

In silico structural characterization of protein targets for drug development against *Trypanosoma cruzi*

Carlyle Ribeiro Lima^{1,2} · Nicolas Carels³ · Ana Carolina Ramos Guimaraes⁴ · Pierre Tufféry¹ · Philippe Derreumaux²

Received: 28 May 2016 / Accepted: 2 September 2016 / Published online: 24 September 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract *Trypanosoma cruzi* is the protozoan pathogen responsible for Chagas disease, which is a major public health problem in tropical and subtropical regions of developing countries and particularly in Brazil. Despite many studies, there is no efficient treatment against Chagas disease, and the search for new therapeutic targets specific to *T. cruzi* is critical for drug development. Here, we have revisited 41 protein sequences proposed by the analogous enzyme pipeline, and found that it is possible to provide structures for *T. cruzi* sequences with clear homologs or analogs in *H. sapiens* and likely associated with trypanothione reductase, cysteine synthase, and ATPase functions, and structures for sequences specific to *T. cruzi* and absent in *H. sapiens* associated with 2,4-

dienoyl-CoA reductase, and leishmanolysin activities. The implications of our structures refined by atomistic molecular dynamics (monomer or dimer states) in their in vitro environments (aqueous solution or membrane bilayers) are discussed for drug development and suggest that all protein targets, except cysteine synthase, merit further investigation.

Keywords *T. Cruzii* · Protein targets · Structures · Molecular dynamics · Drugs

Introduction

According to the World Health Organization [1], 21 Latin American countries are endemic for Chagas disease, which is caused by the parasite *Trypanosoma cruzi* (*T. cruzi*), and about 6–7 million people are infected with *T. cruzi* [2]. Chagas disease has two phases: the initial acute phase lasts for about 2 months with mild or little symptoms, followed by the chronic phase, in which 30 % (10 %) of patients suffer from cardiac (digestive) disorders before death or heart failure [1].

Chagas disease can be cured with benznidazole and nifurtimox if treatment is initiated soon after infection [3]. Both drugs are activated by a NADH-dependent, mitochondrially localized, bacterial-like, type I nitroreductase (NTR) and resistance to these compounds is associated with loss of NTR [4]. It is well established that both drugs are far from being ideal because of their controversial efficacies in chronic phase, severe host toxicity, and long treatment durations [5].

The identification of new protein targets for drug development against Chagas disease is therefore needed and strategies to identify enzymatic functions specific to the particular

Electronic supplementary material The online version of this article (doi:10.1007/s00894-016-3115-9) contains supplementary material, which is available to authorized users.

✉ Pierre Tufféry
pierre.tuffery@univ-paris-diderot.fr

✉ Philippe Derreumaux
philippe.derreumaux@ibpc.fr

¹ Molécules Thérapeutiques in silico, Université Paris Diderot, INSERM UMR-S 973, Sorbonne Paris Cité, 75013 Paris, France

² Laboratoire de Biochimie Théorique, CNRS UPR 9080, Institut de Biologie Physico – Chimique, Université Paris Diderot, Sorbonne Paris Cité, 75005 Paris, France

³ Laboratório de Modelagem de Sistemas Biológicos, National Institute for Science and Technology on Innovation in Neglected Diseases (INCT-IDN), Centro de Desenvolvimento Tecnológico em Saúde (CDTS), Fundação Oswaldo Cruz (FIOCRUZ), Rio de Janeiro, RJ, Brazil

⁴ Laboratório de Genômica Funcional e Bioinformática, Instituto Oswaldo Cruz (IOC), FIOCRUZ, Rio de Janeiro, Brazil

T. cruzi pathogen are highly desirable. Indeed, the enzyme forms that are absent in humans can be used as targets for the development of new drugs that would have a low chance of side effects [6–8].

Recently, several proteins have been proposed as new biotargets against Chagas disease and 3D models were built based on homology modeling followed by molecular dynamics simulations. These *T. cruzi* biotargets include NADH-dependent fumarate reductase [9], ribosomal P0 antigenic protein [10], silent-information regulator 2 proteins or sirtuins [11], and a putrescine–cadaverine permease [12], among others. While these studies provide useful information on specific targets, a large-scale search of promising proteins is desirable.

In this context, the analogous enzyme pipeline (AnEnPi), which compares genomic datasets for analogous enzymes by clustering the primary structures of enzymes with the same described activity, is a tool able to identify enzymatic activities that may be potentially useful for the design of new therapeutic targets. AnEnPi searches for enzymes specific to a parasite and enzymes that are analogous for the same enzymatic function, but structurally distinct in the parasite and its human host [13]. Using AnEnPi with a Blastp similarity raw score of 120 as cut-off, Capriles et al. identified 41 protein sequences classified as analogous or specific for *T. cruzi* compared to *H. sapiens* on the basis of similarity detection by the local alignment program BLAST, and then used automatic 3D structure modeling to infer differences between the host and parasite protein structures [14].

In this study, we revisit these 41 protein candidates both in terms of specificity and 3D modeling by using a more complex strategy and more accurate techniques. Indeed, it is known that profile–profile alignments are more sensitive than the local alignment program PSIBLAST for similarity [15, 16], and automatic 3D modeling methods lead to significant structural errors for sequence identities <80 % [17]. The workflow for *T. cruzi* protein modeling we developed consists of the following steps: sequence clustering, prediction of cell localization and enzyme classification (EC) number and analysis of metabolic pathways, homology search in *H. sapiens* and other organisms, 3D template identification, model generation and refinement by molecular dynamics (MD) simulations refinement.

Overall, we find that it is possible to cluster the 41 protein sequences identified as analogous or specific for *T. cruzi* versus *H. sapiens* into seven groups, and based on the current Protein Data Bank (PDB) structures, it is possible to provide structures for 33 *T. cruzi* sequences associated with five very probable enzymatic activities: ATPase (EC: 3.6.3.6), trypanothione reductase (1.8.1.12), 2,4-dienoyl-CoA reductase (1.3.1.34), cysteine synthase (2.5.1.47), and leishmanolysin (3.4.24.36). The implications of our results for drug development are discussed.

Materials and methods

Our pipeline for the 41 protein sequences is shown in Fig. 1.

Sequence clustering We used the sequence search of the PFAM server to identify the PFAM domains for each of the 41 sequences [18]. Sequences having the same PFAM domains were then grouped into clusters.

Cell localization prediction We used both the EUK-mPloc 2.0 and CELLO v.2.5 servers [19, 20]. Note that the EUK-mPloc 2.0 learning dataset includes 8897 protein sequences (7766 different proteins), classified into 22 eukaryotic subcellular locations.

EC number prediction and metabolic pathway identification Pathway graphs can be thought of as three information layers: (i) the lowest layer is associated with the metabolites that act as substrates (input) and products (output) of enzymatic reactions, (ii) the intermediate layer consists of the enzymes that perform the enzymatic reaction with which the biological engineer intended to interfere through drug inhibition, and (iii) the highest organization level is given by the topology of a metabolic path in which enzymes are inserted. An enzyme within a non-redundant path is a better target than an enzyme in a redundant one since its inactivation will inevitably result in the elimination of the corresponding product. If that path is a key path for the whole parasite metabolism, the target inhibition is expected to have deleterious consequences on the parasite biology. Considering these three organization levels [21], we used KEGG [22, 23] to map the EC number of an enzyme target and MetaCyc [24] to retrieve information about the metabolic reactions of each enzymatic reaction. We also checked that our predictions are consistent with the AnEnPi server [13].

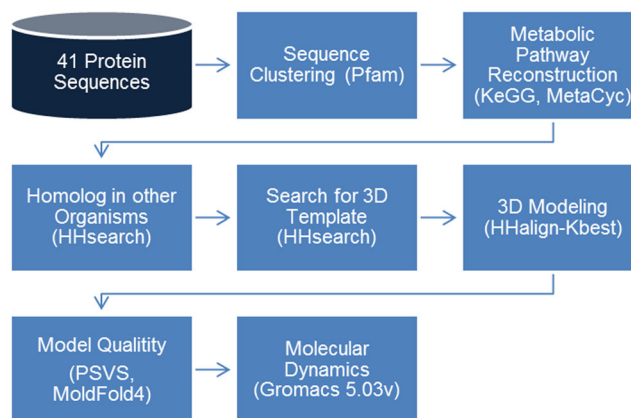


Fig. 1 Workflow used for protein characterization

Homology identification To identify homologs in *H. sapiens* and other organisms, we have used HHpred [25]. A homolog was identified if the sequence identity is > 30 % for an amino acid coverage of more than 90 %.

3D modeling of *T. cruzi* targets We searched for template 3D structures using the HHsearch method hosted on Mobyly against the PDB and SCOP databases [26]. Then 3D models were generated by HHalign-Kbest procedure [27] that perform HHM alignments based on primary sequences and 3D structures. It has to be noted that if HHalign-Kbest identified a hit with a sequence identity >35 % (or < 35 %), 20 (or 50) models were automatically created. Outputs consist of the five best models ranked according their Z-scores awarded by Qmean4 [28] and only the best model was used for homology modeling.

When the template did not cover the full sequence, we examined the putative biological information of the protein (aqueous solution, anchored or inserted into the membrane), and then used the template-based C8-Scorpion [29], the Porter 4.0 [30], and PrDOS [31] servers to determine whether the missing residues are predicted to have secondary structure content or disordered.

