Modeling and Docking of Cysteine Protease-A (CPA) of Leishmania donovani

Sindhuprava Rana[1*], Jay Prakash Mahato[1], Maneesh Kumar[2], Surendra Sarsaiya[1]

[1]Department of Biotechnology, Sri Satya Sai University of Technology & Medical Sciences, Bhopal, Madhya Pradesh, India.
[2]Department of Biotechnology, College of Commerce, Arts & Science, (Magadh University, Bodh Gaya), Patna-800020, Bihar, India.

ARTICLE INFO

ABSTRACT

Leishmaniasis is a tropical disease and causes severe public health issues. The disease caused by parasitic protozoan *Leishmania* species. Many people around the globe are affected in dramatic way, which would be fatal if untreated. Drugs have been made that become resistant to present strains. Cysteine proteases (CPs) are common protein degrading cysteine protease enzymes present in *Leishmania donovani*. In present scenario, to achieve effective enzyme inhibition mechanism, it should be taken into consideration to develop some effective and novel inhibitors to block the binding sites of that protein. Homology modeling of cysteine protease has been done and docked with suitable inhibitors by using various servers and computational tools. The model was designed, compared and validated by DOPE and Verify 3D scores by using DSv3.5. Licochalcone-a alone showed 37 LibDock conformations with 6 different poses were suitably docked at the site 1. The study would help to design the protein with new drugs in respect of resistant one for the treatment of harmful visceral leishmaniasis.

Abbreviations: CHARMM: Chemistry at HARvard Macromolecular Mechanics; **CPs:** Cysteine proteases; **DOPE:** Discrete Optimized Protein Energy; **DSv3.5:** Discovery Studio 3.5; **I-TASSER:** Iterative Threading ASSEmbly Refinement; **NCBI:** National Center for Biotechnology Information; **pBLAST:** protein Basic Local Alignment Search Tool; **PDB:** Protein Data Bank; **RMSD:** Root Mean Square Deviation; **VL:** visceral leishmania.

INTRODUCTION

Visceral Leishmaniasis (VL) is a chronic tropical disease, responsible for severe public health concern around the world. The parasitic species mainly involve in this infection are belong to protozoan *Leishmania donovani* (Das *et al*., 2013). This serious infection is basically mediated by sand flies (*Phleobotomus argentipes*) that transmit the disease to about 12 million people as only reservoir and is endemic in the tropical regions of Africa, America, the Indian subcontinental and subtropical regions of Southeast Asia (Selzer *et al*., 1999).

* Corresponding Author
Surendra Sarsaiya, Department of Biotechnology, Sri Satya Sai
University of Technology & Medical Sciences, Bhopal, Madhya Pradesh,
India. Email: ganeshiitkgp @ gmail.com

The survival of the sand fly vector depends on many biotic (human/ animal blood meals) and abiotic (precipitation and temperature) factors. Such conditional habitat found near Gangetic plains (Ostyn *et al*., 2015). *L. donovani* is an active causative agent that causes human tegumentary and VL that also known as kalazar, in local regions. It would be fatal for human survival if not properly diagnosed and treated well on time (Alvar *et al*., 2012, Silva-Almeida *et al*., 2012). In past decades, many approached have been to create effective vaccine but not much success would achieve in its cure so far (Aït-Oudhia *et al*., 2011, Bhattacharya *et al*., 2016). Cysteine proteases (CPs) are common protein degrading enzymes that cleave various peptides. They are the essential virulence factors in leishmania parasitic. They play key role in many infectious protozoa and viral agents (Mahmoudzadeh-Niknam and McKerrow, 2004; Vicik *et al*., 2006).

Due to its biological importance it would become the potential target to develop new drug to cure the leishmaniasis (Das and Ali, 2014). CPs can also participate in various other cellular metabolisms like excystment, exsheathing and, some parasite immunoevasion and tissue invasion. Many parasitic CPs have some potential immunogenic properties, which have been exploited for diagnostic purposes and vaccine targets (Sajid and McKerrow, 2002). In the *Leishmania*, the mechanism of nucleophilic cysteine thiol is involved catalytic mechanism in catalytic dyad. So, the creation of novel inhibitors is the use of electrophilic moieties that covalently bind to the cysteine residue of the active site. The present aim to achieve effective enzyme inhibition, this should be taken into consideration that this electrophile would incorporate into peptidic or peptidomimetic sequences present in the substrate binding sites. In *L. donovani*, CPs is the virulence factor in human and animals (McKerrow, 1999; Mahmoudzadeh-Niknam and McKerrow, 2004; Rosenthal, 2004).

