DU Journal of Undergraduate Research and Innovation

Volume 1, Issue 3 pp 85-95, 2015

Plasmepsin II as a Potential Drug Target for Resistant Malaria

Megha Gulati*[1], Aruna Narula[2], Raj Vishnu[3], Gunjan Katyal[4], Arti Negi[5], Isbah Ajaz[6], Kritika Narula[7], Gunjan Chauhan[8], Ravi Kant[a], Vanshika Lumb[9], Smriti Babbar[10], Neena R. Wadehra[11], Deepika Bhaskar[b]

*meghagulati94@gmail.com, [1-11]Department of Biochemistry, Shivaji College, University of Delhi 110027, [a] University of Delhi South Campus, [b] Research Council, University of Delhi

ABSTRACT

With drug resistance becoming extensively pervasive in *Plasmodium falciparum* infections, research for alternative drugs is becoming mandatory for prevention and cure of malaria. Increased resistance against anti malarials such as chloroquine and sulfadoxin/pyrimethamine, has resulted in developing new drug therapies . Aspartic proteases called plasmepsin are present in different species of *Plasmodium*. With the use of *in silico* structure-based drug design approach, the differences in binding energies of the substrate and inhibitor were exploited between target sites of parasite and human. The docking studies show several promising molecules from GSK library with more effective binding as compared to the already known inhibitors for the drug targets. Stronger interactions are shown by several molecules as compared to the reference molecules which have shown to be the potential as drug candidates.

Key Words Aspartic protease, Drug resistance, Drug targets, Inhibitors, *in silico* studies, Plasmepsin, *Plasmodium falciparum*

INTRODUCTION

Malaria, a global contagious disease, causes over one million deaths per year. Malaria is caused by a protozoan called *Plasmodium* which has four species namely *P. falciparum, P. ovale, P. malariae and P. vivax.* Out of these four species, *P. falciparum* causes the most lethal type of malaria (2, 4, and 5). This virulent form of *Plasmodium* affects humans with approximately 660,000 deaths per year as per WHO estimates of 2010. Prevention of this disease using anti-malarial vaccine has not been successful. Till date the treatment and prevention of this disease is primarily based on anti-malarial drug administration and anti-vector measures.

Including artemisinin and its derivatives, the parasite has evolved drug resistance against almost all known anti-malarial chemotypes (5). The efficiency of anti-malarial drugs is diminishing due to the ability of *Plasmodium* species to develop drug resistance. So, there is an earnest need to pay attention in inducing the antimalarial drugs that can act through multiple specific mechanisms (5).

For the development of malarial parasite, the degradation of hemoglobin is a necessary step. This degradative process works as an unaffected antimalarial drug (3). Hemoglobin catabolism occurs in the food vacuoles catalyzed by the enzyme aspartic proteases, also famous as Plasmepsins (PMs) (9), previously called as hemoglobinases (6). There are four histo aspartic proteases (PMI, PMII, HAP, and PMIV) out of ten, closely related but vary in specificity, that play a role in catabolism of hemoglobin (2,3,6,7,8). Through their hemoglobin degrading activity, they are the potential target for antimalarial drugs. According to the studies PMII is the effective drug target for treatment of malaria. PMII has two catalytic aspartases D34 and D214. PMII degrade hemoglobin by proteolytic cleavage and also has in strange specificity towards nascent hemoglobin (10).

PMII has its crystal structure easily available in online databatases. Structural, biological, biochemical and inhibition properties of PMII were studied and validated by the crystal structure of our reference PDB ID to proceed our docking studies on the same. PMII shows 33% homology with human protein cadhespin and renin, therefore it was assumed to be an ideal drug target for our studies. We then extracted and screened the antimalarial lead compounds of GSK to get top ligand hits on the basis of ligscore.

To identify the drug molecule against *Plasmodium* three antimalarial drug targets were selected on the basis of their docking procedure, which are essential part of *Plasmodium falciparum* and absent in the host (11). Selected target were plasmepsin, TS -DHFR, and Phospho-ethanolamine methyltranferase. These targets were virtually screened against the GSK library with the filtration of 2um IC50 value.

