STRUCTURAL BIOLOGY COMMUNICATIONS

ISSN 2053-230X

Crystal structure of the MIF4G domain of the *Trypanosoma cruzi* translation initiation factor EIF4G5

Lucca Pietro Camillo dos Santos,[a,b] Bruno Moisés de Matos,[a,c] Brenda Cecilia de Maman Ribeiro,[a] Nilson Ivo Tonin Zanchin[a] and Beatriz Gomes Guimarães[a,b,c]*

[a]Carlos Chagas Institute, Oswaldo Cruz Foundation, Curitiba, Paraná, Brazil, [b]Biosciences and Biotechnology Postgraduate Program, Carlos Chagas Institute, Curitiba, Paraná, Brazil, and [c]Biochemistry Postgraduate Program, Federal University of Paraná, Curitiba, Paraná, Brazil. *Correspondence e-mail: beatriz.guimaraes@fiocruz.br

Received 10 September 2019
Accepted 7 November 2019

Edited by R. C. Garratt, University of São Paulo, Brazil

Keywords: MIF4G domain; translation initiation factor EIF4G5; *Trypanosoma cruzi*.

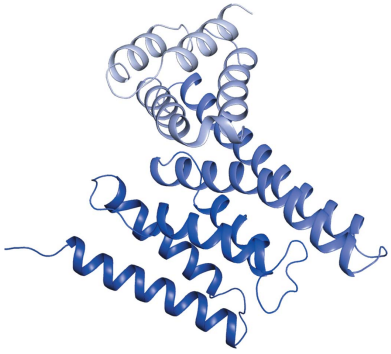
PDB reference: MIF4G domain of EIF4G5, 6ozu

Supporting information: this article has supporting information at journals.iucr.org/f

Kinetoplastida, a class of early-diverging eukaryotes that includes pathogenic *Trypanosoma* and *Leishmania* species, display key differences in their translation machinery compared with multicellular eukaryotes. One of these differences involves a larger number of genes encoding eIF4E and eIF4G homologs and the interaction pattern between the translation initiation factors. eIF4G is a scaffold protein which interacts with the mRNA cap-binding factor eIF4E, the poly(A)-binding protein, the RNA helicase eIF4A and the eIF3 complex. It contains the so-called middle domain of eIF4G (MIF4G), a multipurpose adaptor involved in different protein–protein and protein–RNA complexes. Here, the crystal structure of the MIF4G domain of *T. cruzi* EIF4G5 is described at 2.4 Å resolution, which is the first three-dimensional structure of a trypanosomatid MIF4G domain to be reported. Structural comparison with IF4G homologs from other eukaryotes and other MIF4G-containing proteins reveals differences that may account for the specific interaction mechanisms of MIF4G despite its highly conserved overall fold.

1. Introduction

Eukaryotic initiation factor 4G (eIF4G) is a large and multi-domain protein that plays a scaffold function during translation initiation. The middle domain of eIF4G, termed the MIF4G domain, has been shown to act as a multipurpose adaptor, mediating protein–protein interactions with the translation initiation factors eIF4A and eIF3 and also displaying RNA-binding capabilities (Lamphear *et al.*, 1995; Pestova *et al.*, 1996; Imataka & Sonenberg, 1997). Besides the eIF4F complex, MIF4G-domain-containing proteins are involved in several pathways of mRNA maturation and surveillance, translation and degradation (Ponting, 2000). Some examples are the death-associated protein 5 (DAP5/p97) involved in cap-independent translation mechanisms, which interacts with eIF4A via an MIF4G domain (Virgili *et al.*, 2013), and the exon junction complex protein Upf2, which interacts with its binding partner Upf3 through an MIF4G domain (Kadlec *et al.*, 2004). Additionally, poly(A)-binding protein-interacting protein 1 (PAIP1) contains an MIF4G domain which has been suggested to be responsible for its interaction with eIF4A, with a role in translation regulation (Lei *et al.*, 2011). The MIF4G domain also has a role in the formation of the cytoplasmic deadenylase complex Ccr4–Not (Basquin *et al.*, 2012; Petit *et al.*, 2012; Mathys *et al.*, 2014).



Kinetoplastida, a class of early-diverging eukaryotes that includes pathogenic *Trypanosoma* and *Leishmania* species, display key differences in their translation machinery compared with other eukaryotes. One of these differences involves a larger number of genes encoding eIF4E and eIF4G homologs and the mode of interaction between the translation initiation factors. Trypanosomatids possess six homologs of eIF4E and five homologs of eIF4G, and the specific functions of the different homologs in the context of protein synthesis are not fully understood. Moreover, the interaction between these homologs to form eIF4F-like complexes has been shown to be species-specific (reviewed in Zinoviev & Shapira, 2012; Freire *et al.*, 2017).