Many inhibitors are used for docking purpose that potentially inhibits the cysteine protease of *L. donovani*. Among them licochalcone-A is a novel estrogenic, natural phenolic compound that has antimalarial, anticancer, antitumor and antiviral activities. It also induces apoptosis at some instant that prevents any biological harm to the host cell (Fu *et al.*, 2004). Primin is another naturally organic compound has clear antimicrobial activities (Brondani *et al.*, 2007). Sitamaquine and Miltefosine are the best known drugs used in the treatment of VL through oral route (Sundar and Olliaro, 2007; Loiseau *et al.*, 2011). Amarogentin is an smart organic compound from *Gentiana lutea*, a herbal bitter is generally used in digestive disorders (Keil *et al.*, 2000). Emetine is an effective anti-protozoal alkaloid and induces vomiting. It is also used in the treatment of severe invasive amoebiasis (Kattwinkel, 1949). Fisetin is the active flavonoid, found in various fruits and green vegetables, generally used against numerous harmful diseases. It is an good anti-inflammatory and antioxidant compound (Khan *et al.*, 2013). In this present work, various tools and softwares have been used to understand in silico biological role of cysteine protease of *L. donovani* within the affected host. The structured cysteine protease model would give the proper idea about the protein-ligand interaction. The work clears the selection of some novel drugs which might paved the way to design them for the betterment of human society against leishmaniasis.

METHODS AND MATERIALS

3D Structure Prediction and Model Prediction

The amino acid sequence of cysteine protease (accession: AAL09448) of *Leishmania donovani* has taken from the protein database of NCBI (Sayers *et al.*, 2011). A BLASTP search was done (Altschul *et al.*, 1997) against Protein Data Bank (PDB) (Berman *et al.*, 2000).This protein sequence has used for our homology modeling, computational assessment and to predict the suitable protein-ligand interaction. The 3D structure of the

cysteine protease of *L. donovani* was performed by using various servers online servers like HMM-HMM comparison (HMpred) (Söding, 2005; Remmert *et al.*, 2012), the hierarchical method of protein structure and function (I-TASSER) (Zhang, 2008), the protein structure prediction (Raptor X) (Källberg *et al.*, 2012), the knowledge based approach (Swiss Model) (Arnold *et al.*, 2006).

Protein Stimulation and Active Sites Prediction

The best structure of cysteine protease that was modeled by Swiss model web server, based on the DOPE Score (Shen and Sali, 2006) was selected with lowest value for further study. The actual positions and nature of amino acids within the modeled structure was evaluated by Ramachandran Plot via PROCHECK (Laskowski *et al.*, 1996), ERRAT (Colovos and Yeates, 1993) and Verified 3D (Bowie *et al.*, 1991; Luthy *et al.*, 1992). The protein was also analyzed by verify Protein (Profiles 3D) from DSv3.5. Along with these, other validation like refinement of flexible loop and side chain were done by Discovery Studio version 3.5 (DSv3.5) (Inc, 2012). The refinement was followed by protein stimulation via CHARMM (Brooks *et al.*, 1983) with the help of DSv3.5.

The CHARMM is a dynamic in silico molecular computational program that minimize the energy value during docking while protein has been stimulated under the effect of dielectrics of the atoms of amino acids (Sahoo *et al.*, 2009, Sahoo *et al.*, 2014). The model structure of the cysteine protease with all its computational parameters were further taken for the docking purpose using the same DSv3.5.

