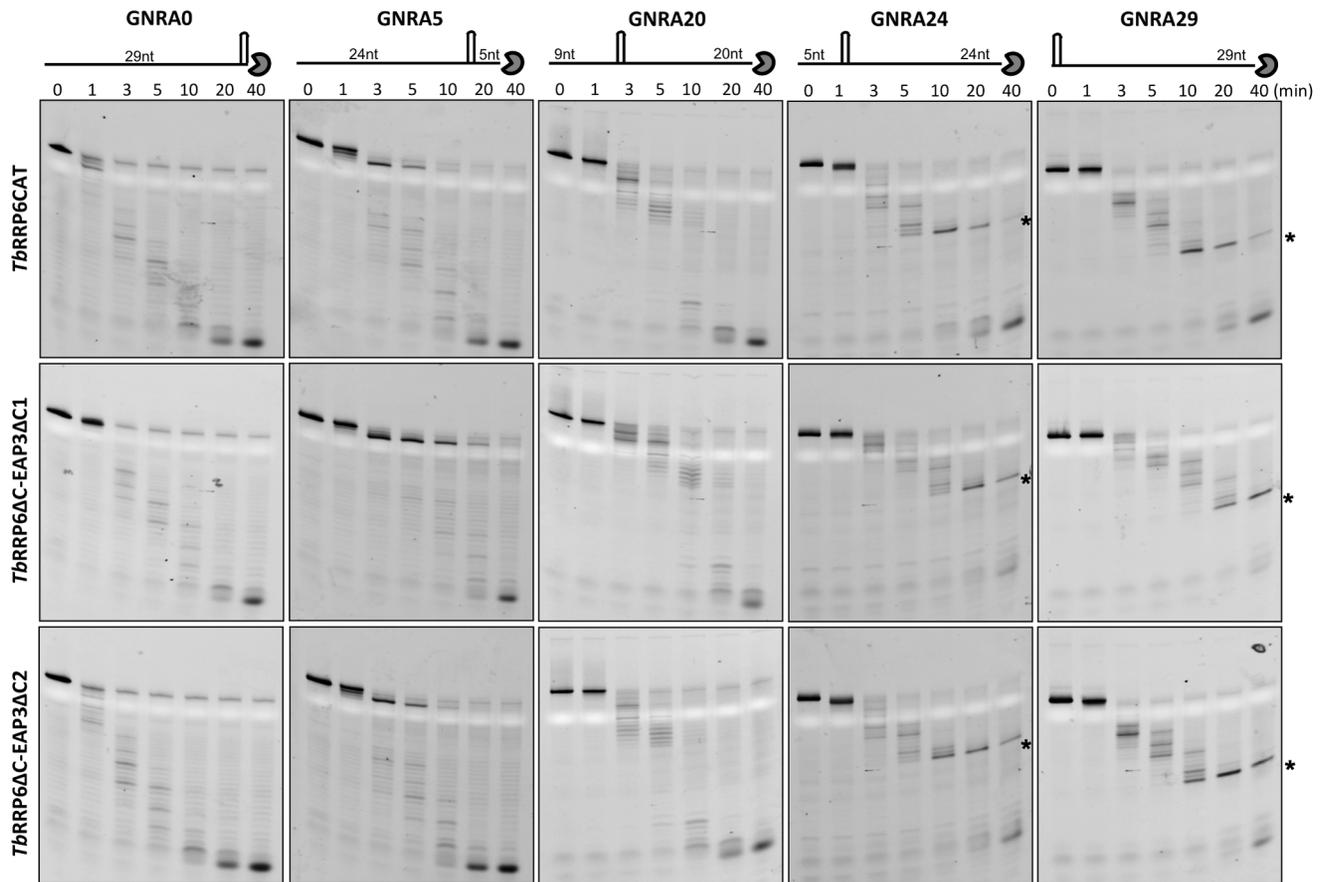


**Figure 5. Exoribonucleolytic activity on single-stranded (ssRNA) and double-stranded (dsRNA) RNA substrates.** Time-course degradation assays were performed with substrates and protein concentration of 0.1  $\mu$ M. Substrates and reaction time points are indicated at the top of the gels and the proteins/complexes are identified on the left.  
doi:10.1371/journal.pone.0089138.g005

magnesium, consistent with the results of Midtgaard and co-workers which suggest that manganese and zinc are specifically required in the active site of Rrp6 [20]. However, the rationale behind the divalent metal ion preferences among the DEDD nucleases is unclear. Structural comparison of apo and holo active sites of *T. brucei*, yeast and human RRP6 and *E. coli* RNase D,

which is more efficient in the presence of magnesium [45], did not allow us to propose a mechanistic basis for the differing effectiveness of the metal ions in catalysis, since the structure of the active sites is highly conserved (not shown).