METHODOLOGY

Various essential metabolic pathways in malarial parasite *Plasmodium falciparum* were studied. The initial 546 proteins were selected on the basis of literature study and most of them were taken from Medicine for Malaria Venture (MMV), www.mmv.org based on relevant information of the listed targets, specifically their 3-D structures in ".pdb" format, essential functions, active sites, active site residues, available ligands & inhibitors, bioinformatics software involved and binding energy value. The targets were shortlisted majorly on the basis of their availability of crystal structure on PDB database and minimum sequence homology with humans; also information of their structural, biological and biochemical interactions with inhibitors was collected as a basis of their selection by using various databases such as PDB, UniPROT, Pubmed and PlasmoDB. Further, 37 putative targets were shortlisted out of 546 proteins which were selected initially from Medicine for Malaria Venture (MMV), www.mmv.org and GSK library for studying homology with human targets using BLAST. Nine targets with homology, less than 40%, were considered for further study. The shortlisted drug targets were - Dihydrofolatereductase (DHFR), Choline kinase, N-methyl transferase,

Plasmepsin 2, Peptide deformylase, Enoyl acyl carrier protein reductase, M1 family aminopeptidase, uridinephosphorylase, and orotatephosphoribosyltransferase- they are part of major metabolic pathways of *Plasmodium falciparum*. Three targets were finally selected to perform docking studies with 13469 leads predicted by GSK against malaria from ChEMBL and TDR targets database v5 using Discovery Studio 2.0. Finally the compounds which were having top hits were validated by AutoDock 1.5.4.

Compounds of the GSK library were extracted from ChEMBL database. Screening of the lead compounds, extracted from ChEMBL, was done using Discovery Studio 2.0 for Plasmepsin II. The top hit compounds were selected on the basis of the LigScore 1 and 2. Validation of the top hit compounds recorded from Discovery Studio was done using AutoDock. The software for validation procedures were **Pymol**: to view the 3D structure of the protein and prediction of binding site, **Chemsketch**: to draw the structure of the ligand/molecule, **Open Babel**: to convert the format of the file from .MOL to .PDB file, **AutoDock1.5.4**: Docking software used for validation, **Cygwin**: to create .glg and .dlg file by running docking algorithm and **UCSF Chimera**: to visualize and analyze H-bonds. The X-ray crystal structures of Plasmepsin II were co-complexed with their inhibitors and cofactors were obtained from pdb. All the docking procedures (validation and library screening) were done using Discovery Studio. The pdb files of Plasmepsin II were taken and their ligand (inhibitors and substrates) was extracted and water molecules were removed. Hydrogen molecules were added and force field was applied until the binding constants of inhibitors matched their reported value. The receptor was minimized and the binding site was identified by extracting the ligand in the pymol and then validating the docking with the same interacting residues. The protein with the characterized binding site was taken for further docking studies and screening the library.

RESULTS AND DISCUSSION

Based on the literature study and information available online, 546 proteins were identified as potential target proteins which are present in the metabolic pathways of Plasmodium falciparum. Further, out of these 546 proteins 37 proteins were shortlisted based on literature study, essentiality for survival of the parasite, structural and functional data available online. Putative targets and proteins with unknown structures were eliminated at this stage. The shortlisted 37 proteins were then tested for homology with similar human targets using NCBI BLAST. Proteins with maximum homology up to only 40% were selected for further *in-silico* studies. Based on the results obtained only 9 proteins were considered as potential drug targets out of which Plasmepsin II was selected for further studies due to its vitality for the survival of parasite.

