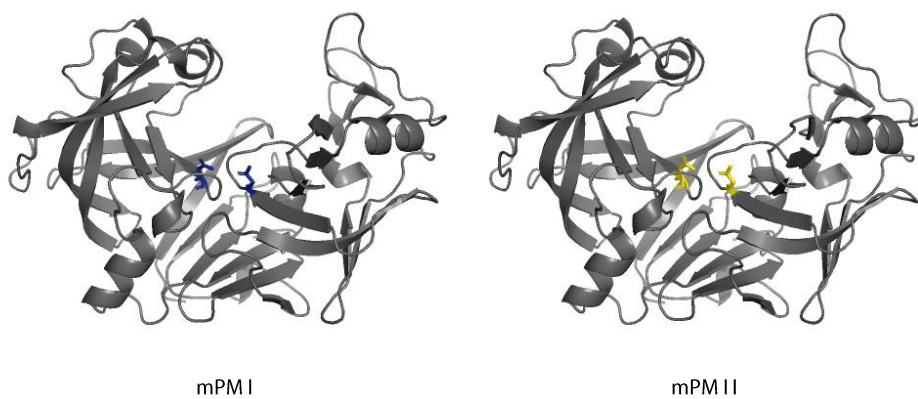


Chapter 5

Molecular interaction studies of *Plasmodium falciparum* specific Plasmepsin I and Plasmepsin II with compounds from *Andrographis paniculata* ethanol and *Commiphora wightii* aqueous extracts



5.1 Introduction

The phytochemicals from the *A. paniculata* EtOH and *C. wightii* AQ extracts with mPM I and mPM II inhibition activity were studied by computational molecular modeling to identify the potential inhibitors from the extracts and to gain insights into their possible mode of action. Approaches to study interactions between molecules such as protein-inhibitor interactions usually include a docking algorithm to predict ligand or inhibitor conformations followed by a method to estimate the binding affinity. AutoDock program is one of the popular tools to evaluate such intermolecular interactions. Several other tools are used in conjunction with a docking program to accomplish the purpose of an *in silico* interaction study. In the current study, structure prediction of mPM I and mPM II enzymes was performed using homology modeling. The docking experiments between the modeled 3D enzymes and the 3D structures of the phytochemicals were performed using AutoDock program. Based on the docking results four such compounds which had an affinity towards the active site of both the enzymes were found. The possible interactions between these compounds and the enzymes were also explored.

5.2 Material and methods

5.2.1 Homology modeling

The protein sequences of PMs PM I and PM II available in the NCBI database (GenBank Id: XM_001348213.1, XM_001348214.1) were used to deduce the protein sequences of the recombinantly expressed *P. falciparum* PMs, mPM I and mPM II. The N-terminal segment which is cleaved off after thrombin treatment was removed from the sequences and the sequences were then subjected to homology modeling. The structures of mPM I and mPM II were predicted by the Phyre2 Protein Fold Recognition Server (Kelley et al., 2015). Normal mode was specified during sequence submission which produces a set of 3D models based on the alignments between the Hidden Markov Models (HMM) of the query sequence and the HMM of known protein structures. Considering model scores (query coverage, confidence, identity) and template features, a model for each enzyme was selected which was validated using PROCHECK (Laskowski et al., 1993), ERRAT (Colovos and Yeates, 1993), PROVE (Pontius et al., 1996) and VERIFY 3D (Bowie et al., 1991; Lüthy et al., 1992) programs available from Structure Analysis and Verification Server version 6 (SAVES v6.0) server of UCLA-DOE LAB. PROCHECK compares the geometry of the

residues in a given protein structure with stereochemical parameters derived from well-refined, high-resolution structures and highlights the regions which need to be examined. It also performs Ramachandran analysis which gives information on the distribution of the residues in the model. ERRAT assesses the non-bonded interactions to identify the incorrectly built regions in protein models. The software PROVE uses deviation from standard atomic volumes as a measure to assess the quality of the structure and VERIFY 3D works by measuring the compatibility of any protein structure with its amino-acid sequence. The 3D structure of the template 2BJU used to build the structures of the enzymes was also subjected to the analysis alongside mPM I and mPM II to obtain data on features that were specific of the structure. The PDB files of the verified models were used for further studies.