Although all five trypanosomatid homologs of eIF4G are predicted to have an MIF4G domain, interactions involving this domain have only recently been described in *Leishmania major* and *Trypanosoma brucei* between the initiation factors EIF4G3 and eIF4A, and between EIF4G4 and eIF4A (Moura *et al.*, 2015). *T. brucei* EIF4G5 was found in a complex with EIF4E6 and an uncharacterized protein of 70.3 kDa, referred to as TbG5-IP, which contains a P-loop NTPase and a guanylyltransferase domain (Freire *et al.*, 2014). However, the interaction motifs involved in such complexes are still unknown. In this work, we have determined the crystal structure of the MIF4G domain of the *T. cruzi* initiation factor EIF4G5 at 2.4 Å resolution, which is the first three-dimensional structure of a trypanosomatid MIF4G domain to be described. This work provides structural data which may contribute to the understanding of the molecular basis for the recognition of specific partners by the extended eIF4G family in these organisms.

2. Materials and methods

2.1. Protein expression and purification

The nucleotide sequence encoding the MIF4G domain of *T. cruzi* EIF4G5 (TcEIF4G5-MIF, comprising residues 142–371) was synthesized by GeneCust after codon optimization for *Escherichia coli* expression and was cloned into the pET-28a vector (Novagen). The recombinant protein was expressed fused with a C-terminal His tag. *E. coli* BL21 Star (DE3) cells were transformed with the expression vector and incubated at 37°C in LB medium containing 100 µg ml^{-1} kanamycin. When the culture reached an OD_{600} of ∼0.6, the temperature was reduced to 18°C and protein expression was induced with 0.2 m*M* isopropyl *β*-D-1-thiogalactopyranoside (IPTG) for approximately 16 h. Cells from 1 l of culture were harvested by centrifugation at 6000*g* for 15 min and suspended and lysed in 20 ml buffer *A* (50 m*M* Tris–HCl pH 8.8, 500 m*M* NaCl, 20 m*M* imidazole) using a microfluidizer (Microfluidics 110L series 0300).

The soluble fraction was isolated by centrifugation at 40 000*g* for 30 min at 4°C. The extract was loaded onto a 5 ml His-Trap HP column (GE Healthcare Life Sciences) equilibrated with buffer *A*. TcEIF4G5-MIF was eluted with a seven column volume (CV) linear gradient from 0 to 15% buffer *B* (buffer *A* + 500 m*M* imidazole), followed by a seven CV linear gradient from 15 to 30% buffer *B*. Fractions containing the target protein were pooled, concentrated to a final volume of 1 ml and loaded onto a Superdex 75 16/60 column (GE Healthcare Life Sciences) equilibrated with 20 m*M* Tris–HCl pH 8.8, 300 m*M* NaCl. The eluted protein was concentrated to approximately 5.5 mg ml^{-1} for crystallization assays.

Expression of selenomethionine-labeled (SeMet) TcEIF4G5-MIF was performed based on the metabolic inhibition method. Briefly, *E. coli* BL21 Star (DE3) cells transformed with the expression vector were incubated at 37°C and 200 rev min^{-1} in M9 minimal medium supplemented with the selection antibiotic as described above. When the culture reached an OD_{600} of ∼0.6, it was supplemented with 10 ml of a sterile mixture containing lysine, phenylalanine and threonine at 10 mg ml^{-1}, isoleucine, leucine and valine at 5 mg ml^{-1} and selenomethionine at 6 mg ml^{-1} diluted into M9 medium without glucose and thiamine. The temperature was reduced to 18°C, and after 30 min of incubation protein expression was induced with 0.2 m*M* IPTG for approximately 16 h. The cells were harvested by centrifugation and SeMet TcEIF4G5-MIF was purified using the same protocol as was used to purify the nonlabeled TcEIF4G5-MIF.

2.2. Crystallization, data collection and processing

TcEIF4G5-MIF was submitted to crystallization trials by the sitting-drop vapor-diffusion method using commercial screens. Optimization of the initial crystallization conditions was performed by varying the precipitant and the buffer. The best crystals were obtained by hanging-drop vapor diffusion at 18°C by mixing the TcEIF4G5-MIF protein at 5 mg ml^{-1} in 20 m*M* Tris–HCl pH 8.8, 300 m*M* NaCl with reservoir solution consisting of 100 m*M* sodium cacodylate pH 6.5, 2 *M* ammonium sulfate. SeMet TcEIF4G5-MIF crystals were grown using the same conditions.

Before data collection, the crystals were cryoprotected by the addition of 25%(*v*/*v*) glycerol to the mother liquor prior to flash-cooling in liquid nitrogen. X-ray diffraction data were collected from TcEIF4G5-MIF and SeMet TcEIF4G5-MIF crystals on the PROXIMA-1 beamline at Synchrotron SOLEIL using a PILATUS 6M detector (Dectris). Diffraction data were processed using the *XDS* package (Kabsch, 2010). The TcEIF4G5-MIF data set was submitted to anisotropy analysis using the *STARANISO* server (Tickle *et al.*, 2018; http://staraniso.globalphasing.org/cgi-bin/staraniso.cgi). An anisotropic cutoff on merged intensity and anisotropic correction were applied to the data. The data statistics are presented in Table 1.