Docking

The accurate identification and characterization of binding sites of the modeled cysteine protease of the *L. donovani* have been identified by using the DSv3.5. The modeled protein was selected and allowed for Protein Preparation under the effect of CHARMM force, followed by cavities identification as active sites where the selected ligands have allowed docking. The three active sites were shown by DSv3.5, in which only the site1 has actively participated for the docking work. Site2 and site3 haven't shown any docked score. After the protein preparation has been performed, the ligand library was made by using different ligands like Licochalcone-A, Primin, Sitamaquine, Miltefosine, Amarogentin, Emetine and Fisetin with their known IC50 value (Ansari *et al.*, 2012) were taken from NCBI PubChem database. the docking were performed by LibDock algorithm inbuild in DSv3.5 (Rao *et al.*, 2007).

RESULTS AND DISCUSSION

The computational analysis using both software and servers of the modeled protein structure enable us the biological properties. The cysteine protease of *L. donovani*, with the hypothetical protein modeled that was prepared and retrieved from Swiss model tool.

Modeled Structure prediction using homology modeling approaches

The proper template identification after pBLAST, followed by modeled protein quality assessment enables to understand the suitable functions, and localization of the sub-cellular protein and their interaction with ligands. The most common and approximate prediction method is homology modeling that provides a proper idea about the protein. The homology model of cysteine protease is shown in **figure 1**. The figure labeled as sequence alpha (α), beta (β) and flexible loops (FL). The cysteine protease consists of seven alpha helixes, ten beta sheet, three η (small helixes) and few flexible loops. The secondary sequence was clearly mentioned the actual amino acids residues involved in making alpha helixes, beta sheet, flexible loops and small helixes (with three amino acids residues) [**Figure 2**]. All the models which were designed from the different

