RESEARCH ARTICLE

Computer aided docking of small molecules to the active site of the malarial aspartic protease, Plasmepsin II

Pradeep Aberathne, Jayamini Mahabaduge and Susil J. Silva[*]
Department of Chemistry, Faculty of Applied Sciences, University of Sri Jayewardenepura, Gangodawila, Nugegoda.

Abstract: Malaria afflicts nearly 500 million people annually and kills about 2 million of them. The pathogenic organisms *plasmodium falciparum* and *Plasmodium vivax* are responsible for most infections and deaths and they are becoming increasingly resistant to drugs, making it essential to identify new antimalarial agents. Proteases are attractive candidates for drug development because they play a vital role in parasite metabolism. Haemoglobin degradation is a metabolic process that is central to the growth and maturation of the malaria parasite. Two aspartic proteases from *Plasmodium falciparum* Plasmepsin (Plm) I and II that initiate degradation of haemoglobin have been identified and characterized. In this paper, the haemoglobin degrading enzyme Plm II is used as the target macromolecule. Computer assisted database search was used to dock molecules into the active site of Plm II using the computer software Dock version 4. Docked compounds were ranked according to their interaction energy score with the active site and the binding energy of some selected compounds were calculated using the computer software MOLARIS. In order to check the selectivity of the selected compounds they were docked into a similar enzyme, cathepsin D found in humans and their binding energies were also calculated. Based on the difference in binding energy values, eighteen compounds were identified as potential antimalarial drug leads.

Keywords: Binding free energy change, Cathepsin D, DOCK v4, malaria, MOLARIS, Plasmepsin II, 3D database

INTRODUCTION

Malaria is by far the most important tropical parasitic disease, and kills more people than any other disease except tuberculosis. In many developing countries including Sri Lanka, malaria exacts an enormous toll in lives and medical costs[1]. The causative agents of human malaria are four species of *Plasmodium*: *P. falciparum*, *P. vivax*, *P. ovale* and *P. malariae*[1]. Of these, *P. falciparum* accounts for the majority of infections and is the most lethal. Sri Lanka currently plays host to both *P. vivax* and *P. falciparum*[2]. Many of the *P. falciparum* infections are now resistant to chloroquine[3]. This is notable because, in the past, there are no records of *P. falciparum* infections in Sri Lanka.

The emergence and spread of resistance to currently available drugs highlight the need for development of novel antimalarial drugs. To discovers new drug leads, structure based design methods with focus on validated macromolecular targets are required.

The digestive vacuole of *P. falciparum* is involved in key processes, including haemoglobin metabolism and detoxification, which represent potential chemotherapeutic targets. *P. falciparum* invades erythrocytes and consumes nearly all the haemoglobin as a source of nutrients during the phase of growth and development. Haemoglobin degradation is mediated by the action of several digestive enzymes. Two aspartic proteases of *P. falciparum* have been implicated in the initial steps of the haemoglobin degradation process. The first protease, Plasmepsin (Plm) I, appears to make a strategic initial clevage that presumably leads to an unfolding of the native haemoglobin structure so that further proteolysis can rapidly proceed. Plm II, the second aspartic protease, is capable of cleaving native haemoglobin but is more active against denatured or fragmented globin, that is produced by the action of Plm I.

There is sufficient evidence to support the validity of the vacuolar proteinases as chemotherapeutic targets: inhibitors of these enzymes have been shown to block haemoglobin degradation and kill the parasite in culture[4].

[*]Corresponding author

So far, no lead compound that justifies further development has emerged from the current screening. Such failure is partly due to the limited collections of compounds tested so far and the lack of real diversity in the compounds from the screened libraries. The main objective of the present study is to identify potent lead compounds that selectively inhibit Plm II through a computer assisted database search. In a recent report, a virtual screening has been carried out against Plm II, HIV protease and homologous human protease cathepsin D (Cat D) using the FlexX docking software with a set of nonstatine inhibitor.