2.3. Structure determination

The structure of TcEIF4G5-MIF was determined by single-wavelength anomalous diffraction using diffraction data from SeMet TcEIF4G5-MIF crystals. The positions of heavy atoms were determined using *SHELXD* (Sheldrick, 2010), the initial phases were calculated using *Phaser* (McCoy *et al.*, 2007) and density modification was performed using *Parrot* (Cowtan,

Table 1
Data-collection and refinement statistics.

Data collection. Values in parentheses are for the outer shell.

	SeMet TcEIF4G5-MIF	TcEIF4G5-MIF†
Diffraction source	PROXIMA-1, SOLEIL	PROXIMA-1, SOLEIL
Wavelength (Å)	0.9790	1.0507
Temperature (K)	100	100
Space group	$P4_12_12$	$P4_12_12$
a, b, c (Å)	70.32, 70.32, 303.06	70.10, 70.10, 303.50
α, β, γ (°)	90, 90, 90	90, 90, 90
Resolution range (Å)	50–2.55 (2.70–2.55)	48.92–2.40 (2.51–2.40)
Total No. of reflections	661195 (103844)	310216 (14619)
No. of unique reflections	47422 (7656)	24867 (1243)
Completeness (spherical) (%)	99.9 (99.4)	80.2 (32.1)
Completeness (ellipsoidal) (%)		94.5 (93.0)
Multiplicity	13.9 (13.6)	12.5 (11.8)
$\langle I/\sigma(I)\rangle$	19.70 (0.96)	21.80 (2.20)
R_{meas} (%)	7.4 (260.7)	6.4 (129.6)
$CC_{1/2}$	100 (65.3)	100 (81.5)
Anomal. Corr.	22 [at 3.12 Å]	
SigAno	0.995 [at 3.12 Å]	

Refinement. Values in parentheses are for the outer shell.

	TcEIF4G5-MIF†
Resolution	48.92–2.40 (2.47–2.40)
R_{work}	0.201 (0.257)
R_{free}	0.243 (0.267)
No. of non-H atoms	
Protein	3789
Heterogen atoms	101
Water	95
R.m.s. deviations	
Bonds (Å)	0.010
Angles (°)	1.13
B factors (Å^2)	
From Wilson plot	66.66
Monomer *A* (main chain)	63.36
Monomer *B* (main chain)	73.95
Water	61.89
Ramachandran plot	
Most favored (%)	98.06
Outliers (%)	0

† The TcEIF4G5-MIF data set was submitted to anisotropy analysis and correction using *STARANISO* (Tickle *et al.*, 2018).

2010). Automatic model building was performed using *Buccaneer* (Cowtan, 2006). At this point, we used the higher resolution data set from the nonlabeled TcEIF4G5-MIF crystal in order to proceed with structure refinement by alternating cycles of *BUSTER* (Bricogne *et al.*, 2017) with visual inspection and manual rebuilding using *Coot* (Emsley *et al.*, 2010). Model validation was performed using the *MolProbity* web server (http://molprobity.biochem.duke.edu/; Chen *et al.*, 2010). Refinement statistics are presented in Table 1. The atomic coordinates and structure factors have been deposited in the Protein Data Bank under accession code 6ozu.

2.4. Structural analysis

Structural comparisons were performed using the *DALI* server (Holm & Laakso, 2016). 3D models of *T. cruzi* initiation factors EIF4G1–EIF4G4 were generated using the *Phyre*2 server (Kelley *et al.*, 2015). Structure-based sequence alignments were performed using *PROMALS3D* (Pei *et al.*, 2008).

3. Results and discussion

3.1. TcEIF4G5-MIF displays local differences despite overall conservation of the MIF4G domain structure

Initial attempts to determine the crystal structure of the MIF4G domain of *T. cruzi* EIF4G5 (TcEIF4G5-MIF) by molecular-replacement methods failed, although MIF4G homolog structures are available in the Protein Data Bank. Therefore, selenomethionine-labeled (SeMet) TcEIF4G5-MIF was produced and structure determination was achieved by the single-wavelength anomalous diffraction method. The initial model obtained from the experimental electron-density map was used to complete building and refinement against a higher resolution native data set. The TcEIF4G5-MIF structure was refined at 2.40 Å resolution to final R_{work} and R_{free} values of 20.1% and 24.3%, respectively (Table 1). The polypeptide chains of the two monomers in the asymmetric unit were clearly defined in the electron density. Superposition of the monomers in the asymmetric unit resulted in an r.m.s.d. of 0.57 Å and we used monomer *A* for structural comparisons.

