Molecular modeling, structural analysis and identification of ligand binding sites of trypanothione reductase from *Leishmania mexicana*

Ozal Mutlu

*Marmara University, Faculty of Arts and Sciences, Department of Biology, Istanbul, Turkey*

**ABSTRACT**

**Background & objectives:** Trypanothione reductase (TR) is a member of FAD-dependent NADPH oxidoreductase protein family and it is a key enzyme which connects the NADPH and the thiol-based redox system. Inhibition studies indicate that TR is an essential enzyme for parasite survival. Therefore, it is an attractive target enzyme for novel drug candidates. There is no structural model for TR of *Leishmania mexicana* (*Lm*TR) in the protein databases. In this work, 3D structure of TR from *L. mexicana* was identified by template-based *in silico* homology modeling method, resultant model was validated, structurally analyzed and possible ligand binding pockets were identified.

**Methods:** For computational molecular modeling study, firstly, template was identified by BLAST search against PDB database. Multiple alignments were achieved by ClustalW2. Molecular modeling of *Lm*TR was done and possible drug targeting sites were identified. Refinement of the model was done by performing local energy minimization for backbone, hydrogen and side chains. Model was validated by web-based servers.

**Results:** A reliable 3D model for TR from *L. mexicana* was modeled by using *L. infantum* trypanothione reductase (*Li*TR) as a template. RMSD results according to C-alpha, visible atoms and backbone were 0.809 Å, 0.732 Å and 0.728 Å respectively. Ramachandran plot indicates that model shows an acceptable stereochemistry.

**Conclusion:** Modeled structure of *Lm*TR shows high similarity with *Li*TR based on overall structural features like domains and folding patterns. Predicted structure will provide a source for the further docking studies of various peptide-based inhibitors.

**Key words** Homology modeling; *Leishmania mexicana*; thiol metabolism; trypanothione reductase

**INTRODUCTION**

Trypanosomatid protozoan parasites comprise genus of *Trypanosoma* and *Leishmania* that cause neglected tropical diseases. Leishmaniasis is seen frequently in 88 countries on four continents (Southeast Asia, Africa, South America including mostly Brazil and Mediterranean countries) and cause 1.6 million new cases annually. Leishmaniasis is transmitted by infected female sandflies to the vertebrate hosts and reproduces within macrophages. Three clinical forms of leishmaniasis are: cutaneous leishmaniasis (*L. major*, *L. tropica* and *L. mexicana*); visceral leishmaniasis (*L. donovani* and *L. infantum*) and mucocutaneous leishmaniasis (*L. braziliensis*) are seen in human hosts.

Like other trypanosomatids, *L. mexicana* has no genes for glutathione reductase (GR), thioredoxin reductase (TrxR), catalase and selenocysteine-containing glutathione peroxidase which are responsible for intracellular thiol redox homeostasis in mammals. Parasites have dithiol trypanothione [T(SH)₂] and trypanothione reductase (TR) in their redox metabolisms which are common for all trypanosomatids but absent in hosts. Trypanosomatids are exposed to reactive oxygen (superoxide anion radical, hydrogen peroxide, peroxynitrite and the hydroxyl radical) and nitrogen molecules that are originated from parasite itself as by-product of aerobic metabolism or from the host macrophage defense system. Elimination of the various hydroperoxidases occurs via an electron transfer cascade which contains the enzymes of trypanothione, trypanothione reductase, tryparedoxin and tryparedoxin peroxidase. These low molecular mass thiols and redox enzymes protect enzymes from reactive derivatives and facilitate adaptation to various metabolic and environmental conditions.

One of the low molecular mass thiol compound is trypanothione [T(SH)₂] which is synthesized from two GSH and one spermidine molecule by glutathionyl spermidine/trypanothione synthetase (TryS) enzyme, an ATP driven reaction. Trypanothione reductase uses NADH as an electron donor for the reduction of trypanothione disulfide to trypanothione. Because of lacking of GR and TrxR in trypanosomatids, TR is a key enzyme which connects the NADPH and the thiol-based redox system. Trypanothione reductase knock out studies on *L. donovani* and *L. major* show decreased capacity to survive inside...
host macrophages\textsuperscript{8, 9}. Conditional knock out studies on \textit{T. brucei} resulted with growth arrest, loss of viability and virulence\textsuperscript{10}. Inhibition studies showed that numerous compounds inhibit trypanosomatid TR but not inhibit host GR\textsuperscript{11–14}. These works indicate that TR is an essential enzyme for parasite survival\textsuperscript{6}. So, TR is an attractive target enzyme for novel drug candidates because of the absence of trypanothione reductase system in mammals\textsuperscript{15–17}.