5.2.2 Docking studies

The molecular interactions between the principle compounds identified from the *A. paniculata* EtOH and the *C. wightii* AQ extract and the enzymes mPM I and mPM II were studied using Autodock 4.2.6 tool (Morris et al., 2009) which is a suite of docking tools freely distributed by The Scripps Research Institute. The entire process of docking that included preparation of molecules for docking, running AutoDock and analyzing the dockings was performed using AutoDock tools, the graphical user interface for AutoDock. Pepstatin A, the established inhibitor of aspartic proteases was included as a positive control. The structural coordinate files for ligands were obtained from Pubchem in SDF format which were then converted to PDB format using Open Babel 2.4.1 software (O'Boyle et al., 2011). To prepare the structures for interaction studies, polar hydrogen and Gasteiger charges were added to both the enzyme structures and the ligand structures and the files were then saved in PDBQT format.

The active site of the enzymes was identified by superposition of the modeled structures over the structure of 2BJU and IH4 ligand complex. AutoGrid was run to precalculate the grid maps of interaction energies for various atom types which were used later during docking. The docking space was defined by a grid box with points $68 \times 68 \times 68$ along the x-, y- and z-axis and the grid center was designated at $-3.588 \times 65.0 \times 6.16$ points along the x-, y- and z- directions. The grid box encompassed both the catalytic aspartic acid residues of the enzymes namely, Asp40 and Asp220 which correspond to the canonical catalytic residues Asp34 and Asp214. Other parameters were kept as default. Docking experiments were performed by employing AutoDock considering macromolecule as rigid and ligand as

flexible. AutoDock run was set to return 25 ligand conformations. Ligand conformations generated have a docking score (or binding energy) associated with them which predicts binding affinity between a ligand and a target, lower score indicating higher binding affinity. The ligand interactions with the active site aspartic acid residues as well as the binding affinity between the ligand and the macromolecule target site were considered to find out the best binding pose for the ligand under consideration. The output ligand poses aligned with the macromolecule structure were analyzed using LigPlot+ (Laskowski and Swindells, 2011), which is a graphical system that generates 2D diagrams of ligand-protein interactions from 3D coordinates.

5.3 Results and discussion

The 3D structures of mPM I and mPM II were constructed using Phyre2 homology modeling (Figure 5.1) and for both mPM I and mPM II, models based on the template c2bjuA (PDB ID: 2BJU, PDB title: Plasmepsin II complexed with a highly active achiral inhibitor) were selected. The query coverage, confidence and identity for the template c2bjuA were found to be 97%, 100% and 73% respectively in the case of mPM I while these parameters were found to be 97%, 100% and 100% respectively in the case of mPM II. The quality and reliability of the models were checked by several structural assessment programs. The Ramachandran plots obtained after PROCHECK analysis showed in the case of mPM I and mPM II the residues in the favorable regions were 89% and 89.9% respectively and in the case of 2BJU they were 90.6% (Figure 5.2). A good quality model is expected to have over 90% of the residues in the most favored regions in the Ramachandran plot. However, the presence of merely 90.6% residues in the favorable region in the template 2BJU which is a well resolved 3D structure suggests that the presence of residues in the regions other than the favorable regions is due to feature of the fold(s) and they do not represent a poor model quality. Thus the mPM I and mPM II models with approximately 90% residues in the favorable regions were considered of good quality based on Ramachandran analysis. The models passed all other quality assessments and they were concluded to represent the native structure of the enzymes with high accuracy.

After the design and validation of the enzyme models, the molecular docking simulations were performed using AutoDock 4.2.6 tool. In total 13 ligands, 5 ligands (Ligand 1-5) from the *A. paniculata* EtOH extract and 8 ligands (Ligand 6-13) from the *C. wightii* AQ extract

