J Mol Model (2012) 18:4089–4100
DOI 10.1007/s00894-012-1411-6

ORIGINAL PAPER

Probing the structure of *Leishmania major* DHFR TS and structure based virtual screening of peptide library for the identification of anti-leishmanial leads

Rajalakshmi Rajasekaran · Yi-Ping Phoebe Chen

Received: 28 November 2011 / Accepted: 19 March 2012 / Published online: 15 April 2012

Abstract Leishmaniasis, a multi-faceted ethereal disease is considered to be one of the World's major communicable diseases that demands exhaustive research and control measures. The substantial data on these protozoan parasites has not been utilized completely to develop potential therapeutic strategies against Leishmaniasis. Dihydrofolate reductase thymidylate synthase (DHFR-TS) plays a major role in the infective state of the parasite and hence the DHFR-TS based drugs remains of much interest to researchers working on Leishmaniasis. Although, crystal structures of DHFR-TS from different species including *Plasmodium falciparum* and *Trypanosoma cruzi* are available, the experimentally determined structure of the *Leishmania major* DHFR-TS has not yet been reported in the Protein Data Bank. A high quality three dimensional structure of *L.major* DHFR-TS has been modeled through the homology modeling approach. Carefully refined and the energy minimized structure of the modeled protein was validated using a number of structure validation programs to confirm its structure quality. The modeled protein structure was used in the process of structure based virtual screening to figure out a potential lead structure against DHFR TS. The lead molecule identified has a binding affinity of 0.51 nM and clearly follows drug like properties.

Keywords Dihydrofolate reductase thymidylate synthase · Docking · Homology modeling · Leishmaniasis · Molecular dynamics · Virtual Screening

R. Rajasekaran · Y.-P. P. Chen (✉)
Department of Computer Science and Computer Engineering,
La Trobe University,
Melbourne, Victoria 3086, Australia
e-mail: phoebe.chen@latrobe.edu.au

Introduction

Leishmaniasis, a parasitic disease caused by the protozoan parasites of Leishmania genus and is transmitted by the bite of an insect vector called the phlebotomine sandfly. Leishmaniasis in each of its three clinical forms such as cutanous, mucosal and visceral, remains as a serious mankind diseases in tropical and subtropical areas of the world. World Health Organization reports that leishmaniasis threatens about 350 million people in 88 countries around the world [who]. In addition the combination of Visceral Leishmaniasis and HIV has been a major threat in southern Europe [1]. Unfortunately due to lack of commercial interest very few new drugs are being introduced against this deadly disease. Between 1975 to 2004 only two of 1556 novel dug compound was intended for leishamaniasis [2]. Until today there were not any effective vaccines developed and control of leishmaniasis primarily relies on chemotherapy.

There have been very few therapeutic arsenals in practice for the treatment of this dreadful infection. The first line of drugs the pentavalent antimony has long been the cornerstone of anti-Leishmania chemotherapy, but the development of resistance against this drug has limited the drug usefulness [3]. The second line of drugs includes pentamide and amphothericin. Toxicity and emerging resistance prevents the usage of pentamide whereas amphothericin B has the potential to induce acute toxicity requiring patient hospitalisation. Unfortunately, other therapeutic formulations such as Meltefosine and Paramomycin have either developed resistance for the parasite or they are prone to host-toxicity and high productive cost, which becomes the therapy to be inaccessible for the infected patients [4]. Since, the available therapies pitch assorted challenges; researchers employ diverse strategies to overcome these barriers and there by develop potent drug for leishmaniasis.

In order to provide a better therapeutic upshot against leishmaniasis, it is necessary to identify suitable and potential drug targets that can specifically target the parasite without affecting the targeted host [5]. A significant effort has to be taken to understand the biological mechanism of the infecting parasite by studying the indispensable metabolic pathways like sterol biosynthetic pathway, glycolytic pathway, purine salvage pathway and folate biosynthesis pathway [6]. Folates are essential vitamins that undergo series of biochemical reactions ranging from nucleotide biosynthesis to re-methylation of homocysteine. Remarkable differences have been identified in the folate metabolism of protozoa parasites and human [7]. One among them is the endogenous folate biosynthesis pathway possessed by protozoan parasites. This pathway is susceptible to antifolate inhibitors. Whereas, humans do not synthesize folate de novo instead use a membrane bound reduced folate carrier to bring the dietary folic acid into the cell [8]. Hence, enzymes related to this metabolism are of high interest as drug targets and the incorporation of anti-folates are believed to provide an idyllic therapy for many diseases [9]. The enzyme dihydrofolate reductase (DHFR) and thymidylate synthase(TS) plays an important role in folate metabolism and therefore are recognized as important targets against leishmaniasis. In most organisms DHFR and TS exist as separate molecular entities. Whereas,in Leishmania, these enzymes are a part of a bi-functional DHFR-TS complex [10]. This enzyme catalyzes both the reductive methylation of 2' − deoxyuridylate and the subsequent reduction of dihydrofolate to yield 2' − deoxythymidylate and tetrahydrofolate at two spatially discrete sites situated on different protein domains [7]. This enzyme plays a major role in both nucleic acid and amino acid biosynthesis and thus is considered to be an important target for anti-bacterial, anti-malarial, anti-cancer and anti-parasitic drugs. Studies have shown that it is possible to achieve inhibition selectivity of the parasite against the human by targeting the DHFR-TS complex of the Leishmania major. Hence, the selectivity issue concerning the parasitic diseases is supposed to be resolved with the inhibition of DHFR-TS enzyme complex.

Understanding the function of proteins is greatly influenced by perspective gained from their three-dimensional structures. Since experimental structures are only available for a limited number of proteins, computational methods for protein structure modeling play an important role. Homology modeling is currently the most accurate method, yielding models suitable for a wide arena of applications, such as structure-based drug design and virtual screening [11]. Comparative or homology modeling is a methodology to predict protein structure based on the general observation that proteins with similar sequences have similar structures. With the experimentally established protein structure (template) in hand, models can be generated for a homologous sequence (target) provided it shares ~30 % or more sequence or structural similarity with the template [12].

Virtual screening has emerged as a high end technology that is gaining increasing use in drug discovery. It is seen as a complementary approach to the experimental high throughput screening (HTS), and when coupled with structural biology, promises to increase the number, and enhance the success, of projects in the lead identification stage of the discovery process [13]. This technique involves the computational docking of large databases into the active site of the receptor to identify potential lead structures [14].

This work aims to develop a quality three dimensional structure of the protein DHFR-TS and identify potential ligands to target the DHFR-TS enzyme. Since, the experimental structure of the L.major DHFR-TS has not yet been resolved. The crystal structure of Trypanasomacruzi DHFR-TS with a ligand bound to it was used as a template to develop the homology model of L.major DHFR TS. We report the successful utilization of the virtual screening protocol to identify the novel DHFR-TS inhibitors as lead structures for the development of anti-leishmanial drugs.