Model refinement by molecular dynamics Simulations were performed with Charmm22* force field [32] at pH 7, i.e., with deprotonated Glu and Asp and protonated Arg and Lys for all proteins, and neutral His with a protonated N ϵ atom for all proteins except leishmanolysin. This all-atom force field has been able to fold properly many soluble proteins [33–35]. The soluble proteins were centered in a cubic box of TIP3P water molecules [36] with a box extended 1.5 nm outside the protein on all sides, and the appropriate numbers of Na⁺ and Cl⁻ ions were added to ensure neutral systems. The leishmanolysin protein simulation includes the Zn²⁺ ion covalently bonded to the three histidines of the protein by application of a harmonic restraint on the bond lengths connecting Zn²⁺ to His. The ATPase transmembrane protein was immersed in a 1-palmitoyl-2-oleoylphosphatidylethano lamine (POPE) bilayer and then centered in a cubic box of TIP3P molecules [37]. Each system was subjected to energy minimization with steepest descent and TNPACK [38], followed by a 1-ns MD simulation in NPT ensemble [39], and a 1-ns NVT MD simulation with a velocity-rescaling thermostat found to sample the canonical ensemble [40].

The GROMACS program (version 5.03) was used to perform the MD simulations with periodic boundary conditions [41]. The bond lengths with hydrogen atoms were fixed with the LINCS algorithm and the equations of motion were integrated with a time step of 2 fs [42]. The electrostatic interactions were calculated using the particle mesh Ewald method and a cutoff of 1.1 nm [43]. A cutoff of 1.2 nm was used for the van der Waals interactions. The nonbonded pair lists were

updated every 10 fs. In what follows, the analysis is based on MD simulation of 100 ns at 300 K. The MD-generated structures were evaluated with respect to the minimized structure using the C α root-mean-square deviation (RMSD), and clustered using a RMSD cutoff of 3 Å.

Model evaluation The quality of the models was evaluated by using the meta-server PSVs 1.5, which analyzes the Ramachandran plot and provides Molprobiy and Procheck scores [44]. Following CASP10 recommendations [45], models were also investigated by ModFOLD4 [46], which returns the estimation of both the global and local (per-residue) quality of 3D protein models.

Hydrophobicity analysis of active sites was performed using the Kyte and Doolittle scale [47], as implemented in Chimera [48], and electrostatics were calculated using the PDB2PQR on-line facility, which performs an analysis of Poisson–Boltzmann electrostatics calculation [49].

For two protein targets, we also examined the vibrational frequency modes of *T. cruzi* and *H. sapiens* models by using the Elnemo [50] and WEBnm@ [51] on-line services.

Results and discussion

Table S1 in Supplementary Materials presents the 41 targets considered in this study, with their identification numbers in *T. cruzi* genome, amino acid lengths, family descriptions, PFAM domains, and predicted EC numbers. Our predicted EC numbers are consistent with the original genome function annotation list from GeneDB [52]. Using our methodology, we are able to cluster these 41 sequences identified by Capriles et al. as analogous or specific for *T. cruzi* versus *H. sapiens* into seven groups based on PFAM analysis and enzymatic activity. The group 4 with 8 sequences related to cruzipain (EC: 3.4.22.51) and the group 5 related to triacylglycerol lipase (EC: 3.1.1.3) with 1 sequence are eliminated after PFAM protein domain analysis because they include large domains (>250 amino acids) of unknown functions (e.g., DUF3586 in cruzipain) that lack any sequence homology with available PDB structures and cannot be modeled de novo in silico methods [53–55] at the moment. The remaining five groups are related to ATPase (EC: 3.6.3.6, two sequences), trypanothione disulfide reductase (EC: 1.8.1.12, two sequences), 2,4-dienoyl-CoA reductase (EC: 1.3.1.34, two sequences), leishmanolysin (EC: 3.4.24.36, 23 sequences), and cysteine synthase (EC: 2.5.1.47, two sequences). In the first three groups, as one sequence is a fragment of the other, we analyzed the sequence of larger size. Both cysteine synthase proteins have 332 residues and share 98 % sequence identity, and one sequence was modeled. The 23 sequences in the leishmanolysin group have between 516

and 621 amino acids, and as explained below, we modeled the structure of one sequence.

Table 1 summarizes the main results of our procedure for the five sequences considered for 3D structure and refinement. For each target, we give its predicted cell localization, whether we found significant homology hit with *H. sapiens* and distant organisms in terms of sequences and PDB structures, and finally the amino acid alignment between the *T. cruzi* and homologous structures. Among these five sequences, only the trypanothione disulfide reductase and the cysteine synthase sequences have clear homologs in *H. sapiens*. We now describe the details of our analysis for each protein and its implication for drug design.

ATPase (predicted EC: 3.6.3.6) The *T. cruzi* ATPase sequence of 898 amino acids has a clear homolog (36 % sequence identity) with an H⁺-transporting ATPase in *Arabidopsis thaliana* (PDB: 3B8C [56], region 7–885, EC 3.6.3.6). The *T. cruzi* sequence has 26 % sequence identity with an ATPase in *Oryctolagus cuniculus* (PDB: 3AR4 [57], region 1–994, EC: 3.6.3.8), and 23 % with a Na⁺/K⁺ transporting ATPase in *Mus musculus* (region 9–851, EC: 3.6.3.9). There is no clear homolog in *H. sapiens*, as there is at most 24 % sequence identity with a calcium transporting ATPase (region 7–872, EC: 3.6.3.8).

Using the X-ray structure from *A. thaliana* (PDB: 3B8C) as a template for *T. cruzi* and *H. sapiens* sequences, there is a Cα

Table 1 Cluster representatives

#	Pfam superfamily	TcruziDB Id.	L	Cell. Localization	Human homologues (EC number, %identity)	Non human homologues (EC number, %identity, i coverage)	Template	Cov.*
1	ATPase	TC00.1047053506649.20	898	Plasma - membrane	-	* <i>A. thaliana</i> (3.6.3.6, 36 %, 7–885) * <i>Mus musculus</i> (3.6.3.9 23 %, 5–861)	3AR4 (2.15A) <i>Oryctolagus cuniculus</i> ATPase1 3.6.3.8 26 %	4–855 / 1–994
2	Trypanothione disulfide reductase	TC00.1047053503555.30	492	Cytoplasmic	Glutathione reductase (1.8.1.7 36 %)	* <i>P. aeruginosa</i> (1.8.1.7 42 %, 2–476) * <i>N. punctiforme</i> (1.8.1.7 39 %, 2–469)	2WPF (1,90A) <i>Trypanosoma brucei</i> Trypanothione reductase 1.8.1.12 82 %	2–492 / 1–488
3	2,4-dienoyl-CoA reductase (NADPH)	TC00.1047053510303.210	717	Cytoplasmic	-	* <i>P. aeruginosa</i> (1.3.1.34 34 %, 1–716) * <i>A. fulgidus</i> (1.6.3.3 28 %, 6–716)	1PS9 (2,20A) <i>Escherichia coli</i> 2,4-dienoyl-CoA reductase 1.3.1.34 31 %	4–717 / 1–671
6	Leishmanolysin	TC00.1047053511211.90	567	Plasma - membrane	-	<i>C. elegans</i> (3.4.24.– 27 %, 69–603)	1LML (1,86A) <i>Leishmania major</i> Leishmanolysin 3.4.24.36 37 %	77–527 / 103–558
7	Cysteine synthase	TC00.1047053507165.50	332	Extracellular	cystathionine β-synthase (4.2.1.22, 33 %)	<i>A. thaliana</i> (2.5.1.47 55 %, 11–320)	3TBH (1,74A) <i>Leishmania donovani</i> cysteine synthase 2.5.1.47 73 %	1–320 / 2–319

For each cluster (#), only one representative sequence is retained, and we report its *T. cruzi* DB identifier (the sequence can be accessed at TcruziDB.org), Pfam superfamily, size (L) in amino acids, predicted cellular localization, identified human/non-human homolog sequences, the PDB template used for 3D modeling, and the limits of the aligned residues (coverage (Cov). – *T. cruzi*/template). For homolog sequences, we report the associated EC number, the sequence identity (%) and the limits of the homologous region in the *T. cruzi* sequence. For the template, we give the resolution, the organism, the Pfam superfamily, the EC. number and the sequence identity (%)

RMSD of 1.2 Å between the *T. cruzi* model and the template, and 1.0 Å between the *T. cruzi* and *H. sapiens* structures. Figure S1 shows the superposition of the *T. cruzi* model on the 3B8C template. However, among five residues important for ACP (phosphomethylphosphonic acid adenylate ester) binding in 3B8C, three residues are conserved (F385 (453), K408 (479), and R441 (522) in *T. cruzi* (*H. sapiens* amino acid numberings), but L443 in *T. cruzi* is changed to G517 in *H. sapiens*, i.e., changing the flexibility of the backbone, and more importantly, the K423 residue in *H. sapiens* is mutated to D359 in *T. cruzi*, i.e., changing the electrostatic feature of the active site. This makes the choice of this template questionable.

Because the X-ray structure from *A. thaliana* has a very low resolution (3.6 Å), the X-ray structure from *Oryctolagus cuniculus* (PDB: 3AR4) with a resolution of 2.15 Å is definitely a better structural template and the sequence alignment shown in Figure S2 was used for the homology modeling and MD refinement. Figure 2a shows the superposition of the *T. cruzi* structure on the 3AR4 template with a C α RMSD of 0.2 Å (free of any minimization) with the location of the six transmembrane α -helices (residues 50–99, 219–284, 614–644, 664–725, 749–799, and 823–854 in *T. cruzi*). Stability of the *T. cruzi* structure was assessed by 100 ns MD simulation at 300 K, which led to a mean RMSD of 5.3 Å from the minimized structure (Fig. 2d). Figure 2e shows the superposition of the most populated cluster representing 85 % of the conformational MD ensemble on the minimized structure. The minimized and all MD generated structures have good but lower quality than the structure arising from comparative modeling using Ramachandran and ModFOLD4 metrics (Fig. S3).