To our knowledge, the ability of *TbRRP6* to degrade double-stranded and 3'-end structured RNA substrates is a new finding



**Figure 6. Exoribonucleolytic activity on structured RNA substrates.** Time-course degradation assays were performed with protein concentration of 0.1  $\mu$ M and 0.1  $\mu$ M of the AU-rich substrate containing a GNRA stem-loop at different positions of the strand (see material and methods). The substrates are named accordingly to the size of the 3' single-stranded extension, as schematically indicated at the top of the gels. The reaction time points are shown in minutes. The asterisks indicate the most stable intermediates observed during degradation.  
doi:10.1371/journal.pone.0089138.g006

for this class of DEDD 3'-5' exonucleases. Degradation of structured RNA/DNA requires an initial step of substrate unwinding before cleavage can occur. Some nucleases are able to degrade structured substrates in the absence of other factors, while others are dependent on association with helicases for this function. RRP6 contains a HRDC (helicase and RNaseD C-terminal) domain but the function of this domain is not completely elucidated. Previous studies revealed that removal of the HRDC domain in yeast Rrp6 disrupts the processing of certain RNA substrates *in vivo* [18]. The recombinant human RRP6 protein lacking the HRDC domain exhibited 100-fold reduction in exoribonuclease activity [21], and disruption of the exonuclease-HRDC domain contact was proposed to prevent proper orientation of certain substrates in the active site [20]. Also, it was previously reported that non-catalytic basic amino acids cooperatively contribute to destabilize double-stranded RNA to degradation by single-strand-preferring ribonucleases [46]. The electrostatic surface of *Tb*RRP6 (Figure 3D) evidences positively charged regions surrounding the active site and a basic path at one side of the HRDC domain. Moreover, analysis of the *Tb*RRP6CAT structure shows basic residues located on the periphery of the catalytic site, most of them conserved in yeast and human orthologues (Figure S4). Further assays are necessary to evaluate the role of the HRDC domain and of individual basic residues in the ability of *Tb*RRP6 to degrade 3'-end structured

RNA, but we speculate that the HRDC domain may function to correctly orientate the substrates, and active site peripheral basic residues may contribute to the destabilization of the double-strand and preparation of the 3'-end for cleavage.

Finally, the differences described for the *T. brucei* exosome compared with the human and yeast complexes, notably, the association of *Tb*RRP6 to both nuclear and cytoplasmic exosomes in stoichiometric amounts and the absence of a Rrp44-like protein in *T. brucei* purified exosome fractions [26,27], could indicate that RRP6 may have a more extensive role in the RNA processing and degradation pathways in trypanosomes, which could be related to its ability to degrade double-stranded and 3'-end structured RNA substrates.

## Supporting Information

**Figure S1 SRCD spectra showing the thermal denaturation of the *Tb*EAP3ΔC1 variant.** Temperature scans (shown in rainbow colors) were performed between 18°C and 78°C, using 3°C steps with 2 minutes settling time.  
(TIF)

**Figure S2 Exoribonucleolytic activity of the mutants *Tb*RRP6CAT-C496S and *Tb*RRP6CAT-C595S on single-stranded (ssRNA) and double-stranded RNA (dsRNA).** Time-course degradation assays were performed with substrates

and protein concentration of 0.1  $\mu\text{M}$ . Substrates and reaction time points are indicated at the top of the gels and the proteins are identified on the left. Comparing with *Tb*RRP6CAT protein (Figure 5) no significant difference in activity was detected. (TIF)

**Figure S3 Structural comparison of the HRDC domains of *Tb*RRP6CAT (red) and *Tb*RRP6CAT-C496S mutant (yellow).** The disulfide bond is coloured in green. On the right, 2Fo-Fc electron density maps of native (top) and mutant (bottom) proteins are shown in gray and contoured at 1.2  $\sigma$ . (TIF)

**Figure S4 Basic residues surrounding *T. brucei* RRP6 active site.** *Tb*RRP6CAT surface is represented in gray with arginine and lysine residues highlighted in blue. The active site cavity is coloured in red. *Tb*RRP6 residues K374, K335 and R439 are conserved in human (K417, K377 and K480, respectively) and yeast (K342, R302 and R400, respectively) orthologues; K374 is conserved in yeast Rrp6 (K343). The residues R327 and R331 are not conserved in the human and yeast sequences but the side chains of K479 and R399 (human and yeast, respectively) are

orientated in such a way as to occupy similar positions to the *Tb*RRP6CAT side chains. (TIF)

**Table S1 Primers used for site-directed mutagenesis.** Mutated bases are represented in red. (DOCX)

## Acknowledgments

This work benefited from the facilities and expertise of the IMAGIF Structural and Proteomic Biology Unit of the Centre de Recherche de Gif (<https://www.imagif.cnrs.fr>). We particularly thank Christophe Velours from the Biophysics facility and David Cornu from the SICaPS facility. We gratefully acknowledge Nilson Zanchin for critical reading of the manuscript.