Methods

Computational details

The computational modeling in this work was performed using a number of modules offered by Discovery Studio 1.7 from Accelrys. Inc. Protein structure construction through homology modeling technique was carried out using the Modeller 8.2 that comes inbuilt with DS 1.7 and the docking studies were conducted using the LIGANDFIT module. Structure evaluation servers, Verify3D, ERRAT, PSQS and PROSA, were used to validate the quality of the modeled protein structure. All *in silico* experiments were done on a Pentium 4 core2 Duo workstation using a Windows XP operating system.

Homology modeling

Any structure based drug design approach relies on quality three dimensional structures of the target protein [15]. In this study, the experimentally determined structure of *Leishmania major* DHFR TS was not available in the PDB and hence the structure was modeled using homology modeling approach. Very recently, several other protein structures of Leishmania species such as *L. donovani chagasi* DHFR-TS [16] and *Leishmania* MAPK [17]were successfully modeled through this comparative/homology modeling

method. A BLAST search against the PDB database revealed the crystal structure of DHFR-TS of *Trypanosomacruzi* (PDB ID: 2H2Q) to be the best template with 66 % identity. As the first step toward model building, the target and the template sequences were aligned with Align2D that is available within the MODELLER 8.2 of DS 1.7. The alignment was performed using BLOSUM62 matrix [18]. The structurally conserved regions were identified based on the Needleman-Wunsch algorithm

Followed by the sequence alignment, the *Leishmania major* DHFR TS structure was predicted based on the coordinates from the template structure, *T. cruzi*. High quality 3D structure of proteins were obtained by satisfying spatial restraints [19]. The modeled protein structure was energy minimized using the CHARMM module in the standard dynamics cascade protocol in Accelrys D.S.1.7. All the protein atoms were parameterized using the CHARMM force field with a non-bonded atom cut-off of 14 Å. At 300 K, 5000 steps of energy minimization were conducted using Adopted basis Newton-Raphson method (ABNR) with a root mean square gradient of 0.05 kcal mol^{-1} $Å^{-1}$.

The preliminary assessment of the modeled protein structure quality was carried out by accessing the Ramachandran plot [20] facility offered by DS 1.7. Further, the modeled structure of the *Leishmania major* DHFR TS protein was evaluated by a few other popular structure validation servers including Verify3D [21] and ERRAT [22] that are offered by the UCLA-DOE Institute of Genomics and Proteomics, CA, PSQS [23] and ProSA, online protein structure analyses servers [24].

Docking studies

Docking was carried out using the LigandFit docking protocol present in DS 1.7.LigandFit method employs a cavity detection algorithm for detecting invaginations in the protein as candidate active site regions. A shape comparison filter is combined with a Monte Carlo conformational search for generating ligand poses consistent with the active site shape [25]. The first step is the cavity detection to identify the active site of the protein which is followed by the ligand docking. The default attributes while running ligand fit is altered by choosing a minimization after docking. The docking algorithm itself works by conformational searching of flexible ligands and protein is chosen to be flexible. The attribute pose saving is changed to save the maximum number of poses and the docking interaction filters are changed. In the advanced attribute the force field used is set to CHARMm and the final minimization is set to be full potential to allow the whole complex to be minimized. The scoring method is chosen to be the LUDI1, LUDI2, and LUDI3. The other values are left to be default.

Binding site analysis using known ligand docking

The binding site is a small region or a pocket, where ligand molecules can best fit or bind to trigger the receptor and produce the desirable effect. Thus, recognizing the binding site or the active site residues in the target structure is of high importance [26]. It was comprehended that very often the ligand is bound in the largest cleft in over 83 % of the protein [27]. Thus in this work, the binding site of the target protein was defined from the volume of ligand approach using the binding site module in DS 1.7. To characterize the binding site of the protein for further studies the modeled structure was docked with the deoxy-uridin5'- monophosphate(DU) extracted from the crystal structure of *T.cruzi*. The accuracy of the method was predicted using the root mean square deviation (r.m.s.d) of the docked pose and the crystal structure of the ligand. The residues within $4A^0$ were characterized as the binding site residues. Further the residues were compared with the experimental determined binding site residues from the crystal structure (Table 1).

The docked ligand poses has been analyzed primarily by their scores. The aim of scoring the poses was to identify the correct binding pose by ranking the protein-ligand complexes according to their binding affinities. The inhibitory action of the ligand can be expressed as binding constant, dissociation constant, inhibition constant etc. Dissociation constant (K_d) or inhibition constant (K_i) is commonly used to describe the affinity between a ligand and a protein and represents how tightly a ligand binds to a particular protein. The smaller the K_i value, the more tightly the ligand is bound to the target. DS provide different scoring functions like LigScore 1, LigScore2, PLP1, PLP2, Jain, PMF and Ludi scores. Though all of them have been used for consensus scoring, Ludi 3 score has been considered for binding affinity calculations, $K_d$is calculated using the formula,

$$\text{Ludi score} = -100 \log K_d$$

Table 1 Docking energy and hydrogen bonding residues in the DU-L.*major* complex and the DU-T.*cruzi* complex

Protein-ligand complex	Docked energy (kcal mol^{-1})			Contacting residues (up to 4.0 Å) in final docked position*
	Steric	Electrostatic	Total	
DU- DHFR TS Leishmania	−19.47	−77.78	−97.25	L392,C395,**H396**, Q416,**R417,S418**, **D420**,G424,**N428**, H458,Y460
DU- DHFR TST.cruzi	−55.5	−35.0	−90.5	**Y343**,C403,**H404**, Q422,**R423**,S424, D426,G430,**N434**, H464

*Residues given in bold are involved in the hydrogen bonding interactions

Virtual screening

Virtual screening, or *in silico* screening, is a technique which is widely used in the pharmaceutical industry as a productive and cost-effective technology in the search for novel lead compounds [28]. Virtual screening of Leishmania DHFR TS was done against LigandFit/Cap database using Ligand fit docking protocol. The chemical available for purchase (CAP) Database provides access to details on reagents and screening compounds available from leading vendors around the world. This database contains 126923 molecules which were subsequently screened by various stages of docking. A flowchart depicting the virtual screening protocol is shown in Fig. 1. The initial screening was done based on the binding affinity of the ligand where a flexible docking was carried out. The top scoring compounds with K_d value less than 20 nm were selected for the next level of screening. The next level of screening was performed by docking the selected compounds to the active site of the receptor. This time the receptor protein was kept rigid and the initial poses were refined by rigid body minimization of the ligand using Dreiding force field [29]. The top scoring 21 compounds were selected and the Candidate poses were minimized using all atom CHARMM force field and smart minimization methods until the r.m.s gradient for potential energy was less than 005 kcal mol^{-1} $Å^{-1}$.The binding free energies of all the energy minimized poses were calculated using the interaction energy protocol. The binding free energy, K_d value, interacting residues, dock score, and the hydrogen bonding residues of the top scoring compounds were tabulated for further screening. In the next stage the compounds were subjected to Lipinski rule of five