Trypanothione reductase is a member of FAD-dependent NADPH oxido-reductase protein family and shares close structural similarities with GR, lipoamide dehydrogenase and eukaryotic TrxRs\textsuperscript{5}. It has a mass of 52 kDa, comprises of a FAD and NADPH binding, central and interface domains and has two active sites. Structural features of TR from trypanosomatids and crystal structures of TR from \textit{T. cruzi}\textsuperscript{18, 19}, \textit{Crithidia fasciculata}\textsuperscript{20} and \textit{L. infantum}\textsuperscript{21} were solved. However, there is no 3D model for TR of \textit{L. mexicana} (\textit{LmTR}) in the protein databases. In this work, \textit{in silico} homology modeling approach was used to determine 3D model of TR from \textit{L. mexicana}, model was validated by web-based tools, structure was analyzed and possible drug target pockets were identified.

**MATERIAL & METHODS**

**Template selection and multiple sequence alignment**

The National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov) database was used to obtain amino acid sequence of the target TR from \textit{L. mexicana} (\textit{LmTR}) (MHOM/GT/2001/U1103) with the accession number of CBZ23392. Template identification was done by NCBI BLAST tool and two types of BLAST (Basic Local Alignment Search Tool), protein-protein BLAST and PSI-BLAST (Position Specific Iterated-BLAST) were performed for template selection against database Protein Data Bank (PDB) proteins. Homolog structure having the best score was selected as a template protein after comparative searching. Template protein PDB file and amino acid sequence in FASTA format was downloaded from the PDB (http://www.rcsb.org/pdb/home/home.do). Multiple sequence alignment was performed by both Molsoft and ClustalW\textsuperscript{22} from the European Bioinformatics Institute.

**Homology model development and validation**

A 3D model for the \textit{LmTR} enzyme was built by Molsoft ICM. Knowledge from the target sequence, alignment and 3D structure of the template protein which was obtained from the PDB was used to construct the model. Refinement of the obtained model was achieved by performing local energy minimization for backbone, hydrogen and side chains. Secondary structure of the protein was estimated by web-based program SOPMA\textsuperscript{23}.

Validation of the resultant protein model was achieved by checking stereochemistry, energy profile and residue environment. RAMPAGE Server\textsuperscript{24} was used to check stereochemistry of the model. Model quality was analyzed by ProSA\textsuperscript{25} and Protein Structure Validation Suite—PSVS 1.3 (http://psvs-1_4-dev.nesg.org). Statistics of non-bonding interactions was performed by ERRAT Server v.2.0\textsuperscript{26}. Physical and chemical parameters for the modeled protein (molecular weight, theoretical pl, amino acid composition, extinction coefficient, aliphatic index and grand average of hydropathicity) were performed by ProtParam tool\textsuperscript{27}.

**Ligand binding site identification**

Possible ligand binding sites/pockets on the 3D structure of protein were identified by Molsoft-ICM PocketFinder, web-based servers of CASTp-computer atlas of surface topography of proteins\textsuperscript{28} and GHECOM 1.0: Grid-based HECOMi finder\textsuperscript{29}.