The *T. cruzi* sequence is predicted to be a P-type H⁺-exporting ATPase (EC: 3.6.3.6) found in the oxidative phosphorylation pathway and allowing proton transportation across the plasma membrane to generate the electrochemical potential gradient of cells. P-type ATPases catalyze the selective active transport of ions like H⁺, Na⁺, K⁺, Ca²⁺, Zn²⁺, and Cu²⁺ across diverse biological membrane systems, and several molecules have been shown to inhibit ATPase activity to different degrees [58]. For instance, artemisinin was found to inhibit growth of cultured *T. cruzi* at concentrations in the low micromolar range and inhibit Ca²⁺-dependent ATPase activity in *T. cruzi* membrane [59]. Also, miltefosin was found to inhibit the Na⁺-ATPase. This compound was, however, also found to inhibit the protein kinase C present in the plasma membrane of *T. cruzi* [60].

The alignment of the *T. cruzi* and *H. sapiens* structures (RMSD of 1.2 Å) using an *O. cuniculus* template suggests that the machinery for the dynamic and rotational motion of the cytoplasmic-compartment sector with respect to the membrane sector, driven by the H⁺ electrochemical potential gradient, is conserved. This is supported by normal

mode analysis of the two systems, free of ATP ligand, since the lowest frequency modes are almost identical (data not shown). Figure 2b zooms the active sites in the template, *T. cruzi* and *H. sapiens* and shows their hydrophobic and electrostatic surfaces. As seen, while the hydrophobic surface is slightly changed (bottom panel), the electrostatics potential (middle panel) is completely changed from *T. cruzi* to *H. sapiens*, being more negative in *H. sapiens* than in *T. cruzi*. Figure 2c shows that four residues of the active site crucial for ATP binding are strictly conserved among the template, *T. cruzi* and *H. sapiens*, i.e., F385, K408, R441, and L443 in *T. cruzi* amino acid numbering. As seen in Fig. 2b (bottom panel), the mutated residues between *T. cruzi* and (*H. sapiens*) are D359 (K423), L361 (T425), and T392 (M460). Interestingly, these three mutations observed in *T. cruzi* are strictly conserved in *T. rangeli*, *T. vivax*, *T. brucei*, *T. congolense*, *B. saltans*, *L. major*, *T. grayi*, and *G. theta*, indicating a strong evolutionary pressure of these positions in related species.

Whether these three mutations in or near the active site that change the electrostatic potential can govern the selectivity of the ATP-driven pumps for a drug in *T. cruzi* with respect to *H. sapiens*, and make the ATPase a suitable candidate, as suggested by the chemogenomics resource for neglected tropical diseases (TDR) [61] and cell surface proteome analysis of human-hosted *T. cruzi* life stages [62], remains to be explored.

Trypanothione reductase (predicted EC: 1.8.1.12) The 492-amino-acid trypano-thione reductase (TryR) in *T. cruzi* has a clear homolog in *Trypanosoma brucei* (PDB: 2WPF [63], EC: 1.8.1.12) with a sequence identity of 82 % over the amino acid region 1–488. The *T. cruzi* sequence also has 42 and 39 % sequence identities with two glutathione disulfide reductase with EC: 1.8.1.7 in *Pseudomonas aeruginosa* and *Nostoc punctiforme*, and 36 % sequence identity with a glutathione reductase (EC: 1.8.1.7) in *H. sapiens*. As a result, the X-ray structure (PDB: 2WPF), consisting of a homo-dimer with 2 × 488 amino acids, was selected as structural template and the sequence alignment shown in Figure S4 was used for the homology modeling of the *T. cruzi* and *H. sapiens* sequences. The superposition of the 3D models created by HAlign-Kbest for *H. sapiens* and *T. cruzi* on the template leads to small RMSD deviations of 1 and 0.6 Å, respectively (Fig. 3a, b). Analysis of the 100-ns MD simulation of the *T. cruzi* model at 300 K shows that the system is very stable (Fig. 3d), with one unique cluster and a mean RMSD of 2.5 Å from the full homo-dimer minimized structure (Fig. 3e). As for ATPase models, the MD generated structures of trypanothione reductase have lower quality than the structure obtained by comparative modeling as evaluated by the Ramachandran and ModFOLD4 metrics (Fig. S3). Under the harmonic approximation, the *T. cruzi* and *H. sapiens*

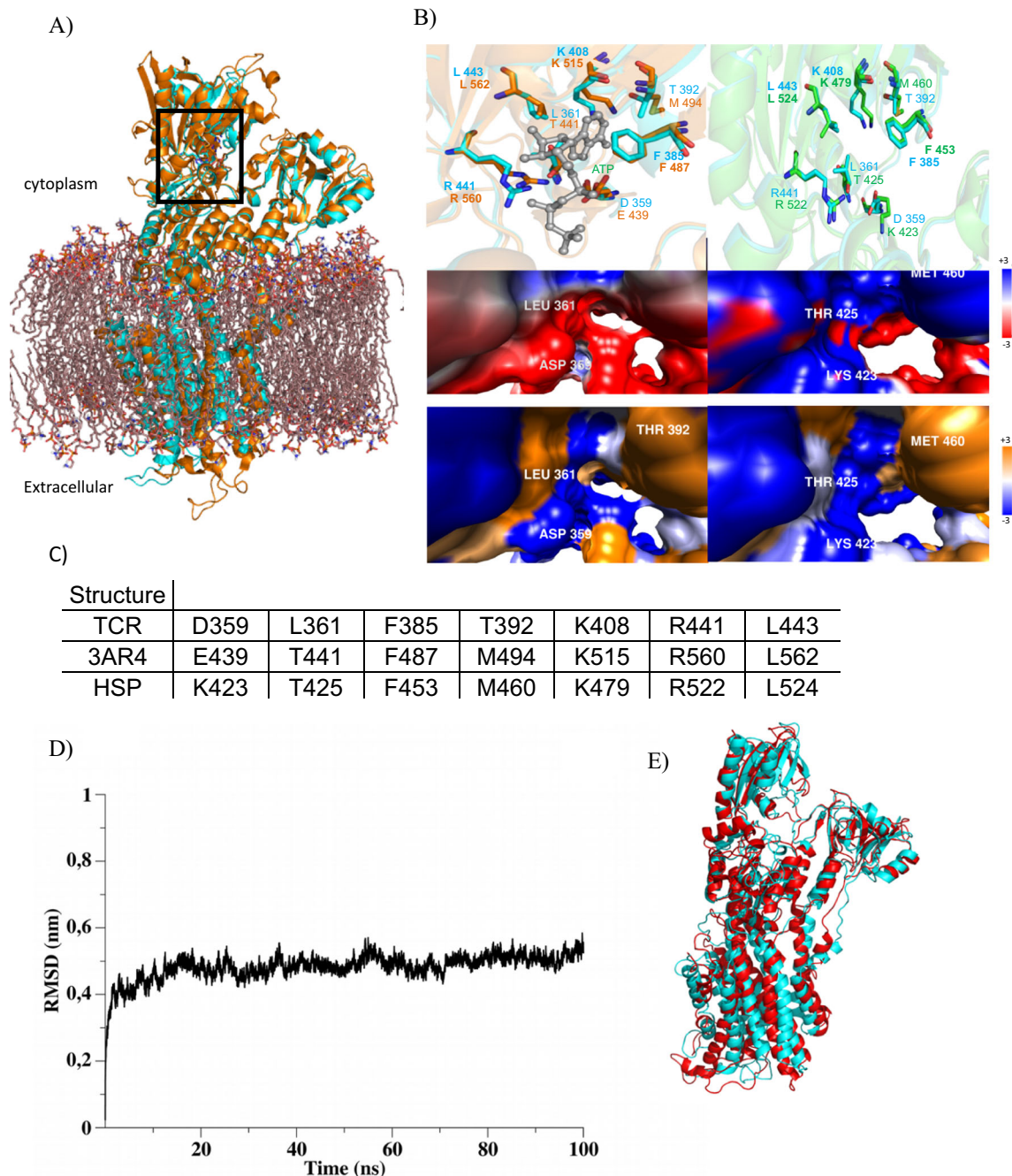


Fig. 2 **a** ATPase (EC 3.6.3.6). *T. cruzi* model (cyan), superimposed on *O. cuniculus* structure (PDB: 3AR4, orange) in explicit membrane environment. **b** *Top*: Focus on the active site using the template (orange) with ATP (gray) of *T. cruzi* model (cyan) in right and *H. sapiens* model (green). *Middle*: Electrostatic potential (Kb.T.ec⁻¹) – blue: negative, red: positive. *Bottom*: Hydrophobicity – orange:

hydrophobic, blue: hydrophilic. **c** Varying amino acids of the active site in the three species. **d** RMSD variation as the function of time (ns), relative to the initial structure. **e** Superimposition of the initial structure (cyan) and most populated cluster (red) over a 100-ns period molecular dynamics

structures show identical atomic fluctuations and deformation energies projected onto the calculated normal modes along the amino acid sequence (data not shown).

