

Structural and Functional Characterization of the Egress and Invasion Machinery of the Malaria Parasite: Proposing a New Way Forward in Malaria Therapeutics from an Atomistic Perspective

Miss Geraldene Munsamy

216039011



A thesis submitted to the College of Health Sciences, University of KwaZulu-Natal
Westville, in fulfilment of the requirements of the degree of Doctor of Philosophy

Supervisor

Prof. Mahmoud Soliman

Durban

2019

Structural and Functional Characterization of the Egress and Invasion Machinery of the Malaria Parasite: Proposing a New Way Forward in Malaria Therapeutics from an Atomistic Perspective

Miss Geraldene Munsamy

216039011

2018

A thesis submitted to the College of Health Sciences, University of KwaZulu-Natal, Westville,
in fulfilment of the requirements of the degree of Doctor of Philosophy

This is to certify that the contents of this thesis are the original research work of Miss
Geraldene Munsamy

As the candidate's supervisor, I have approved this thesis for submission.

Supervisor:

Signed:..... Name:

Date:.....

PREFACE

The thesis is divided into eight chapters, including this one:

Chapter 1

This chapter includes insight into the background, rationale and relevance of the study as well as the projected aim and objectives. The general outline and structure of the thesis concludes this chapter

Chapter 2

Provides a general overview of the malaria parasite and the life cycle of the *Plasmodium falciparum* parasite. This chapter also highlights the role of the aspartic protease family of enzymes known as Plasmepsins (Plm) as formidable targets for the effective treatment of malaria, elucidating their integral role in the egress and invasion of the malaria parasite.

Chapter 3

This chapter provides a brief introduction into computational chemistry, integrating molecular modeling and molecular dynamic simulation protocols and their applications in the drug design and discovery process. The computational techniques that have been theoretically explained, in this chapter have been applied in this study.

Chapter 4: (Published work- this chapter is presented in the required format of the journal and is the final version of the published manuscript)

This chapter includes the homology protocol integrated in this study for the generation of the predicted structures of the latter aspartic protease family members expressed within the *Plasmodium falciparum* species namely; Plm VI-X. Structural and functional analysis was performed on PIX and PX due to the crucial role each enzyme plays in erythrocyte egress and invasion within the human host. Structural and sequence analysis was performed on PIX and PX to accentuate the importance of designing inhibitors specific to each enzyme. The flap dynamics of PIX and PX were also explored in this study integrating a defined set of parameters specific to aspartic proteases, to gain insight into the biological activity of PIX and PX respectively. This article has been published in RSC Advances (IF = 3.096).

Chapter 5: (Published work- this chapter is presented in the required format of the journal and is the final version of the published manuscript)

This chapter presents the binding landscape of Plm IX and X in complex with the experimental peptidomimetic competitive inhibitor known as 49c. This study exhibits the supremacy of a hydroxyl-ethyl-amine scaffold as a potential therapeutic inhibitor of the aspartic protease family of enzymes. Thermodynamic analysis integrating per-residue energy decomposition analysis were performed to accentuate the key amino acid residues of PIX and PX that are responsible for binding as well as key moieties of 49c that render structural restraints to the flap and hinge domain crucial for activity in PIX and PX respectively. This article has been published in the Journal of Cellular Biochemistry (IF=3.469)

Chapter 6: (Submitted work-this chapter is presented in the required format of the journal and is the final version of the submitted manuscript)

This chapter, “Unveiling a new era in Malaria therapeutics: A tailored molecular approach towards the design of Plm IX inhibitors” presents the integration of bio-molecular computational techniques for the design of tailored inhibitors explicitly to PIX. A pharmacophore model approach was employed to screen for potential hits specific to PIX utilizing the virtual screening protocol. In this study, comparative molecular dynamic simulations were performed on the apo and screened hits bound to the protein to reveal the molecular mechanism and structural conformation of the flap domain of PIX, thus aiding in the design of effective inhibitors against this parasitic viral target. This manuscript has been submitted to The Protein Journal (submission ID: JOPC-D-18-00182R2)

Chapter 7: (Submitted work-this chapter is presented in the required format of the journal and is the final version of the submitted manuscript)

