Pharmacophore
(An International Research Journal)
Available online at http://www.pharmacophorejournal.com/
Original Research Paper

COMPARATIVE METABOLOMIC INVESTIGATION OF LEISHMANIA MAJOR AND HOMOSAPIENS: AN IN-SILICO TECHNIQUE TO DISCOVER AND DEVELOP NOVEL DRUG TARGETS

R J Meshram[1]*, S N Jangle[2]

[1]* Bioinformatics, Center for Biotechnology,
Pravara Institute of Medical Sciences, Loni,
Taluka: Rahata, District: Ahmednagar,
Maharashtra, India
[2] Biochemistry Department, Rural Medical College,
Pravara Institute of Medical Sciences, Loni,
Taluka: Rahata, District: Ahmednagar,
Maharashtra, India

ABSTRACT

Visceral leishmaniasis is a parasitic disease caused by the protozoan scrounger *Leishmania donovani*. This disease is responsible for high rates of mortality and morbidity, especially in tropical regions of the world including India. With increasing problems due to resistance and clinical efficacy, the drugs currently used to treat this disease is becoming increasingly less effective, resulting in the urgent need for finding novel drug targets and new drug candidates in this area. Presented study describes comparative metabolomic approach through which entire metabolomes of *Homo sapiens* and *Leishmania major* are compared to detect unique biochemical reactions in them and further analysis of developing a structural model for enzyme holding key position in parasite‟s metabolic network using homology modeling approach. Here we report 97 potential target proteins and molecular structure of one target (*Trypanothione reductase*) from two subspecies (*Leishmania donovani* and *Leishmania major*) is described and compared. Results from current study can be used as background for developing new drug candidates against novel drug target described herein.

Keywords: *Leishmania donovani*, Comparative metabolomics, Homology modeling.

INTRODUCTION

Leishmaniasis is a vector-borne disease. This disease can give different clinical symptoms including cutaneous, mucosal, and visceral forms.[1] Both the cutaneous and mucosal forms can cause severe deformities to patients, including ulcerative skin lesions and the destruction of mucous membranes, in some cases

leading to permanent disfigurement. *Visceral leishmaniasis* due to *Leishmania donovani* is the most severe form of Leishmania infections. The annual incidence of *visceral leishmaniasis* is estimated to be 500,000 cases. In India, the reported number of cases is around 20,000 and number of deaths about 200 per year. Estimated number of cases is much higher. Bihar state is the worst affected with 33 districts endemic. It is also found in the neighboring states of West Bengal with ten districts affected, Jharkhand with five districts endemic and Uttar Pradesh with four. The overall prevalence of visceral leishmaniasis around the globe is estimated to be 12 million people, and the population at risk is 350 million.[2] The visceral form of the disease, caused principally by *L. donovani*, *L. infantum*, and *L. chagasi*, represents the greatest threat to human health, with symptoms ranging from fever and weight loss to hepatosplenomegaly, leading to death in untreated cases.[3] The current treatment strategies based on chemotherapy, for these parasitic diseases are limited and are not ideal, as they are often associated with severe side effects. The emergence of drug resistant parasites presents an additional and major problem. All these facts underline the urgent need for finding newer and effective drug targets and development of novel, cheap, safe, and easy-to-administer molecules for the treatment of these infectious diseases. An ideal drug target should possess at least two qualities. First, drug target protein should be unique to pathogen genome; second, it should have vital function in pathogen physiology. Present work initially gyrate around comparing metabolomes of Humans and Leishmania to decipher unique and vital target protein molecule in pathogen metabolome and later developing a structural model for identified target protein.

MATERIALS AND METHODS

Rahnuma [4] is hypergraphs based tool for prediction and analysis of metabolic pathways and comparison of entire metabolic networks that identifies biochemical reactions that are present in one organism which are absent in other, in other words it can find unique reactions, hence unique enzyme thus solving first criterion of target protein. The second norm of vitality is solved by comparing metabolic pathways of host and pathogen from KEGG database [5] and literature search. All available KEGG pathway maps of human and Leishmania were subject to standard comparative analysis using full network comparison mode on Rahnuma interface. Multiple Sequence Alignment was performed using Clustal X 1.81 [6] and Alignment was visualized in GeneDoc 2.7.[7] The computational technique that generates the structure from sequence using structural information from its homolog is called as homology modeling. This technique relies on the fact that proteins are evolutionary related and closely related proteins are known to possess similar 3D structure. Modeller [8] is stand-alone, freely available package for homology modeling. Modeller performs homology modeling by method called „satisfaction of spatial restraints" that uses a set of restraints derived from the alignment and expressed as Probability Density Functions for different types of restraints, finally the model is obtained by minimizing the violations to these restraints. Studies have proved that Modeller outperforms most of other homology modeling suits, it"s fast, reliable and freely available and hence we selected it in our study.[9,10]

RESULTS AND DISCUSSION

Comparative Metabolomics of Leishmania and Human

We have identified 1627 common reactions, 97 were found unique to pathogen metabolome, while 1036 were exclusive to humans and remaining 492 were common to both (Fig1A). These 97 unique reactions participate in 42 metabolic pathways encircling amino acid, lipid, carbohydrates, nucleotide, and cofactor metabolism, pathways that include glycolysis and TCA and other metabolic conversions