Trypanothione reductase is an essential enzyme of the unique trypanothione-based thiol metabolism of Trypanosomatidae. Trypanothione is an unusual form of

glutathione containing two molecules of glutathione joined by a spermidine (polyamine) linker. TryR is a flavoenzyme protein that catalyzes the reaction trypanothione + NADP⁺ ↔ trypanothione disulfide + NADPH + H⁺ [64]. Trypanosomatids lack both glutathione reductase and thioreductase, and as a result, TryR is the only connection between NADPH-

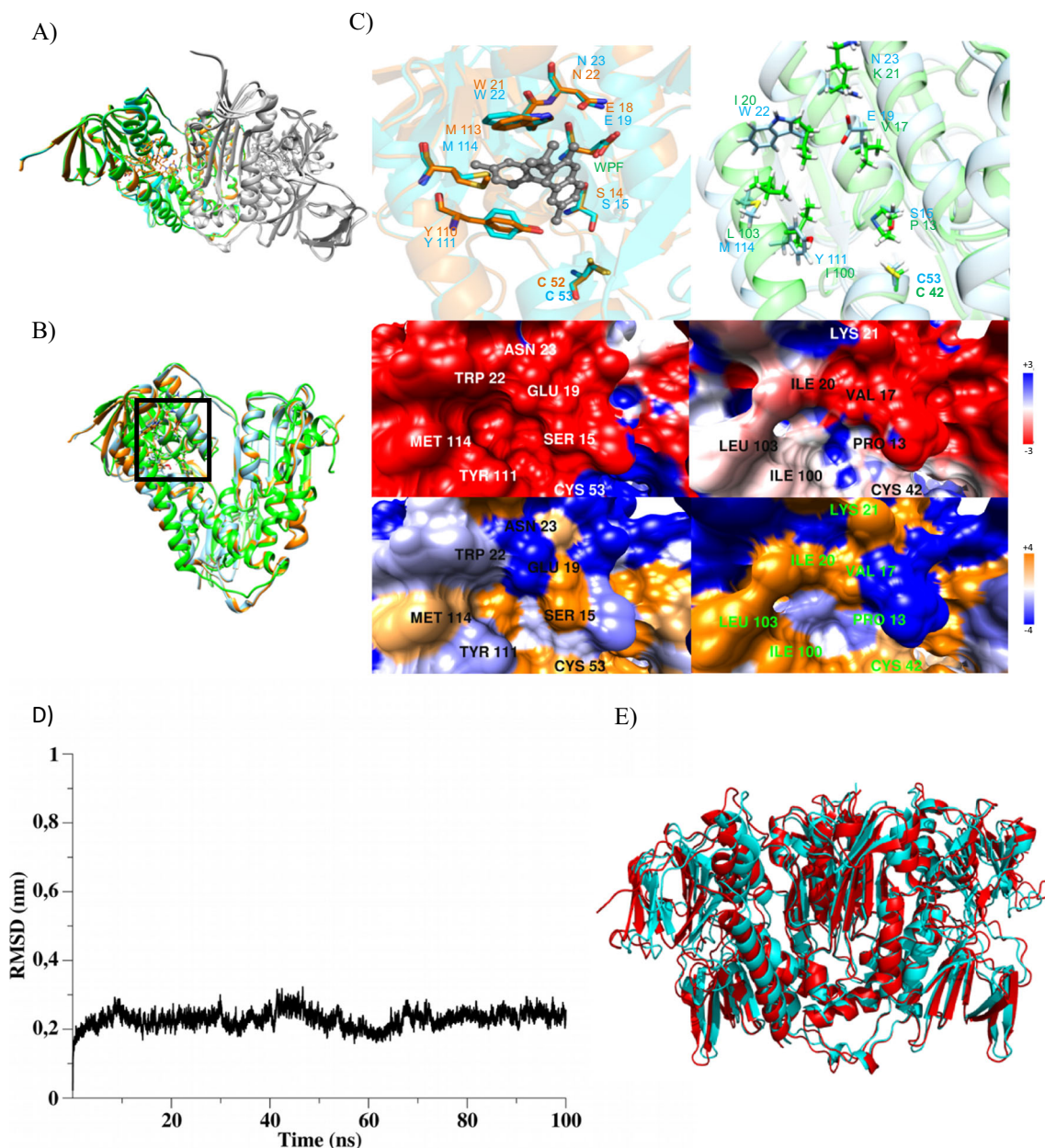


Fig. 3 **a** Dimer structure of the trypanothione disulfide reductase (EC 1.8.1.12). *T. cruzi* model (cyan), superimposed on *T. brucei* structure (PDB: 2WPF, orange) and *H. sapiens* model (green). Gray depicts the second monomer. **b** Focus on only chain A of *T. cruzi* and *H. sapiens* superimposed on template structure. **c** Focus on the active site. *Top left*: Template (orange) with WPF inhibitor (gray) and *T. cruzi* (cyan). *Top right*: *T. cruzi* and *H. sapiens* (green). *Middle*: Electrostatic potential

(Kb.T.ec⁻¹) – blue: negative, red: positive. *Left*: *T. cruzi*, *right*: *H. sapiens*. *Bottom*: Hydrophobicity, orange: hydrophobic, blue: hydrophilic. *Left*: *T. cruzi*, *right*: *H. sapiens*. **d** RMSD variation as the function of time (ns), relative to the initial structure. **e** Superimposition of the initial structure (cyan) and most populated cluster (red) over 100-ns period molecular dynamics

and thiol-based redox systems, the latter substituting for many antioxidant functions [65].

The active sites in *T. cruzi* and *T. brucei* have both hydrophilic and hydrophobic characteristics, while the active site in *H. sapiens* is highly hydrophobic (Fig. 3c, bottom). Further analysis of the active sites shows that all residues binding the WPF ligand (3,4-dihydroquinazoline inhibitor) are conserved between *T. cruzi*

and *T. brucei*, while only a single cysteine at positions 53 and 42 is conserved between *T. cruzi* and *H. sapiens* (Fig. 3c, top), changing therefore drastically the electrostatic potential surface (Fig. 3c, middle). Clearly, the low amino acid conservation of the catalytic site between Trypanosomatidae and humans, and the importance of this enzyme for protozoan parasites, make this protein target very attractive for drug development.

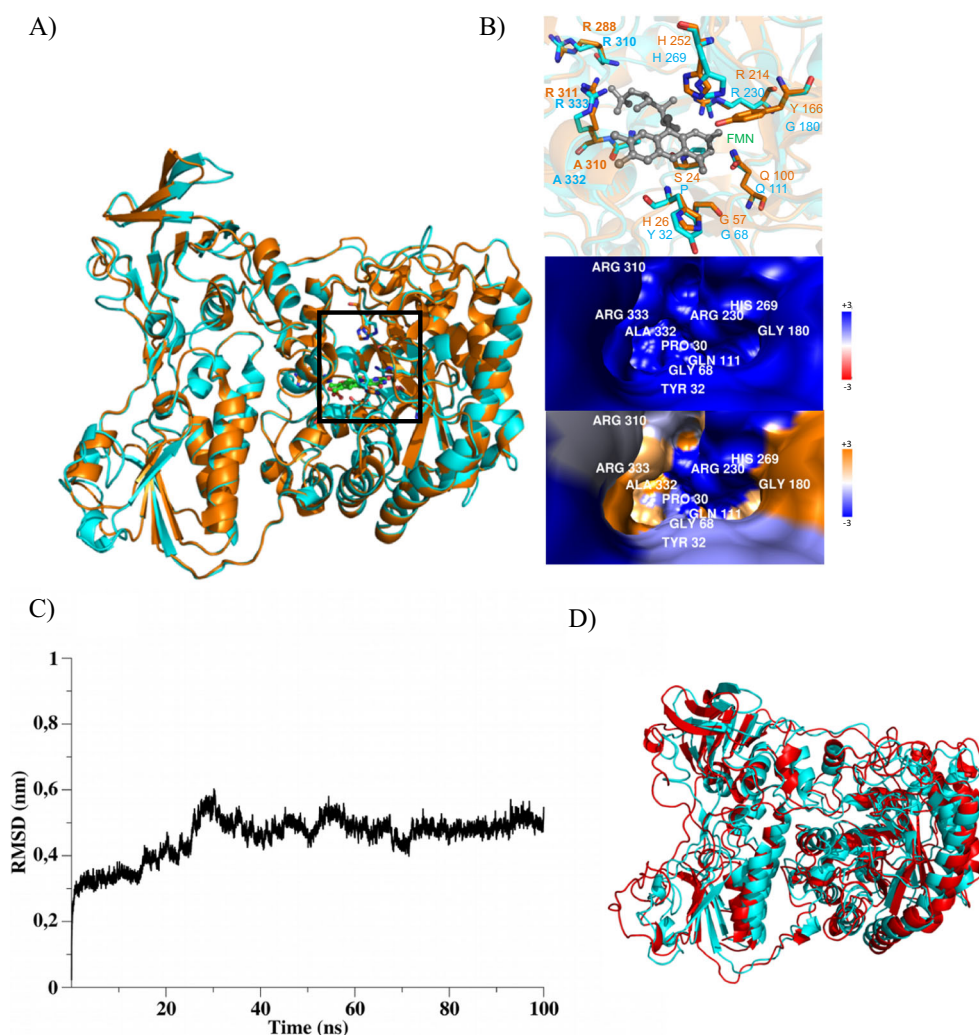
It has to be emphasized that previous studies have reported *T. cruzi* trypanothione reductase inhibitors such as the antimicrobial chlorhexidine and a piperidine derivative [66]. Recently, Lavorato et al. described the antitrypanosomal activity and cytotoxicity profile of 20 novel 1,3-bis(aryloxy)propan-2-amine derivatives as new candidates for further development as potential anti-trypanosomal agents [67]. Also, 82 novel TryR inhibitors down to the nM range were identified by using a combined in vitro and in silico screening approach [68].

2,4-dienoyl-CoA reductase (predicted EC: 1.3.1.34) The 717-amino-acid *T. cruzi* protein shares 31 % sequence identity with the 2,4-dienoyl-CoA reductase (DECRI1, EC: 1.3.1.34) from *E. coli* (region 1–671) with a structure determined by X-ray (PDB: 1PS9) at a 2.2 Å resolution [69]. The cellular localization and EC number in *T. cruzi* are predicted to be identical to those in *E. coli*. The DECRI1 enzyme participates in the beta-oxidation and metabolism of polyunsaturated fatty enoyl-CoA esters. DECRI1 catalyzes the reaction *trans*-2,3-

dehydroacyl-CoA + NADP⁺ ↔ *trans,trans*-2,3,4,5-tetrahydro acyl-CoA + NADPH + H⁺. The *T. cruzi* sequence has also 34 % sequence identity with a protein in *Pseudomonas aeruginosa* with the same EC number (1.3.1.34). Searching in the proteome of *H. sapiens*, the *T. cruzi* sequence shares only 19 % sequence identity with two proteins differing by EC numbers (4.2.1.22 and 4.1.1.17) and covering only the amino acid region 391–717.