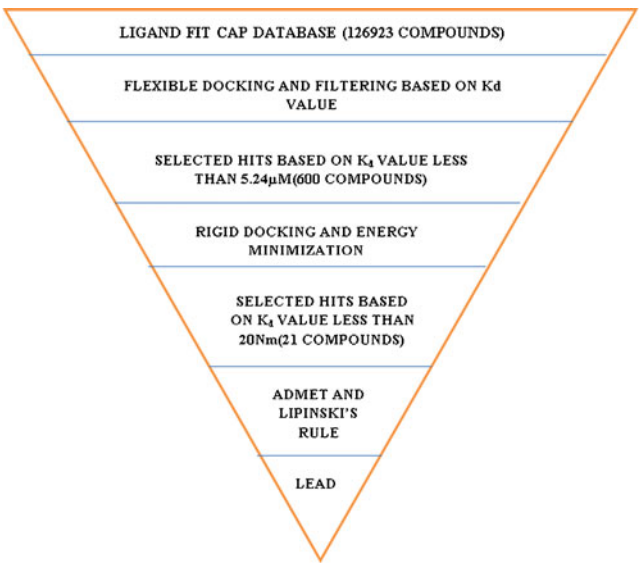

Fig. 1 Flowchart depicting the various stages of structure based virtual screening of the peptide library

[30] where the molecules with molecular weight less than 500, number of H-bond donors less than 5, and acceptors less than 10 were selected. Finally the best lead structure was selected based on the binding free energy and ADMET analysis of the top scoring compounds.

Molecular dynamics simulation of a fully hydrated complex of Leishmania DHFR TS with inhibitor

In order to check the overall stability of *Leishmania*DHFR TS complex with inhibitor in the presence of explicit solvent, a molecular dynamics using the CHARMM module has been carried out on a fully hydrated *Leishmania* complex with inhibitor. The first step has been the energy minimization of the hydrated model during which the backbone of the model has been kept fixed. The minimized, hydrated complex has then been subjected to a molecular dynamics simulation in three stages. In the first stage the temperature of the system has been raised from 0-300 K over 20 ps of simulation time. Next, the system has been equilibrated over 20 ps, and finally the production run has been carried out over another 125 ps.

Prediction of selectivity of identified inhibitor against human DHFR TS

Selectivity of the identified lead compound is one of the major issue in parasitic diseases.The leishmania DHFR TS are closely related to the vertebrate enzyme and studies show that there is no significant difference between the the human and leishmania DHFR TS domains [10]. Thus proving the selectivity of the inhibitor identified becomes vital. In Human the DHFR and TS domain exist separately and the TS sequence of Human is found to be 61 % similar to the Leishmania DHFR TS. In order to prove the selectivity of the lead compound the 3D structure of the Human TS was obtained from PDB(1JU6) and was docked with the identified lead compound. The same scoring and minimization methods described above were used to identify the Kd value.

Results and discussion

Homology modeling and structure quality assessment

T.cruzi DHFR TS was identified as the most appropriate template sequence through NCBI-BLAST. Although DHFR-TS protein sequences from various organisms exhibited extensive sequence similarity against the target sequence, T.cruzi DHFR TS sequence (PDB ID: 2H2Q) displayed 66 % of identity against the target. Further, the template possessed a good crystal structure resolution with

an inhibitor, deoxy-uridin bound to it which will aid the binding site recognition of the target. The alignment of the target and the template sequence showed that there were highly conserved regions. Following the alignment, the homology model of the target protein structure was generated based on the coordinates from the template structure. The obtained modeled structure was refined through energy minimization and molecular dynamics simulations. The modeled protein structure (Fig. 2) was superimposed against the template structure, so as to observe the level of similarities among them. The superimposed structure is given in Fig. 3. The positional root mean square deviation (RMSD) value between the *Leishmania major* DHFR TS and the template DHFR TS was found to be 0.52 Å for all the C_α atoms over 511 residues. This indicated that the structural alignment among the two structures was good and the modeled structure can now be validated for quality.

The modeled and minimized protein structure was thoroughly checked for quality through several levels of computational assessments. At the foremost level, Ramachandran plot (Fig. 4) of the predicted model was assessed to verify the torsion angle of backbone carbon in the model. It was observed that more than 99 % of non- glycine residues were in the allowed region or additionally allowed region and only four non-glycine residues were found in the disallowed region, but still, all of these amino acid residues were far from the binding site and hence had less effect in the structure quality. The modeled structure was verified using several other structure evaluation servers, operating on different strategies, before proceeding toward further analysis. ProSA is an online server that employs knowledge based potentials mean fields to analyze the quality of protein folds. ProSA produces Z-scores and energy graph as a function of amino acid sequence position to check the accuracy of the protein folds.

Fig. 3 Superimposed 3D structures of the modeled *L.major* DHFR TS and the *T.cruzi* template The RMSD between the superimposed structures was found to be 0.52 Å for all the C_α atoms over 511 residues

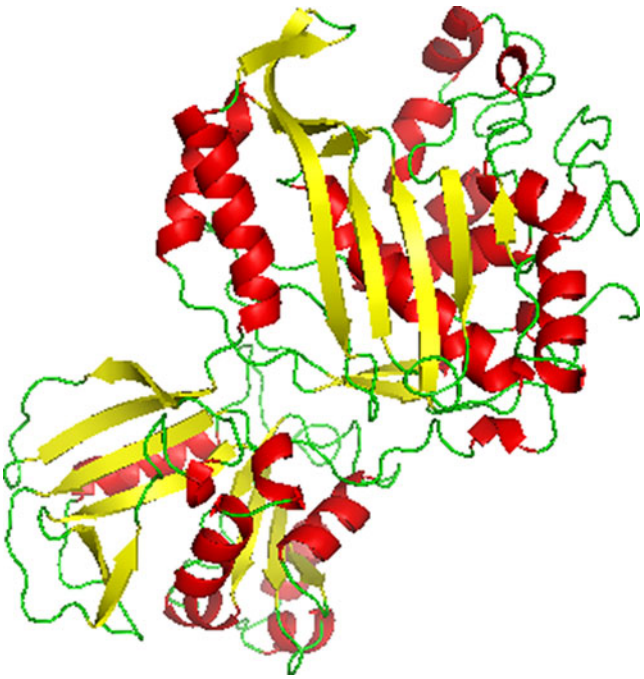

Fig. 2 Modeled 3D structure of the *Leishmania major* DHFR TS protein with the α-helix, β-sheet and turns representation

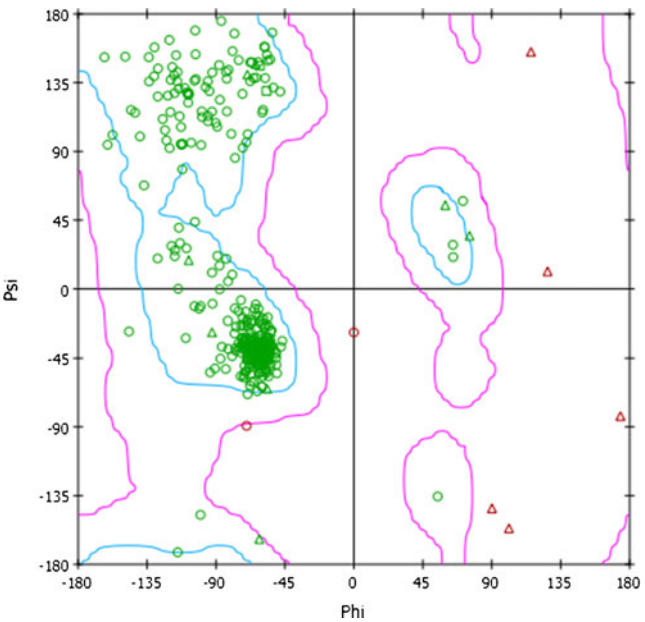

Fig. 4 Ramachandran plot of the modeled *L.major* DHFR TS structure obtained using the Modeller 8.2 in DS 1.7

ProSA analysis of the *L. major* DHFR TS showed that the Z-score value was −9.89 (Fig. 5a) that was within the range of native conformations of the crystal structures while the Z-score of the template structure was found to be −10.74. ProSA energy plot of L. major DHFR TS is shown in Fig. 5b in which the thick line represents the plot calculated by averaging energy over each residue and finally assigned to the central residue and whereas, the thin line displays the same energy graph with a smaller window size of ten residues.