(Fig1B). Reactions from amino acid metabolism specifically comprise of „Alanine, Aspartate and Glutamate metabolism", „Cyanoamino acid" and „Seleno aminoacid" metabolism, „Cysteine and Methionine metabolism", „Tyrosine metabolism", „Glycine, Serine and Threonine metabolism", „Phenylalanine metabolism", „Lysine" biosynthesis and degradation, „Tryptophan metabolism" and finally „Arginine and Proline metabolism". (Fig1C). Unique reactions from Lipid Metabolism encompassed pathways like Glycerophospholipid metabolism, Ubiquinone and other terpenoid-quinone biosynthesis, Steroid biosynthesis, Glycerolipid metabolism, Limonene and pinene degradation, Fatty acid biosynthesis, Lipoic acid metabolism and alpha-Linolenic acid metabolism. (Fig1D). Nucleotide metabolism also had four unique reactions both from Purine and Pyrimidine metabolism, similarly Carbohydrate metabolism differed in two organisms in terms of reactions involved in Pentose phosphate pathway, Pentose and glucuronate interconversions, Galactose, Fructose and mannose metabolism and „Starch and Sucrose Metabolism". Another major difference was observed concerning Cofactor metabolism that included reactions from pathways like Vitamin B6, Biotin, and Thiamine, Nicotinate and nicotinamide metabolism, 66 % of Glutathione metabolism from Leishmania differed from human (fig1e). All other remaining unique reaction that can not be broadly classified into these groups were clustered in "Other group" that incorporated reactions from Nitrogen and Sulfur metabolism, Ascorbate and Aldarate metabolism, Butanoate metabolism, Phenylpropanoid biosynthesis and biosynthesis of secondary metabolites. Next, enzymes that can serve as potential target were scanned manually by comparing the metabolic pathways from KEGG.

Trypanothione

NADP Oxidoreductase more commonly called Trypanothione reductase (TR) is one of the key enzymes of glutathione metabolism which is unique to pathogen carrying out oxidation and reduction of Disulfide Bridge in trypanothione simultaneously generating reducing power in form of NADPH and is found to be closely associated with enzyme Ascorbate: Hydrogen peroxide Oxidoreductase that is involved in detoxification of Hydrogen peroxide to water. Inhibiting this enzyme can prove dual attack on Pathogen, firstly by depleting a local source of NADPH resulting imbalances in energy generation and secondly by accumulation of toxic metabolite hydrogen peroxide, that is known to cause oxidative stress to lipids, Protein and even DNA (Fig 2).

Modeling of Trypanothione Reductase

Template identification and sequence analysis

The sequence of TR enzyme of *L. donovani* and *L. major* was retrieved from SWISS-PROT [11] database with accession number P39050 and Q4QJG7 respectively. A template selection search was performed using BLAST-P [12] against PDB [13] database from NCBI interface, simultaneously "Template Identification Tool" at Swiss-Model interface [14, 15] provided by Swiss Institute of Bioinformatics was utilized for template selection. . In results for both target sequences we got 10 significant hits with E-value zero, the best of these comprise of model 2jk6 from species *Leishmania infantum* sharing 98 % sequence identity with TR from *Leishmania donovani* and 95 % with TR from *Leishmania Major* . Similar results were obtained from NCBI BLAST-P server.

Elementary Sequence analysis of TR from *L.donovani* revealed presence of a redox active disulfide bridge at the N terminal domain between Cysteine 52 and 57 residues (Fig 6 A). Active site is located at the interface of two interacting monomers around residue His 461 that comprise residues from both the chains (Fig 6 B). The FAD binding Domain was found on N terminal Domain from residue ASP 35 to CYS 52 (Fig 6 C). The active site is sequentially

conserved in TR of *L donovani* and *L major*, FAD binding domain of TR in *L major* is almost conserved, except a mutant motif FFA is present instead of ALV in *L donovani* and LFA in *L. infantum* at position 44 to 46. It is worth to note that all the changes are conservative (Fig 3).

Target template alignment

The Align2D command [16] from modeler was utilized to align query sequence to template structures. Align2D command is a modification of classic Needleman-Wunsch Dynamic programming algorithm [17] as it considers structural information from the template when constructing an alignment, which is accomplished through a variable gap penalty function that tends to place gaps in solvent-exposed and curved regions, outside secondary structure segments, and between two positions that are close in space.

Model generation

Single Model was generated implementing standard Automodel Class from Modeller.

Energy minimization

Energy minimization was done *in vacuo* using Force field approach by means of GROMOS96 43Bl parameters set [18], without reaction field applying 20 steps of Steepest Descent Algorithm within DEEP VIEW 3.7. [19]

Model refinement

The initial model generated was subject to geometrical and stereochemical analysis. Total of 79 checks were performed using What-if program, the models so produced performed well on almost all the tests except 171 and 237 „Abnormally short interatomic distances" (Bumps) were reported in Structure of TR from *L.donovani* and *L.major* respectively which on 10 iterative refinements on What If Web Interface (http://swift.cmbi.ru.nl/servers/html/index.html), were neatly reduced to 54 and 69 respectively Fig 4.