TcEIF4G5-MIF conserves the overall structure of the MIF4G domain, consisting of five tandem HEAT repeats, resulting in a solenoid-type structure (Marcotrigiano *et al.*, 2001). Each HEAT repeat consists of two parallel α-helices, named 'a' and 'b', connected by a loop. The so-called 'a' and 'b' helices form the concave and convex sides of the MIF4G domain structure, respectively (Fig. 1). Structural comparison using the *DALI* server (Holm & Laakso, 2016) confirms the

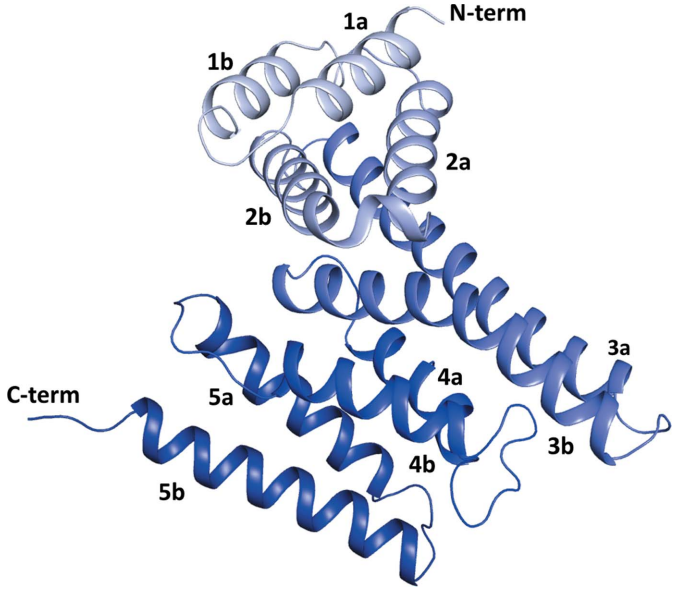


Figure 1
Ribbon representation of the MIF4G domain of *T. cruzi* EIF4G5. The HEAT repeats, numbered from 1 to 5, comprise α-helices 'a' and 'b' and are colored from light to dark blue. The N- and C-termini are indicated.

structural similarity between TcEIF4G5-MIF and other MIF4G domains despite their low sequence identity (Table 2). However, local structural differences are observed (Fig. 2). We initially investigated the structural similarity between TcEIF4G5-MIF and MIF4G domains from initiation factors 4G through a comparison with the available structures of human EIF4G3 (formerly called eIF4GII; Marcotrigiano *et al.*, 2001) and *Saccharomyces cerevisiae* EIF4G1 (Schütz *et al.*, 2008). The three homologs feature a long HEAT repeat 3; however, TcEIF4G5-MIF features longer helices 3a and 3b, whereas extended flexible loops connecting these helices (not modeled in the structures) are present in human and *S. cerevisiae* eIF4G (Fig. 2*a*). Structural comparison with other MIF4G-containing proteins shows that the MIF4G domains from human Upf2 (Kadlec *et al.*, 2004) and PAIP1 (Lei *et al.*, 2011) feature shorter helices forming HEAT repeat 3 when compared with TcEIF4G5-MIF (Fig. 2*b*).

The TcEIF4G5-MIF structure also shows a distinct orientation of the fifth HEAT repeat. Helices 5a and 5b in TcEIF4G5-MIF show a significant twist relative to the homolog structures (Fig. 2). Helices 4b and 5b participate in the intermolecular interface of the asymmetric unit dimer (not shown), which could explain the conformational change involving HEAT repeat 5 of TcEIF4G5-MIF. However, the reorientation of HEAT repeats may also suggest a molecular mechanism for interaction with partners in multi-protein complexes (as discussed below). The MIF4G domain has been shown to act as a multipurpose adaptor; thus, despite its overall structural conservation, local differences are likely to be the key to the recognition of distinct partners among the various MIF4G-containing complexes.

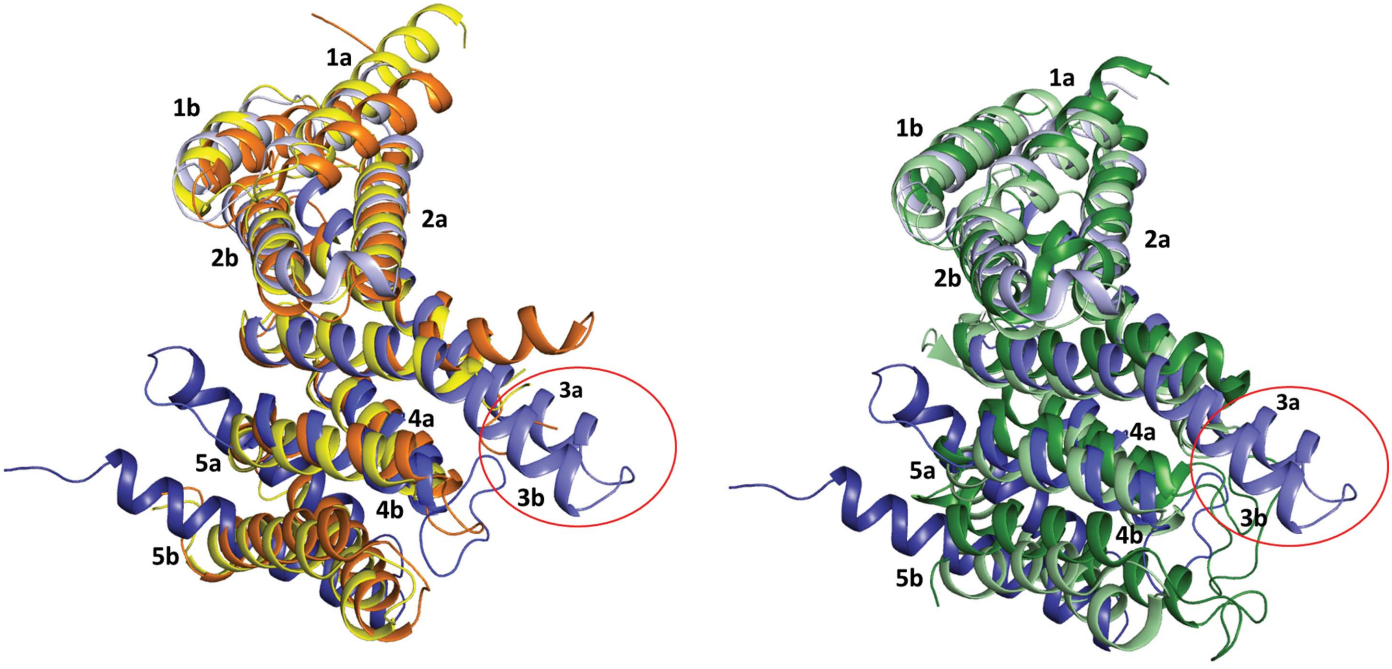
Table 2
Structural comparison of the MIF4G domain of *T. cruzi* EIF4G5 against the Protein Data Bank using the *DALI* server.