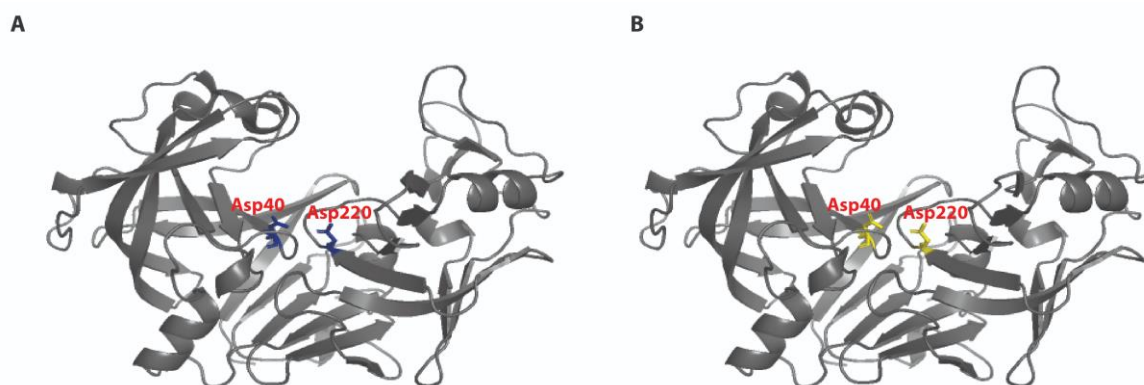


Figure 5.1: Structures of mPM I and mPM II.

(A) Three-dimensional structure of mPM I predicted by Phyre2 visualized using PyMol. The catalytic aspartic acid residues Asp40 and Asp220 are shown as sticks in blue color.

(B) Three-dimensional structure of mPM II predicted by Phyre2 visualized using PyMol. The catalytic aspartic acid residues Asp40 and Asp220 are shown as sticks in yellow color.

were studied for their inhibitory potential and interactions towards the active site of the enzymes mPM I and mPM II (Table 5.1). An AutoDock run for an enzyme-ligand pair generated 25 conformations of a ligand with different docking scores. The lowest possible binding energy conformation of the ligand that appeared to interact with the catalytic aspartic acid residues of the enzyme was considered as the best pose of the ligand. Whereas the minimum binding energy conformation was considered as the best pose for the ligand if it was found to bind away from the catalytic dyad of the enzyme. The control ligand pepstatin A best pose was found to dock with mPM I and mPM II with docking scores of -6.95 and -10.42 respectively. According to the docking scores of the best conformations, of all the test ligand compounds 2-amino-2-ethyl-4-(methylsulfonimidoyl)butanoic acid (Ligand 1), andrographolide (Ligand 3), safrole (Ligand 11) and sorbitol hexaacetate (Ligand 12) were suggested to have the strongest binding affinity for mPM I as well as mPM II. In the case of these four ligands, out of the total 25 conformations obtained after execution of docking, more than 3 conformations that docked at the site near to the docking site of the best pose were obtained, which suggested a strong affinity between the ligand and the enzyme at that site. Diethyl (2-methylallyl) phosphate (Ligand 4) although achieved a good docking score when docked with mPM I, was not explored further as its binding site was found to be outside the active site of the enzyme, away from the catalytic dyad and it thus had the least possibility to interfere with the catalytic action of the enzyme.