MATERIALS AND METHODS

Preparation of Plasmepsin II (Plm II) and Cathepsin D (Cat D): The first step in the search for new ligands is to obtain the three-dimensional structure of the target macromolecule. The X-ray crystal structure of the *P. falciparam* Plm II and human Cat D bound to a peptidyl inhibitor, pepstatin, were determined by previous studies[6,7] (Figure 1). The coordinates of the atomic positions of both Plm II and Cat D were obtained from the Protein Data Bank[8] (codes 1m43A and 1lyb). The visualization and analysis of the protein structure were done using the RASMOL[9a,9b] package. All water molecules were removed prior to docking. The atom charges, based on Gasteiger *et al.*[10a] partial charge method, were added using the software package BABEL[10b].
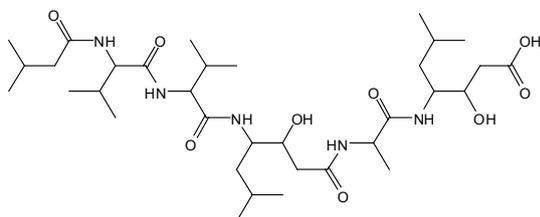

Figure 1: X -ray crystal structure of the *P. falciparum* Plm II and human Cat D bound to pepstatin[6,7]

The hydrogen atoms were added using the program Charmm[11]. After adding hydrogen atoms and prior to docking procedure, an energy minimization using the Charmm27 all atom force field was carried out to remove any abnormal strains in the resultant structure. The energy minimization of the proteins was performed using 100 steps of steepest decent followed by 250 steps of adopted basis Newton-Raphson methods keeping all heavy atoms fixed.

The compounds from two commercially available databases, Maybridge and Chembridge[12] were screened using the software package DOCK 4.0[13] for potential Plm II inhibitors. The three dimensional Cartesian coordinates of the compounds in the above two databases in MOL2 format required by the DOCK 4.0 were provided by the Department of Pharmaceutical Sciences, University of Maryland, Baltimore, MD, USA. Alex. The default method that DOCK uses for the site point generation involves creating an inverse surface of the binding site. This is defined by the set of overlapping spheres that fill the binding site and touch the molecular surface at only two points. A sphere cluster containing 51 spheres was selected for this study by carefully visualizing them in the active site. The sphere centres are used as site points. These spheres were used in the initial screening of all compounds in the database for the identification of the prospective ligands that complement the active site. The molecular surface is generated using the Connolly surface software, MS[14]. All compounds initially selected contained ten or less rotatable bonds and five sixty heavy atoms.

The binding site of interest on the target is labeled with a set of spheres generated by DOCK that is complementary to the site. Electrostatic and van der Waals potential fields generated over the binding site by the enzyme are pre-calculated using a grid. In both docking calculations, a grid box with dimensions 30.947x38.031x36.228 Å was employed. The spacing of the grid was 0.3 Å. This grid box enclosed all the residues in the active site and extra margin of 5 Å was used in all directions. Each ligand was then flexibly docked into the site of interest, oriented to maximize the overlap with the sphere centres and followed by a grid-based energy scoring step to rank the ligands.

Plm II and Cat D were aligned using the programme PROFIT[15]. The same procedure as in Plm II was followed to dock selected compounds to the active site of Cat D.

Binding free energy calculation: The binding free energy was calculated using the MOLARIS programme[16]. MOLARIS uses the scaled Protein Dipoles Langevin Dipoles[17] (PDLD/S) method and the Linear Response Approximation[18] (LRA) version of this approach (PDLD/S-LRA)[19] to calculate the absolute binding free energy.

When calculating the binding free energy of a ligand to an enzyme using the MOLARIS programme, the entire system is divided into four regions. Region I contains the atoms of the inhibitor. Region II holds the residues making up the active site. In the present study, region II is defined as the residues within 16 Å from the centre of the inhibitor. Region III is the sphere of Langevin dipoles around Region II. The radius of the region III was 25 Å. In order to obtain meaningful results for the contribution of the bulk region (that surrounds Region III), it is

important not to have any Region II group extending out of this region. Region IV is the bulk region around Region III. This region is treated as a continuum with the dielectric constant of the given solvent. In order to perform the MOLARIS calculations we have to define residual charges for the atoms in region I and for the surrounding macromolecule. These charges were calculated within the PM3[20] semi-empirical Hamiltonian using the programme AMSOL[21].