Using the sequence alignment shown in Fig. S5 for homology modeling, the superposition of the *T. Cruzii* model (sequence 4–717) on the 1PS9 structure (sequence 1–671) has a RMSD of 0.3 Å prior to energy minimization (Fig. 4a). Analysis of the 100 ns MD simulation at 300 K shows that the *T. cruzi* model fluctuates about 5 Å from its minimized structure (Fig. 4c), and the large RMSD deviation comes from the long loop regions that need to be introduced from residues 666 to 717 (Fig. S5). Analysis of the MD-generated structures leads to two clusters, with the first cluster representing 90 % of the ensemble (Fig. 4d) and high quality as assessed by ModFOLD4 values (Fig. S3).

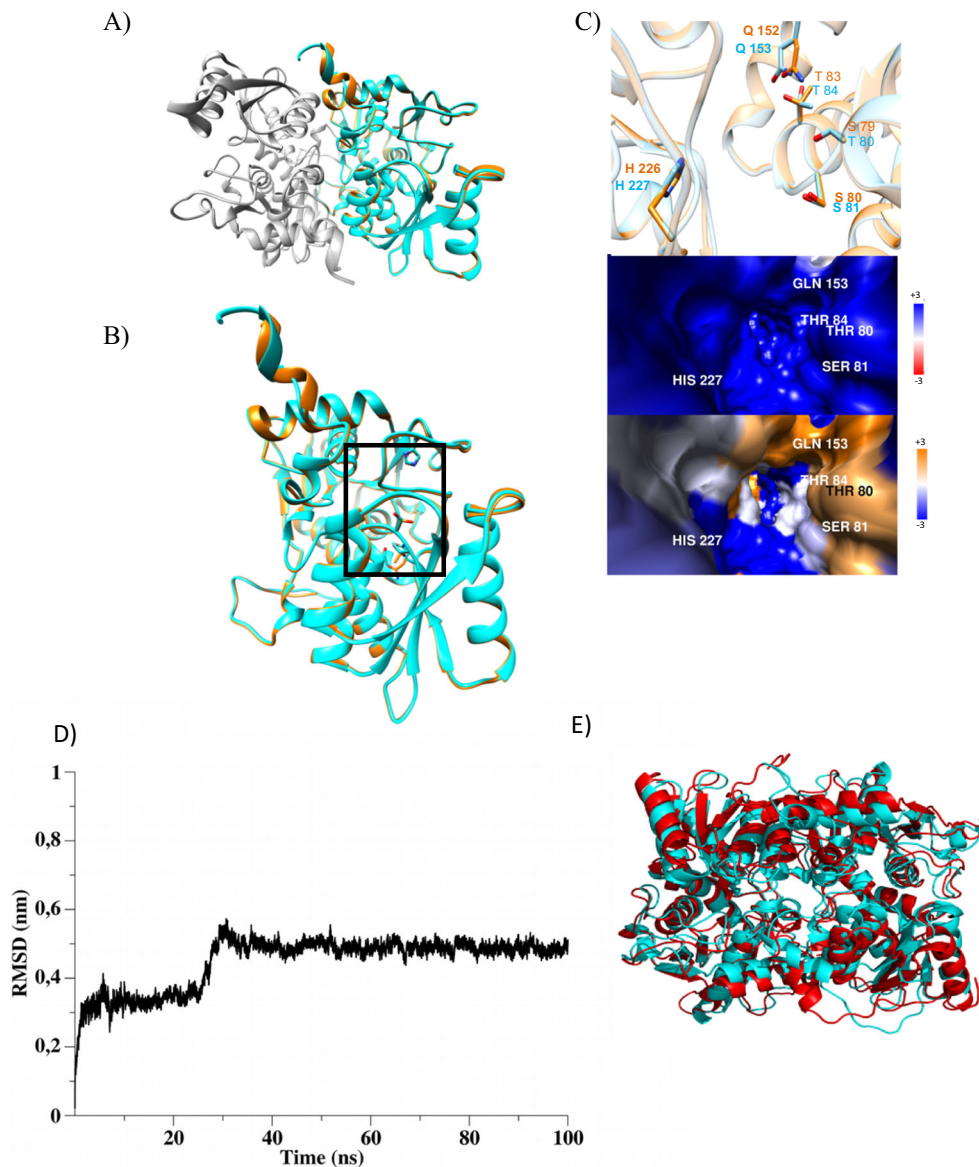
Fig. 4 **a** 2,4-dienoyl-CoA reductase (NADPH) (EC: 1.3.1.34). *T. cruzi* model (*cyan*), superimposed on *E. coli* structure (PDB: 1PS9, *orange*). **b** Focus on the active site. *Top*: Template (*orange*) with FMN inhibitor (*gray*) and *T. cruzi* (*cyan*). *Middle*: Electrostatic potential (Kb.T.ec⁻¹) – *blue*: negative, *red*: positive *Bottom*: Hydrophobicity – *orange*: hydrophobic, *blue*: hydrophilic. **c** RMSD variation as the function of time (ns), relative to the initial structure. **d** Superimposition of the initial structure (*cyan*) and most populated cluster (*red*) over a 100-ns period molecular dynamics



There are currently no chemical compounds associated to the *T. cruzi* gene, as reported in the TDR database. Most catalytic residues that allow FMN (Flavin mononucleotide) substrate binding are conserved between the *T. cruzi* and *E. coli* sequences (G57, Q100, R214, R288, A330, and R331 in *E. coli* vs. G68, Q111, R230, R310, A332, and R333 in *T. cruzi*) and lead to a positive electrostatic field (Fig. 4b). The two residues S24 and H26 in *E. coli* are changed into P30 and Y32 in *T. cruzi* (Fig. 4b), but these two mutations are not sufficient to alter the polar features of the binding site, located in an inner enzyme region (Fig. 4c). Interestingly, humans have an enzyme for the same function, but with a completely 3D structure (homotetramer, each of 302 amino acids, PDB 1W6U), and a distinct catalytic center [70], making DECR1 very suitable for the design of drugs specific to *T. cruzi*.

Cysteine synthase (predicted EC: 2.5.1.47) The *T. cruzi* sequence of 332 amino acids has a clear homolog (73 % sequence identity) with a cysteine synthase of 324 amino acids from *Leishmania donovani* (region: 2–319, EC: 2.5.1.47, PDB: 3TBH at a 1.74-Å resolution [71]). The *T. cruzi* sequence also shares 55 % sequence identity with a cysteine synthase in *Arabidopsis thaliana* (EC: 2.1.1.47) and 33 % sequence identity with a human protein (region 6–332 with EC: 4.2.1.22, a cystathionine β -synthase). As a result, the homodimer structure from *Leishmania donovani* was selected as structural template and the sequence alignment shown in Figure S6 was used for the homology modeling of the *T. cruzi* protein. Figure 5a and b show the superposition of the predicted HHalign-Kbest model on the *Leishmania donovani* structure with a RMSD of 0.2 Å. Analysis of the 100-ns MD simulation at 300 K of the homodimer of *T. cruzi* shows an

Fig. 5 **a** Dimer structure of the cysteine synthase (EC: 2.5.1.47). *T. cruzi* model (cyan), superimposed on *L. donovani* structure (PDB: 3TBH, orange). Gray depicts the second monomer. **b** Focus on only chain A of *T. cruzi* superimposed on template structure. **c** Focus on the active site. *Top*: Template (orange) and *T. cruzi* (cyan). *Middle*: Electrostatic potential (Kb.T.ec^{-1}) – blue: negative, red: positive *Bottom*: Hydrophobicity – orange: hydrophobic, blue: hydrophilic. **d** RMSD variation as the function of time (ns), relative to the initial structure. **e** Superimposition of the initial structure (cyan) and most populated cluster (red) over 100-ns period molecular dynamics



averaged RMSD deviation of 4.5 Å with respect to the minimized structure (Fig. 5d), with the most populated cluster, representing 70 % of the conformational ensemble (Fig. 5e) and with high quality as measured by the PSVS metrics (Fig. S3).

There is currently no chemical compound associated to the *T. cruzi* gene, as reported in the TDR database [61]. Cysteine synthase is a key constituent for the survival of trypanosomatids, and thus vital to the survival of *T. cruzi* in its vertebrate hosts. It constitutes one of the key pathways in the parasite defense against oxidative stress. Two different routes for cysteine biosynthesis have been described: reverse-transsulfuration (RTS) and de novo pathways. RTS has been demonstrated in fungi and mammals and includes the complete process leading to cysteine from methionine via the intermediary formation of cystathionine. These reactions are catalyzed by two enzymes: CβS (cystathionine β-synthase) and CGL (cystathionine γ-lyase). The de novo pathway, catalyzed by two steps starting with serine acetyltransferase to form *O*-acetylserine from L-serine and acetyl-coenzyme A, is found in plants, bacteria, and some protozoa, such as *Entamoeba histolytica*, *Leishmania major*, and *Leishmania donovani*, but is absent in mammals [72].