**RESULTS & DISCUSSION**

**Template selection and multiple sequence alignment**

BLAST searching of the TR from \textit{L. mexicana} (\textit{LmTR}) was performed against PDB proteins. The four top scores were from \textit{L. infantum} (PDB number: 2JK6, Max score: 932, Identity: 91%, E-value: 0); \textit{C. fasciculata} (PDB number: 1FEA, Max score: 832, Identity: 80%, E-value: 0); \textit{T. cruzi} (PDB number: 1NDA, Max score: 686, Identity: 68%, E-value: 0); and \textit{T. brucei} (PDB number: 2WBA, Max score: 686, Identity: 66%, E-value: 0). Evolutionary tree showed these proteins have closely related amino acid compositions. Sequence identity above 50% is ideal for structure-based drug design and target assessment, site directed mutagenesis and assignment of protein function\textsuperscript{30}. Therefore, according to the identity, max score and E-value, the closest enzyme sequence was from \textit{L. infantum} (PDB 2JK6) that was selected as a template for \textit{in silico} homology modeling. Figure 1 shows multiple sequence alignments of the TR from \textit{L. mexicana} (\textit{LmTR}), \textit{L. infantum} (\textit{LiTR}), \textit{C. fasciculata} (\textit{CfTR}), \textit{T. brucei} (\textit{ TbTR}) and \textit{T. cruzi} (\textit{TcTR}) enzymes based on evolutionary tree. Multiple sequence alignment results showed that identity between \textit{LmTR} and \textit{LiTR} is 90%, \textit{CfTR} is 80%, \textit{TcTR} is 67%, and \textit{ TbTR} is 65% which were nearly same as that of BLAST search results.

**Homology model development and validation**

Trypanothione reductase from \textit{L. mexicana} was mod-
Fig. 1: Multiple alignments of trypanothione reductases from *L. infantum* (Li), *L. mexicana* (Lm), *C. fasciculata* (Cf), *T. cruzi* (Tc) and *T. brucei* (Tb) based on evolutionarily conserved amino acid sequences. Green columns show consensus residues among all species and yellow columns indicate at least three consensus residues. Red columns show catalytic site residues (C52, C57 and H461) which have conserved all enzymes.
eled by using the parameters of the template structure and target amino acid sequence. Number of models were generated and then RMSD (root mean square deviation) values between target and template structures were calculated based on C-alpha, visible atoms and backbone traces. Lower RMSD indicates stereochemically similar template and target structures, and the model with the best RMSD score was selected for further model validation study. Root mean square deviation results according to C-alpha, visible atoms and backbone were 0.809 Å, 0.732 Å and 0.728 Å respectively (Fig. 2). Selected model was checked for stereochemistry, energy profiles, potential errors, non-bonding interactions and some physical and chemical properties. ProtParam results show that the protein model is about 53.2 kDa, pI is 5.58, has 60 negatively and 49 positively charged residues, aliphatic index is 85.58 and grand average of hydropathicity is –0.095. According to the RAMPAGE results, 96.7% of residues (472) are in favoured region, 3.3% (16) of them are in allowed and no residue in the outlier region in Ramachandran plot. ERRAT was used to predict non-bonding interactions by examining the statistics of pair wise atomic interactions (CC, CN, CO, NN, NO, and OO). As a result, 93.562 as overall quality factor was obtained which means the model has good resolution structure and only four residues were indicated as rejected. ProSA calculates overall protein quality based on the scores of all experimentally determined protein chains available in PDB. The program gives a z-score and a plot of the residue energies. Z-score of LmTR was –10.94 which indicates that the model was within the range of the typically found proteins of similar size and the structure is non-problematic according to the energy plot. Protein structure validation suite (PSVS 1.3) was performed for analyzing overall model quality based on the global quality scores of four different structure validation programs (Verify3D, ProsaII, Procheck and MolProbity). All quality check results show the model is close to the native form and the environment profile of the structure is good. PDB file of the resultant model has been deposited in the protein model database (http://mi.caspur.it/PMDB) with the ID number of PM0078340.

Fig. 2: (a) Superimposition of the target (LmTR) and the template (LiTR) enzymes based on carbon backbone (RMSD: 0.728 Å). Green color denotes LmTR and blue is for LiTR. Arrows indicate beta sheets and turns show alpha helices; (b) Ribbon imaging of LmTR shows FAD binding domain (red), NADPH binding domain (green) and interface domain (cyan). Superimposition of the FAD (c), NADPH (d) and interface domains (e) of the trypanothione reductases from L. infantum (orange), L. mexicana (red), C. fasciculata (green) and T. cruzi (blue).