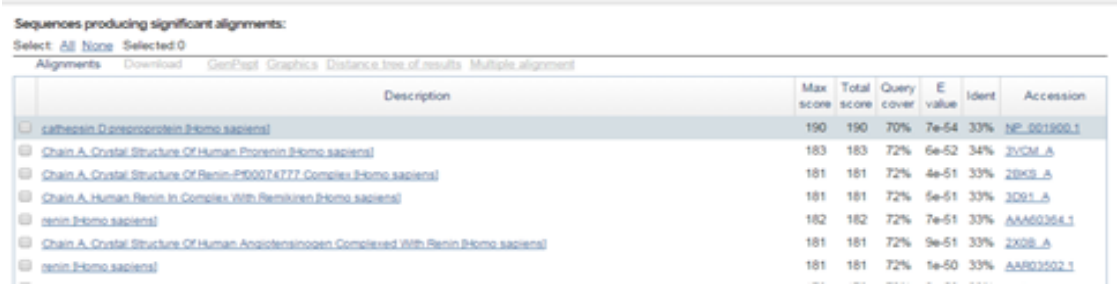

Figure 1 Homology of Plasmepsin II with *Homo sapiens*.

Fig 1 shows the result of homology studies performed against *Homo sapiens*. The above figure comprises of different columns like Max score, Total score, E value, Identity and Accession. 9 proteins were selected on the basis of percentage of identity and based on the attributes like structural, biological, biochemical and inhibition properties of PMII.

Table 1 Plasmepsin II attributes used *in silico* studies [2, 12, 13, 14, 15]

Protein structure and biology	Description
Protein ID	2IGY, IJ8J, 1LEE, 1LF2, 1LF3, 1LF4, 1ME6, 1PFZ, 1SME, 1W6H, 1W6I, 1XDH 1XE5, 1XE6, 2BJU, 2BL3, 2IGX, 2R9B, 3F9Q, 4CKU
Uniprot ID	P46925- PLM2_PLAFA
Sub cellular location	food vacuole of plasmodium(2)
Protein name	Plasmepsin 2
Length of protein	329(13)
Protein family	aspartic protease(14)
No. of subunits	2(14)
No. of subunits	2(PDB ID 2IGY)
Length of subunit	A-158, B-171(13)
No. of binding site	2(14)
Number of binding sites	2(PDB ID 2IGY)
Protein function	Hydrolysis of the bonds linking particular hydrophobic residues in hemoglobin or globin present in host cell. Can also cleave small molecules substrates such as Ala-Leu Glu-Arg-Thr-Phe-|-Phe(NO2)-Ser-Phe-Pro

	Thr(14)
Name of the coenzyme/prosthetic group	None
Essential for survival or virulence?	essential for survival because it can provide main source of amino acids for growth and maturation of the plasmodium by host cell Hb degradation(12)
Action mechanism	cleaves in the B helix of the alpha chain of human hb between phe 33 and leu 34 that leads to proteolytic cleavage(13)
Interacting partners	falciparins(12) falcilicin(12) dipeptidyl aminopeptidase1(12)
Type of interaction	protein-protein(14)
X-RAY/NMR/Model	X-RAY(14)
Key active site residues	Asp34, Asp121(2)
Name of the ligand if present	A2T(PDB ID 2IGY)
Functional unit	Both chain a and b(PDB ID 2IGY)
Link of paper where structure reported	http://www.rcsb.org/pdb/explore/explore.do structureId=2igy
Resolution	2.60 A(14)
Name of natural substrates	hemoglobin(15)
Total no. of ligands	2(PDB ID 2IGY)
Name of the bound inhibitor	A2T(PDB ID 2IGY)
Liganded/ Unliganded	liganded

Table I lists the protein biology and structural characteristics of Plasmepsin II. After the docking score (delta G and Ki values) was validated with the already published data for the inhibitors for the proteins, the GSK library with 13469 chemically synthesized compounds was screened against the 3D structures of the proteins. These compounds having more than 80% inhibition for *Plasmodium falciparum* were downloaded from https://www.ebi.ac.uk/chemblntd

in SMILES format. Before the screening of compounds, target proteins were validated using Discovery Studio and RMSD values were confirmed as < 2. After this, the GSK compounds were made to run against the protein in Discovery Studio for their screening. Leads or ligands with the best hit and docking score with these proteins were selected (Table-II). Lig Score 1 was set as standard score due to higher accuracy in predicting ligand- protein interaction energy for different types of proteins.