Model evaluation

Models after refinement were subject to evaluation using standard tools like Procheck[20], ERRAT[21] and Verify3D[22] along with WhatCheck/WhatIf [23] online servers. Procheck results of TR Model for *L major* indicated 93.1 % of Non-Glycine and Non-Proline residues to be present in Core region, 6.4 % in additionally allowed region, and 0.2 % in generously allowed region of the Ramachandran plot. In other words 98.8 % of Non-Glycine and Non-Proline lied in allowed region and only single residue was found to be in disallowed region. There were 20 Proline residues and all of them were reported to lie in core region, where as only 7 Glycine residues out of 50 were reported to be present in disallowed region. In short 98.38 of all residues lie in allowed region of Ramachandran plot (Fig 5A), in case of TR of *Leishmania Donovani*, 93.8 % of Non-Glycine and Non-Proline residues to be present in Core region, 5.7 % in additionally allowed region, and 0.2 % in generously allowed region of the Ramachandran plot. In other words 98.8 % of Non-Glycine and Non-Proline lied in allowed region and only single residue was found to be in disallowed region. There were 19 Proline residues and all of them were reported to lie in core region, where as only 7 Glycine residues out of 50 were reported to be present in disallowed region. In short 98.38 of all residues lie in allowed region of Ramachandran plot (Fig 5B). All these figures indicate excellent quality of model in terms of covalent geometry and Dihedral angles. Non-bonded atom-atom interactions were checked using ERRAT , overall quality factor was found to be 94.32 for *L. major* and 90.37 for *L. donovani*, most of residues have lower error values, again indicating good quality in terms of Non-bonded atom-atom interactions. The compatibility of an atomic model (3D) with its own amino acid sequence (1D) was analyzed using Verify 3DProgram. The returned 3D-1D profile for the most part for models of both organisms stayed above 0.2 again demonstrating

good quality of model. The resulting TR Models (Fig 6) were visualized in Deep View 4.0 and submitted to PMDB Database [24] with PMDB ID PM0077277 and PM0077278 for *L. donovani* and *L major* respectively.

CONCLUSION

Present study illustrates a novel approach of comparative metabolomics for finding an efficient drug target. We have identified 97 unique metabolic reactions and thus 97 potential target proteins against *Leishmania donovani*. One new target and its molecular structure is also described here in ,the resulting good quality model of trypanothione reductase can further provide a springboard for research in area of structure based drug design against *Leishmaniasis* and results of this study can be effectively used in identifying novel drug candidates using Molecular Docking and Virtual Screening approaches.

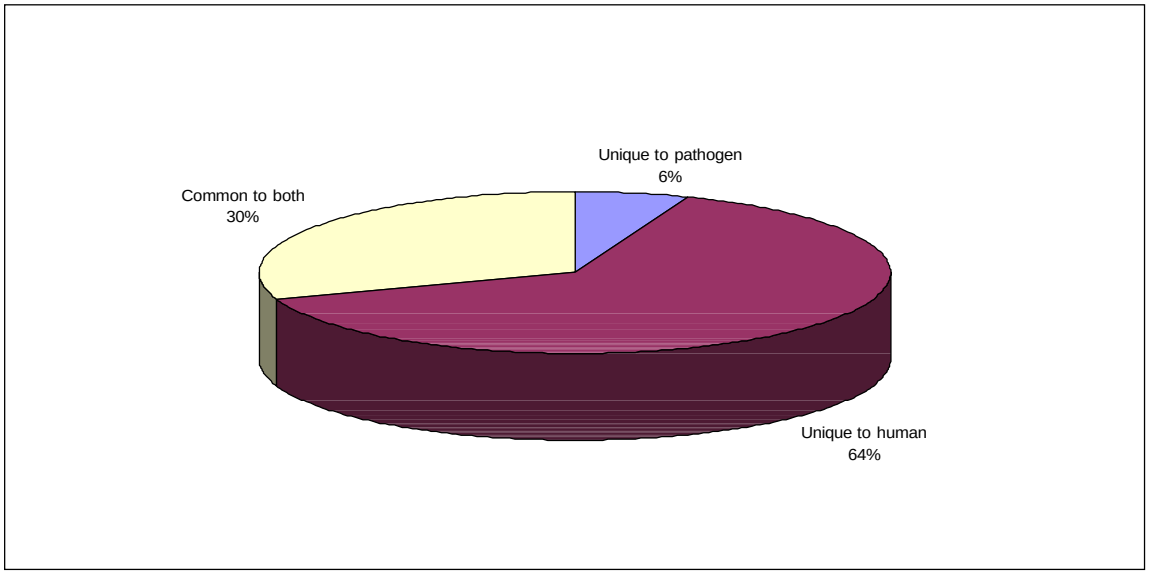

Figure1 A: Distribution of reactions in pathogen and host.

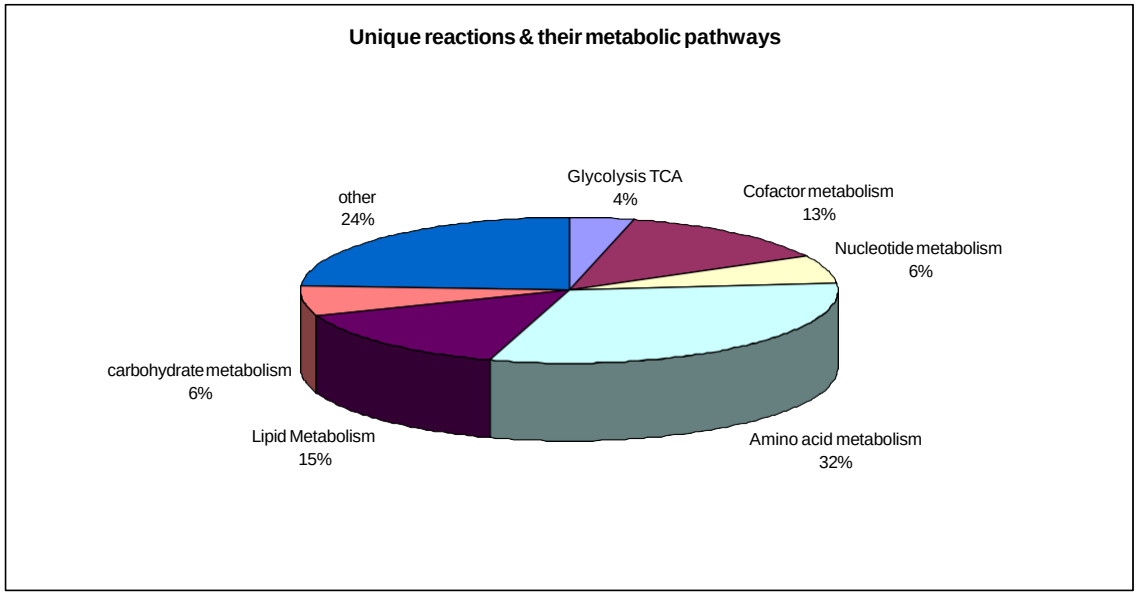

Figure1 B: Unique reactions in pathogen sorted according to metabolic pathways.