This chapter is entitled “Fundamental class of inhibitors activates “fireman’s grip”: An enhanced binding analysis in search of effective malarial therapeutics”. This study demonstrates comparative structural analysis based on pharmacokinetic profiling and bio-molecular analysis of three main classes of aspartic protease inhibitors. The insight extracted from this study reveal the structural superiority of a single class of inhibitors known as the aminohydantoin class of inhibitors, as they were able to effectively stabilize the binding domain employing the use of the “Fireman’s grip” This study further affirms the experimental finding of CWHM-117, an aminohydantion derived inhibitor which exhibited nanomolar

inhibitory activity against *Plasmodium* aspartic protease PIX and PX. From the investigations carried out in this study, the experimental inhibitor CWHM-117 exhibited structural inhibitory superiority, thus advancing the understanding of aminohydantoins in malaria therapy. This manuscript has been submitted to RSC advances journal (submission ID: RA-ART-03-2019-002290).

ABSTRACT

The past decade has witnessed numerous efforts to control the invasive tactics of the malarial parasite, including focused research towards selective malarial inhibitors of *Plasmodium falciparum*, the most lethal strain of the *Plasmodium* species. The recent discovery of the key mediators of egress and erythrocyte invasion of the malaria parasite has opened a new avenue that may be harnessed for the development of effective therapeutics that may permanently eradicate the malaria virus. These new parasitic targets of *P. falciparum* are PIX and PX and have gained considerable attention in drug discovery pipelines however, the absence of crystal structures of these enzymes evidenced a lack in structural information, as there is currently little known regarding the structural dynamics, active site domains and the mechanism of inhibition of these enzymes. This has therefore led to the modeling of the 3D protein structure of each enzyme to gain a fundamental understanding regarding the structural and functional characteristics that may be visualized from an atomistic perspective. The emergence of new drug targets has led to the integral use of computational techniques including molecular modeling, molecular docking, virtual screening protocols and molecular dynamic simulations which allow chemists to evaluate and assess millions of compounds and thus funnel out potential lead drugs. These *in silico* techniques further justify the current use of Computer-Aided Drug Design as a cost-effective approach to fast track the drug discovery process. The above-mentioned techniques, amongst a vast range of other computational tools were integrated in this study to provide insight into conformational changes that elucidate potential inhibitory mechanisms, identification of the active site cleft, characterization and pharmacophoric features leading to novel small molecule inhibitors. This study focused on analysing the flap dynamics specific to the aspartic protease family of enzymes using a defined set of parameters to map out the binding domain for the design of potential antimalarial drugs. To gain a molecular perspective of the conformational binding of two proposed experimental drugs which showed substantial inhibitory activity against PIX and PX molecular dynamic simulations were performed and further evaluated employing *in silico* thermodynamic analysis to provide insight into the proposed binding of mode of each inhibitor, highlighting the key moieties required for binding. A pharmacophoric model was also generated using *in silico* tools to screen for tailored inhibitors specific to PIX. The aim of this study was to generate fundamental insight into the structural and functional characterization of two prominent targets that play an indispensable role in survival of the malaria virus. The implementation of the

information extracted from this study, may provide a structural outline for molecular biologists, and pharmaceutical scientists to aid in the design of novel antimalarial therapeutics.

DECLARATION I -PLAGIARISM

I, Geraldene Munsamy, declare that

1. The research reported in this thesis, except where otherwise indicated, is my original research.
2. This thesis has not been submitted for any degree or examination at any other university.
3. This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.
4. This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
 - a. Their words have been re-written, but the general information attributed to them has been referenced.
 - b. Where their exact words have been used, then their writing has been placed in italics and inside quotation marks and referenced.
5. This thesis does not contain text, graphics or tables copied and pasted from the internet, unless specifically acknowledged, and the source being detailed in the thesis and in the References sections.

A detailed contribution to publications that form part and/or include research presented in this thesis is stated (include publications submitted, accepted, in press and published).

Signed: **G.Munsamy**

DECLARATION II- LIST OF PUBLICATIONS

1. Geraldene Munsamy, Pritika Ramharack and Mahmoud Soliman (2018) Egress and invasion machinery of malaria: an in-depth look into the structural and functional features of the flap dynamics of PIX and PX, RSC Advances, 8(39), pp.21829-21840 (*published*)

Contribution:

Geraldene Munsamy: contributed to the project by performing all the experimental work and manuscript preparation and writing.