T. cruzi sequence is a clear homolog of an *O*-acetylserine sulfhydrylase from *L. donovani*. The active site with the Asp-Gly-Ser-Gly-Ile ligand is fully conserved in terms of amino acids and hydrophilic character from *T. cruzi* to *L. donovani*, with only a single amino acid change (T80 changed to S79) (Fig. 5c). As shown in Fig. S6, the residues involved in the active sites with cysteine synthase and cystathionine β-synthase activities are conserved in *T. cruzi*, *L. donovani*, and *H. sapiens* and the differences in the physico-chemical properties of the residues flanking the active sites are minor (Fig. S7). Clearly, the high conservation of the active site and the presence of this enzyme in humans do not make this protein a good target for drug design specific to *T. cruzi*. However, there have been several reports of inhibitor screening and development against this molecule from different organisms such as *Entamoeba histolytica* and *Mycobacterium tuberculosis*, using an off-catalytic site strategy, and it was suggested that among the open form, the intermediate state and the closed form of the enzyme, the intermediate state of the enzyme might be the ideal target for the design of very effective high-affinity inhibitors [73].

Leishmanolysin (predicted EC: 3.4.24.36) The superfamily of leishmanolysin consists of 23 *T. cruzi* sequences with 516–621 amino acids and cross-sequence identities varying between 37 and 96 %. Multiple sequence alignment using ClustalW [74] shows that 22 sequences (except the smallest of 516 amino acids with a truncated N-terminus) have a common domain of 565 amino acids with > 30 % sequence identity and, among these, there are seven sequences of 567 amino

acids and seven sequences with > 570 amino acids displaying either longer N-terminus (two sequences), longer C-terminus (four sequences), or longer N- and C-termini (1 sequence). The N-terminus can be extended by up to 56 amino acids.

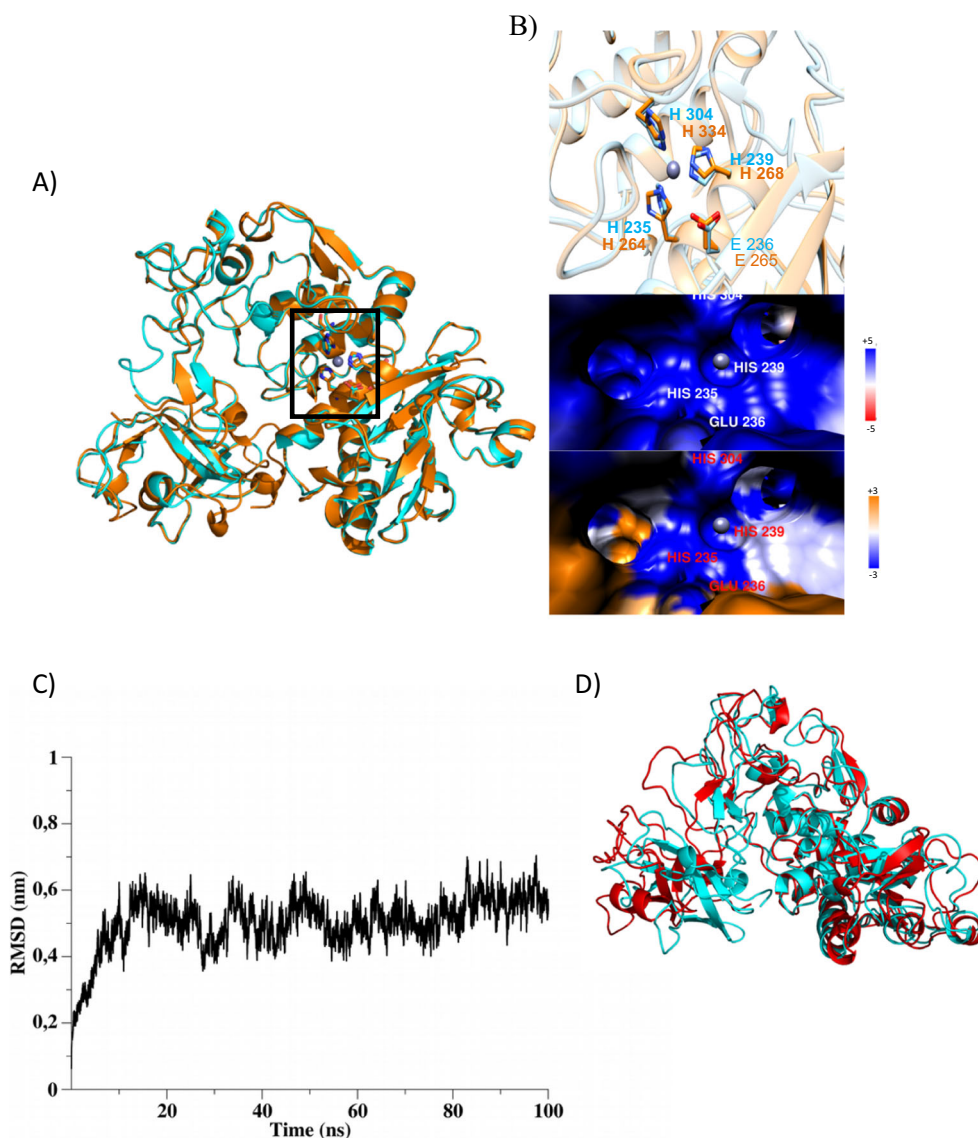
Taking the TC00.1047053511211.90 sequence of 567 amino acids as a representative of the 23 *T. cruzi* sequences, the region 77–527 shares 37 % sequence identity with a leishmanolysin from *Leishmania major* (region: 103–558, EC: 3.4.24.36, PDB: 1LML at a 1.8-Å resolution [75]). This sequence has at most a sequence identity of 25 % with a protein from *H. sapiens* with EC = 4.2.1.22 (region 69–603), but the alignment shows a gap percentage of 16 %. Based on this result, the X-ray structure 1LML was selected as structural template of the region 77–527, and it remained to determine whether the missing residues 1–76 and 528–567 could be modeled by other means. The N-terminus cannot be easily modeled since leishmanolysin is a membrane-bound zinc proteinase, active in situ, but the exact anchorage of the N-terminus to the membrane is unknown. Using the best secondary structure prediction methods, i.e., C8-Scorpion and Porter 4.0, the region 528–567 is predicted to be free of any secondary structure. Similarly, the C-terminus region is predicted as disordered using PrDOS, preventing therefore its 3D modeling. Note that this intrinsic disorder of the C-terminus is shared by the other 21 sequences.

The sequence alignment shown in Figure S8 was used for the homology modeling, and the superposition of the *T. cruzi* structure on its template (with a RMSD of 0.4 Å) is shown in given in Fig. 6a. Fig. 6B zooms the active site in the template and *T. cruzi* and shows the high hydrophilic character and negative electrostatic field of the active site. The three histidines essential for Zn²⁺ binding are conserved in both species, as well as the residues Glu and Met. Stability of the *T. cruzi* model complexed by Zn²⁺ was assessed by 100-ns MD simulation at 300 K. It is found that the system has an average RMSD of 5.5 Å from the minimized structure (Fig. 6c), the most populated cluster representing 50 % of the conformational MD (Fig. 6d) having very good quality using PSVS and ModFOLD4 metrics (Fig. S3).

Leishmanolysin is the predominant glycoprotein surface antigen of promastigotes of various species of *Leishmania*. The crystal structure has shown three domains, two of which had novel folds at that time, and the N-terminal domain has a similar structure to the catalytic modules of zinc proteases, revealing therefore that leishmanolysin is a member of the metzincin class of zinc proteinases [75]. The similarity of the active site structure and amino acid composition in *T. cruzi* model to previously well-characterized metzincin-class zinc proteinases suggests to test known zinc-metalloproteinase inhibitors, such as 1,10-phenanthroline or peptide-based inhibitors [76].

We would like to emphasize that for the five targets, we have evaluated the stability of the residues in the binding site

Fig. 6 **a** Leishmanolysin (EC: 3.4.24.36). *T. cruzi* model (cyan), superimposed on *L. major* structure (PDB: 1LML, orange). **b** Focus on the active site. *Top*: Template (orange) with Zinc Ion (gray) and *T. cruzi* (cyan). *Middle*: Electrostatic potential (Kb.T.ec^{-1}) – blue: negative, red: positive, gray: Zn ion. *Bottom*: Hydrophobicity – orange: hydrophobic, blue: hydrophilic. **c** RMSD variation as the function of time (ns), relative to the initial structure. **d** Superimposition of the initial structure (cyan) and most populated cluster (red) over 100-ns period molecular dynamics



in bound state by calculating the $C\alpha$ RMS deviations of the active site residues in the minimized structure and the most populated (best) MD cluster with respect to their starting positions. All residues in the active site remain at their native positions although the ligands (ATP in ATPase, WPF in trypanothione reductase, FMN in NADPH and Asp-Gly-Ser-Gly-Ile in cysteine synthase) are absent in the simulations, with most residues deviating by less than 0.3 \AA and at most by 1 \AA . Also we have checked that we have reasonable contacts between the designed protein receptors and their natural ligands by analyzing the $C\alpha$ and $C\beta$ deviations of the active site residues in the *T. cruzi* minimized structure and most populated MD cluster with respect to the PDB structure used for homology without any minimization. Small values of $0.1\text{--}0.2 \text{ \AA}$ indicate that the contact maps for the designed *T. cruzi* minimized proteins binding to their corresponding ligands are very similar to the contacts maps of the homolog proteins. The

largest values are expected and observed for the designed *T. cruzi* most populated MD cluster, but all remain below 1.5 \AA . Finally, we have verified that the dynamics does not change the hydrophobic/hydrophilic character and the electrostatic potential of the active sites; these properties being essentially conserved from the minimized structure to the most populated MD cluster.

Conclusions

The identification of specific *T. cruzi* protein targets for drug development is desirable. In recent years, drug discovery for Chagas disease has evolved from limited compound testing in manual assays to sophisticated in vitro and in vivo assays. Drug discovery must consider that *T. cruzi* is a genetically diverse organism with phenotypically variations, and inhibitor

susceptibility is a well-established phenotypic variation both in vitro and in vivo [77]. In the next 5 years, 200,000 people living with Chagas disease will die from heart disease and related complications, and the BENEFIT trial pursues a global initiative of diagnosis, treatment, and research [78].