Table II Leads for Plasmepsin II selected by docking

Compound ID	Lig score
TCMDC-136645	5.22
TCMDC-138112	5.26
TCMDC -137540	5.29
TCMDC-138112	5.29
TCMDC-124216	5.31
TCMDC-124125	5.36
TCMDC-135138	5.38
TCMDC-124181	5.43
TCMDC-125201	5.43
TCMDC-133642	5.45
TCMDC-136645	5.45
TCMDC-137699	5.56
TCMDC-137699	5.97

Table II shows different Leads for Plasmepsin II selected by docking against Plasmepsin II. These compounds were found to have more effective binding than the already bound inhibitor at the same binding site and higher score than reference compounds. The top hits screened in Discovery Studio were validated in AutoDock by performing docking procedures. For that separate PDB of the best hit compounds in its best pose has to be docked with the reference protein with no bound ligand using AutoDock tools. Validation of AutoDock tools was done using reference PDB IDs (For Plasmepsin II it is 2IGY) of the respective selected target proteins. The reference PDB structures of proteins were downloaded with already attached ligands from Protein Data Bank.

Table III Protein inhibitor data of Plasmepsin II [16]

Protein ID	2IGY
Inhibitor ID	A2T
Ligand SMILE	CCCCCc1ccc(cc1)C(=O)N(Cc1ccc(cc1)N(C)c1ccncc1)C1CCN(CCC(C C)CC1 (16)
Ligand InChi	InChI=1S/C35H48N4O/c1-5-6-7-8-29-9-13-31(14-10-29)35(40)39(34 20-25-38(26-21-34)24-19-28(2)3)27-30-11-15-32(16-12-30)37(4)33-17 22-36-23-18-33/h9-18,22-23,28,34H,5-8,19-21,24-27H2,1-4H3 (16)
Mol Wt	540.78 g/mol
Binding with target(active/allosteric)	active
Interacting target residues	trp 41B, tyr 192B, ile 123B, PHE 111B, ile 32B, ser 218b

Table III lists already known inhibitors of Plasmepsin II [16] with its inhibitor ID, name and SMILE structure.

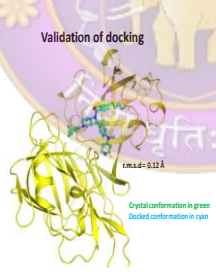

Figure 2 Comparison between binding interactions of reference compound and the compound identified with Plasmepsin II.

In case of Plasmepsin II there were nine top hits. Reference Compound with PDB id is 2IGY with A2T that is the known inhibitor. The interacting energy as well as interacting residues were also reported (Figure. 2)

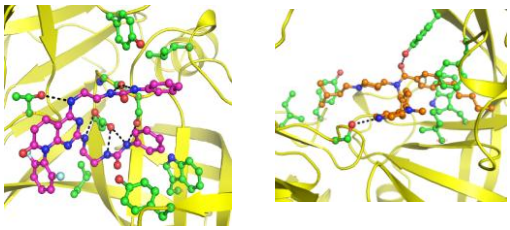

Figure 3 Overall comparison of identified and reference compound

Overall comparison shows that the lead compound identified against Plasmepsin II showed better LigScore (5.97) than the reference compound (5.22). Reference compound had total interaction energy of -76.73 kcal/mol which was greater than the total interaction energy of the lead compound (-82.46). This was a result of the greater interaction of the lead compound with the protein. Ile2, Asp34, Trp41, Val42, Pro43, Met75, Tyr77, Val105, Phe111, Thr114, Tyr115, Phe120, Asp121, Ile123, Tyr192, Ile212, Ser218 and Ile300 were the residues of Plasmepsin II that interacted with reference compound, while the residues of Plasmepsin II that interacted with the identified compound were found to be Met15, Ile32, Asp34, Gly36, Ser37, Trp41, Met75, Tyr77, Ile123, Leu131, Tyr192, Ile212, Asp213, Gly216, Thr217, Ser218 and Ile300. Of the total interacting residues, 8 were found to be common in both the interactions.