Pritika Ramharack: contributed by performing post dynamic analysis, writing of manuscript and creation of all graphs and images.

Mahmoud E.S. Soliman: Supervisor

Appendix A: Pdf version of the publication

2. Geraldene Munsamy, Clement Agoni and Mahmoud Soliman (2018) A dual target of PIX and PX: Unveiling the atomistic superiority of a core chemical scaffold in malaria therapy, Journal of Cellular Biochemistry, 120(5), pp.7876-7887 (*published*)

Contribution:

Geraldene Munsamy: contributed to the project by performing all the experimental work and manuscript preparation and writing.

Clement Agoni: contributed to the project by creating all images and graphs and writing.

Mahmoud E.S. Soliman: Supervisor

Appendix B: Pdf version of the publication

3. Geraldene Munsamy and Mahmoud Soliman (2018) Unveiling a new era in Malaria therapeutics: A tailored molecular approach towards the design of PIX inhibitors, The Protein Journal (*submitted*)

Contribution:

Geraldene Munsamy: contributed to the project by performing all the experimental work and manuscript preparation and writing

Mahmoud E.S. Soliman: Supervisor

Appendix C: Pdf version of the submitted manuscript

4. Geraldene Munsamy, Clement Agoni, Pritika Ramharack and Mahmoud Soliman (2019) Fundamental class of inhibitors activates “fireman’s grip”: An enhanced binding analysis in search of effective malarial therapeutics, *RSC Advances(submitted)*

Contribution:

Geraldene Munsamy: contributed to the project by performing manuscript preparation and writing.

Clement Agoni: contributed to the project by performing experimental work

Pritika Ramharack : contributed by performing post dynamic analysis and creation of all graphs and images.

Mahmoud E.S. Soliman: Supervisor

Appendix D: Pdf version of the submitted manuscript

RESEARCH OUTPUT

A- LIST OF PUBLICATIONS

1. Geraldene Munsamy, Pritika Ramharack and Mahmoud Soliman (2018) Egress and invasion machinery of malaria: an in-depth look into the structural and functional features of the flap dynamics of PIX and PX, *RSC Advances*, 8(39), pp.21829-21840.
2. Geraldene Munsamy, Clement Agoni and Mahmoud Soliman (2018) A dual target of PIX and PX: Unveiling the atomistic superiority of a core chemical scaffold in malaria therapy, *Journal of Cellular Biochemistry*, 120(5), pp.7876-7887.

- IN PRESS

1. Geraldene Munsamy and Mahmoud Soliman (2018) Unveiling a new era in Malaria therapeutics: A tailored molecular approach towards the design of PIX inhibitors, *The Protein Journal*.
2. Geraldene Munsamy, Clement Agoni, Pritika Ramharack and Mahmoud Soliman (2019) Fundamental class of inhibitors activates “fireman’s grip”: An enhanced binding analysis in search of effective malarial therapeutics, *RSC Advances*.

B- CONFERENCES

1. Poster Presentation “Egress and invasion machinery of malaria: an in-depth look into the structural and functional features of the flap dynamics of PIX and PX”- 12th National CHPC Conference, Century City Cape Town, South Africa, 2-6 December 2018.

ACKNOWLEDGEMENTS

Psalms 23:4 *“Even when I walk through the darkest valley, I will not be afraid for you are close beside me”*

My Parents

Thank you for your unwavering encouragement, support, prayers and love. Thank you for raising me to be the person that I am today, I am eternally grateful to you both. This work would never have been possible without your support. You will always be my shining stars in the darkest of nights. I love you both!

Professor M.E.S Soliman

Thank you for your valuable guidance, advise, understanding and motivation throughout my journey as a MSc and Ph.D. student.

The Molecular Bio-Computation & Drug Design Group

Thank you all for the constant laughter, your company, pleasant nature, friendships and being my home away from home.

My Loved ones, siblings and Friends

Thank you all for your concern, kind words and love, I truly appreciate it all. Ronald, thank you for choosing to be a part of my journey, you are my greatest friend and happy ending.

University of KwaZulu-Natal College of Health Sciences and National Research Foundation

Financial support as postgraduate student.