ERRAT is another popular structure evaluation server that works by analyzing the statistics of non-bonded interactions between different atom types within the protein. The overall quality factor observed in ERRAT for the modeled structure was 88.17 % which was very close to the quality factor of the template, 95.52 %. The modeled structure was further verified using the PSQS, protein structure quality score, server that showed that the predicted structure possessed the PSQS value of −0.2259 while the PSQS value of the template structure was −0.2449. Finally, VERIFY 3D analyses also confirmed the quality of the modeled structure. A comparative table of different scores obtained from these structure valuation servers for the modeled Structure and the template structure is given (Table 2). All assessments showed that the quality of the modeled structure of *L. major* DHFR TS was very close to that of the experimentally determined template structure.

Binding site analysis

To identify the active site of the predicted model the known inhibitor, deoxy uridin (DU) in the template structure was extracted and docked with the modeled structure using the "define binding site based on selected ligand" feature of the DS.

In this approach, the LIGANDFIT searches for appropriate chemical space in the protein structure based on the position of the selected ligand. The amino acid residues within $4A^0$ from the docked ligand, in the target structure, are selected and overlaid on the binding site of the crystal structure of the template so as to validate the predicted active site. The target binding site and the template binding site were found to be highly conserved (Fig. 6) with most of the residues in common, and the root mean square deviation of DU-*Leishmania* DHFR TS complex with that of DU-*T.Cruzi* DHFR TS was observed to be 0.37 Å. Studies state that RMSD with less than 1.5 Å from the experimentally crystallized geometry is a characteristic of a good docking function [31]. The sequence alignment of the target and the template also showed no significant difference and most of the residues in the binding site are found to be similar in both the cases.Literature strongly suggest that there is a high level of similarity in the binding site of the protozoans [32, 33]

Hydrogen bonds pave efficient means for effective communication between the drug and the proteins and therefore are very attractive to drug scientists. Hence, in this work, we have also analyzed the hydrogen bonding interaction between the modeled *L.major* DHFR TS and the docked DU inhibitor (Table 1). The predicted binding mode of DU in the active site of L. major DHFR TS displayed that there were significant hydrogen bonding interaction between the DU inhibitor and the amino acid residues of the modeled protein structure such as His396, Gln416, Arg417, Asp420, Asn428 and Ser418 (Fig. 6). Whereas the hydrogen bonding residues in the template structure were His404, Arg423 and Asn434. Thus, this indicates that the binding site residues and the significant hydrogen bonding patters in the modeled *L.major* DHFR TS protein structure are agreeing very well with that of the template crystal structure, *T. Cruzi*. Analyzing the surface of the binding site showed the binding site to be a hydrophobic pocket lined with hydrophilic residues (Fig. 7). The binding site residues predicted agrees very well with the reported binding site residues of

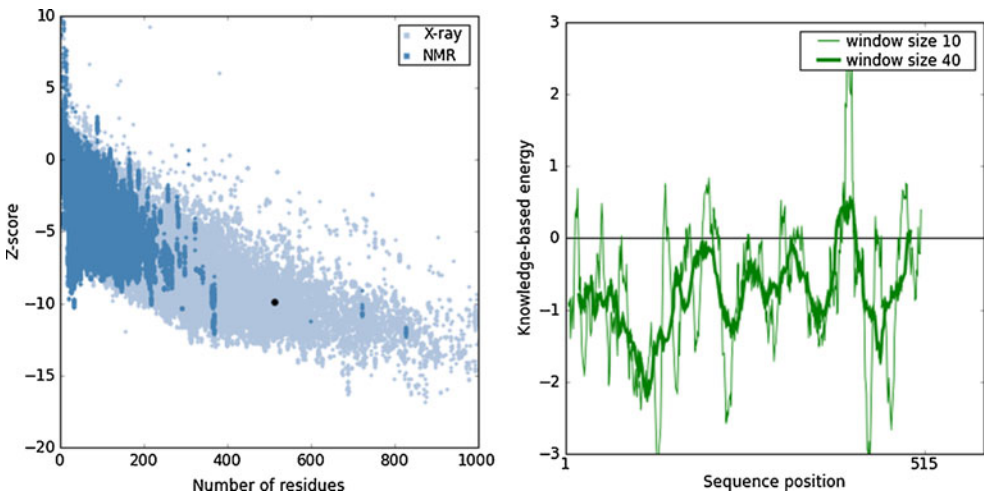

Fig. 5 **a** Z-score graph of the modeled *L.major* structure obtained from ProSA web server based on all the X-ray crystallography protein structures and the NMR determined structures. The Z-score of the modeled structure was found to be within the range of experimental structures. **b** Energy profile diagram acquired from ProSA

Table 2 Comparative quality scores obtained from ERRAT, PSQS and ProSA for the modeled *L.major* DHFR TS structure and the template structure, *T.cruzi* DHFR TS

Structure	Prosa Z- score[a]	Errat score[b] (in %)	PSQS[c]
Modeled *Leishmania major* DHFR TS	−9.89	88.166	−0.2259
T.Cruzi DHFR TS	−10.74	95.524	−0.2449

[a] Z-score of a protein structures are observed to be within the range of scores for the native proteins of same size

[b] Errat scores in percentage of calculated error values for the protein structures were below the 95 % rejection limit

[c] The average PSQS score for any PDB structures is −0.27

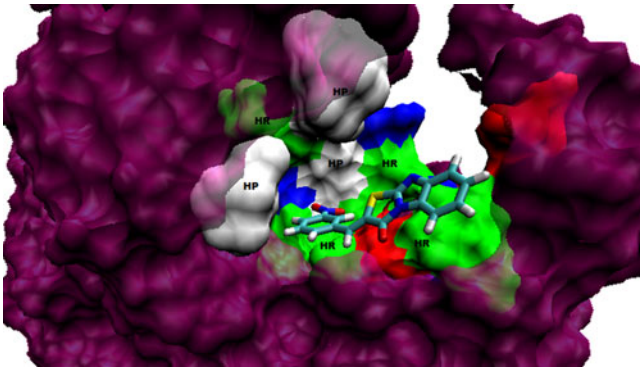

Fig. 7 The active site surface where the ligand is bound showing the hydrophobic residues and hydrophilic residues

leishmania major DHFR for different species [33]. A high degree of similarity was reported within the active site of the protozoans, where the residues tyr335,cys403,his396, arg417,ser418,asp420,asn428,his458 of leishmanial active site are highly conserved in other protozoans.