Figure 3 shows that the lead compound identified against Plasmepsin II showed better LigScore and more negative interaction energy than the reference compound.

Reference compound had total interaction energy at -76.73 kcal/mol which was greater than the total interaction energy of the lead compounds. This depicts greater interaction of the lead compound with the protein.

The comparison also shows that the lead compounds identified against malaria had better inhibition than already known inhibitor present in the crystal structure of Plasmepsin II (PDB ID 2IGY). Figure 4 shows the comparison between the docked conformations of reference compound and identified compound with Plasmepsin II and their interaction patterns with different residues at same binding pocket.

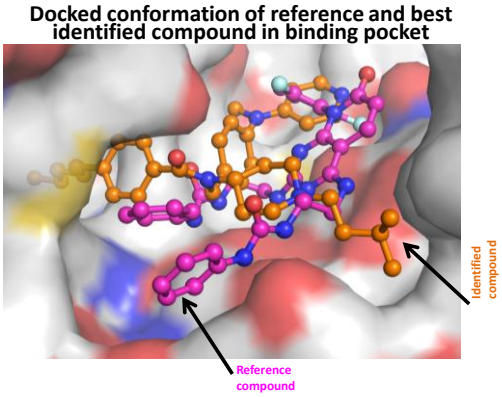

Figure. 4 Interaction pattern of docked conformation in reference compound and identified compound.

Figure 5a shows the interacting residues and hydrogen bonds formed after docking between TCMDC-135138 and Plasmepsin II. For Plasmepsin II with TCMDC – 136645, the Best Run was 1 with binding energy -7.68 and RMSD value at 3.28 (Fig 5b). For Plasmepsin II with TCMDC – 137699, the Best Run was 4 with binding energy -6.47 and RMSD value at 3.68 (Fig. 5c). For Plasmepsin II with TCMDC – 133642, the Best Run was 3 with binding energy – 5.39 and RMSD value at 3.96 (Fig. 5d). For Plasmepsin II with TCMDC – 124125, the Best Run was 3 with binding energy - 7.79 and RMSD value at 3.11 (Fig. 5e). For Plasmepsin II with TCMDC – 124181, the Best Run was 2 with binding energy - 6.95 and RMSD value at 4.96 (Fig. 5f). For Plasmepsin II with TCMDC – 125201, the Best Run was 6 with binding energy -7.34 and RMSD value at 2.66 (Fig. 5g). For Plasmepsin II with TCMDC – 138112, the Best Run was 1 with binding energy -2.77 and RMSD value at 3.75 (Fig. 5h).