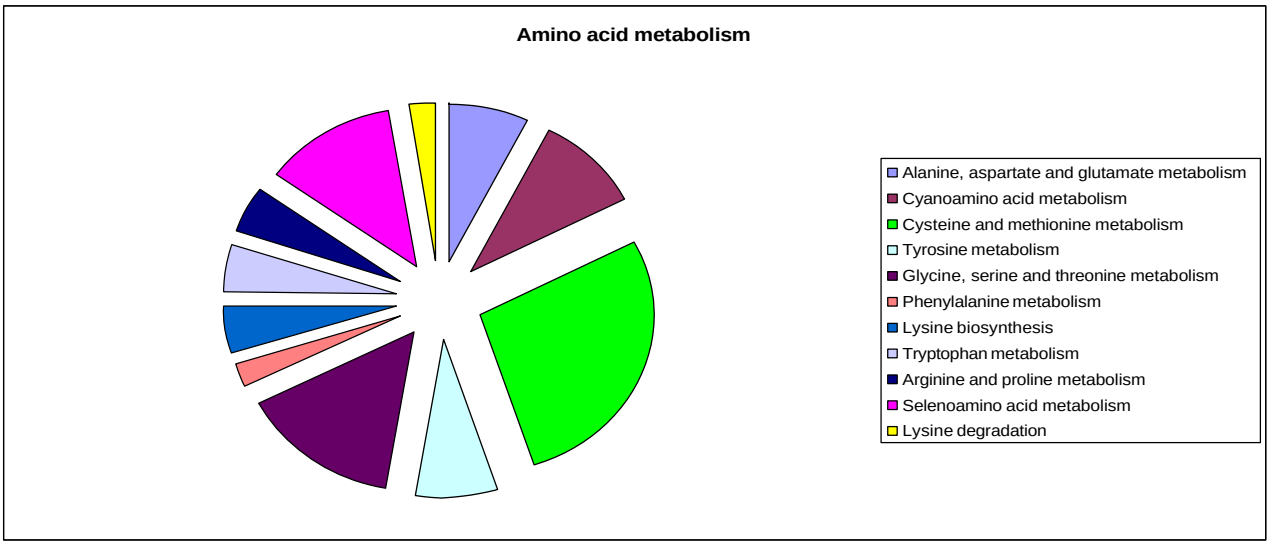

Figure1 C: Unique reactions in pathogen involved in amino acid metabolism.

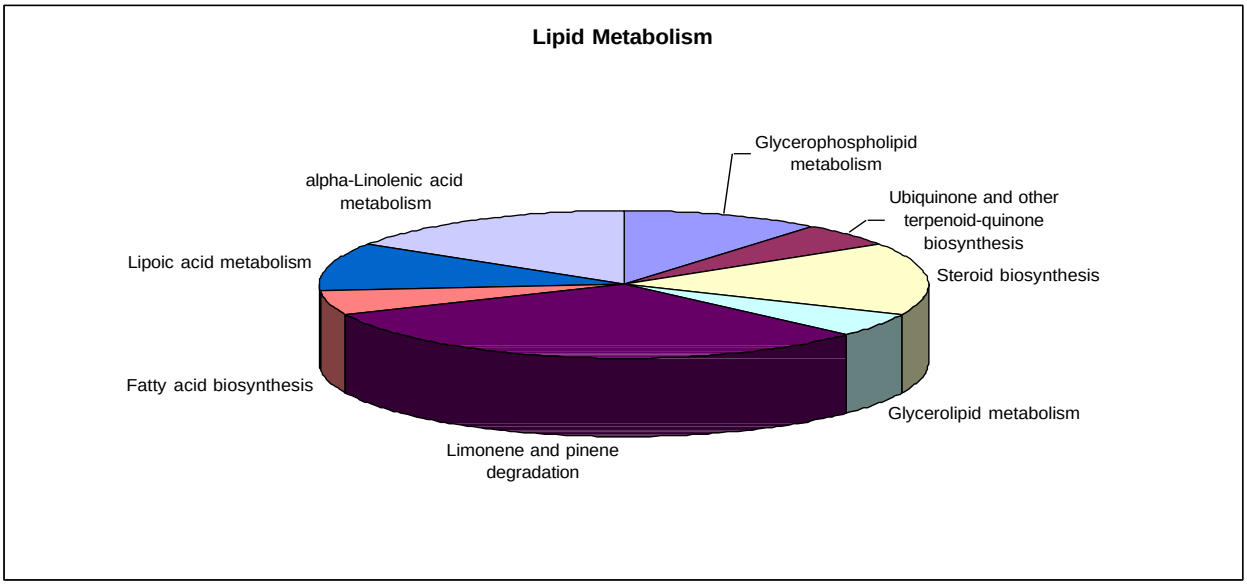

Figure1 D: Unique reactions in pathogen involved in lipid metabolism.