In this study, by following a complex bioinformatics procedure, we have proposed structural models for five *T. cruzi* enzymes with ATPase, trypanothione reductase, 2,4-dienoyl-CoA reductase, cysteine synthase, and leishmanolysin functions. Based on sequence and active-site similarity with several organisms and notably *H. sapiens*, we show that these five protein targets, except cysteine synthase, are very attractive for drug development against *T. cruzi*. This development based on computer simulations is not straightforward, however, because it is important to explore a much larger dynamic time-scale of the enzymes and enzymes/inhibitors using advanced simulation methods so as to evaluate the correct binding pose [79, 80] and time-consuming free energy procedures in order to predict accurate binding affinity that goes beyond simple scores [81–83].

Acknowledgments C.L. thanks CAPES/CsF and Science without Borders PhD grant program from Brazil, BEX Process 11917/13-0.

Compliance with ethical standards

Conflict of interest The authors declare no financial conflicts of interest regarding the publication of this manuscript.

References

1. WHO (World Health Organization). http://www.who.int/topics/chagas_disease/en/
2. Noireau F, Diosque P, Jansen AM (2009) *Trypanosoma cruzi*: adaptation to its vectors and its hosts. *Vet Res* 40(2):1–23. doi:10.1051/vetres/2009009
3. Viotti R, Vigliano C, Lococo B, Alvarez MG, Petti M, Bertocchi G, Armenti A (2009) Side effects of benznidazole as treatment in chronic Chagas disease: fears and realities. *Expert Rev Anti-Infect* 7(2):157–163. doi:10.1586/14787210.7.2.157
4. Wilkinson SR, Taylor MC, Horn D, Kelly JM, Cheeseman I (2008) A mechanism for cross-resistance to nifurtimox and benznidazole in trypanosomes. *Proc Natl Acad Sci U S A* 105(13):5022–5027
5. De Castro SL, Batista DGJ, Batista MM, Batista W, Dailyri A, Souza EM, Menna-Barreto RFS, Oliveira GM, Salomão K, Silva CF, Silva PB, Soeiro MNC (2011) Experimental chemotherapy for Chagas disease: a morphological, biochemical and proteomic overview of potential *Trypanosoma cruzi* targets of amidines derivatives and naphthoquinones. *Mol Biol Int* 306928
6. Galperin MY, Walker DR, Koonin EV (1998) Analogous enzymes: independent inventions in enzyme evolution. *Genome Res* 8:779–790
7. Guimarães AC, Otto TD, Alves-Ferreira M, Miranda AB, Degraive WM (2008) In silico reconstruction of the amino acid metabolic pathways of *Trypanosoma cruzi*. *Genet Mol Res* 7:872–882
8. Alves-Ferreira M, Guimarães ACR, Capriles PVSZ, Dardenne LE, Degraive WM (2009) A new approach for potential drug target discovery through in silico metabolic pathway analysis using *Trypanosoma cruzi* genome information. *Mem Inst Oswaldo Cruz* 104:1100–1110
9. Merlino A, Vieites M, Gambino D, Coitiño EL (2014) Homology modeling of *T. cruzi* and *L. major* NADH-dependent fumarate reductases: ligand docking, molecular dynamics validation, and insights on their binding modes. *J Mol Graph Model* 48:47–59. doi:10.1016/j.jmgm.2013.12.001
10. Gomez Barroso JA, Aguilar CF (2014) Chagas disease: a homology model for the three-dimensional structure of the *Trypanosoma cruzi* ribosomal P0 antigenic protein. *Eur Biophys J* 43(8–9):361–366
11. Sacconay L, Angleviel M, Randazzo GM, Marçal Ferreira Queiroz M, Ferreira Queiroz E, Wolfender J-L et al (2014) Computational studies on sirtuins from *Trypanosoma cruzi*: structures, conformations and interactions with phytochemicals. *PLoS Negl Trop Dis* 8(2):e2689
12. Soysa R, Venselaar H, Poston J, Ullman B, Hasne M-P (2013) *Biochem J* 452(3):423–432. doi:10.1042/BJ20130350
13. Otto TD, Guimarães ACR, Degraive WM, Miranda AB (2008) AnEnPi: identification and annotation of analogous enzymes. *BMC Bioinformatics* 9:544
14. Capriles PVSZ, Guimarães ACR, Otto TD, Miranda AB, Dardenne LE, Degraive WM (2010) Structural modelling and comparative analysis of homologous, analogous and specific proteins from *Trypanosoma cruzi* versus *Homo sapiens*: putative drug targets for Chagas' disease treatment. *BMC Genomics* 11:610
15. Altschul S, Madden T, Schäffer A, Zhang J, Zhang Z, Miller W, Lipman D (1997) Gapped blast and PSI-Blast: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402
16. Soding J (2015) Protein homology detection by HMM-HMM comparison. *Bioinformatics* 21(7):951–960. doi:10.1093/bioinformatics/bti125
17. Taylor TJ, Tai CH, Huang YJ, Block J, Bai H, Kryshchak O, Montelione GT, Lee B (2014) Definition and classification of evaluation units for CASP10. *Proteins* 82(Suppl 2):14–25. doi:10.1002/prot.24434
18. Finn RD et al (2014) Pfam: the protein families database. *Nucleic Acids Res* 42(D1):D222–D230
19. Chou K-C, Shen H-B (2010) A New method for predicting the subcellular localization of eukaryotic proteins with both single and multiple sites: Euk-mPLoc 2.0. *PLoS ONE* 5:e9931
20. Chou K-C, Shen H-B (2008) Cell-PLoc: a package of web-servers for predicting subcellular localization of proteins in various organisms. *Nat Protoc* 3:153–162
21. Altman T et al (2013) A systematic comparison of the MetaCyc and KEGG pathway databases. *BMC Bioinformatics* 14:112
22. Kanehisa M, Goto S (2000) KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28:27–30
23. Okuda S et al (2008) KEGG Atlas mapping for global analysis of metabolic pathways. *Nucleic Acids Res* 36(Web Server issue):W423–W426
24. Caspi R et al (2012) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res* 40(Database issue):D742–D753
25. Söding J, Biegert A, Lupas AN (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* 33(suppl 2):W244–W248
26. Néron B et al (2009) Mobylye: a new full web bioinformatics framework. *Bioinformatics* 25:3005–3011
27. Yu J et al (2015) HHalign-Kbest: exploring sub-optimal alignments for homology comparative modeling. *Bioinformatics*. doi:10.1093/bioinformatics/btv441