servers were compared and validated by DOPE scores by using DSv3.5. The suitable template model that retrieved from the Swiss model has been taken for our study which has the lowest DOPE value of -43031.174938 as the best-modeled structure. The model structure was proved by Verify 3D that showed 83.06 % value. The protein verify by using DSv3.5, shows the 111.74 verify score which lies in between verify expected high score (139.723) and verify expected low score (62.8754). The model validation PROCHECK tool was used to determine Ramachandran plot [**Figure 3**] that assures the quality model. The result of Ramachandran plot showed the amount of amino acid residues i.e. 85% favored regions, 13.1% additional allowed regions, 1.5% generously allowed regions and 0.4% disallowed regions favored represented as a reliable and good quality model. The reliability of the modeled protein was also checked by using ERRAT that showed 83.162 overall model quality [**Figure 4**].
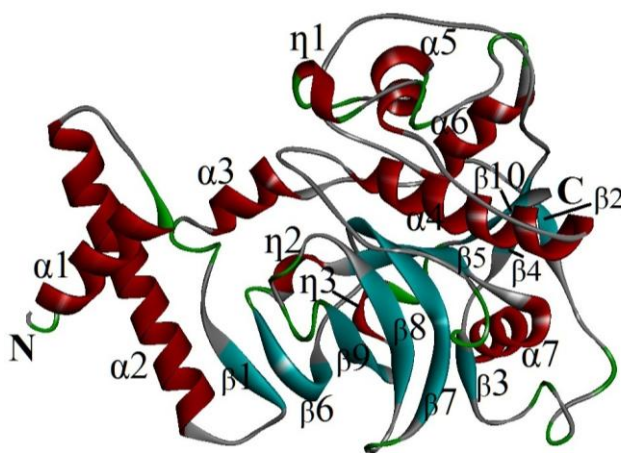
Fig 1: 3D Model Cysteine Protease was produced by DSv3.5 having α helices are in red, β sheets are in blue and flexible loops are in grey.
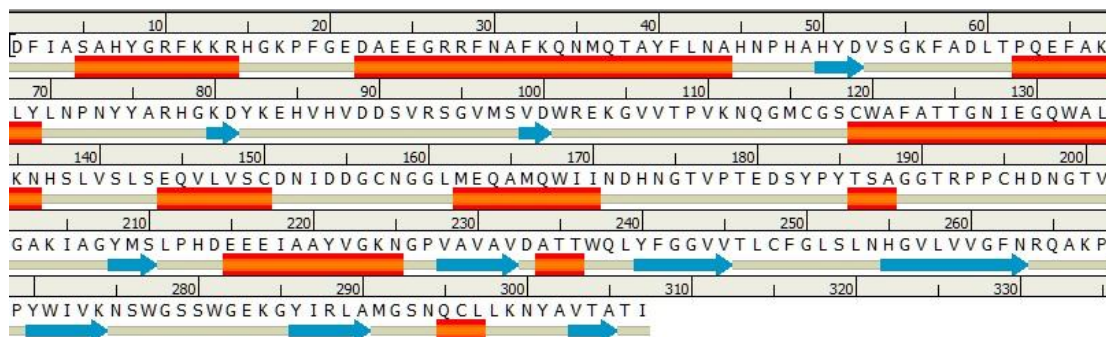
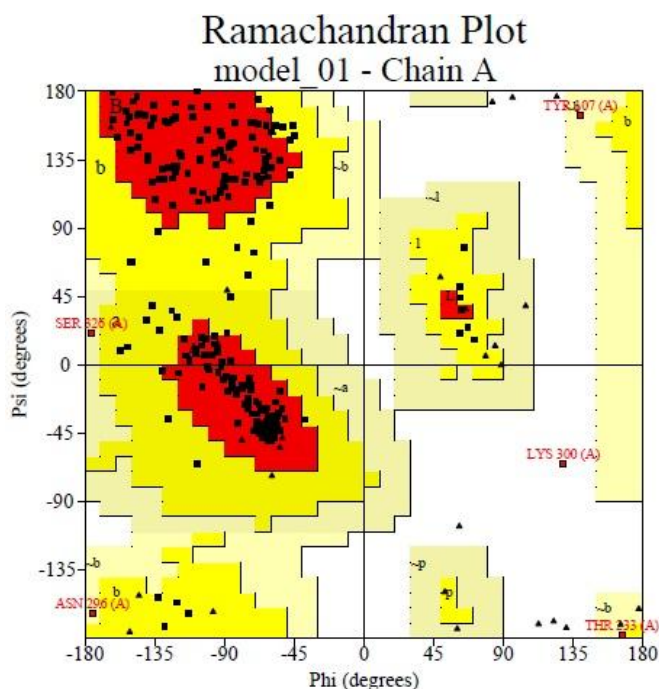**Fig 2:** Secondary Structure of the modeled cysteine protease by DSv3.5.
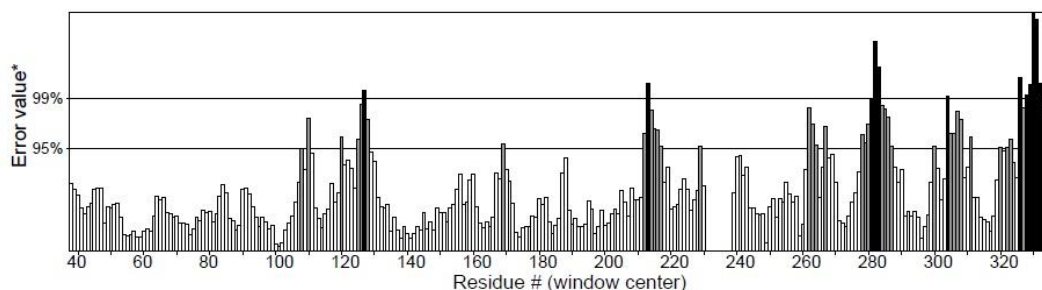
Fig 3: Ramachandran Plot by PROCHECK.


Fig 4: Overall quality checked by ERRAT.