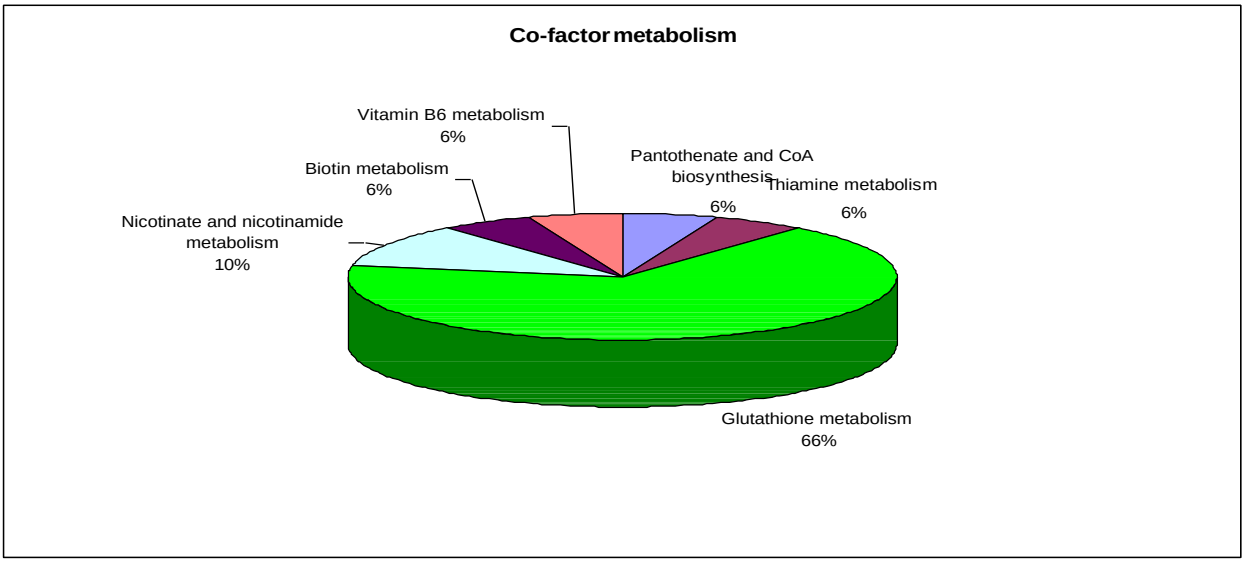

Figure1 E: Unique reactions in pathogen involved in co-factor metabolism.

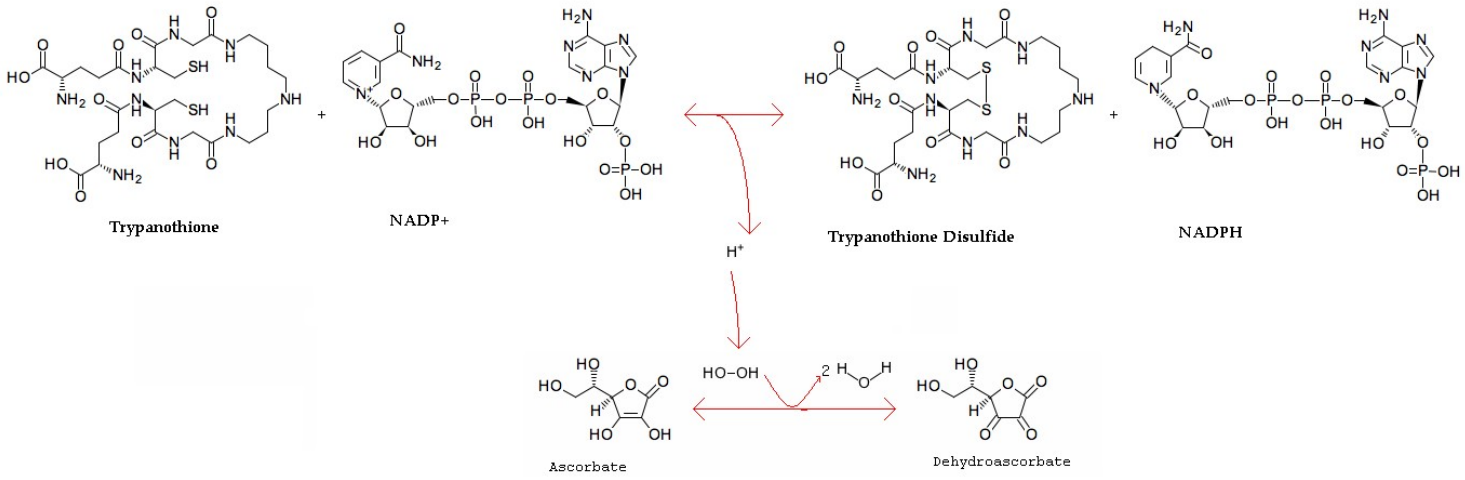

Figure2: Association of TR reaction with ascorbate peroxidase reaction.

Figure3: Alignment of FAD Binding Domain shaded by Physiochemical properties showing conservative changes (Black shading for Hydrophobic and Yellow for Small amino acid Residues).