YAHWEH SHAMMAH

LIST OF ABBREVIATIONS

3D	Three-Dimensional
Å	Amperes
ACT's	Artemisin-based combination therapies
DCCM	Dynamic Cross Correlation
DNA	Deoxyribonucleic acid
FDA	Food and Drug Administration
GAFF	General amber Force Field
HAP	Histo aspartic protease
K	Kelvin
MC	Monte Carlo
MD	Molecular Dynamics
MM	Molecular Mechanics
MM/GBSA	Molecular Mechanics/Generalized Born Surface Area
MM/PBSA	Molecular Mechanics/Poisson-Boltzmann Surface Area
NMR	Nuclear Magnetic Resonance
Ns	nanoseconds
<i>P. falciparum</i>	<i>Plasmodium falciparum</i>
<i>P. malariae</i>	<i>Plasmodium falciparum</i>
<i>P. ovale</i>	<i>Plasmodium ovale</i>
<i>P. vivax</i>	<i>Plasmodium vivax</i>
PBVS	Pharmacophore-based virtual screening
PCA	Principal Component Analysis
PDB	Protein data bank
PES	Potential energy surface
PEXEL	<i>Plasmodium</i> export element
PI	plasmepsin I
PII	plasmepsin II

PIII	plasmepsin III
PIV	plasmepsin IV
PIX	plasmepsin IX
PME	Particle-mesh Ewald method
PRED	Per Residue Decomposition
PV	plasmepsin V
PVI	plasmepsin VI
PVII	plasmepsin VII
PVIII	plasmepsin VIII
PX	plasmepsin X
QM	Quantum Mechanics
RBC	Red blood cell
RdRp	RNA Dependent RNA polymerase
RESP	Restrained Electrostatic Potential
RMSD	Root Mean Square Deviation
RMSF	Root Mean Square Fluctuation
RoG	Radius of Gyration
RSCB	Research Collaboratory for Structural Bioinformatics
SASA	Solvent accessible surface area
SUB1	Subtilisin-like serine protease
VS	Virtual screening
WHO	World Health Organization
A	Alpha
B	Beta
ΔG	Free Binding energy

LIST OF AMINO ACIDS

Three Letter Code	Amino Acid
Ala	Alanine
Arg	Arginine
Asn	Asparagine
Asp	Aspartic Acid
Cys	Cysteine
Gln	Glutamine
Glu	Glutamic Acid
Gly	Glycine
His	Histidine
Ile	Isoleucine
Leu	Leucine
Lys	Lysine
Met	Methionine
Phe	Phenylalanine
Pro	Proline
Ser	Serine
Thr	Threonine
Trp	Tryptophan
Tyr	Tyrosine
Val	Valine

TABLE OF CONTENT

PREFACE.....	III
ABSTRACT.....	VI
DECLARATION I -PLAGIARISM.....	VIII
DECLARATION II- LIST OF PUBLICATIONS.....	IX
RESEARCH OUTPUT.....	XI
ACKNOWLEDGEMENTS	XII
LIST OF ABBREVIATIONS.....	XIII
LIST OF AMINO ACIDS	XV
CHAPTER 1	1
1.1 Background and Rationale.....	1
1.2 Aims and Objectives	4
1.3 Novelty and significance.....	5
1.4 References.....	6
CHAPTER 2	11
2.1 Introduction.....	11
2.2 Life Cycle of the Malaria parasite	12
2.3 Aspartic Proteases of <i>Plasmodium falciparum</i>	14
2.4 Plasmepsins of <i>Plasmodium falciparum</i>	15
2.5 New targets for Malaria chemotherapy: Erythrocyte Rupture and Invasion	16
2.6 References.....	18
CHAPTER 3	24
3.1 An Introduction into Computational Chemistry	24
3.2 Quantum mechanics.....	24
3.3 The Schrödinger equation.....	25
3.4 The Born-Oppenheimer approximation.....	26