Superposition data for the top five hits are presented.

Protein	PDB code	*Z*-score	No. of aligned residues	R.m.s.d. (Å)	Identity (%)
S. cerevisiae EIF4G1	2vsx	14.6	187	3.9	16
H. sapiens PAIP1	3rk6	13.6	177	3.4	20
H. sapiens CWC22	6icz	12.7	177	3.4	15
H. sapiens Upf2	4cek	12.0	167	3.2	15
H. sapiens CBP80	3fey	10.4	174	3.7	12

3.2. The residues involved in the eIF4G–eIF4A interaction are not conserved in *T. cruzi* EIF4G5-MIF

The middle domain of eIF4G has been shown to interact with the RNA helicase eIF4A in the formation of the translation initiation complex eIF4F (Imataka & Sonenberg, 1997). The crystallographic structure of the MIF4G domain of *S. cerevisiae* EIF4G1 in complex with eIF4A shows that the interaction involves HEAT repeats 1, 2 and 5 of MIF4G (Schütz *et al.*, 2008; Fig. 3). The sites of interaction with eIF4A were shown to be structurally conserved in the eIF4G homolog DAP5 (Virgili *et al.*, 2013). In particular, a conserved motif observed in HEAT repeat 1 of ScEIF4G1-MIF, 611-KSLLNKLTLEMF-622, was shown to be essential for the interaction with eIF4A (Schütz *et al.*, 2008). Structural



(*a*) (*b*)

Figure 2
Structural superposition of the MIF4G domains of *T. cruzi* EIF4G5 and other MIF4G-domain-containing proteins. (*a*) Superposition of TcEIF4G5-MIF (light to dark blue) and the MIF4G domains of *Homo sapiens* EIF4G3 (yellow; PDB entry 1hu3; former name eIF4GII; Marcotrigiano *et al.*, 2001) and *S. cerevisiae* EIF4G1 (orange; PDB entry 2vsx; Schütz *et al.*, 2008). The HEAT repeats are identified. (*b*) Superposition of TcEIF4G5-MIF (light to dark blue) and the MIF4G domains of *H. sapiens* Upf2 (dark green; PDB entry 1uw4; Kadlec *et al.*, 2004) and *H. sapiens* PAIP1 (light green; PDB entry 3rk6; Lei *et al.*, 2011). Figs. 2(*a*) and 2(*b*) are shown in the same orientation. The helices 3a and 3b extension in TcEIF4G5-MIF is highlighted in red.

superposition of TcEIF4G5-MIF reveals that whereas the residues involved in hydrophobic interactions which stabilize the packing of α-helices are conserved, four residues whose side chains make direct contacts with eIF4A in ScEIF4G1 (Ser612, Asn615, Thr618 and Glu620) are replaced by glycines in *T. cruzi* EIF4G5 (Fig. 3*b*). In addition, the conformational change of HEAT repeat 5 in TcEIF4G5-MIF causes helices 5a and 5b to move away from the interface with eIF4A when compared with the homolog from *S. cerevisiae* (Fig. 3*a*). Interestingly, this structural difference is mainly owing to a