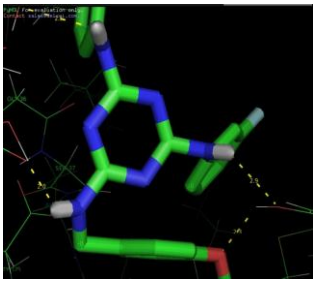

Fig 5 a

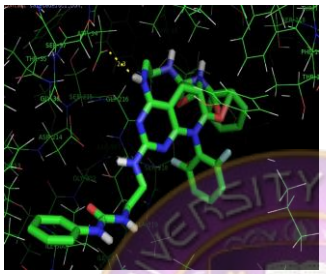

Fig 5 b

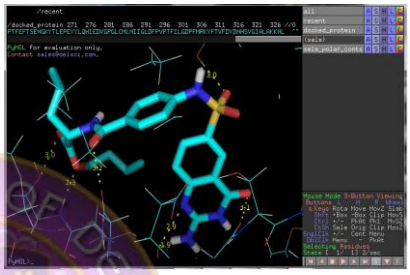

Fig 5 c

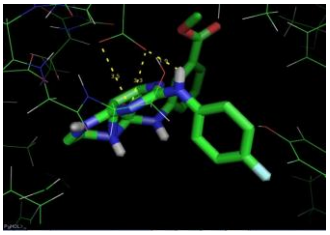

Fig 5 d

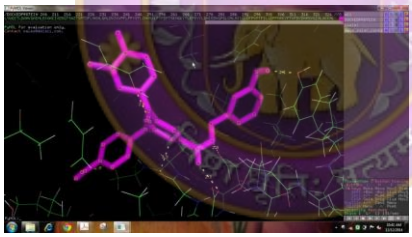

Fig 5 e

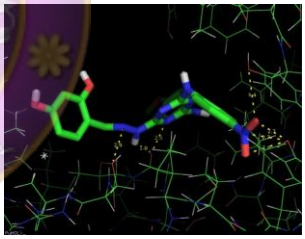

Fig 5 f

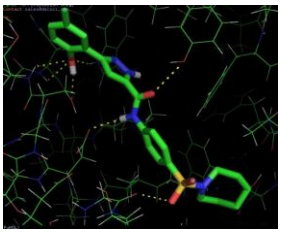

Fig 5 g

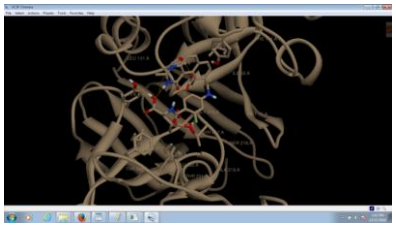

Fig 5 h

Figure 5 Hydrogen bonds formed after docking between Plasmepsin II and (a) TCMDC-135138, (b) TCMDC-136645, (c) TCMDC-137699, (d) TCMDC-133642, (e) TCMDC-124125, (f) TCMDC-124181, (g) TCMDC-125201, (h) TCMDC-138112.

Lig Score 1 was set as standard score due to higher accuracy in predicting ligand protein interaction energy for different types of proteins. Comparison of top hit compounds was done

with reference compounds, that is, known inhibitors of Plasmepsin II. Lig Score greater than the reference compounds indicate a better fit of ligand/affinity for the target site. For Plasmepsin II, the LigScore of reference compound was 5.22 and total interaction energy was found to be – 76.73 kcal/mol which is more negative than the total interaction energy of top hits, Validation of top hits in Auto dock tools – for e.g. TCMDC 137699 was found to have total interaction energy of -82.46 kcal/mol.

Auto Dock tools give the protein- Ligand interaction in terms of Root Mean Square Deviation (RMSD) and Binding Energy. Top hits from Discovery Studio were validated on Auto Dock and comparison of results was done with data of reference compound. Top hit compounds with RMSD less than or equal to the reference compound indicate a better affinity for the protein. RMSD value equal to 2 was taken as the maximum value for selection of compounds. For Plasmepsin II, RMSD of reference compound was 0.12 A though none of the top hits selected for Plasmepsin II had RMSD less than or equal to 2 but many of them had better binding energy. Many of the top hits compounds of Plasmepsin II did not follow the criteria of selection (RMSD less than or equal to 2.0). These top hit compounds can be considered as positive results and can be used for in vitro studies.

CONCLUSION AND FUTURE DIRECTION

Plasmepsin can be considered as the promising drug target for future studies. Work on the lead generation and lead optimization strategies that will include ADMET studies, 2D Visualization will be conducted for further studies.

ACKNOWLEDGMENTS

The group thanks University of Delhi for providing grant for this Innovation Project without which it would not have been possible to carry out this study. The timely help and suggestions of the mentors Dr. Manoj, AIIMS and Dr. Anshu Bhardwaj, Open Source Drug Discovery have been a guiding force in reaching to the observations and results of this study.