## Author Contributions

Conceived and designed the experiments: RLB BGG. Performed the experiments: RLB BGG FW BP. Analyzed the data: RLB BGG FW PL AT. Contributed reagents/materials/analysis tools: PL BP FW AT BGG. Wrote the paper: RLB BGG.

## References

- Clayton CE (2002) Life without transcriptional control? From fly to human and back again. *EMBO J* 21: 1881–1888.
- Palencharm JB, Bellofatto V (2006) Gene transcription in trypanosomes. *Mol Biochem Parasitol* 146: 135–141.
- Clayton C, Shapira M (2007) Post-transcriptional regulation of gene expression in trypanosomes and leishmanias. *Mol Biochem Parasitol* 156: 93–101.
- Fernández-Moya SM, Estévez AM (2010) Posttranscriptional control and the role of RNA-binding proteins in gene regulation in trypanosomatid protozoan parasites. *Wiley Interdiscip Rev RNA* 1: 34–46.
- Mitchell P, Petfalski E, Shevchenko A, Mann M, Tollervey D (1997) The exosome: a conserved eukaryotic RNA processing complex containing multiple 3'–5' exoribonucleases. *Cell* 91: 457–466.
- Liu Q, Greimann JC, Lima CD (2006) Reconstitution, activities, and structure of eukaryotic RNA exosome. *Cell* 127: 1223–1237.
- Makino DL, Baumgärtner M, Conti E (2013) Crystal structure of an RNA-bound 11-subunit eukaryotic exosome complex. *Nature* 495: 70–75.
- Dziembowski A, Lorentzen E, Conti E, Séraphin B (2007) A single subunit, Dis3, is essentially responsible for yeast exosome core activity. *Nat Struct Mol Biol* 14: 15–22.
- Allmang C, Petfalski E, Podtelejnikov A, Mann M, Tollervey D, et al. (1999) The yeast exosome and human PM-S1 are related complexes of 3'–5' exonucleases. *Genes Dev* 13: 2148–2158.
- Lorentzen E, Basquin J, Tomecki R, Dziembowski A, Conti E (2008) Structure of the active subunit of the yeast exosome core, Rrp44: diverse modes of substrate recruitment in the RNase II nuclease family. *Mol Cell* 29: 717–728.
- Schneider C, Leung E, Brown J, Tollervey D (2009) The N-terminal PIN domain of the exosome subunit Rrp44 harbors endonuclease activity and tethers Rrp44 to the yeast core exosome. *Nucleic Acids Res* 37: 1127–1140.
- Bonneau F, Basquin J, Ebert J, Lorentzen E, Conti E (2009) The yeast exosome functions as a macromolecular cage to channel RNA substrates for degradation. *Cell* 139: 547–559.
- Wasmuth EV, Lima CD (2012) Exo- and endoribonucleolytic activities of yeast cytoplasmic and nuclear RNA exosomes are dependent on the noncatalytic core and central channel. *Mol Cell* 48: 133–144.
- Steitz TA, Steitz JA (1993) A general two-metal-ion mechanism for catalytic RNA. *Proc Natl Acad Sci USA* 90: 6498–6502.
- Zuo Y, Deutscher MP (2001) Exoribonuclease superfamilies: structural analysis and phylogenetic distribution. *Nucleic Acids Res* 29: 1017–1026.
- Van Dijk EL, Schilders G, Pruijn GJ (2007) Human cell growth requires a functional cytoplasmic exosome, which is involved in various mRNA decay pathways. *RNA* 13: 1027–1035.
- Briggs MW, Burkard KT, Butler JS (1998) Rrp6p, the yeast homologue of the human PM-Scl 100-KDa autoantigen, is essential for efficient 5.8 S rRNA 3' end formation. *J Biol Chem* 273: 13255–13263.
- Callahan KP, Butler JS (2008) Evidence for core exosome independent function of the nuclear exoribonuclease Rrp6p. *Nucleic Acids Res* 36: 6645–6655.
- Stead JA, Costello JL, Livingstone MJ, Mitchell P (2007) The PMC2NT domain of the catalytic exosome subunit Rrp6p provides the interface for binding with its cofactor Rrp47p, a nucleic acid-binding protein. *Nucleic Acids Res* 35: 5556–5567.
- Midtgaard SF, Assenholt J, Jonstrup AT, Van LB, Jensen TH, et al. (2006) Structure of the nuclear exosome component Rrp6p reveals an interplay between the active site and the HRDC domain. *Proc Natl Acad Sci USA* 103: 11898–11903.
- Januszky K, Liu Q, Lima CD (2011) Activities of human RRP6 and structure of human RRP6 catalytic domain. *RNA* 17: 1566–1577.