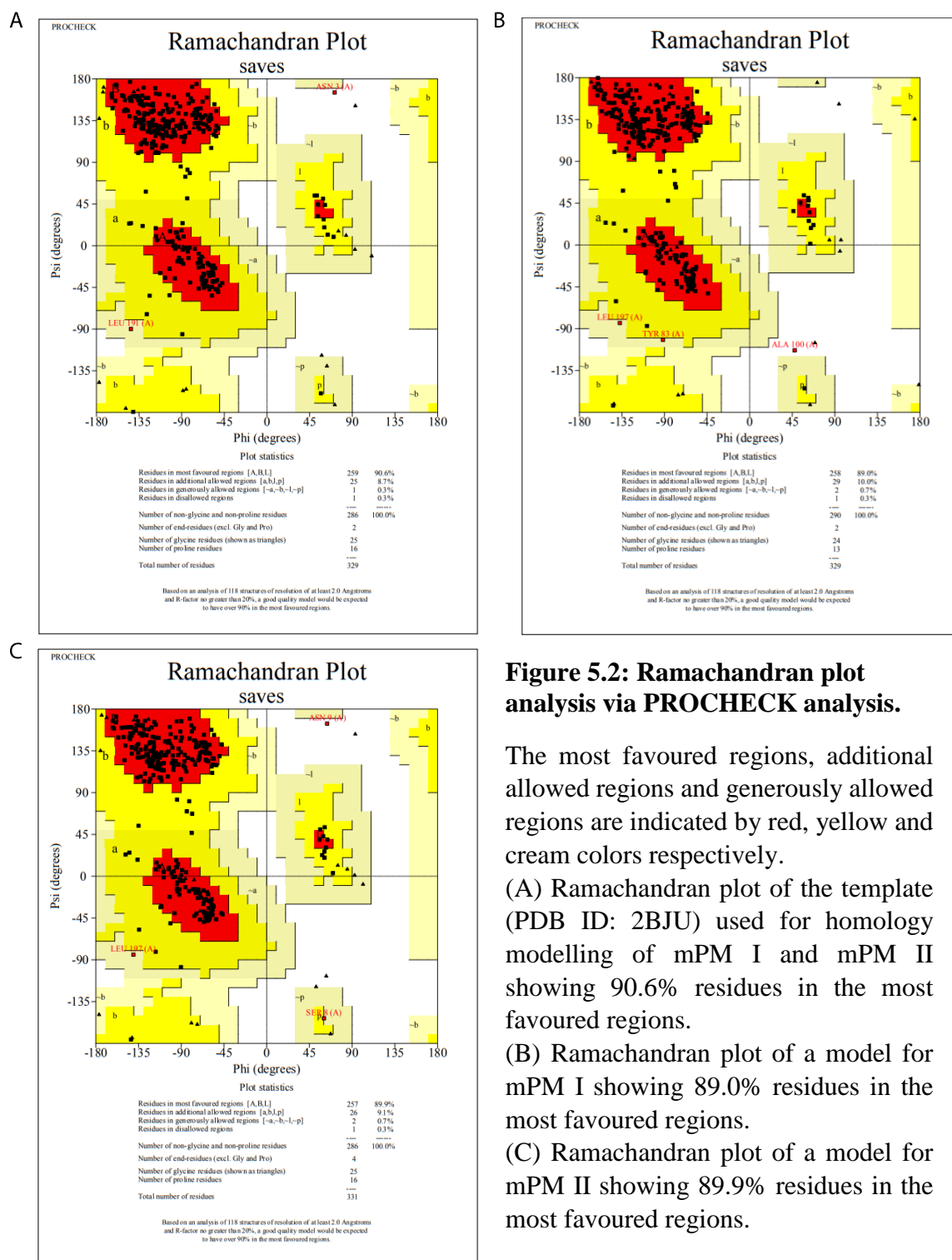


Figure 5.2: Ramachandran plot analysis via PROCHECK analysis.

The most favoured regions, additional allowed regions and generously allowed regions are indicated by red, yellow and cream colors respectively.

(A) Ramachandran plot of the template (PDB ID: 2BJU) used for homology modelling of mPM I and mPM II showing 90.6% residues in the most favoured regions.

(B) Ramachandran plot of a model for mPM I showing 89.0% residues in the most favoured regions.

(C) Ramachandran plot of a model for mPM II showing 89.9% residues in the most favoured regions.

Table 5.1: Molecular docking simulation between the 3D models of mPM I and mPM II and compounds identified from *A. paniculata* EtOH and *C. wightii* AQ extracts.

S. No.	Ligand	DS	
		mPM I	mPM II
	Pepstatin A	-6.95	-10.42
1	2-amino-2-ethyl-4-(methylsulfonimidoyl)butanoic acid	-7.45	-6.47
2	14-deoxy-11,12-didehydroandrographolide	-4.29	-5.02
3	Andrographolide	-7.41	-7.16
4	Diethyl (2-methylallyl) phosphate	-7.91	-5.23
5	Dimethylaminopropanol	-5.01	-5.14
6	Anhalidine	-6.07	-4.99
7	Anhalolidine	-6.35	-4.76
8	Methyl cinnamate	-3.67	-3.45
9	Myo-inositol	-3.37	-4.02
10	Quinic acid	-6.33	-4.99
11	Safrole	-7.29	-8.76
12	Sorbitol hexaacetate	-7.29	-7.95
13	Valine	-4.07	-4.00

Pepstatin A was used a positive control.