REFERENCES

1. Nadlla Alves Bispo, Richard Culleton, Lourival Almeida Silva, Pedro Cravo **(2013)**, A Systematic In Silico Search for Target Similarity Identifies Several Approved Drugs with Potential Activity against the Plasmodium falciparum Apicoplast, . *PLoS ONE* **8(3)**: e59288. doi:10.1371/journal.pone.0059288

2. Bhaumik P, Gustchina A, Wlodawer A, **(2011)**. Structural studies of vacuolar plasmepsins, 10.1016/j.bbapap.2011.04.008.

3. Moura PA, Dame JB, Fidock DA.**(2009)**. Role of Plasmodium falciparum digestive vacuole plasmepsins in the specificity and antimalarial mode of action of cysteine and aspartic protease inhibitors, **53(12)**:4968-78.

4. Spitzmuller A, Mestres J **(2013)**. Prediction of the P. *falciparum* Targt Space Relevent to Malaria Drug Discovery. *PloS Comput Biol* **9(10)**:e1003257.

5. Daria Van Tyne,Baba Dieye,clarisssa Valim, Rachel F Daniels et al, **(2013)**. Changes in drug sensitivity and anti malarial drug resistance mutation over time among Plasmodium falciparum arasites in senegal. *Malaria Journal* .12:441

6. Francis SE, Gluzman IY, Oksman A, Knickerbocker A, Mueller R, Bryant ML, Sherman DR,Russell DG, Goldberg DE. **(1994)**. Molecular characterization and inhibition of a *Plasmodium falciparum* aspartic hemoglobinase. *EMBO J*. **(13):**306–317.

7. Banerjee R, Liu J, Beatty W, Pelosof L, Klemba M, Goldberg DE.**(2002)** Four plasmepsins are active in the *Plasmodium falciparum* food vacuole, including a protease with an active-site histidine. Proc. *Natl. Acad. Sci. USA.*; **(99):**990–995.

8. Coombs GH, Goldberg DE, Klemba M, Berry C, Kay J, Mottram JC. **(2001)**. Aspartic proteases of *Plasmodium falciparum* and other parasitic protozoa as drug targets. *Trends Parasitol*.**(17)**:532–537.

9. Francis SE, Sullivan DJ Jr. Goldberg DE. **(1997)** Hemoglobin metabolism in the malaria parasite *Plasmodium falciparum*. *Annu. Rev. Microbiol.* **(51)**:97–123.

10. Frederick Christopher Chullai,Evanora Nongrum,Rajnikant Namdeo, **(2011)**. Proposing de-novo Generated, Iteratively Optimized new lead molecules Targeting Plasmodium falciparum Plasmepsin – II, *International Journal of ChemTech Research CODEN(USA):IJCRGG* ISSN : **0974-4290 (Vol. 3)**,pp 1538-1547.

11.Bhaskar D, Wadehra N R, Gulati M, Narula A, Vishnu R, Katyal G, **(2015)**. *In silico* Studies on Selected Drug Targets for Combating Drug Resistance in *Plasmodium falciparum*. World Academy of Science, Engineering and Technology, *International Journal of Medical, Health, Biomedical and Pharmaceutical Engineering* (Vol: 9, No:2): 188-192

12. AlejandroGilL, Pedro A.Valiente, PedroG.Pascutti, and TirsoPons ,**(2011)**. Computational Perspectives into Plasmepsins Structure—Function Relationship: Implications to Inhibitors Design. *Journal of Tropical Medicine*. (Article ID 657483), 1-15

13. Jun Liu, Eva S. Istvan and Daniel E. Goldberg, **(2006)**. Hemoglobin-degrading PlasmepsinII Is Active as a Monomer, *J. Biol. Chem.*, **(281):**38682-38688

14. Boss, C., Corminboeuf, O., Grisostomi, C., Meyer, S., Jones, A.F., Prade, L., Binkert, C., Fischli, W., Weller, T., Bur, D. **(2006)**. Achiral, cheap, and potent inhibitors of Plasmepsins I, II, and IV. *Journal: Chemmedchem* (1): 1341-1345.

15. www.uniprot.org

16. www.rcsb.org