Protein-Ligand Interaction

After assessment of suitable cysteine protease protein model detecting by different computational tools and servers, some active binding sites of the model cysteine protease were created, then we tried to analyze the specific binding affinity of some suitable substrate ligands that were effectively docked with the 3D modeled structure. There are three different binding sites having charged amino acids, were detected by using receptor cavities protocol of DSv3.5. The Licochalcone-A has a highest LibDock score of 71.728 also showed the absolute energy value of 86.5400 for the binding site 1 (the position value of the site 1: -28.787, 22.2787, 20.1339, 19.6) of the model protein. Licochalcone-A has showed 37 poses of LibDock conformations and 5 hits, in which only 6 different poses were suitably docked at the site 1. Licochalcone-A along with some others ligands which are also perfectly bound at this site 1 are providing the LibDock scores [**Table 1**], along with their LibDock Score. They are Primin, Sitamaquine, Miltefosine, Amarogentin, Emetine and Fisetin also exhibited high binding activities with the receptor protein. The binding between model cysteine protease protein and

ligand Licochalcone-A binding was shown in **Figure 5**. The subcellular localization of the cysteine protease protein is present within the extracellular regions and some part of periplasmic and cytoplasmic also. The proper Protein-Ligand interaction is shown in **Figure 6.** Here, the frequently involved amino acid of model protein that forms hydrogen bonds with the ligand is His 81. The group i.e.: ND1 of His 81 interacts with :O4 of Licochalcone-A. Our docking result and the structure analysis suggest that the model protein having active binding site to ligand properly. The study would help to design the new drugs in respect of resistant one for the cure of leishmaniasis.

**Table 1:** Comparative LibDock Scores of different effective ligands

Ligands	LibDock Score	Absolute Value	Relative Value
Licochalcone-A	71.728	86.5400	13.3502
Primin	71.5692	27.9008	09.4004
Sitamaquine	70.4248	63.1020	17.5402
Miltefosine	69.6103	27.3418	12.7868
Amarogentin	68.3701	90.1020	18.6526
Emetine	67.4750	81.6668	16.2672
Fisetin	66.8878	40.5343	00.0039
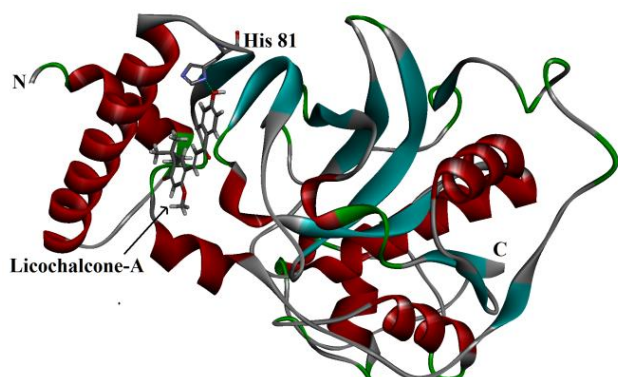
Fig 5: 3D Model of cysteine protease of L. donovani was produced by DSv3.5
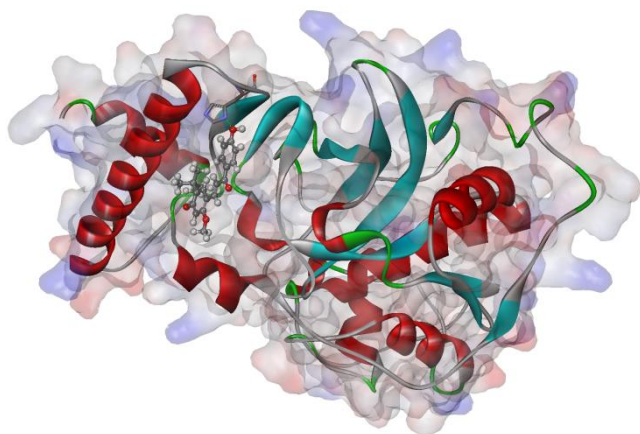
Fig 6: 3D representation of subcellular localization of the model protein within cell.
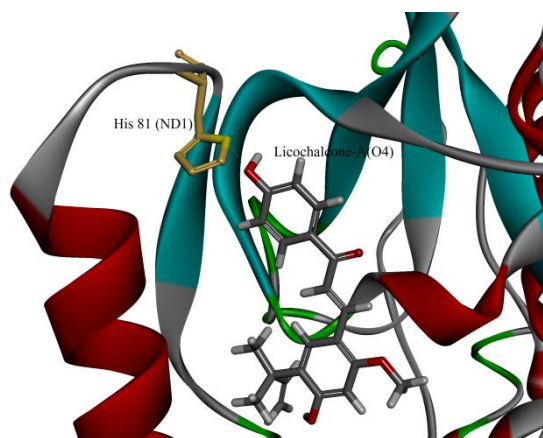
Fig 7: The Close Interaction of Amino Acids of cysteine protease :ND1 (orange) of His 81, and :O4 (Green) of Licochalcone-A

ACKNOWLEDGMENTS

I thank all my colleagues from Biotechnology Department for their humble guidance and support during my work.