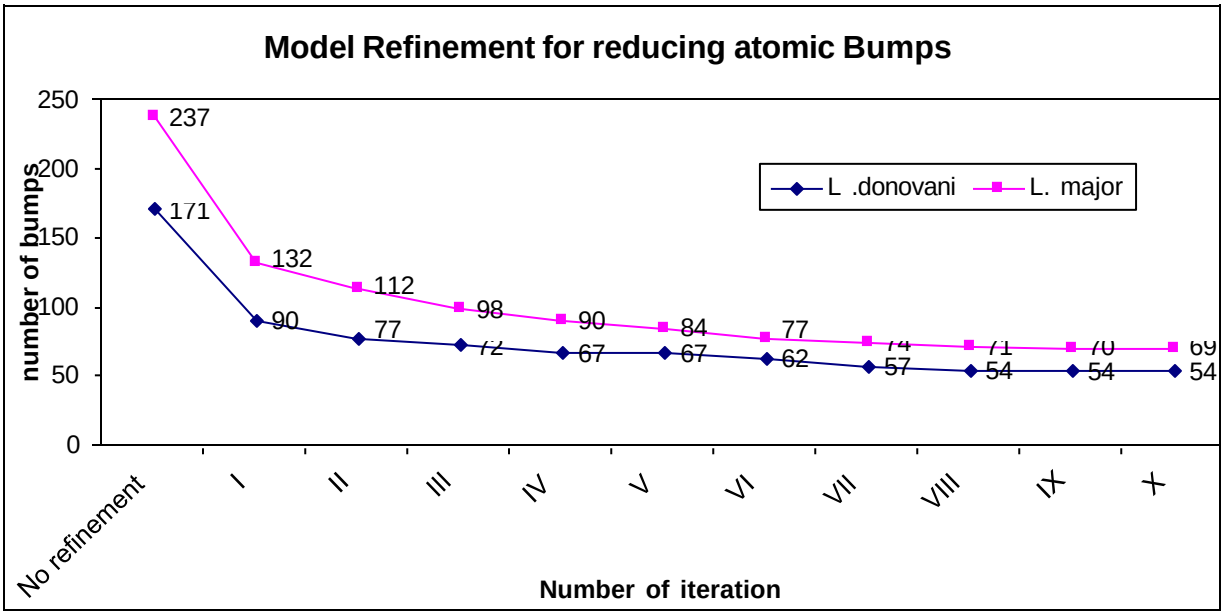

Figure4: Model refinement for reducing atomic bumps.