- Mitchell P, Petfalski E, Houalla R, Podtelejnikov A, Mann M, et al. (2003) Rrp47p is an exosome-associated protein required for the 3' processing of stable RNAs. *Mol Cell Biol* 23: 6982–6992.
- Schilders G, van Dijk E, Pruijn JM (2007) C1D and hMtr4p associate with the human exosome subunit PM/Scl-100 and are involved in pre-rRNA processing. *Nucleic Acids Res* 35: 2564–2572.
- Costello JL, Stead JA, Feigenbutz M, Jones RM, Mitchell P (2011) The C-terminal region of the exosome-associated protein Rrp47 is specifically required for box C/D small nucleolar RNA 3'-maturation. *J Biol Chem* 286: 4535–4543.
- Estévez AM, Kempf T, Clayton C (2001) The exosome of *Trypanosoma brucei*. *EMBO J* 20: 3831–3839.
- Estévez AM, Lehner B, Sanderson CM, Ruppert T, Clayton C (2003) The roles of intersubunit interactions in exosome stability. *J Biol Chem* 278: 34943–34951.
- Haile S, Cristodero M, Clayton C, Estévez AM (2007) The subcellular localization of trypanosome RRP6 and its association with the exosome. *Mol Biochem Parasitol* 151: 52–58.
- Cristodero M, Böttcher B, Diepholz M, Scheffzek K, Clayton C (2008) The *Leishmania tarentolae* exosome: Purification and structural analysis by electron microscopy. *Mol Biochem Parasitol* 159: 24–29.
- Giuliani A, Jamme F, Rouam V, Wien F, Giorgetta JL, et al. (2009) DISCO: a low-energy multipurpose beamline at synchrotron SOLEIL. *J Synchrotron Radiat* 16: 835–841.
- Refregiers M, Wien F, Ta HP, Premvardhan L, Bac S, et al. (2012) DISCO synchrotron-radiation circular-dichroism endstation at SOLEIL. *J Synchrotron Radiat* 19: 831–935.
- Whitmore L, Wallace BA (2008) Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers* 89: 392–400.
- Gateiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, et al. (2005) Protein identification and analysis tools on the ExPASy Server. In: John MW, editors. *The proteomics protocols handbook*. Humana Press. pp. 571–607.
- Wien F, Wallace BA (2005) Calcium fluoride micro cells for synchrotron radiation circular dichroism spectroscopy. *Appl Spectrosc* 59: 1109–1113.
- Lees JG, Smith BR, Wien F, Miles AJ, Wallace BA (2004) CDtool-An integrated software package for circular dichroism spectroscopic data processing, analysis and archiving. *Anal Biochem* 332: 285–289.
- Lobley A, Whitmore L, Wallace BA (2002) Dichroweb: an interactive website for the analysis of protein secondary structure from circular dichroism spectra. *Bioinformatics* 18: 211–212.
- Kabsh W (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J Appl Crystallogr* 26: 795–800.
- McCoy AJ, Grosse-kunstleve RW, Adams PD, Winn MD, Storoni LC, et al. (2007) Phaser crystallographic software. *J Appl Crystallogr* 40: 658–674.
- Bricogne G, Blanc E, Brandl M, Flensburg C, Keller P, et al. (2009) BUSTER, version 2.9.1. Global Phasing Ltd., Cambridge, UK.
- Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60: 2126–2132.

40. Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, et al. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* 35: W375–W383.
41. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci USA* 98: 10037–10041.
42. Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA (2004) PDB2PQR: an automated pipeline for the setup, execution, and analysis of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res* 32: W665–W667.
43. Feigenbutz M, Jones R, Besong TMD, Harding SE, Mitchell P (2013) Assembly of the yeast exoribonuclease Rrp6 with its associated cofactor Rrp47 occurs in the nucleus and is critical for the controlled expression of Rrp47. *J Biol Chem* 288: 15959–15970.
44. Fadda A, Färber V, Droll D, Clayton C (2013) The roles of 3'-exoribonucleases and the exosome in trypanosome mRNA degradation. *RNA* 19: 937–947.
45. Cudny H, Zaniewski R, Deutscher MP (1981) Escherichia coli RNaseD. *J Biol Chem* 256: 5633–5637.
46. Sorrentino S, Naddeo M, Russo A, D'Alessio G (2003) Degradation of double-stranded RNA by human pancreatic ribonuclease: crucial role of noncatalytic basic amino acid residues. *Biochemistry* 42: 10182–10190.