A Dell OptiPlex Pentium III computer was used in this study. The average CPU time required to dock a compound from the database was 25 s. The binding free energy calculation using MOLARIS took 51 s of CPU time (average) per compound.

RESULTS

Docking database compounds into Plm II

A total of 125000 compounds from two databases were screened for compounds showing van der Waals and electrostatic complementarity with the active site of malarial Plm II using the program DOCK 4.0. In this method, each compound from the database was docked into the active site and the orientation corresponding

to the minimum energy score was obtained. To test the efficiency of this method, the DOCK was used to regenerate pepstatin-Plm II complex structure. The difference between the X-ray crystallographic and docked orientations of pepstatin was determined by calculating the root mean square deviation (RMSD) for 48 heavy atoms. The RMSD of the docked pepstatin with respect to the X-ray crystal structure was 1.1 Å (for 48 heavy atoms). The orientation of both X-ray and docked structures of pepstatin is shown in Figure 2. From this initial search, compounds having favourable interaction energies with the Plm II binding site were collected. Energy scores of those compounds ranged from -56 to –24 kcal/mol. The total interaction energy is calculated as the sum of van der Waals and electrostatic energy components. The compounds with highest negative energy scores are presented in Table 1.

Docking study using Cat D

Cat D and Plm II have 116 common amino acids corresponding to 35% identity. The amino acid sequences of Cat D and Plm II were aligned according to the conserved amino acids. The RMSD between the two aligned structures were found to be 1.54 Å (for all heavy atoms). The backbone atoms of the aligned structures

Table 1: Binding and interaction energies of selected compounds in kcal/mol

	Compound ID	Plasmepsin II		Cathepsin D		
		Binding free energy change (MOLARIS)	Interaction Energy (DOCK)	Binding free energy change (MOLARIS)	Interaction Energy (DOCK)	$\Delta\Delta G_{bind}$ (MOLARIS)
1	Cb004885	-25.86	-40.89	-8.38	-41.93	-16.48
2	MB40016	-20.33	-49.09	-5.81	-45.03	-14.51
3	MB39857	-20.23	-46.52	-9.54	-46.45	-10.69
4	MB40015	-19.36	-46.86	-5.63	-46.99	-13.73
5	MB40277	-16.69	-46.63	-7.89	-46.2	-8.8
6	MB39853	-16.64	-46.81	10.2	-50.61	-26.84
7	MB42058	-15.45	-52.37	0.91	-47.22	-16.36
8	Cb016629	-16.7	-46.08	0.52	-38.38	-17.22
9	MB32187	-14.92	-46.43	-8.11	-50.47	-6.81
10	Cb006048	-14.85	-52.94	-7.21	-50.46	-7.64
11	MB07954	-14.69	-50.39	-6.73	-58.9	-7.96
12	MB01546	-11.62	-46.25	-3.8	-46.52	-7.82
13	Cb000916	-11.18	-52.29	-5.8	-54.6	-5.38
14	Cb030175	-10.97	-51.17	-5.07	-48.67	-5.9
15	Cb049425	-6.11	-51.77	5.9	-48.11	-12.01
16	Cb009746	-5.28	-52.65	3.89	-55.51	-9.17
17	MB10480	-4.83	-47.68	4.33	-48.3	-9.16
18	Cb000656	-4.79	-51.79	0.86	-43.77	-5.65