Conflict of Interests: There are no conflicts of interest.

REFERENCES

Aït-Oudhia K, Gazanion E, Vergnes B, Oury B Sereno D. Leishmania antimony resistance: what we know what we can learn from the field. Parasitology research, 2011; 109(5): 1225.

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic acids research, 1997; 25(17): 3389-3402.

Alvar J, Vélez ID, Bern C, Herrero M, Desjeux P, Cano J, Jannin J, den Boer M, Team WLC. Leishmaniasis worldwide and global estimates of its incidence. PloS one, 2012; 7(5): e35671.

Ansari MY, Dikhit MR, Sahoo GC, Das P. Comparative modeling of HGPRT enzyme of L. donovani and binding affinities of different analogs of GMP. International journal of biological macromolecules, 2012; 50(3): 637-649.

Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics, 2006; 22(2): 195-201.

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The protein data bank. Nucleic acids research, 2000; 28(1): 235-242.

Bhattacharya P, Dey R, Dagur PK, Joshi AB, Ismail N, Gannavaram S, Debrabant A, Akue AD, KuKuruga MA, Selvapandiyan A. Live Attenuated Leishmania donovani Centrin Knock Out Parasites Generate Non-inferior Protective Immune Response in Aged Mice against Visceral Leishmaniasis. PLoS Negl Trop Dis, 2016; 10(8): e0004963.

Bowie JU, Luthy R, Eisenberg D. A method to identify protein sequences that fold into a known three-dimensional structure. Science, 1991; 253(5016): 164-170.

Brondani D, Nascimento C, Moreira DDM, Leite AL, de Souza I, Bieber L. Synthesis and Antitumour Activity of the Primin (2-methoxy-6-n-pentyl-1, 4-benzoquinone) and Analogues. Medicinal Chemistry, 2007; 3(4): 369-372.

Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. Journal of computational chemistry, 1983; 4(2): 187-217.

Colovos C, Yeates TO. Verification of protein structures: patterns of nonbonded atomic interactions. Protein science, 1993; 2(9): 1511-1519.

Das A, Ali N. Combining cationic liposomal delivery with MPL-TDM for cysteine protease cocktail vaccination against Leishmania donovani: evidence for antigen synergy and protection. PLoS Negl Trop Dis, 2014; 8(8): e3091.

Das P, Alam MN, Paik D, Karmakar K, De Chakraborti T. "Protease inhibitors in potential drug development for Leishmaniasis". Indian J Biochem Biophys, 2013; 50 (5): 363-76.

Fu Y, Hsieh T-C, Guo J, Kunicki J, Lee MY, Darzynkiewicz Z, Wu JM. Licochalcone-A, a novel flavonoid isolated from licorice root (Glycyrrhiza glabra), causes G2 and late-G1 arrests in androgen-independent PC-3 prostate cancer cells. Biochemical and biophysical research communications, 2004; 322(1): 263-270.

Inc A. Discovery studio modeling environment, release 3.5. Accelrys Discovery Studio. Accelrys Software Inc, San Diego, 2012.

Källberg M, Wang H, Wang S, Peng J, Wang Z, Lu H, Xu J. "Template-based protein structure modeling using the RaptorX web server." Nature protocols, 2012; 7(8): 1511-1522.

Kattwinkel EE. Death due to cardiac disease following the use of emetine hydrochloride in conditioned-reflex treatment of chronic alcoholism. New England Journal of Medicine, 1949; 240(25): 995-997

Keil M, Härtle B., Guillaume A, Psiorz M. Production of amarogentin in root cultures of Swertia chirata. Planta medica, 2000; 66(05): 452-457.