DS: Docking Score

The control ligand pepstatin A and the four compounds with strong binding affinity to mPM I and mPM II were aligned in their best pose with the enzymes and were subjected to LigPlot+ program. When docked with mPM I, the control ligand pepstatin A formed hydrogen bonds (H-bonds) with the catalytic aspartic acid residue Asp220 and Tyr223 (Figure 5.3A). The catalytic residues have been underlined for emphasis at some places in the study. Other amino acid residues Tyr83, Ile129, Try198, Gly 222, Ile218 and Ile306 were involved in hydrophobic interactions. With mPM II pepstatin A formed H-bonds with Asp220, Ser224 and Thr223 (Figure 5.4A). While it formed hydrophobic contacts with Gly42, Tyr198, Gly222, Ile296, Leu298 and Ile306. Pepstatin A H-bond interactions with the catalytic aspartic acid residues were observed in the co-crystal structure of PM II and pepstatin A (PDB ID: 1W6I).

The Ligand 1 was found to make H-bonds with the catalytic aspartic acid residue Asp40 and Gly222 in mPM I (Figure 5.3B) and hydrophobic contacts with Val20, Met21, Ile38, Tyr83, Phe117, Ala120, Phe126, Ile129, Ser224 and Thr223. It formed H-bonds with Ser224 and Tyr198 in mPM II active site (Figure 5.4B). In addition, it formed hydrophobic contacts with the residues Asp40, Gly42, Asp220, Gly222, Thr223, Ile296, Phe300 and Ile306.