Figure 2: Orientation of pepstatin in the active site of Plm II in X-ray crystallographic and docked structures. The pictures were prepared using the programme PyMOL[22]
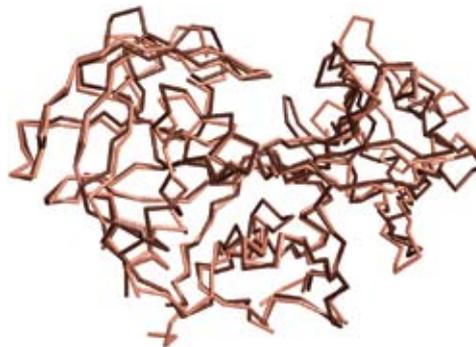

Figure 3: Structure based alignment of Cat D (tan) and Plm II (brown)

Cat D	Ala13	Gln14	Val31	**Asp33**	**Gly35**	Ser36	Ser37	Ile76
Plm II		Met15	Ile32	**Asp34**	**Gly36**	**Ser37**		Met75
Cat D	His77	Tyr78	Gly79	Ser80	Gly81		Thr125	Phe126
Plm II	Asn76	**Tyr77**	Val78	**Ser79**		Phe111	**Thr114**	
Cat D	Ala129	Phe131	Ile134	Ile142	Tyr205	Ile229	**Asp231**	**Gly233**
Plm II			**Ile123**	Leu131	**Tyr192**		**Asp214**	**Gly216**
Cat D	**Thr234**	**Ser235**	Leu236	Val238			Met309	Ile311
Plm II	**Thr217**	**Ser218**	Ala219	Thr221	Asn288	Ile290	Leu292	
Cat D	Pro312	Leu318	Ile320					
Plm II			**Ile300**					

Figure 4: Structure based alignment of the active site amino acids of Cat D and Plm II. Conserved amino acids are in boldface. Blank means no corresponding amino acid found in the active site.
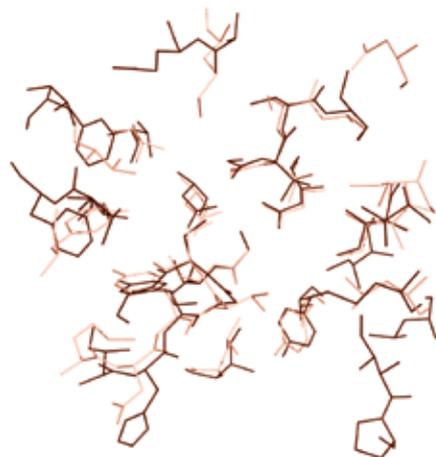

Figure 5: Structure based alignment of the amino acids in the active sites of Plm II (tan) and Cat D (brown).
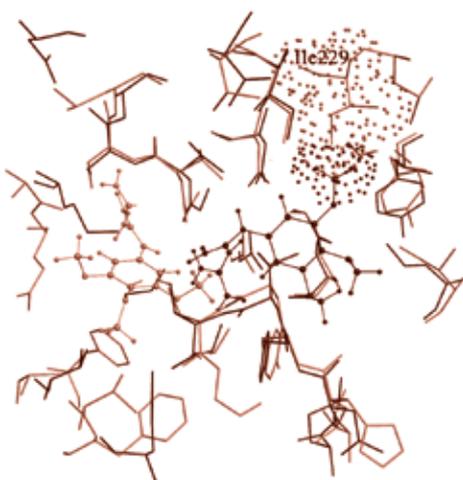

Figure 6: Binding sites of MB39853 in Plm II (brown) and in Cat D (tan). Active site residues of Plm II (brown) and Cat D (tan) are drawn in stick figures.
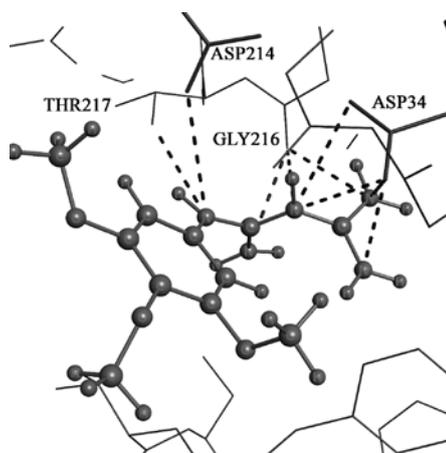