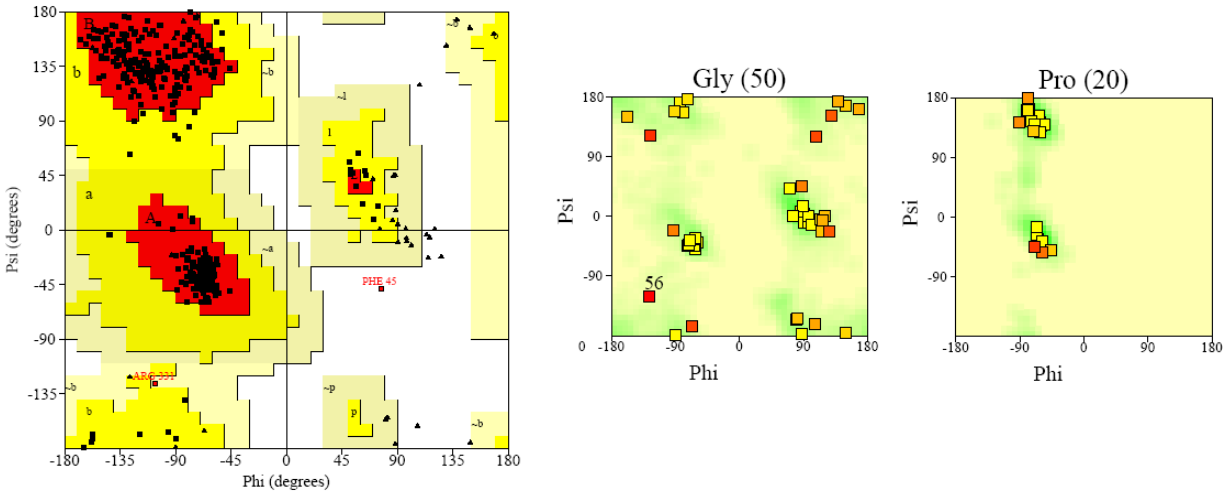

Figure5 A: Procheck results for TR from *L major*

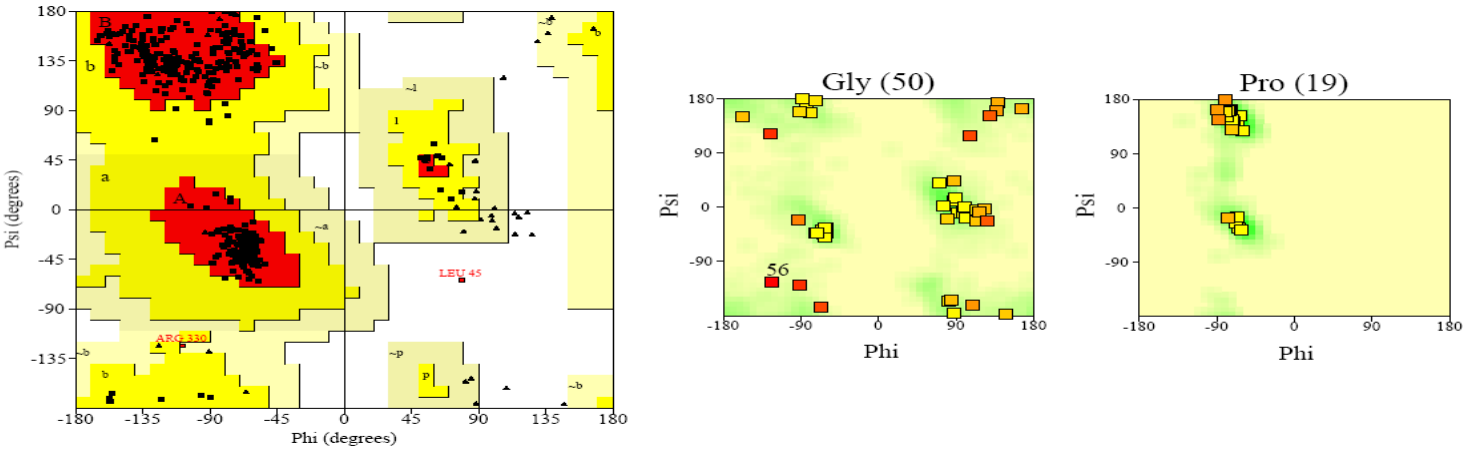

Figure5 B: Procheck results for TR from *L donovani*

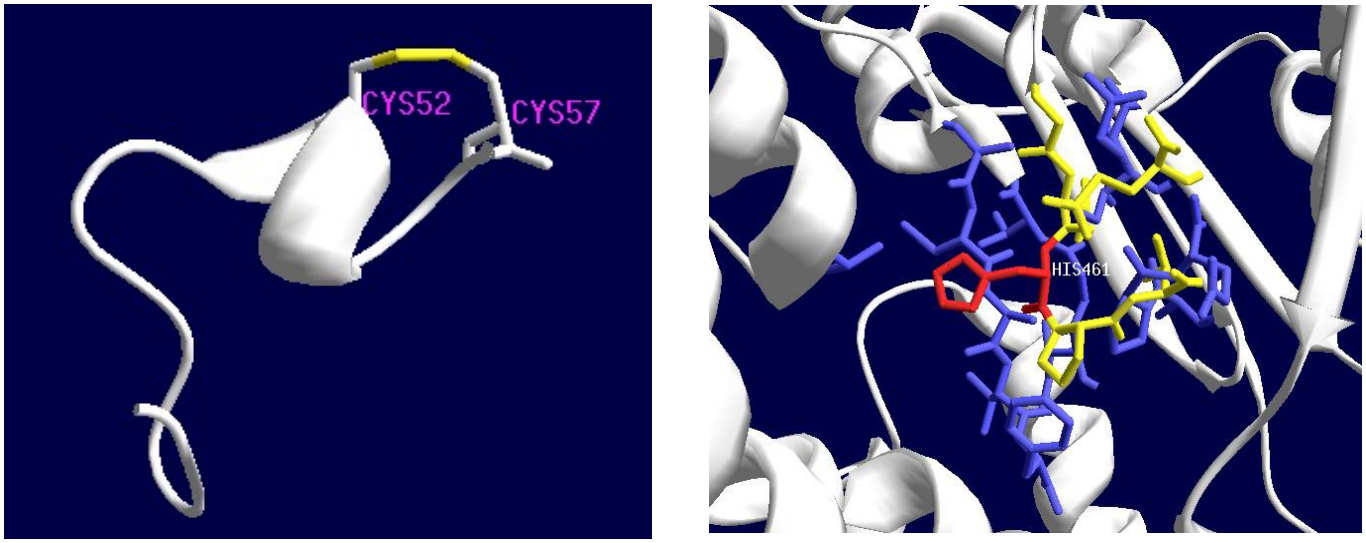

Figure6 A (Left): Redox active disulfide bond (yellow) between cysteine 52 and cysteine 57.

Figure6 B (Right): Active site around histidine 461 (red) at interface of chain a (blue residues) and chain b (yellow residues).

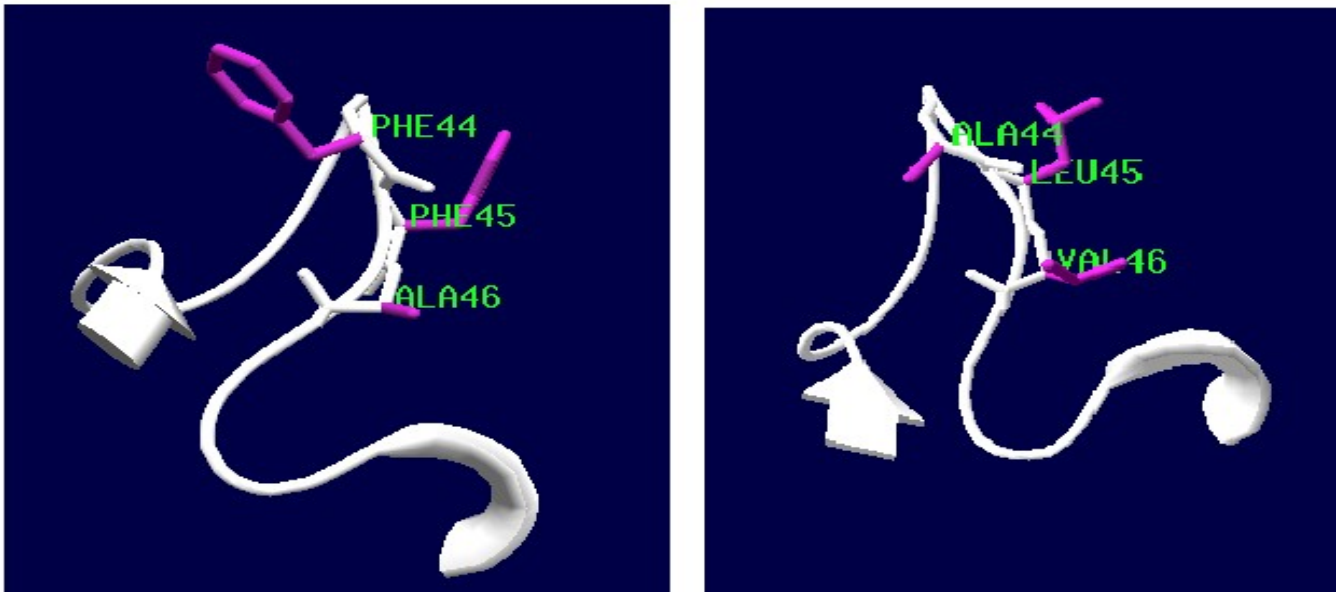

Figure6 C: Comparison of mutant region (pink) of FAD binding domain (white ribbons).