Virtual screening

Virtual screening of Leishmania DHFR TS against LigandFit/Cap database using LigFit docking protocol. From the initial screening of 126,923 compounds the top hits were selected based on the K_d value of the compounds. It is observed that the docking of the known inhibitor with the modeled receptor structure reported a K_d value of 5.24× 10-6 μM. Hence, the compounds exhibiting a K_d value less than this threshold were selected for further screening. In the next level of screening where the selected compound were rigidly docked with the receptor the top hits showing K_d value less than 21 nm were selected which resulted in the total of 21 lead structures.

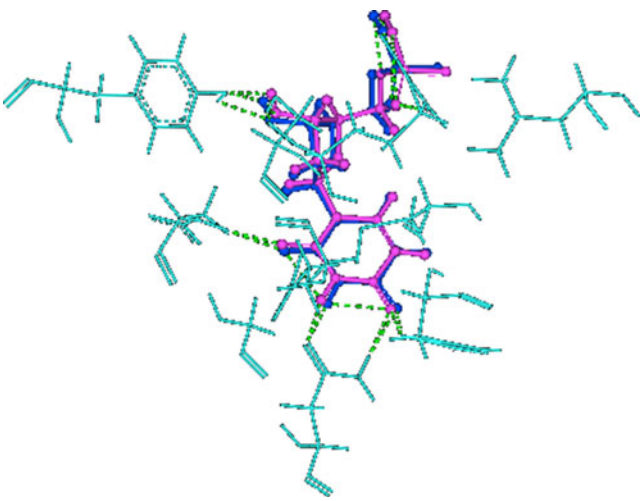

Fig. 6 Superimposition of the inhibitor binding sites of *T.cruzi* and the modeled *L.major* those are within its 5 Å region. Hydrogen bonding interactions between the active site residues of *L.major* and DU inhibitor are shown as dashed lines

For all the selected 21 docked complex the binding energies were calculated and tabulated (Table 3). The interacting residues and the corresponding hydrogen bonding residues were also tabulated. Further screening was done based on the Lipinski's rule of five. Of the selected 21 lead structures (Fig. 8) the molecular weight of the compounds with the cap key 567788,552014,73228,366370 were greater than 500 which makes them not suitable for oral drugs and were not considered for further analysis. The remaining compounds showed acceptable hydrogen bond donors and acceptors.

The top scoring inhibitor with the ID 571633 reported a K_d value of 0.51 nm this value indicates that there is significant interaction between the inhibitor and the protein. This predicted binding affinity is better than the reported binding affinities of the DHFR TS inhibitor trimetrexate, K_d value is 4.68 nm [34]. Further, the calculated binding free energy of the selected ligand is much lower compared to the docking energy of the Ligand DU. The compound exhibits strong interaction with receptor through hydrogen bonds and Van der walls forces. The N terminal of the inhibitor forms a hydrogen bond with the side chain residue of tyrosine335 and the carboxyl group of the inhibitor forms a series of hydrogen bond with the side chain residues of arginine417. The position of the docked ligand and the hydrogen bonding residues are detailed in Figs. 9 and 10.

The ADMET prediction of the selected lead compound showed good oral bioavailability, absorption and permeation. The predicted logP value was 4.55 and the molecular weight of the compound is 323 Da which considerably falls within the acceptable range.Thus based on the binding affinity and the ADMET properties the compound with the cap key id 571633 was identified to be a potential lead structure.

Table 3 List of lead molecules with their corresponding scores and interacting residues and hydrogen bonding residues shown in bold

Cap key	LUDI 3	K_d (nM)	Docking score	Molecular weight	Docking energy	Interacting residues
571633	929	0.51	45.672	323.33	−109.25	L392,C395,H396,Q416,R417,S418, D420,G424,N428,H458,Y460,Y335
567788	892	1.2	16.132	713.62	−60.281	L392,C395, H396, Q416, R417,S418, D456,L246,Y335, H458,Y460,D248, G424
77793	871	1.94	51.72	389.19	−44.292	Y335,C395,L392,S418,H396,N428, G424,L423,D420,I308,N312
554448	869	2.04	42.007	398.23	−39.612	I308,N312,L392,Y335,H396,G424, N428,D420,Y460,C395,N417,S418
62186	849	3.24	45.371	418.41	−30.845	I308,Y335,H396,L392,N417,D248, S418,C395,N428,G424,D420,H458, W309,N249
552014	836	4.37	26.474	667.83	−36.22	D420,Y460,H458,D248,N249,L392, N417,S418,C395,C419,H396,N428
587191	835	4.47	48.483	446.4	−41.591	G424,D420,C395,S418,H458,R417, Y335,L423,I308,N428,C419,H396
369123	829	5.13	50.512	477.94	−38.493	G424,D420,Y460,H458,D248,N249, L392,N417,S418,C395,C419,H396, N428
580658	828	5.25	59.451	435.45	−44.984	D420,Y460,H458,C419,S418,R417, P393,L392,C395,H396,N428,G424, I308
299194	818	6.61	43722	337.38	−37.005	I308,Y335,H396,C395,S418,C419, Q416,D420,N428,G424,L423,V279
569663	811	7.76	48774	387.76	−45.328	D420,Y460,H458,C419,S418,R417, P393,L392,C395,H396,N428,G424, I308
69386	807	8.51	39.644	385.42	−36.482	Y335,L392,H396,N428,G424,C395, S418,C419,V425,D248,N312,I308, W309,D420
570546	798	10.4	55.063	346.34	−38.956	I308,Y335,H396,G424,D420,S418, N428,C419,L392,H458,Y460,C395
50811	793	11.75	38.26	311.43	−37.194	I308,L423.D420,Y460,S418,C395, C419,Y335,H396,N428,G424,
14443	792	12.02	56.659	490.6	−47.151	I308,N428,C419,C395,L392,R417, S418,H458,D248,Y460,N462,H463, L423,G424,D420
73228	790	12.59	61.536	517.42	−40.827	W309,N428,H396,Y335,C395,L392, D420,Y460,H463,L423,G424,I308, F427,
569397	787	13.49	41.608	330.39	−39.992	I308,L392,L423,S418,D420,G424, N428,Q416,V425,C419,H396,
366370	784	14.45	50.186	548.68	−36.401	I308,W309G424,C395,H396,L423, D420,N428,F427,Y460,D248,L392, S418
568058	781	15.48	51.991	344.38	−31.962	F427,I308,W309,H396,C395,C395, S418,H458,C419,N428,G424,D420
49872	781	15.48	54.4	410.43	−43.076	I308,N312,L392,Y335,D248,R417, S418,Y460,C395,D420,G424,N428, H396,G424
367843	778	16.59	52.133	436.21	−48.127	D420,G424,N428,S418,Q416,H396, N312,C395,H458,L392