Figure 7: Hydrogen bonds formed between MB39853 and active site residues in Plm II
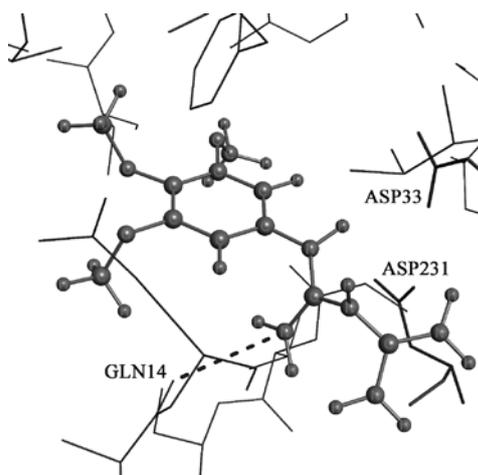
Figure 8: Hydrogen bonds formed between MB39853 and amino acid residues in the active site of Cat D

are shown in Figure 3. Considerably higher degree of homology (52%) at the active site was found by comparing Cat D and Plm II. The amino acids within 5.0 Å from the pepstatin in both enzymes were considered as active site residues. The number of active site residues in Cat D and Plm II are 32 and 25 respectively. These amino acids are shown in Figure 4. They were also aligned according to conserved amino acids and the RMSD was found to be 0.44 Å (for heavy atoms). The aligned active sites are shown in Figure 5. These results are important in the design process to find inhibitors that specifically inhibit Plm II but not Cat D.

The 150 highest energy score compounds screened against Plm II using the compounds from two chemical databases were docked into the active site of the Cat D using the program DOCK 4.0. In this method each compound was docked into the active site and the orientation corresponding to the minimum energy score was obtained. To test the efficiency of this method, the DOCK was used to regenerate pepstatin-Cat D complex structure. The difference between the X-ray crystallographic and docked orientations of pepstatin was determined by calculating the RMSD. The RMSD of the docked pepstatin with respect to the X-ray crystal structure was 0.43 Å (for 48 heavy atoms). The total interaction energy between a docked molecule and the Cat D is calculated as the sum of van der Waals and electrostatic energy components and given in Table 1.

Calculation of binding free energy

One of the promising methods for calculating the absolute binding free energies is the semi-microscopic version of the protein dipoles Langevin dipoles[19] (PDLD/S) method. The programme MOLARIS developed by Warshel *et al.*[16] provides tools to evaluate binding free energies using the PDLD/S method. The binding free energy changes, ΔG_{bind}, of 20 docked molecules were calculated using the PDLD/S method and the results are tabulated in Table 1. The negative value of ΔG_{bind} indicates spontaneous formation of ligand-protein complex. The binding energy results showed that most of the compounds bind more favourably with Plm II than Cat D. The highest negative binding free energy change of −25.86 kcal/mol was obtained for Cb004885 in the binding site of Plm II. The same compound binds to the Cat D with binding free energy of −8.38 kcal/mol.
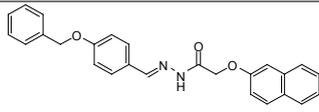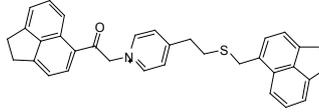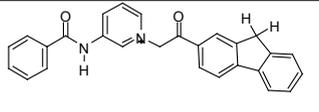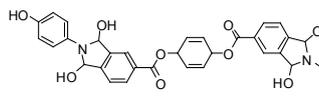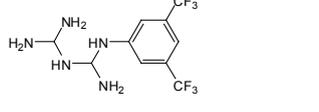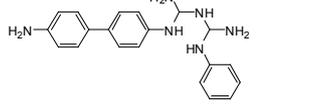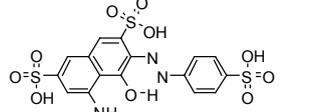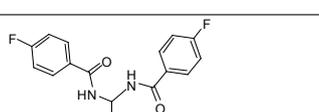
The difference in binding free energy change, $\Delta\Delta G_{bind}$, between Plm II and Cat D was calculated and is listed in Table 1. This quantity can be considered as a measure of the specificity of a docked compound towards Plm II over Cat D. Among the 20 compounds investigated in the present study, MB39853 showed the highest negative $\Delta\Delta G_{bind}$.