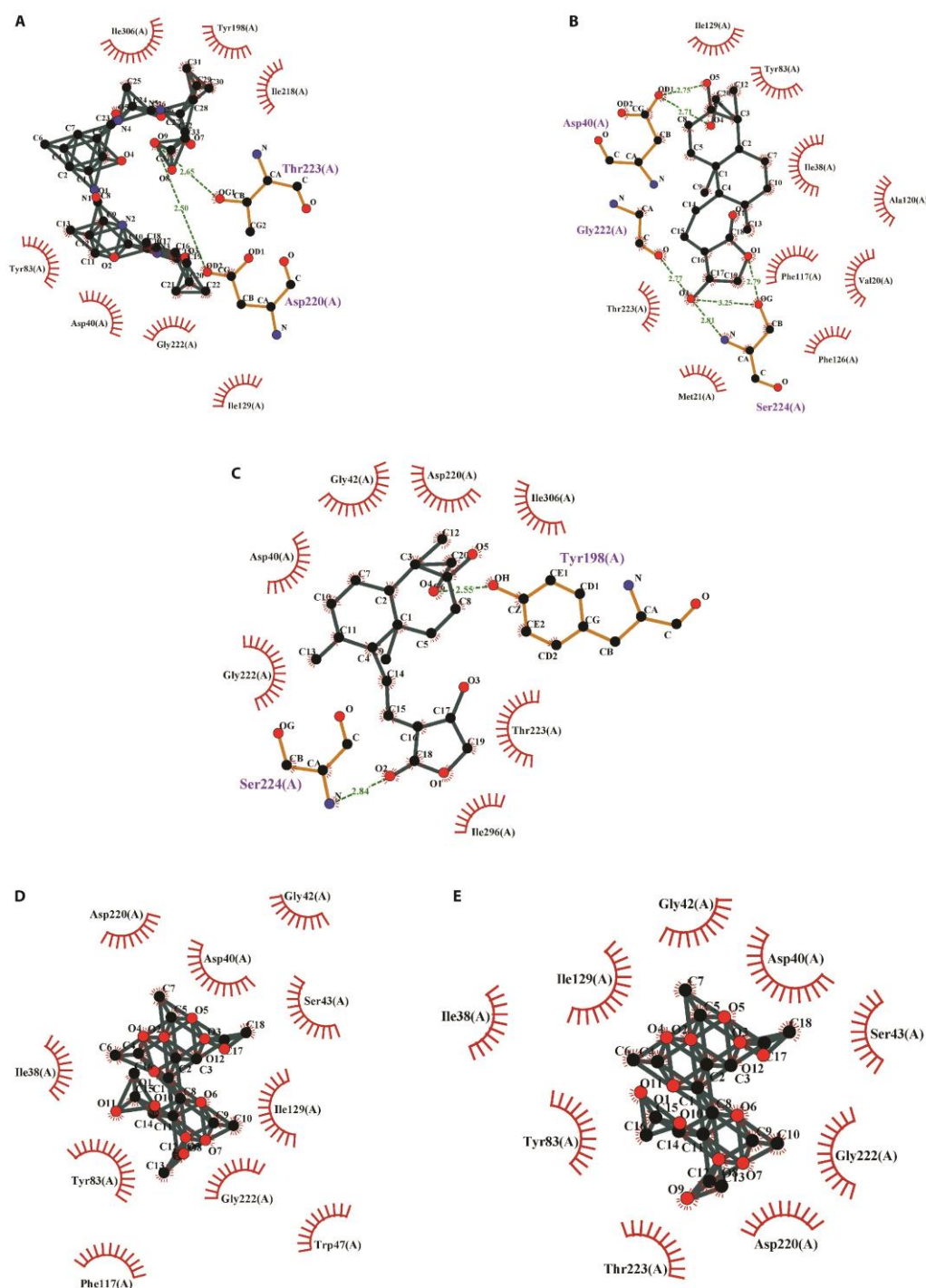


Figure 5.3: Intermolecular interactions between mPM I and ligands.

The interactions between mPM I and the ligands were analyzed by LigPlot+. The ligands and the enzyme side chains are shown in a ball-and-stick representation, with the ligand bonds colored in grey. The H-bonds are shown as green dotted lines, while the spoked arcs represent the enzyme residues and the ligand atoms involved in the hydrophobic contacts. (A) Positive control pepstatin A; (B) 2-amino-2-ethyl-4-(methylsulfonimidoyl)butanoic acid; (C) Andrographolide; (D) Safrole; (E) Sorbitol hexaacetate.