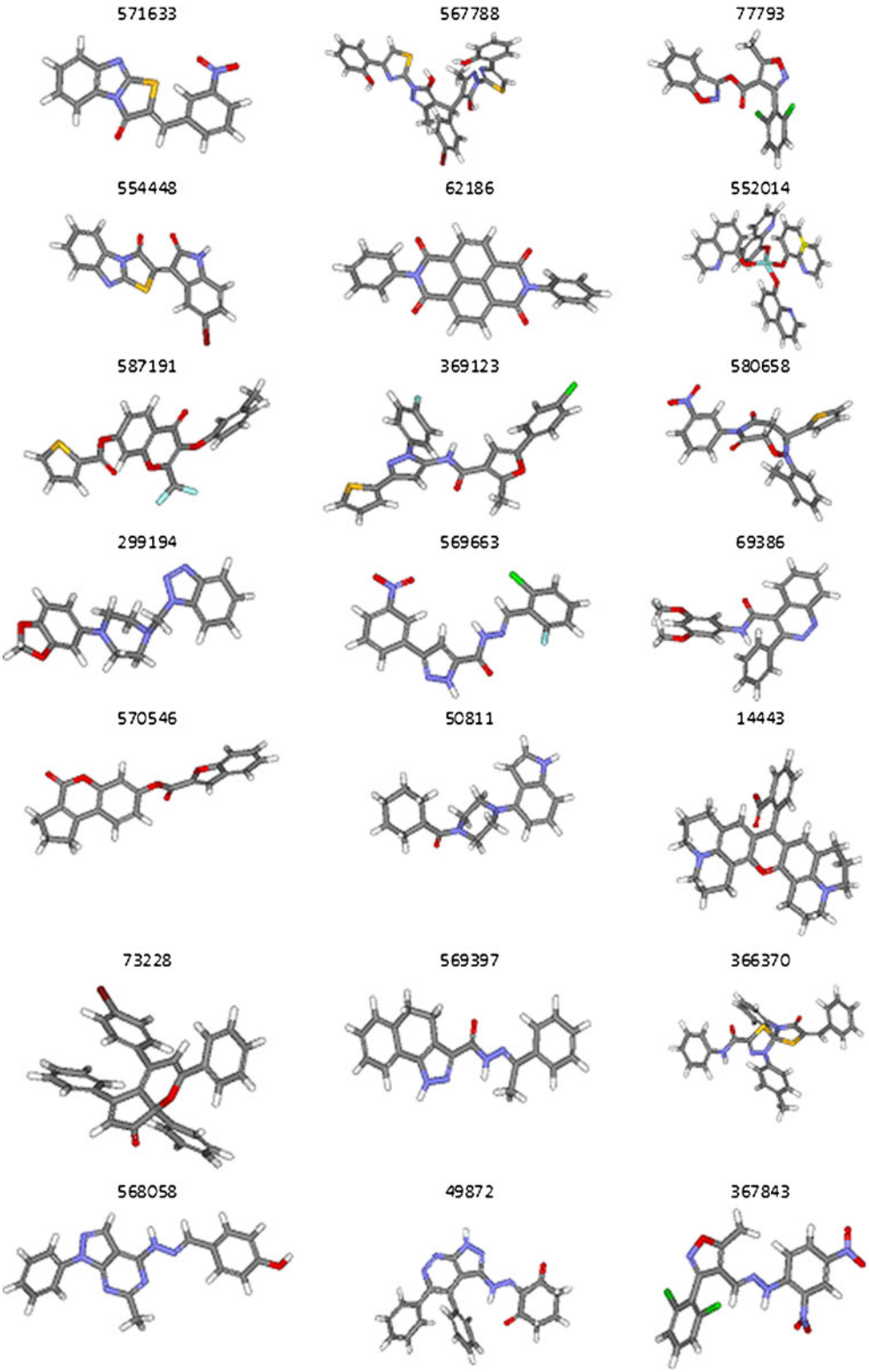

Fig. 8 2D Structure of top 21 hits obtained through rigid docking

Molecular dynamics

The effect of solvent on the *Leishmania* DHFR TS model bound with compound 571633 was studied by a molecular dynamics simulation of fully hydrated complex as described in methods. During the production phase of 125 ps following the initial heating and equilibration phase, the total energy and the simulation temperature were found to remain

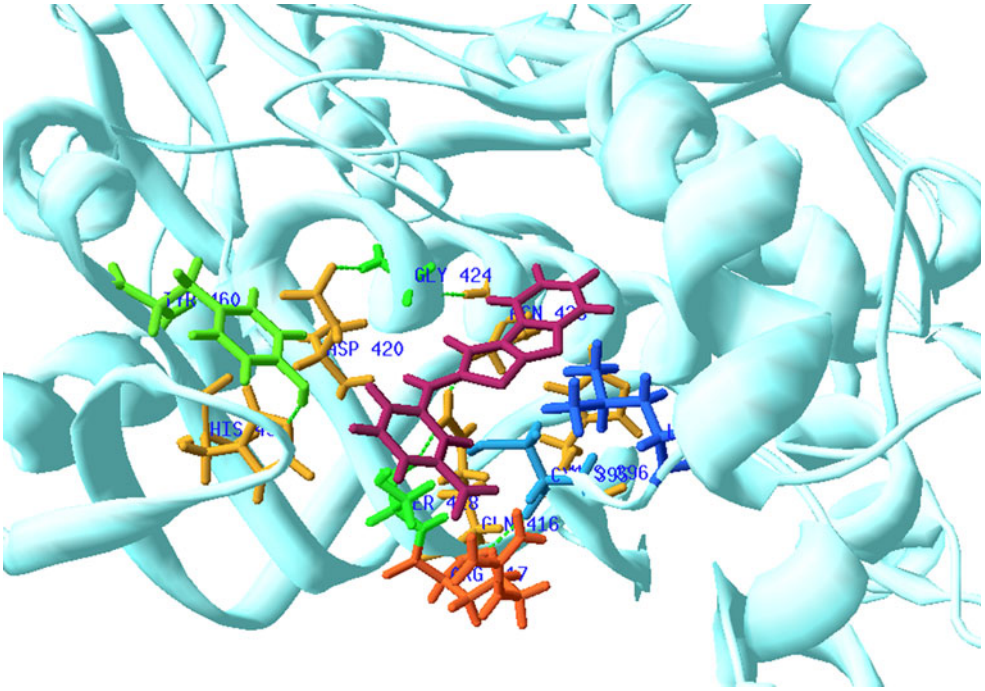

Fig. 9 Docking position of inhibitor 571633 with modeled leishmania DHFR TS showing interacting residues

steady with only little fluctuations. The snapshots of the dynamics trajectory at a 0,25,50,75,100,125 ps of the production run are shown in Fig. 11. While corresponding interaction energies are given in Table 4. The results indicate