In order to understand the different binding modes of compounds at the active sites of Plm II and Cat D, a detailed structural analysis of MB39853 complexed with Plm II and Cat D was carried out at molecular level. As shown in Figure 6, MB39853 occupies two different binding sites in Plm II and Cat D. The Plm II and Cat D are aspartic proteases and they contain two aspartates in the active site that are responsible for the cleavage of the peptide bond. The results revealed that MB39853 forms hydrogen bonds with both aspartate residues (Asp 34 and Asp 216) in Plm II (Figure 7) but not with Cat D (Figure 8). The hydrogen bond interactions between −NH of MB39853 and aspartates in Plm II is similar to the central statin hydroxyl group in the pepstatin A[6], which binds to the catalytic aspartates and is thought to be the principal reason for the inhibition. As shown in Figure 7, MB39853 also formed hydrogen bonds with Gly 216 and Thr 217 in the active site of Plm II. However, only a single hydrogen bond with Gln 14 was observed in Cat D (Figure 8).

One might speculate about the reason why MB39853 in Cat D did not occupy the same site as in Plm II. It is possible that the steric hindrance between the Ile229 (this residue is absent in Plm II, see Figure 6) and one of the OCH_3 groups in MB39853 prevents it binding to the Plm II binding site. As shown in Figure 6, the repulsion due to overlaping of two van der Waals spheres of Ile229 and OCH_3 group prevents MB39853 entering the same binding site that it occupies in Plm II.

Table 2: Chemical structures and corresponding IUPAC names of compounds

ID	Structure	IUPAC Name[a]
CB004885		*N*-(amino(phenylamino)methyl)-*N,N*-dimethylmethanetriamine
MB40016		*N*-(diaminomethyl)-*N*-(2,3-dihydrobenzo[*b*][1,4] dioxin-6-yl) methanetriamine
MB39857		4,4'-(1*E*,1'*Z*)-(4,4'-methylene *bis*(4,1-phenylene)*bis*(azan-1-yl-1-ylidene)) *bis*(methan-1-yl-1-ylidene)*bis*(2,6-bis(dihydroxyamino)phenol)
MB40015		1-benzyl-2,4,6-trihydroxy-5-(1-(phenylamino)ethyl)pyrimidine-1,3-diium
MB40277		(*E*)-*N*'-(3-methoxybenzylidene)-2-(naphthalene-2-yloxy)acetohydrazide
MB39853		*N*-(diaminomethyl)-*N*-(3,4,5-trimethoxyphenyl)methanetriamine
MB42058		(*NE,NE*)-*N,N*'-bis(4-methoxybenzylidene)-1-(4-methoxyphenyl) methanediaminum
CB016629		*N*-(diaminomethyl)-*N*-(4-methoxybenzyl)methanetriamine
MB32187		1-(2-(9*H*-fluoren-2-yl)-2-oxoethyl)-3-benzamidopyridinium
CB00648		*N*-(diaminomethyl)-*N,N*'-diphenylmethanetriamine
MB07954		4-(2-(1 *H*-indol-3-yl)ethyl)-1-(2-(4-fluorophenyl)-2-oxoethyl)pyridinium