REFERENCES

1. Bern, C; Maguire, JH and Alvar, J (2008), "Complexities of Assessing the Disease Burden Attributable to *Leishmaniasis*", ***PLoS Neglect. Trop. Dis,*** Vol.2, 313.
2. Desjeux, P (2004), "*Leishmaniasis*: current situation and new perspectives", ***Comp. Immunol. Microbiol. Infect. Dis.***, Vol.27, 305-318.
3. Reithinger, R (2008), "*Leishmaniases* burden of disease: ways forward for getting from speculation to reality", ***Plos Neglect. Trop. Dis***, Vol.2, 285.
4. Mithani, A; Preston, GM and Hein, J (2009), "Rahnuma: hypergraph-based tool for metabolic pathway prediction and network comparison", ***Bioinformatics***, Vol. 25 (14), 1831-1832
5. Kanehisa, M; Goto, S; Hattori, M and Aoki-Kinoshita, KF *et al.* (2006), "From genomics to chemical genomics: new developments in KEGG", ***Nucleic Acids Res.***, Vol.34, 354–357.
6. Jeanmougin, F; Thompson, JD; Gouy, M and Higgins, DG *et al.* (1998), "Multiple sequence alignment with Clustal X", ***Trends Biochem Sci***, Vol.23, 403-405.
7. Nicholas, KB; Nicholas, HBJ and Deerfield, DWI (1997), "GeneDoc: analysis and visualization of genetic variation", ***EMBnet NEWS***, Vol. 4, 14.
8. Sali, A and Blundell, TL (1993), "Comparative modelling by satisfaction of spatial restraints", ***J. Mol. Biol***, Vol.234, 779-815.
9. Wallner, B and Elofsson, A (2005), "All are not equal: A benchmark of different homology modeling programs", ***Protein Sci.***, Vol.14 (5), 1315-1327.
10. Xiang, Z (2006), "Advances in Homology Protein Structure Modeling", ***Curr Protein Pept Sci.***, Vol.7 (3), 217-227.
11. Boeckmann, B; Bairoch, A; Apweiler, R and Blatter, MC *et al.* (2003), "The Swiss-Prot Protein Knowledgebase and its supplement TrEMBL in 2003", ***Nucleic Acids Res.***, Vol.31, 365-370.
12. Altschul, SF; Gish, W; Miller, W and Myers, EW *et al.* (1990), "Basic local alignment search tool." ***J Mol Biol.***, Vol.215 (3), 403-410.
13. Westbrook, J; Feng, Z; Chen, L and Yang, H *et al.* (2003), "The Protein Data Bank and

structural genomics." ***Nucleic Acids Res.***, Vol.31, 489-491.

14. Arnold, K; Bordoli, L; Kopp, J and Schwede, T (2006), "The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling", ***Bioinformatics***, Vol.22, 195-201.
15. Schwede, T; Kopp, J; Guex, N and Peitsch, MC (2003), "SWISS-MODEL: an automated protein homology-modeling server", ***Nucleic Acids Res.***, Vol.31, 3381-3385
16. Madhusudhan, MS; Martí-Renom, MA; Sanchez, R and Sali, A (2006), ***Prot. Eng. Des. & Sel.*** Vol.19, 129-133.
17. Needleman, SB and Wunsch, CD (1970), "A general method applicable to the search for similarities in the amino acid sequence of two proteins", ***J. Mol. Biol.***, Vol.48, 443-453.
18. Van Gunsteren, WF; Brunne, RM; Gros, P and van Schaik, RC *et al.* (1994), "Accounting for molecular mobility in structure determination based on nuclear magnetic resonance spectroscopic and X-ray diffraction data", ***Methods Enzymol.***, Vol.239, 619-654.
19. Guex, N and Peitsch, MC (1997), "SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling", ***Electrophoresis***, Vol.18, 2714-2723.
20. Laskowski, RA; MacAurther, MW; Moss, DS and Thornton, JM (1993), "Procheck-a program to check the stereochemical quality of protein structures", ***J Appl Cryst***, 26, 47-60.
21. Colovos, C and Yeates, TO (1993), "Verification of protein structures: patterns of nonbonded atomic interactions", ***Protein Sci. Sep***, Vol.2 (9), 1511-1519.
22. Eisenberg, D; Lüthy, R and Bowie, JU (1997), "VERIFY3D: assessment of protein models with three-dimensional profiles", ***Methods Enzymol.***, Vol. 277, 396-404.
23. Hooft, RW; Vriend, G; Sander, C and Abola, EE (1996), "Errors in Protein Structures" ***Nature***, Vol.381 (6580), 272.
24. Castrignano, T; De Meo, PD; Cozzetto, D and Talamo, IG *et al.* (2006), "The PMDB Protein Model Database", ***Nucleic Acids Res.***, Vol.34, D306-D309