28. Eswar N et al (2006) Comparative protein structure modeling using modeller. *Curr Protoc Bioinformatics* 54:5.6.1–5.6.37
29. Yaseen A, Li Y (2014) Template-based C8-SCORPION: a protein 8-state secondary structure prediction method using structural information and context-based features. *BMC Bioinformatics* 15(Suppl 8):S3. doi:10.1186/1471-2105-15-S8-S3
30. Mirabello C, Pollastri G (2013) Porter, PaleAle 4.0: high-accuracy prediction of protein secondary structure and relative solvent accessibility. *Bioinformatics* 29(16):2056–2058. doi:10.1093/bioinformatics/btt344
31. Ishida T, Kinoshita K (2007) PrDOS: prediction of disordered protein regions from amino acid sequence. *Nucleic Acids Res* 35(Web Server issue):W460–W464
32. Piana S, Lindorff-Larsen K, Shaw DE (2011) How robust are protein folding simulations with respect to force field parameterization? *Biophys J* 100(9):L47–L49. doi:10.1016/j.bpj.2011.03.051
33. Lindorff-Larsen K, Piana S, Dror RO, Shaw DE (2011) *Science* 334(6055):517–520. doi:10.1126/science.1208351
34. Hoang Viet M, Derreumaux P, Nguyen PH (2015) *J Chem Phys* 143(2):021101. doi:10.1063/1.4926535
35. Zhang T, Nguyen PH, Nasica-Labouze J, Mu Y, Derreumaux P (2015) *J Phys Chem B* 119(23):6941–6951. doi:10.1021/acs.jpcc.5b03381
36. Jorgensen WL et al (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79(2):926–935
37. Ng HW, Laughton CA, Doughty SW (2014) Molecular dynamics simulations of the adenosine A2a receptor in POPC and POPE lipid bilayers: effects of membrane on protein behavior. *J Chem Inf Model* 54:573–581
38. Derreumaux P, Zhang GH, Schlick T et al (1994) A truncated Newton minimizer adapted for CHARMM and biomolecular applications. *J Comput Chem* 15(5):532–552
39. Nosé S, Klein ML (1983) Constant pressure molecular dynamics for molecular systems. *Mol Phys* 50(5):1055–1076
40. Bussi G, Donadio D, Parrinello M (2007) Canonical sampling through velocity rescaling. *J Chem Phys* 126(1):014101
41. Pronk S et al (2013) GROMACS 4.5: a high-throughput and highly parallel open-source molecular simulation toolkit. *Bioinformatics* 29:845–854
42. Hess B et al (1997) LINCS: a linear constraint solver for molecular simulations. *J Comput Chem* 18(12):1463–1472
43. Darden T, York D, Pedersen L (1993) Particle mesh Ewald: an $N \log(N)$ method for Ewald sums in large systems. *J Chem Phys* 98:10089e10092
44. Bhattacharya A, Tejero R, Montelione GT (2007) Evaluating protein structures determined by structural genomics consortia. *Proteins Struct Funct Bioinformatics* 66(4):778–795
45. Kryshchafovich A, Moulton J, Bales P, Bazan JF, Biasini M, Burgin A, Chen C, Cochran FV, Craig TK, Das R, Fass D, Garcia-Doval C, Herzberg O, Lorimer D, Luecke H, Ma X, Nelson DC, van Raaij MJ, Rohwer F, Segall A, Seguritan V, Zeth K, Schwede T (2014) *Proteins* 82(Suppl 2):26–42. doi:10.1002/prot.24489
46. Roche DB, Buenavista MT, McGuffin LJ (2014) *Methods Mol Biol* 1137:83–103. doi:10.1007/978-1-4939-0366-5_7
47. Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 157(1):105–132
48. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem* 25(13):1605–1612
49. 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
50. Suhre K, Sanejouand YH (2004) Elnemo: a normal mode web server for protein movement analysis and the generation of templates for molecular replacement. *Nucleic Acids Res* 32(Web Server issue):W610–W614
51. Tiwari SP, Fuglebakk E, Hollup SM, Skjærven L, Cragolini T, Grindhaug SH, Tekle KM, Reuter N (2014) WEBnm@ v2.0: web server and services for comparing protein flexibility. *BMC Bioinformatics* 15:427
52. El-Sayed NM, Myler PJ, Bartholomeu DC, Nilsson D, Aggarwal G, Tran AN, Ghedin E, Worthey EA, Delcher AL, Blandin G et al (2005) The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. *Science* 309(5733):409–415
53. Yang J, Zhang Y (2015) Protein structure and function prediction using I-TASSER. *Curr Protoc Bioinformatics* 52:5.8.1–15
54. Leaver-Fay A, Tyka M, Lewis SM, Lange OF, Thompson J, Jacak R, Kaufman K, Renfrew PD, Smith CA, Sheffler W, Davis IW, Cooper S, Treuille A, Mandell DJ, Richter F, Ban YE, Fleishman SJ, Corn JE, Kim DE, Lyskov S, Beronzo M, Mentzer S, Popović Z, Havranek JJ, Karanicolas J, Das R, Meiler J, Kortemme T, Gray JJ, Kuhlman B, Baker D, Bradley P (2011) *Methods Enzymol* 487:545–574. doi:10.1016/B978-0-12-381270-4.00019-6
55. Shen Y, Maupetit J, Derreumaux P, Tufféry P (2014) Improved PEP-FOLD approach for peptide and miniprotein structure prediction. *J Chem Theory Comput* 10(10):4745–4758
56. Pedersen BP, Buch-Pedersen MJ, Morth JP, Palmgren MG, Nissen P (2007) Crystal structure of the plasma membrane proton pump. *Nature* 450(7172):1111–1114
57. Toyoshima C, Yonekura S, Tsueda J, Iwasawa S (2011) Trinitrophenyl derivatives bind differently from parent adenine nucleotides to Ca^{2+} -ATPase in the absence of Ca^{2+} . *Proc Natl Acad Sci U S A* 108:1833–1838
58. Yatime L, Buch-Pedersen MJ, Musgaard M, Morth JP, Lund Winther AM, Pedersen BP, Olesen C, Andersen JP, Vilsen B, Schiøtt B, Palmgren MG, Møller JV, Nissen P, Fedosova N (2009) P-type ATPases as drug targets: tools for medicine and science. *Biochim Biophys Acta* 1787(4):207–220
59. Mishina YV, Krishna S, Haynes RK, Meade JC (2007) Artemisinin inhibits *Trypanosoma cruzi* and *Trypanosoma brucei* rhodesiense in vitro growth. *Antimicrob Agents Chemother* 51(5):1852–1854
60. Saraiva VB, Wengert M, Gomes-Quintana E, Heise N, Caruso-Neves C (2009) Na^{+} -ATPase and protein kinase C are targets to 1-*O*-hexadecylphosphocoline (miltefosine) in *Trypanosoma cruzi*. *Arch Biochem Biophys* 481(1):65–71. doi:10.1016/j.abb.2008.10.018
61. Magariños MP, Carmona SJ, Crowther GJ, Ralph SA, Roos DS, Shanmugam D, Van Voorhis WC, Agüero F (2012) TDR targets: a chemogenomics resource for neglected diseases. *Nucleic Acids Res* 40(Database issue):D1118–D1127
62. Queiroz RML, Charneau S, Bastos IMD, Santana JM, Sousa MV, Roepstorff P, Ricart CAO (2014) Cell surface proteome analysis of human-hosted *Trypanosoma cruzi* life stages. *J Proteome Res* 13(8):3530–3541
63. Patterson S et al (2011) Dihydroquinazolines as a novel class of *Trypanosoma brucei* trypanothione reductase inhibitors: discovery, synthesis, and characterization of their binding mode by protein crystallography. *J Med Chem* 54:6514–6530
64. Krieger S, Schwarz W, Ariyanayagam MR, Fairlamb AH, Krauth-Siegel RL, Clayton C (2000) Trypanosomes lacking trypanothione reductase are avirulent and show increased sensitivity to oxidative stress. *Mol Microbiol* 35:542–552
65. Krauth-Siegel RL, Meiering SK, Schmidt H (2003) The parasite-specific trypanothione metabolism of *Trypanosoma* and *Leishmania*. *Biol Chem* 384:539–549
66. Meiering S, Inhoff O, Mies J, Vincek A, Garcia G, Kramer B, Dormeyer M, Krauth-Siegel RL (2005) Inhibitors of

- Trypanosoma cruzi* trypanothione reductase revealed by virtual screening and parallel synthesis. *J Med Chem* 48(15):4793–4802
67. Lavorato SN, Sales Júnior PA, Murta SM, Romanha AJ, Alves RJ (2015) In vitro activity of 1,3-bisaryloxypropanamines against *Trypanosoma cruzi*-infected L929 cultures. *Mem Inst Oswaldo Cruz* 110(4):566–568
 68. Beig M, Oellien F, Garoff L, Noack S, Krauth-Siegel RL, Selzer PM (2015) Trypanothione reductase: a target protein for a combined in vitro and in silico screening approach. *PLoS Negl Trop Dis* 9(6):e0003773. doi:10.1371/journal.pntd.0003773
 69. Hubbard PA et al (2003) The crystal structure and reaction mechanism of *Escherichia coli* 2,4-dienoyl-CoA reductase. *J Biol Chem* 278:37553–37560
 70. Alphey MS, Yu W, Byres E, Li D, Hunter WN (2005) Structure and reactivity of human mitochondrial 2,4-dienoyl-CoA reductase: enzyme-ligand interactions in a distinctive short-chain reductase active site. *J Biol Chem* 280(4):3068–3077
 71. Rose PW, Prlić A, Bi C, Bluhm WF, Christie CH, Dutta S, Green RK, Goodsell DS, Westbrook JD, Woo J, Young J, Zardecki C, Berman HM, Bourne PE, Burley SK (2015) The RCSB protein data bank: views of structural biology for basic and applied research and education. *Nucleic Acids Res* 43(Database issue):D345–D356. doi:10.1093/nar/gku1214
 72. Romero I, Téllez J, Yamanaka LE, Steindel M, Romanha AJ, Grisard EC (2014) Transsulfuration is an active pathway for cysteine biosynthesis in *Trypanosoma rangeli*. *Parasit Vectors* 7:197. doi:10.1186/1756-3305-7-197
 73. Mazumder M, Gourinath S (2016) Structure-based design of inhibitors of the crucial cysteine biosynthetic pathway enzyme *O*-acetyl serine sulfhydrylase. *Curr Top Med Chem* 16(9):948–959
 74. Larkin MA C, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Bioinformatics 23(21):2947–2948. doi:10.1093/bioinformatics/btm404
 75. Schlagenhauf E et al (1998) The crystal structure of the *Leishmania major* surface proteinase leishmanolysin (gp63). *Structure* 6:1035–1046
 76. Corradin S, Ransijn A, Corradin G, Bouvier J, Delgado MB, Fernandez-Carneado J, Mottram JC, Vergères G, Mauël J (2002) Novel peptide inhibitors of *Leishmania* gp63 based on the cleavage site of MARCKS (myristoylated alanine-rich C kinase substrate)-related protein. *Biochem J* 367(Pt 3):761–769
 77. Moraes CB, Franco CH (2016) Novel drug discovery for Chagas disease. *Expert Opin Drug Discovery* 11(5):447–455
 78. Pecoul B, Batista C, Stobbaerts E, Ribeiro I, Vilasanjuan R, Gascon J, Pinazo MJ, Moriana S, Gold S, Pereiro A, Navarro M, Torrico F, Bottazzi ME, Hotez PJ (2016) The BENEFIT trial: where do we go from here? *PLoS Negl Trop Dis* 10(2):e0004343
 79. Derreumaux P, Schlick T (1998) The loop opening/closing motion of the enzyme triosephosphate isomerase. *Biophys J* 74(1):72–81
 80. Chebaro Y, Jiang P, Zang T, Mu Y, Nguyen PH, Mousseau N, Derreumaux P (2012) *J Phys Chem B* 116(29):8412–8422. doi:10.1021/jp2118778
 81. Tuffery P, Derreumaux P (2012) Flexibility and binding affinity in protein-ligand, protein-protein and multi-component protein interactions: limitations of current computational approaches. *J R Soc Interface* 9(66):20–33. doi:10.1098/rsif.2011.0584
 82. Bueren-Calabuig JA, Pierdominici-Sottile G, Roitberg AE (2014) Unraveling the differences of the hydrolytic activity of *Trypanosoma cruzi* trans-sialidase and *Trypanosoma rangeli* sialidase: a quantum mechanics-molecular mechanics modeling study. *J Phys Chem B* 118(22):5807–5816
 83. Daligaux P, Bernadat G, Tran L, Cavé C, Loiseau PM, Pomel S, Ha-Duong T (2016) Comparative study of structural models of *Leishmania donovani* and human GDP-mannose pyrophosphorylases. *Eur J Med Chem* 107:109–118. doi:10.1016/j.ejmech.2015.10.037