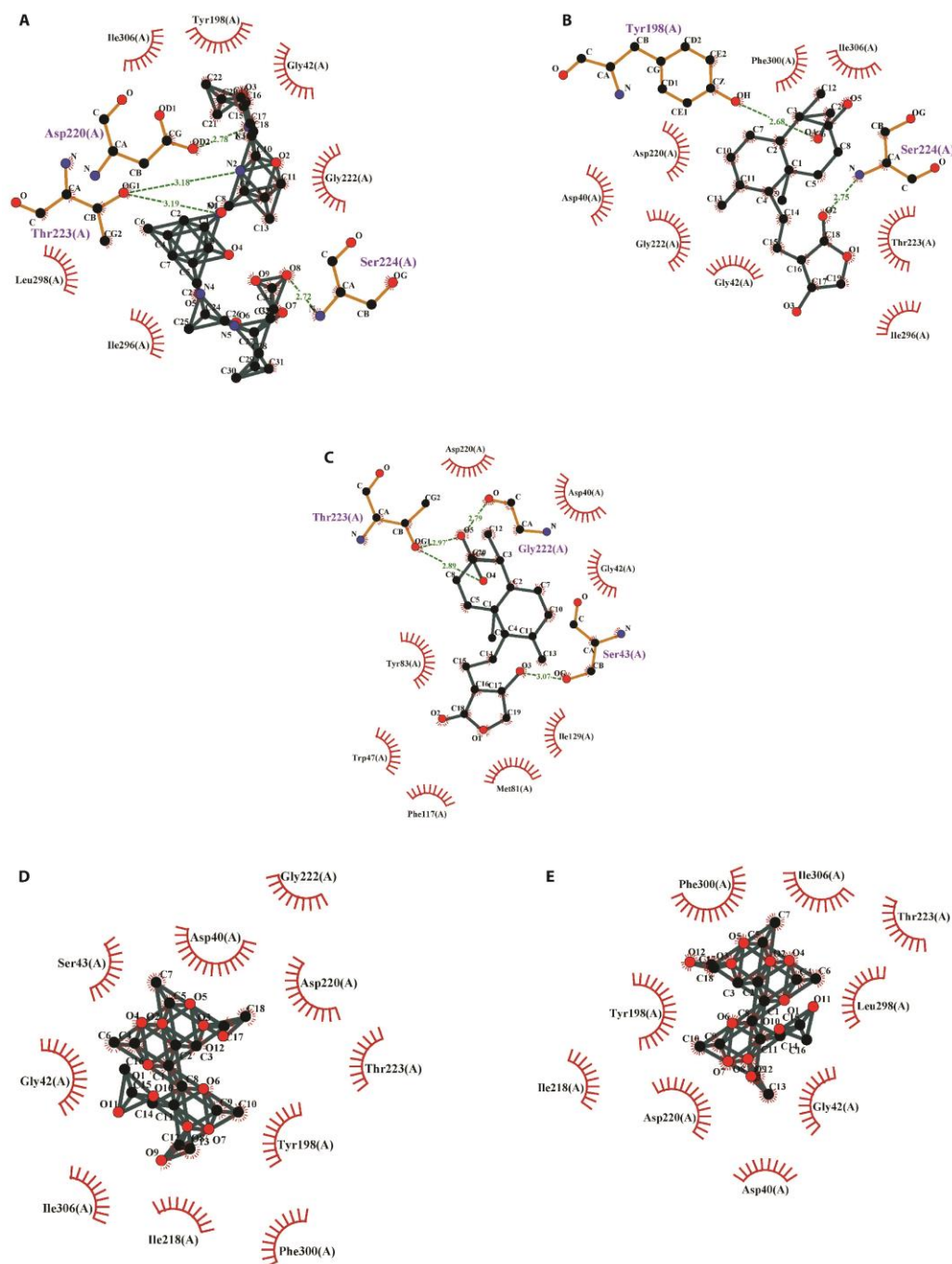


Figure 5.4: Intermolecular interactions between mPM II and ligands.

The interactions between mPM II and the ligands were analyzed by LigPlot+. The ligands and the enzyme side chains are shown in a ball-and-stick representation, with the ligand bonds colored in grey. The H-bonds are shown as green dotted lines, while the spoked arcs represent the enzyme residues and the ligand atoms involved in the hydrophobic contacts. (A) Positive control pepstatin A; (B) 2-amino-2-ethyl-4-(methylsulfonimidoyl)butanoic acid; (C) Andrographolide; (D) Safrole; (E) Sorbitol hexaacetate.

In the case of Ligand 3, the residues Ser224 and Tyr198 of mPM I were involved in H-bond interactions while the ligand made hydrophobic contacts with the residues Asp40, Gly42, Asp220, Gly222, Thr223, Ile296 and Ile306 (Figure 5.3C, Figure 5.5A). This ligand formed H-bonds with the residues Ser43, Gly222 and Thr223 of mPM II enzyme (Figure 5.4C, Figure 5.5B). Additionally, its hydrophobic contacts with Asp40, Gly42, Trp47, Tyr83, Met81, Phe117, Ile129 and Asp220 were found. Andrographolide H-bonding interaction with the catalytic aspartic acid residues in PM II has earlier been reported from *in silico* studies (Megantara et al., 2015).

The Ligand 11 was found to interact with mPM I by forming hydrophobic contacts to Ile38, Asp40, Gly42, Ser43, Trp47, Tyr83, Phe117, Ile129, Asp220 and Gly222 (Figure 5.3D). In the case of mPM II also, this ligand formed hydrophobic interactions with the enzyme residues Gly42, Ser43, Asp40, Tyr198, Ile218, Asp220, Gly222, Thr223, Phe300 and Ile306 (Figure 5.4D).

Lastly, the Ligand 12 was found to make hydrophobic contacts with Ile38, Asp40, Gly42, Ser43, Tyr83, Ile129, Asp220, Gly222 and Thr223 residues of mPM I (Figure 5.3E). In mPM II it formed hydrophobic contacts with Asp40, Gly42, Tyr198, Asp220, Ile218, Thr223, Leu298, Phe300 and Ile306 (Figure 5.4E).

Altogether all the four ligands with strong binding affinity were found to bind the enzymes in the vicinity of the catalytic aspartic acid residues of the enzymes Asp40 and Asp220. The high affinity with the active site of the enzymes and the H-bond interactions or the possible hydrophobic interactions with the catalytic aspartic acid residues indicate the blocking of the active site residues as the possible mechanism of mPM I and mPM II inhibition by the four ligands which was observed *in vitro*. Frequently observed in protein-ligand complex H-bond plays an important role in interaction as it confers rigidity to the protein structure and specificity to intermolecular interactions (Hubbard and Kamran Haider, 2010). This suggests that 2-amino-2-ethyl-4-(methylsulfonimidoyl)butanoic acid (Ligand 1) and andrographolide (Ligand 3) can form more stable complexes with the enzymes as compared to safrole (Ligand 11) and sorbitol hexaacetate (Ligand 12) and therefore can act as more promising inhibitors against mPM I and mPM II.