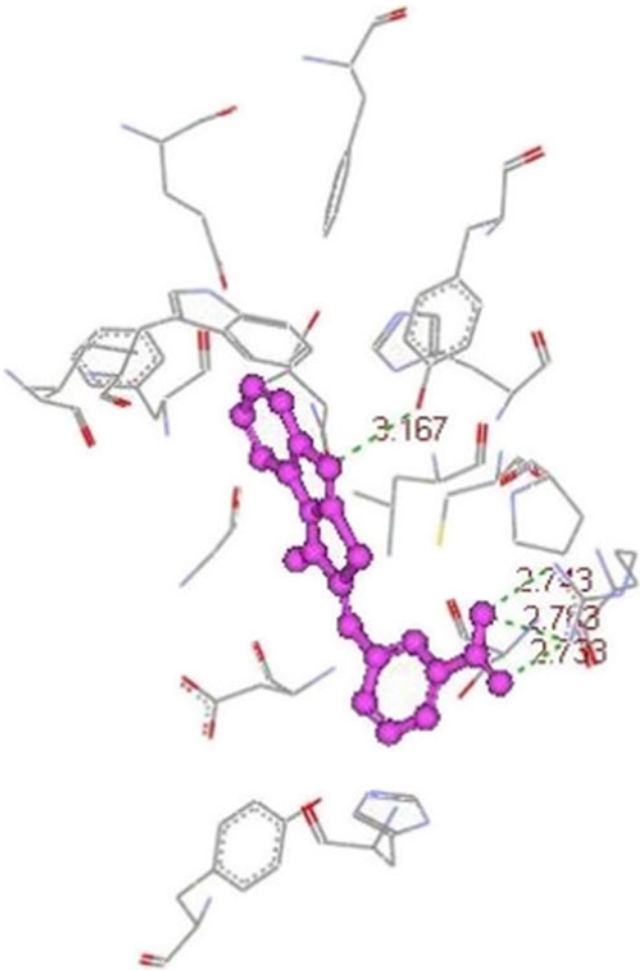

Fig. 10 Interaction set of designed inhibitor 571633 H-bonds are shown as dotted lines

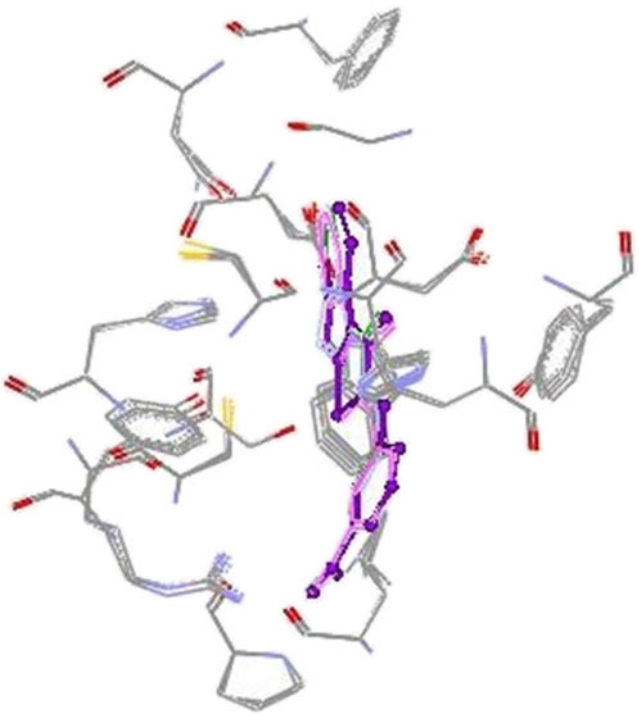

Fig. 11 Molecular dynamics trajectory of the modeled complex of *Leishmania* DHFR TS with identified inhibitor (571633). Snapshots of designed inhibitor (ball and stick) and selected *Leishmani* DHFR TS binding site conformers (lines) are extracted from production dynamics trajectory at time intervals of 25, 50, 75, 100 and 125 ps

Table 4 Fully hydrated dynamics simulation results

	Docked energy of ligand (kcal mol^{-1})						
	With solvent and enzyme			With enzyme only			
Time (ps)	Steric	Electrostatic	Total	Steric	Electrostatic	Total	Contacting residues (upto 4.0 Å) (hydrogen bonded residues are highlighted in bold)
25	−17.35	−58.46	−75.81	−17.35	−58.46	−75.81	E287,W309,**Y335**,L392,C393,H395,R396,**R417**,C418,D419,H420, G424,F427,N428,H458,Y460
50	−17.81	−58.46	−75.81	−19.05	−59.72	−78.77	E287,W309,**Y335**,L392,C393,H395,R396,**R417**,C418,D419,H420, G424,F427,N428,H458,Y460
75	−21.46	−53.31	−74.77	−21.46	−53.31	−74.78	E287,W309,**Y335**,L392,C393,H395,R396,**R417**,C418,D419,H420, G424,F427,N428,H458,Y460
100	−19.96	−51.43	−71.39	−19.96	−51.44	−71.39	E287,W309,Y335,L392,C393,H395,R396,**R417**,C418,D419,H420, G424,F427,N428,H458,Y460

that the ligand moves into a stable position with a lower docked energy. There was no significant difference in the binding mode and the hydrogen bonds observed was retained after 100 ps of production phase. The backbone of the ligand which was allowed to vary was fairly stable and the interactions were retained. Thus, it is shown that the compound is bound to the receptor in presence of explicit solvent.

Selectivity of identified inhibitor

The structure of human TS(1ju6) was obtained from the PDB and the protein was prepared and the binding site was identified using the bound ligand in the crystal structure. The LIGANDFIT protocol was used for docking the identified lead molecule 571633.The corresponding Kd value of the inhibitor was identified to be 4.8 Nm (Table 5) which when compared with the Kd value of the identified lead compounds clearly indicates that the identified lead compound is 100 times selective .

Conclusions

Leishmaniasis remains as endemic in several parts of the world and is a serious health problem in numerous underdeveloped countries. Due to lack of commercial interest very few new drugs are being introduced against this deadly disease. Identifying suitable drug target is essential for effective drug development. In this study we have applied the concept of structure based drug design to identify a potential drug target against this neglected tropical disease and develop a lead molecule against the identified target DHFR TS.

With the absence of an experimentally determined structure of L. major DHFR TS, we bring forward a computationally derived three dimensional structure of this protein through homology modeling approach. The energy minimized model structure was cautiously verified for its quality using a number of structure validation servers and all of the analyses confirmed the high level quality of the modeled structure whose RMSD value to that of the experimental template structure was 0.52 Å. The binding site of the modeled protein was predicted by performing docking studies with the known inhibitor deoxy uridin from the template structure. The active site residues were identified and it showed high similarity with the template binding site. Virtual screening was carried out to identify a lead compound that can effectively inhibit the Leishmanial DHFR TS through a series of docking protocols. After careful screening and refinement the inhibitor with the Cap Key ID 571633 was identified as a potent lead structure based on the binding affinity. The K_d value reported was 0.51 nm which is much less compared to that of the binding affinity of the known inhibitors that are reported so far. The compound also exhibited strong interaction with the receptor through hydrogen bonds and van der waals interactions. Further analysis of the ADMET properties and Lipinski's rule were also positive and thus the molecule 571633 and its analogues could be a promising lead compound for the development of antileishmanial drug.