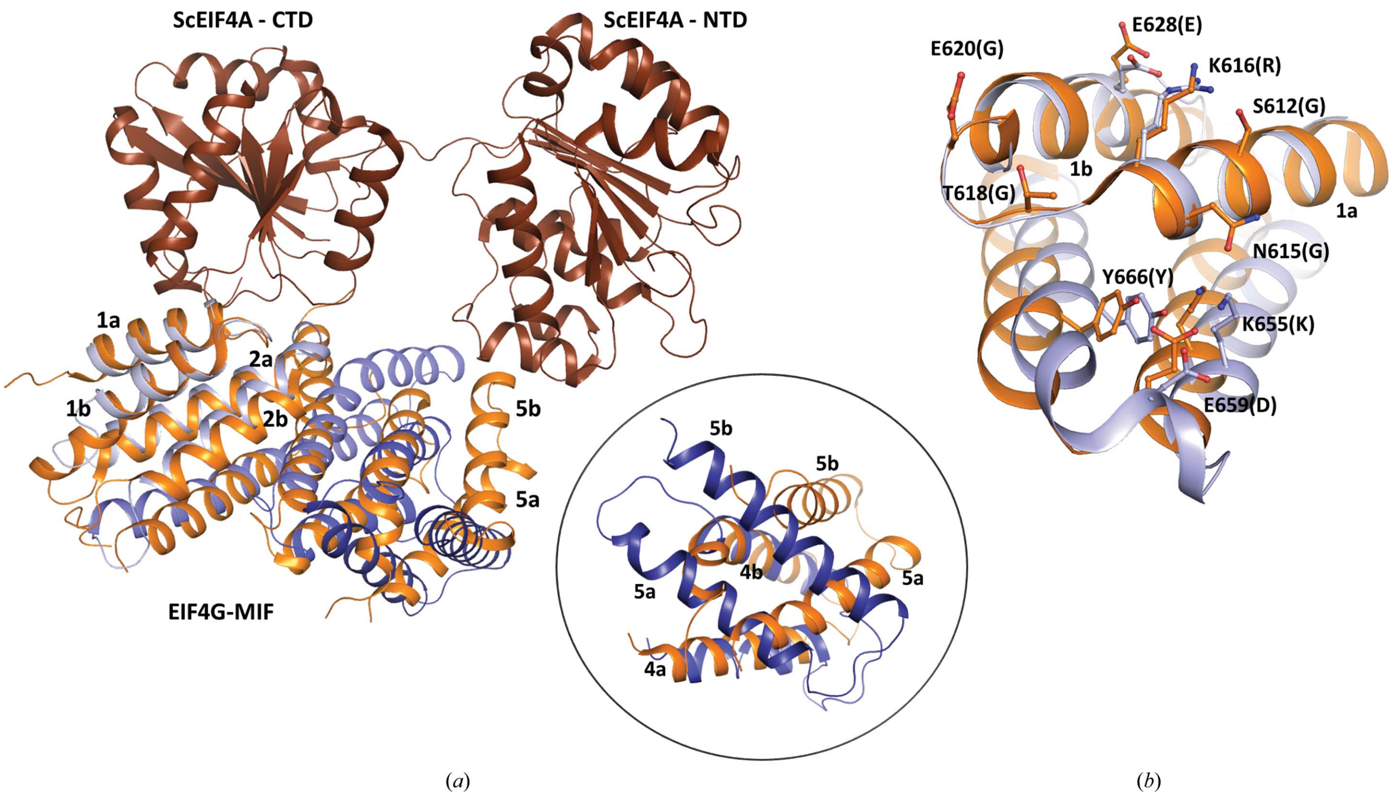


Figure 3
Superposition of TcEIF4G5-MIF (blue) and *S. cerevisiae* EIF4G1-MIF (ScEIF4G1-MIF; orange) in complex with eIF4A (brown; PDB entry 2vsx). (*a*) Overall superposition. The HEAT repeats 1, 2 and 5 involved in the ScEIF4G1–eIF4A interaction are indicated. The inset highlights the orientation change of helices 5a and 5b between the two structures. (*b*) Detail of the MIF4G residues involved in the ScEIF4G1–eIF4A interaction. Helices 1a and 1b are indicated. The corresponding residues in TcEIF4G5-MIF are indicated in parentheses.

```
                             *  ** *   *      *                             *     *     *
ScEIF4G1  ANRWVPKTELLDKDEVERKMKSLLNKLT--LEMFDAISSEILAIANISVWETNGETLKAVIEQIFLKACD--EPHWSSMYAQLCGKVVKELNPDITDETN 686
TcEIF4G1  CSLRGGRTRAKPVRLTHHKVMSILSRIT--PAKYDVLLDEMMLLPL---RQLEDTELLEICKVVFEKAVQ--EPAYSGLYARLAQDVCSMKDGGRETEP- 269
TcEIF4G2  -----------------HLVISILNRVSNDTEKYREVKNELLRLPI---PEANDAMLAKIVDVFFLKAVR--EQHFSNLYADLVSALCKVPGGQRIVGD- 345
TcEIF4G3  ----------------ERLVQGSLNKLS--TANFDDMVAKLQTEVI----FSTKETLNIAVRIIFKKALD--EPECSKSYAGVCYKLAEFEVGLTAAKQK 136
TcEIF4G4  -----------------RMVQSTLNKLT--ESNFDLIAKAILVPEI----ILNANVVADVVRLIYDKALV--EPVFAGLYARLCYMIVRYEYDYRSTQSG 162
TcEIF4G5  -------------RKLHNEVLGILGRVG--EGNLPKMKEELTNLPI---RQSTDKEIQDVIQVIFNKSIQPEDSIFVPYYVKLVVSLINDIGEG------ 219
                           α1a                α1b                    α2a                      α2b

ScEIF4G1  ETKTGPKLVLHYLVARCHAEFDKGWTD------------KLPMSEEYYAAASAKRRGLGLVRFIGFLYRLNLLTGKMMFECFRRLMKDLT---------- 776
TcEIF4G1  RDQSFSRRFRRMLIASCEAQFQRPLQLSADELVDRVTGVPLGEEEVEGRRTRLKDRLVGNIRFVAELFKIGFLTERVIEDVIRILVRDYNP------DEP 363
TcEIF4G2  KQQSLEYRMRQQLLHRCQMEFRQLEEQSIERAE--RGGEDRASQGAEESFKERRDRVCGIVRFVGELFLRHIVTEKVIGNILVTACTGLYGREGLCVPPP 443
TcEIF4G3  AEGKKYSKLRNAVVSIAQEEFLSRRKMP--------SMEGLTEEEMELRRTTFMRRKRANMKFIGELFMHKVLSCNTMMNIIQTIMQEAE--------RG 220
TcEIF4G4  EVDAQKSEVRVAIVEKCQQMFDGATKAV-------------RETASEEEAEKVRKRNVNNIKFAGELFLKSLITQKIIDRILGEKLFNI----------- 238
TcEIF4G5  --EPAGRFIRNAIIRQCQRTFENAEEAQAQLER---EIANLPEEEAEQRRLIFAGKQKANINFLGLLFTHGLVREKVVLHVLEWLLYGTER-------KR 307
                        α3a                                   α3b                       α4a

                                                                  *  *
ScEIF4G1  DSPSEETLESVVELLNTVGEQFETDSEGSQLLDSLFGILDNIIQTA-KISSRIKFKLIDIKELRHDKNWN------ 853
TcEIF4G1  TAKEEAIFEVFQTLLRHTGAVLKEH--MPQLLAQSLGIAKTIELS--HPKPRVVFLMMDLSDLNR--ANG------ 427