MB01546		(*E*)-*N'*-(4-(benzyloxy)benzylidene)-2-(naphthalene-2-yloxy)acetohydrazide
CB000916		1-(2-(1,2-dihydroacenaphthylen-5-yl)-2-oxoethyl)-4-(2-((1,2-dihydroacenaphthylen-5-yl)methylthio)ethyl)pyridinium
CB030175		1-(2-(9H-fluoren-2-yl)-2-oxoethyl)-3-benzamidopyridinium
CB049425		cyclohexa-2,5-diene-1,4-diyl bis(1,3-dihydroxy-2-(4-hydroxyphenyl)isoindoline-5-carboxylate
CB009746		*N*-(3,5-bis(trifluoromethyl)phenyl)-*N*-(diaminomethyl)methanetriamine
MB10480		*N*-(amino(4'-aminobiphenyl-4-yl amino)methyl)-*N*-phenylmethanetriamine
CB000656		(*E*)-5-amino-4-hydroxy-3-((4-sulfophenyl)diazenyl)naphthalene-2,7-disulfonic acid
CB010056		*N*,*N'*-((4-fluorophenyl)methylene)bis(4-fluorobenzamide)

^a IUPAC names were generated using the software package CHEMDRAW[23]

DISCUSSION

An initial set of 20 potential inhibitors of Plm II has been identified *via* computer-based screening of compounds from two chemical databases against the active site. All compounds were initially screened against Plm II using the computationally efficient scoring approach based on the interaction energy. The RMSD of 1.1 Å compared to the X-ray crystal structure of pepstatin indicated the reliability of the docking procedure that has been used in the present study to identify lead compounds. The quality of the procedure was also assured by calculating the hydrogen bond contacts between the pepstatin and the active site of the Plm II. The initial set of compounds selected based on energy score values showed some chemical diversity. Their chemical structures and IUPAC names are given in Table 2.

The compounds with highest energy score against Plm II were docked into human Cat D. The binding free energy changes of those compounds with both enzymes, Plm II and Cat D, were also determined. Although the total interaction energies of a given compound in both enzymes were very close, considerable differences in binding free energies were obtained. The difference in binding free energy may indicate the difference in binding of a given compound at the active sites of Plm II and Cat D. As expected, MB39853 occupies two different binding pockets in Plm II and Cat D. The compounds with

negative binding free energy changes ($\Delta\Delta G_{bind}$) obtained here may not be considered as having potential drugs for malaria but they can still be used as the starting point for synthesis of the more potent inhibitors of Plm II. The work presented here will be useful for the further study of new compounds for drug design against malaria.

Acknowledgement

References

1. World Health Organization (1998). *Malaria fact sheet number 94*. WHO, Avenue Appia 20, CH-1211, Geneva 27, Switzerland.

2. Gunawardena D.M., Wickremasinghe A.R., Muthuwatta L., Weerasingha S., Rajakaruna I., Senanayaka T., Kota P.K, Attanayake N., Carter R. & Mendis K.N. (1998). Malaria risk factors in an endemic region of Sri Lanka, and the impact and cost implications of risk factor-based interventions. *American Journal of Tropical Medicine and Hygine* **58**: 533-542.

3. Handunnetti S.M., Jayasinghe S., Pathirana P.P.S.L., Fernando R., Sheriff M.H.R. & Mondis K.N. (1994). Chloroquine and sulphadoxine pyrimethamine resistant *Plasmodium falciparum* infection in Sri Lanka. *Ceylon Medical Journal* **39**: 45-46.

4. Rosenthal P.J., Wollish W.S., Palmer J.T. & Rasnick D. (1991). Antimalarial effects of peptide inhibitors of a Plasmodium falciparum cysteine proteinase. *Journal of Clinical Investigation* **88**:1467-1472.

5. Weik S., Luksch T., Evers A., Bottcher J., Sotriffer C.A., Hasilik A., Loffler H., Klebe G. & Rademann J. (2006). The potential of P1 site mutations in peptidomimetic protease inhibitors explored *via* the use of CC-coupling linker reagents. *ChemMedChem* **1**: 445.

6. Silva A.M., Lee A.Y., Gulnik S.V., Majer P., Collins J., Bhat T.N., Collins P.J., Cachaue R.E., Luker K.E., Gluzman I.Y., Francis S.E., Oksman A., Goldberg D.E & Erickson J.W. (1996). Structure and inhibition of plasmepsin II, a hemoglobin-degrading enzyme from *Plasmodium falciparum. Proceedings of the National Academy of Sciences of the United States of America* **93**: 10034.