Table 5 Selectivity of lead compound against human DHFR TS

	Target	Corresponding K_d value at 298 K (nM)	Interacting residues
1	DHFR TS-Leishmania	0.51	L392,C395,H396,Q416, R417,S418,D420,G424, N428,H458,Y460,Y335
2	TS-human	4.8	Y109, N112, M311, I108, L221,V135,

References

1. Desjeux P (2003) Leishmania/HIV co infections : epidemiology in Europe. Ann Trop Med Parasitol 97 (1):3 %U http://ci.nii.ac.jp/naid/10018519041/en/
2. den Boer ML, Alvar J, Davidson RN, Ritmeijer K, Balasegaram M (2009) Developments in the treatment of visceral leishmaniasis. Expert Opin Emerging Drugs 14(3):395–410. doi:10.1517/14728210903153862
3. Sundar S (2001) Drug resistance in Indian visceral leishmaniasis. Trop Med Int Health 6(11):849–854. doi:10.1046/j.1365-3156.2001.00778.x
4. Peacock CS, Seeger K, Harris D, Murphy L, Ruiz JC, Quail MA, Peters N, Adlem E, Tivey A, Aslett M, Kerhornou A, Ivens A, Fraser A, Rajandream M-A, Carver T, Norbertczak H, Chillingworth T, Hance Z, Jagels K, Moule S, Ormond D, Rutter S, Squares R, Whitehead S, Rabbinowitsch E, Arrowsmith C, White B, Thurston S, Bringaud F, Baldauf SL, Faulconbridge A, Jeffares D, Depledge DP, Oyola SO, Hilley JD, Brito LO, Tosi LRO, Barrell B, Cruz AK, Mottram JC, Smith DF, Berriman M (2007) Comparative genomic analysis of three Leishmania species that cause diverse human disease. Nat Genet 39(7):839–847
5. Chen YPP, Chen F (2008) Identifying targets for drug discovery using bioinformatics. Expert Opin Ther Targets 12(4):383–389. doi:10.1517/14728222.12.4.383
6. Chen YPP, Chen F (2008) Using bioinformatics techniques for gene identification in drug discovery and development. Curr Drug Metabol 9(6):567–573
7. Anderson AC (2005) Targeting DHFR in parasitic protozoa. Drug Disc Today 10(2):121–128
8. Tenter AM, Heckeroth AR, Weiss LM (2000) Toxoplasma gondii: from animals to humans. Int J Parasitol 30(12–13):1217–1258
9. McLuskey K, Gibellini F, Carvalho P, Avery MA, Hunter WN (2004) Inhibition of Leishmania major pteridine reductase by 2,4,6-triaminoquinazoline: structure of the NADPH ternary complex. Acta Crystallogr Sect D 60(10):1780–1785. doi:10.1107/S0907444904018955
10. Beverley SM, Ellenberger TE, Cordingley JS (1986) Primary structure of the gene encoding the bifunctional dihydrofolate reductase-thymidylate synthase of Leishmania major. Proc Natl Acad Sci 83(8):2584–2588
11. Kopp J, Schwede T (2004) Automated protein structure homology modeling: a progress report. Pharmacogenomics 5(4):405–416. doi:10.1517/14622416.5.4.405
12. Cavasotto CN, Phatak SS (2009) Homology modeling in drug discovery: current trends and applications. Drug Disc Today 14 (13–14):676–683
13. Lyne PD (2002) Structure-based virtual screening: an overview. Drug Disc Today 7(20):1047–1055
14. Desai PV, Patny A, Sabnis Y, Tekwani B, Gut J, Rosenthal P, Srivastava A, Avery M (2004) Identification of novel parasitic cysteine protease inhibitors using virtual screening. 1. The ChemBridge Database. J Med Chem 47(26):6609–6615. doi:10.1021/jm0493717
15. Anderson AC (2003) The process of structure based drug design. Chem Biol 10(9):787–797
16. Maganti L, Manoharan P, Ghoshal N Probing the structure of <i>Leishmania donovani chagasi</i> DHFR-TS: comparative protein modeling and proteinâ€“ligand interaction studies. J Mol Model 16 (9):1539-1547. doi:10.1007/s00894-010-0649-0
17. Awale M, Kumar V, Saravanan P, Mohan C Homology modeling and atomic level binding study of <i>Leishmania</i> MAPK with inhibitors. J Mol Model 16 (3):475-488. doi:10.1007/s00894-009-0565-3
18. Henikoff S, Henikoff JG (1992) Amino acid substitution matrices from protein blocks. Proc Natl Acad Sci 89(22):10915–10919
19. Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 234(3):779–815
20. Ramachandran GN, Ramakrishnan C, Sasisekharan V (1963) Stereochemistry of polypeptide chain configurations. J Mol Biol 7(1):95–99
21. Luthy R, Bowie JU, Eisenberg D (1992) Assessment of protein models with three-dimensional profiles. Nature 356 (6364):83–85
22. Colovos C, Yeates TO (1993) Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci 2(9):1511–1519. doi:10.1002/pro.5560020916
23. Jaroszewski L, Pawlowski K, Godzik A (1998) Multiple model approach: exploring the limits of comparative modeling. J Mol Model 4(10):294–309. doi:10.1007/s008940050087
24. Wiederstein M, Sippl MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 35(suppl 2):W407–W410. doi:10.1093/nar/gkm290
25. Venkatachalam CM, Jiang X, Oldfield T, Waldman M (2003) LigandFit: a novel method for the shape-directed rapid docking of ligands to protein active sites. J Mol Graph Model 21(4):289–307
26. Kalyaanamoorthy S, Chen Y-P Structure-based drug design to augment hit discovery. Drug Discovery Today 16 (17-18):831-839
27. Laskowski RA, Luscombe NM, Swindells MB, Thornton JM (1996) Protein clefts in molecular recognition and function. Protein Sci 5(12):2438–2452
28. Shoichet BK (2004) Virtual screening of chemical libraries. Nature 432(7019):862–865
29. Mayo SL, Olafson BD, Goddard WA (1990) DREIDING: a generic force field for molecular simulations. J Phys Chem 94 (26):8897–8909. doi:10.1021/j100389a010
30. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (2001) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 46(1–3):3–26
31. Vieth M, Hirst JD, Brooks CL (1998) Do active site conformations of small ligands correspond to low free-energy solution structures? J Comput Aided Mol Des 12(6):563–572. doi:10.1023/a:1008055202136
32. Ian HG (2002) Inhibitors of dihydrofolate reductase in leishmania and trypanosomes. Biochim Biophys Acta (BBA) – Mol Basis Disease 1587 :249-257
33. Zuccotto F, Martin ACR, Laskowski RA, Thornton JM, Gilbert IH (1998) Dihydrofolate reductase: A potential drug target in trypanosomes and leishmania. J Comput Aided Mol Des 12(3):241–257. doi:10.1023/a:1016085005275
34. Schormann N, Senkovich O, Walker K, Wright DL, Anderson AC, Rosowsky A, Ananthan S, Shinkre B, Velu S, Chattopadhyay D (2008) Structure-based approach to pharmacophore identification, in silico screening, and three-dimensional quantitative structure–activity relationship studies for inhibitors of Trypanosoma cruzi dihydrofolate reductase function. Proteins Struct Funct Bioinf 73 (4):889–901. doi:10.1002/prot.22115