TcEIF4G2  YTPTESQMDETITLLGIVDAAFFRTPLGSSMLVECTQFLQHWSLR--HPVSRIRFLLMSAVERLQ--KLVEERNA- 514
TcEIF4G3  GYPTSEDIEFLTELFLTIGESLDAVPELRVRLNGYFELLEKLKDQKEVYPPRIRFKMLDLIELRS----------- 285
TcEIF4G4  -TPNDMDLEVVVNLLEVVGKMYEEK--HPAAQTALWKLLGELQEER-RFSNRIRFLLQNLIERRN----------- 299
TcEIF4G5  RFPADYELIHFMNLLLTCGKSFSKE--GQEFVPKFRAVLEELMHM--HPQRRMQFLLLNTVETID-NNW------- 371
                      α4b                α5a                  α5b
```

Figure 4
Structure-based sequence alignment of the MIF4G domains of *S. cerevisiae* EIF4G1 (ScEIF4G1) and *T. cruzi* eIF4G homologs. The alignment was performed with *PROMAL3D* using the ScEIF4G1 and TcEIF4G5 crystal structures and the homology models generated for EIF4G1 to EIF4G4 from *T. cruzi*. Sequences are colored according to amino-acid identity (yellow background) or similarity (gray background). The residues forming the HEAT repeats are highlighted in red and the corresponding α-helices are labeled. The residues participating in the interaction of ScEIF4G1 with eIF4A are indicated with stars.

bend in the C-terminal portion of helix 4b of *S. cerevisiae* EIF4G1 which is not observed in TcEIF4G5-MIF (see the inset in Fig. 3*a*).

The interaction between initiation factors eIF4G and eIF4A in trypanosomatids has been described to involve the EIF4G3 and EIF4G4 homologs (Moura *et al.*, 2015). In *L. major*, the amino-acid triplet LNK (belonging to the conserved MIF4G motif described above) is essential for eIF4A binding (Moura *et al.*, 2015). In order to analyze the conservation of the eIF4A-interacting residues within *T. cruzi* eIF4G homologs, we generated homology models for the MIF4G domains of *T. cruzi* EIF4G1 to EIF4G4 and performed a structure-based sequence alignment with the MIF4G domains of *T. cruzi* EIF4G5 and *S. cerevisiae* EIF4G1 (Fig. 4). Similarly to *L. major*, the helix 1a LNK triplet is conserved in *T. cruzi* EIF4G3 and EIF4G4, suggesting a conserved interaction pattern with eIF4A, which is partially conserved in TcEIF4G2. The frequency of glycine residues in helix 1a appears to be characteristic of EIF4G5 and indicates a disruption in this homolog of the interaction surface of HEAT repeat 1 with EIF4A.

The protein-binding roles of MIF4G are specific to the cellular context and partners and are likely to be determined by specific molecular features within its conserved overall structure. Structural differences in MIF4G domains convey significant differences in shape and chemical properties to their surfaces, and are likely to contribute to their function as selective scaffolding domains.