Fig. 3: SOPMA server results for the secondary structure of the model LmTR.
Overall structure

Secondary structure of TR from *L. mexicana* was predicted by SOPMA server. As a result, random coils (33.2%) and alpha helices (33.6%) are dominant features. Extended strands and beta turns have found as 24.85 and 8.35% respectively (Fig. 3). Crystal structure of TR from *L. infantum*\(^{21}\), *C. fasciculata*\(^{20}\) and *T. cruzi*\(^{18, 19}\) were solved. According to these studies all the TRs share three domains; FAD binding domain, NADPH binding domain and interface domain which are responsible in forming dimer state\(^{21}\) (Fig. 2). Catalytic site comprises redox active disulfide bridge (Cys52–Cys57 in *L. infantum* and Cys53–Cys58 in *T. cruzi*) and an active site base (His461) of the interface domain\(^{19, 21}\). In *Ln*TR redox active residues, active site bases are found in the positions Cys52–Cys57 and His461 like those found in *L. infantum* TR and Cys52–Cys57 form a disulfide bond about 2.2 Å distance (2.05 Å in *Li*TR) (Fig. 4). FAD domain consists five stranded parallel beta sheets (7–10, 31–35, 119–124, 155–158 and 322–324), three stranded antiparallel beta sheets (126–129, 134–138 and 149–152) and four alpha helices (14–26, 50–78, 92–116 and 335–349). NADPH domain contains three stranded antiparallel, three parallel beta sheets and two alpha helices. Interface domain located after the residue 360 to the end of the structure which contains four stranded antiparallel beta sheets and one alpha helix. NADPH binding domain residues (Tyr221, Arg222, Arg228/229, Tyr198, Ile199, Met333, Ala365 and Glu202/203) found in *Li*TR and *Tc*TR well-preserved in the *Ln*TR. These analyses show that the overall fold of *Ln*TR highly resemble to those of other TRs’ crystal structures. According to the active site and pocket prediction results, eight different pockets were identified by Molsoft ICM and 10 were from CASTp-computer atlas of surface topography of proteins and 13 were from GHECOM 1.0: Grid-based HECOMi finder. Results from Molsoft are given in Fig. 4 and Table 1. Pockets (Number 1 and 4) occupy the area of NADPH.

![Fig. 4:](image-url)
Table 1. Identified volume, area and residue information of possible ligand binding pockets.

<table>
<thead>
<tr>
<th>Pocket No.</th>
<th>Volume (AAA)</th>
<th>Area(Å²)</th>
<th>Pocket residues</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>324.355</td>
<td>304.9</td>
<td>G66, A67, Y69, M70, L73, R74, R85, L88, S90, G209, R213</td>
</tr>
<tr>
<td>3</td>
<td>234.547</td>
<td>278.676</td>
<td>F396, P398, H401, K409, F411, G431, S433, I438, H461, S464, E466, E467</td>
</tr>
<tr>
<td>4</td>
<td>224.235</td>
<td>269.759</td>
<td>L167, V194, I199, Y221, R222, L227, R228, S254, P255, A284, R287</td>
</tr>
<tr>
<td>6</td>
<td>142.982</td>
<td>171.104</td>
<td>G229, F230, H359, V362, C364, T374, T378, H428</td>
</tr>
<tr>
<td>7</td>
<td>129.093</td>
<td>138.48</td>
<td>S14, L17, E18, W21, V53, Y110, M113, T335, I339</td>
</tr>
<tr>
<td>8</td>
<td>116.639</td>
<td>145.139</td>
<td>K61, V64, T65, Q68, F367, K639, P371, S433, P435, E436</td>
</tr>
</tbody>
</table>

Residues with bold type indicate active site (Cys52 and His461) and NADPH binding (Y198, I199, M333, Y221, R222 and R228) residues.

CONCLUSION

Although all Leishmanias share a unique thiol-based metabolism which is common to all trypanosomatid family members, human hosts lack this antioxidant mechanism. Hence, enzymes of thiol-metabolism are subjected to oxidative stress. Leishmania–macrophage interactions: Insights into the redox biology. Free Radic Biol Med 2003; 19(2): 320–8.

REFERENCES

Vector Borne Dis 2013; 50: March 2013


Correspondence to: Ozal Mutlu, Marmara University, Faculty of Arts and Sciences, Department of Biology, Goztepe Campus, 34722, Goztepe, Istanbul, Turkey.
E-mail: ozal.mutlu@marmara.edu.tr

Received: 28 September 2012 Accepted in revised form: 30 January 2013