7. Baldwin E.T., Bhat T.N., Gulink S., Hosur M.V., Sowder R.C., Cachau R.E., Collins J., Silva A.M. & Erickson J. W. (1993). Crystal structures of native and inhibited forms of human cathepsin d: implications for lysosomal targeting and drug design. *Proceedings of the National Academy of Sciences of the United States of America* **90**:6796.

8. Bernstein F.C., Koetzle T.F., Williams G.J.B., Meyer E.F., Brice M.D., Rodgers J.R., Kennard O., Shimanouchi T. & Tasumi M. (1997). The Protein Data Bank: a computer based archival file for macromolecular structures. *Journal of Molecular Biology* **112**: 535.

9. (a) RasMol version 2.7.2.1, Roger Sayle and E. James Milner-White (1995). RasMol: biomolecular graphics for all. *Trends in Biochemical Sciences* **20**:(9) 374.
   (b) Herbert J. Bernstein (2000). Recent changes to RasMol, recombining the variants. *Trends in Biochemical Sciences* **25**:(9) 453-455.

10. (a) Gasteiger J. & Marsili M. (1980). *Tetrahedron* **36**: 3219
    (b) BABEL, *joplin.biosci.Arizona.edu:/pub/Babel*

11. Brooks B.R., Bruccoleri R.E., Olafson B.D. , States D.J., Swaminathan S. & Karplus M. (1983). CHARMM: A programme for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry* **4** (2): 187-217.

12. (a) Maybridge. *www.maybridge.com*.
    (b) Chembridge. *www.chembridge.com*.

13. (a) Kuntz I., Blaney J.M., Oarley S.J., Langridge R. & Ferrin T.E. (1982). A geometric approach to macromolecule-ligand interactions. *Journal of Molecular Biology* **161**: 269.
    (b) Meng E.C., Shoichet B.K. & Kuntz I. (1992). Automated docking with grid-based energy evaluation. *Journal of Computational Chemistry* **13**:505.
    (c) Makino S., Kuntz I. (1997). Automated flexible ligand docking method and its application for database search. *Journal of Computational Chemistry* **18**:1812.

14. Connolly M.L. (1983). Analytical molecular surface calculation. *Journal of Applied Crystallography* **16**:548.

15. ProFit. *http://www.bioinf.org.uk/software/profit/.*

16. Chu Z.T., Villà J., Strajbl M., Schutz C.N., Shurki A. & Warshel A. *MOLARIS version alpha 9.06*. University of Southern California, California, USA.

17. Madura J.D., Nakajima Y., Hamilton R.M., Wierzbicki A. & Warshel A. (1996). Calculations of the electrostatic free energy contributions to the binding free energy of sulfonamides to carbonic anhydrase. *Structural Chemistry* **7**:131.

18. Lee F.S., Chu Z.T., Bolger M.B. & Warshel A. (1992). Calculations of antibody-antigen interactions: microscopic and semi-microscopic evaluation of the free energies of binding of phosphorylcholine analogs to McPC603. *Protein Engineering* **5**: 215.

19. Muegge I., Tao H. & Warshel A. (1998). A fast estimate of electrostatic group contributions to the free energy of protein inhibitor binding. *Protein Engineering* **10**:1363.

20. Stewart J.J.P. (1989). A fast estimate of electrostatic group contributions to the free energy of protein inhibitor binding. *Journal of Computational Chemistry* **10:** 209.

21. Hawkins G.D., Giesen D.J., Lynch G.C., Chambers C.C., Rossi I., Storer J.W., Li J., Liotard D.A., Cramer C.J. & Truhlar T.G. *AMSOL version 6.5.2*. University of Minnesota, Minnesota, USA.

22. The PyMOL Molecular Graphics System on World Wide Web. *http://www.pymol.org*.

23. Chemdraw. *www.chembridge.com*.