4. Conclusion

We have described the first three-dimensional structure of an MIF4G domain belonging to a trypanosomatid translation initiation factor eIF4G. Despite the overall structural conservation, the MIF4G domain of *T. cruzi* EIF4G5 features specific differences which may play a role in the mechanism of partner recognition and may account for the distinct functions of trypanosomatid eIF4G homologs in translation or regulatory mechanisms. Moreover, given the diversity of MIF4G interaction partners, detailed atomic information on MIF4G domains contributes to illustrate the spectrum of MIF4G structural specificities behind their interaction properties.

Acknowledgements

The authors acknowledge the Platform for Protein Purification and Characterization of the FIOCRUZ Technical Platform program. The authors are grateful to Synchrotron SOLEIL for beam time awarded at the PROXIMA-1 beamline (Proposal 20170736) and for all of the support that they have benefited from as SOLEIL users.

Funding information

This work was supported by Fundação Araucária (grant 05/2016), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES Financing Code 001), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, NITZ 312195/2015-0 BGG 304027/2015-4) and Fundação Oswaldo Cruz.

References

Basquin, J., Roudko, V. V., Rode, M., Basquin, C., Séraphin, B. & Conti, E. (2012). *Mol. Cell*, **48**, 207–218.
Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Sharff, A., Smart, O. S., Vonrhein, C. & Womack, T. O. (2017). *BUSTER* v.2.10.2. Global Phasing Ltd, Cambridge, UK.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Cowtan, K. (2006). *Acta Cryst.* D**62**, 1002–1011.
Cowtan, K. (2010). *Acta Cryst.* D**66**, 470–478.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Freire, E. R., Malvezzi, A. M., Vashisht, A. A., Zuberek, J., Saada, E. A., Langousis, G., Nascimento, J. D., Moura, D., Darzynkiewicz, E., Hill, K., de Melo Neto, O. P., Wohlschlegel, J. A., Sturm, N. R. & Campbell, D. A. (2014). *Eukaryot. Cell*, **13**, 896–908.
Freire, E. R., Sturm, N. R., Campbell, D. A. & de Melo Neto, O. P. (2017). *Pathogens*, **6**, 55.
Holm, L. & Laakso, L. M. (2016). *Nucleic Acids Res.* **44**, W351–W355.
Imataka, H. & Sonenberg, N. (1997). *Mol. Cell. Biol.* **17**, 6940–6947.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Kadlec, J., Izaurralde, E. & Cusack, S. (2004). *Nat. Struct. Mol. Biol.* **11**, 330–337.
Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. (2015). *Nat. Protoc.* **10**, 845–858.
Lamphear, B. J., Kirchweger, R., Skern, T. & Rhoads, R. E. (1995). *J. Biol. Chem.* **270**, 21975–21983.
Lei, J., Mesters, J. R., Brunn, A. & Hilgenfeld, R. (2011). *Biochem. Biophys. Res. Commun.* **408**, 680–685.
Marcotrigiano, J., Lomakin, I. B., Sonenberg, N., Pestova, T. V., Hellen, C. U. T. & Burley, S. K. (2001). *Mol. Cell*, **7**, 193–203.
Mathys, H., Basquin, J., Ozgur, S., Czarnochi-Cieciura, M., Bonneau, F., Aartse, A., Dziembowski, A., Nowotny, M., Conti, E. & Filipowicz, W. (2014). *Mol. Cell*, **54**, 751–765.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Moura, D. M., Reis, C. R. S., Xavier, C. C., da Costa Lima, T. D., Lima, R. P., Carrington, M. & de Melo Neto, O. P. (2015). *RNA Biol.* **12**, 305–319.
Pei, J., Kim, B. H. & Grishin, N. V. (2008). *Nucleic Acids Res.* **36**, 2295–2300.
Pestova, T. V., Shatsky, I. N. & Hellen, C. U. (1996). *Moll. Cell. Biol.* **16**, 6870–6878.
Petit, A. P., Wohlbold, L., Bawankar, P., Huntzinger, E., Schmidt, S., Izaurralde, E. & Weichenrieder, O. (2012). *Nucleic Acids Res.* **40**, 11058–11072.
Ponting, C. P. (2000). *Trends Biochem. Sci.* **25**, 423–426.
Schütz, P., Bumann, M., Oberholzer, A. E., Bieniossek, C., Trachsel, H., Altmann, M. & Baumann, U. (2008). *Proc. Natl Acad. Sci. USA*, **105**, 9564–9569.
Sheldrick, G. M. (2010). *Acta Cryst.* D**66**, 479–485.
Tickle, I. J., Flensburg, C., Keller, P., Paciorek, W., Sharff, A., Vonrhein, C. & Bricogne, G. (2018). *STARANISO*. Global Phasing Ltd, Cambridge, UK.
Virgili, G., Frank, F., Feoktistova, K., Sawicki, M., Sonenberg, N., Fraser, C. S. & Nagar, B. (2013). *Structure*, **21**, 517–527.
Zinoviev, A. & Shapira, M. (2012). *Comp. Funct. Genomics*, **2012**, 813718.