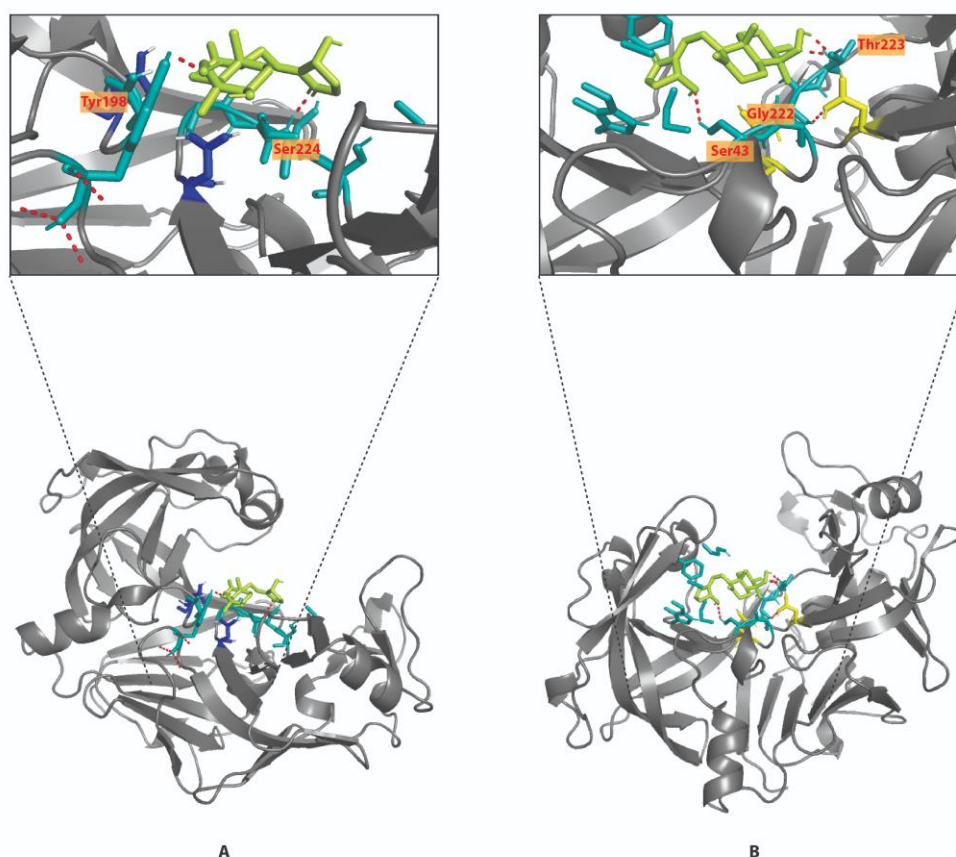


Figure 5.5: Intermolecular interactions between the enzymes mPM I and mPM II and the ligand andrographolide.

The enzyme-ligand complexes were visualized using PyMOL. Andrographolide is shown in green and the interacting enzyme side chains are shown in cyan as sticks. The H-bonds are represented by red dotted lines. The insets provide a magnified view of the interactions. (A) mPM I and andrographolide complex, catalytic acid residues are depicted as sticks in blue; (B) mPM II and andrographolide complex, catalytic residues are depicted as sticks in yellow.

In conclusion, four compounds viz., 2-amino-2-ethyl-4-(methylsulfonimidoyl)butanoic acid and andrographolide from the *A. paniculata* EtOH extract and safrole and sorbitol hexaacetate from the *C. wightii* AQ extract were identified with the ability to inhibit the enzymes mPM I and mPM II by *in silico* studies. The interaction of the compounds with the catalytic dyads marks them as very important PM blocking agents. The compounds 2-amino-2-ethyl-4-(methylsulfonimidoyl)butanoic acid and andrographolide can be taken as lead structures to develop new antiplasmodial drugs with inhibition of PM I and PM II as their mechanism of action. *In vitro* and *in vivo* studies can be performed using the pure compounds to evaluate their synergistic effect and thus assess their candidature for combinational therapy regimen.