Khan N, Syed DN, Ahmad N, Mukhtar H. Fisetin: a dietary antioxidant for health promotion. Antioxidants & redox signaling, 2013; 19(2): 151-162.

Laskowski RA, Rullmann JAC, MacArthur MW, Kaptein R, Thornton JM. AQUA and PROCHECK-NMR: programs for checking the

quality of protein structures solved by NMR. Journal of biomolecular NMR, 1996; 8(4): 477-486.

Loiseau P, Cojean S, Schrével J. Sitamaquine as a putative antileishmanial drug candidate: from the mechanism of action to the risk of drug resistance. Parasite, 2011; 18(2): 115-119.

Luthy R, Bowie JU, Eisenberg D. Assessment of protein models with three-dimensional profiles. Nature, 1992; 356(6364): 83.

Mahmoudzadeh-Niknam H, McKerrow J. Leishmania tropica: cysteine proteases are essential for growth and pathogenicity. Experimental parasitology, 2004; 106(3): 158-163.

McKerrow JH. Development of cysteine protease inhibitors as chemotherapy for parasitic diseases: insights on safety, target validation, and mechanism of action. International journal for parasitology, 1999; 29(6): 833-837.

Ostyn B, Uranw S, Bhattarai NR, Das ML, Rai K, Tersago K, Pokhrel Y, Durnez L, Marasini B, Van der Auwera G . Transmission of Leishmania donovani in the hills of Eastern Nepal, an outbreak investigation in Okhaldhunga and Bhojpur districts. PLoS Negl Trop Dis, 2015; 9(8): e0003966.

Rao SN, Head MS, Kulkarni A, LaLonde JM Validation studies of the site-directed docking program LibDock. Journal of chemical information and modeling, 2007; 47(6): 2159-2171.

Remmert M, Biegert A, Hauser A, Söding J. HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. Nature methods, 2012; 9(2): 173-175.

Rosenthal PJ. Cysteine proteases of malaria parasites. International journal for parasitology, 2004 34(13): 1489-1499.

Sahoo GC, Dikhit MR, Rani M, Das P. Homology modeling and functional analysis of LPG2 protein of Leishmania strains. J Proteomics Bioinform, 2009; 2: 32-50.

Sahoo GC, Ansari MY, Dikhit MR, Kannan M, Rana S, Das P. Structure prediction of gBP21 protein of L. donovani and its molecular interaction. Journal of Biomolecular Structure and Dynamics, 2014; 32(5): 709-729.

Sajid M, McKerrow JH. Cysteine proteases of parasitic organisms. Molecular and biochemical parasitology, 2002; 120(1): 1-21.

Sayers EW, Barrett T, Benson DA, Bolton E, Bryant SH, Canese K, Chetvernin V, Church DM, DiCuccio M, Federhen S. Database resources of the national center for biotechnology information. Nucleic acids research 39(suppl 1), 2011; D38-D51.

Selzer PM, Pingel S, Hsieh I, Ugele B, Chan VJ, Engel JC, Bogyo M, Russell DG, Sakanari JA, McKerrow JH. Cysteine protease inhibitors as chemotherapy: lessons from a parasite target. Proceedings of the National Academy of Sciences, 1999; 96(20): 11015-11022.

Shen MY, Sali A. Statistical potential for assessment and prediction of protein structures. Protein science, 2006; 15(11): 2507-2524.

Silva-Almeida M, Pereira BAS, Ribeiro-Guimarães ML, Alves CR. Proteinases as virulence factors in Leishmania spp. infection in mammals. Parasites & vectors, 2012; 5(1): 160.

Söding J. Protein homology detection by HMM–HMM comparison. Bioinformatics, 2005; 21(7): 951-960.

Sundar S, Olliaro PL. Miltefosine in the treatment of leishmaniasis: clinical evidence for informed clinical risk management. Therapeutics and clinical risk management, 2007; 3(5): 733.

Vicik R, Busemann M, Baumann K, Schirmeister T. Inhibitors of cysteine proteases. Current topics in medicinal chemistry, 2006; 6(4): 331-353.

Zhang Y. I-TASSER server for protein 3D structure prediction. BMC bioinformatics, 2008; 9(1): 40.