3.5 Potential energy surface	27
3.6 Molecular mechanics	28
3.7 Force fields.....	29
3.8 Molecular dynamics.....	30
3.9 Molecular Dynamics Post-Analysis.....	31
3.10 System Stability of Simulated Systems	32
3.10.1 System Convergence.....	32
3.10.2 Root Mean Square Deviation (RMSD).....	32
3.10.3 Radius of Gyration (RoG).....	33
3.11 Thermodynamic calculations	33
3.11.1 Binding free-energy calculations	33
3.12 Conformational Features of System	35
3.12.1 Root Mean Fluctuation (RMSF)	35
3.13 Other Computer-Aided Drug Design Techniques Utilized in the Study.....	35
3.13.1 Molecular docking	35
3.13.2 Homology modeling	36
3.13.3 Virtual Screening (VS)	37
3.14 References.....	38
CHAPTER 4.....	48
4.1 Graphical Abstract	49
4.2 Abstract	50
4.3 Introduction.....	51
4.4 Computational Methodology	53
4.4.1 In silico Modelling.....	53
4.4.2 System preparation.....	54
4.4.3 Molecular dynamics and post dynamic analysis.....	54
4.4.4 Parameters used in this study.....	55

4.5 Results and discussion	56
4.5.1 Exclusive Structural Features of Plasmepsins	56
4.5.2 Mechanistic Flexibility of plasmepsin IX and X	57
4.5.3 Plasmepsin Residue Fluctuations.....	58
4.5.4 Distance [d1] and Dihedral Angle [Φ] Analysis.....	61
4.5.5 Tric α angles θ_1 and θ_2	63
4.5.6 Radius of Gyration (ROG).....	66
4.6 Conclusions.....	67
4.7 References.....	69
CHAPTER 5.....	74
5.1 Abstract.....	75
5.2 Introduction	76
5.3 Computational Methods	78
5.3.1 Systems Preparation.....	78
5.3.2 Molecular Docking	79
5.3.4 Binding Free Energy Calculations	80
5.3.5 Dynamic cross correlation	81
5.4 Results and Discussion.....	82
5.4.1 Assessment of structural stability of systems	82
5.4.2 49c impedes residue fluctuation in plasmepsin IX and X.....	82
5.4.3 Inhibitory impact of 49c on the flap motion of plasmepsin IX and X.....	84
5.4.4 Exploration of structural compactness of plasmepsin IX and X upon binding of 49c	88
5.4.5 Analogous binding characteristics of 49c on plasmepsin IX and X	89
5.5 Conclusion.....	93
5.6 References.....	95
CHAPTER 6.....	100

6.1 Abstract	101
6.2 Introduction.....	102
6.3 Computational Methodology	105
6.3.1 Pharmacophore model generation and Screening.....	105
6.3.2 Molecular docking	106
6.4 System preparation.....	106
6.4.1 Molecular dynamics.....	106
6.4.2 Post-Dynamic Analysis.....	108
6.4.3 Per-Residue Decomposition Analysis & Binding Free Energy Computation	108
6.5 Molecular dynamic analysis and Post MD Analysis	109
6.5.1 Screened compounds facilitate the restriction of PIX structural flexibility.....	109
6.5.2 Distance analysis between the flap tip and hinge residue.....	110
6.5.3 Radius of Gyration.....	112
6.5.4 Binding free energy calculations and Per-residue energy decomposition analysis	113
6.6 Conclusion	117
6.7 References.....	119
CHAPTER 7	125
7.1 Abstract.....	126
7.2 Background	127
7.3 Characterizing Existing Malarial Aspartic Protease Inhibitors	128
7.4 Assessing the paradigm of Aminohydroxy acids in malaria therapy	130
7.5 Plasmeprin and CWHM-117 acquisition and preparation.....	131
7.5.1 Molecular Docking	131
7.5.2 Molecular Dynamic (MD) Simulations	132
7.6 Post-Dynamic Analysis.....	133
7.6.1 Binding Free Energy Calculations	133

7.7 Results and Discussion	133
7.7.1 Pharmacokinetic Profiling of CWHM-117 demonstrates favorable drug-likeness	134
7.7.2 Comparative bioinformatic analysis of Plasmepsins demonstrate greater interactions with Aminohydantoins.....	134
7.7.3 Sensitivity of PIX and PX to CWHM-117 result in transformed conformational ensembles.....	135
7.7.4 Enhanced binding prowess of CWHM-117 to PIX over PX	139
7.5 Conclusion	142
7.6 References.....	144
CHAPTER 8	148
8.1 Conclusion	148
8.2 Future perspectives	149
SUPPLEMENTARY INFORMATION	150
APPENDICES	161

TABLE OF FIGURES

Figure 1.1 The complex life cycle of the <i>Plasmodium falciparum</i> parasite. Liver, erythrocytic, transmission, and mosquito stages ¹⁴	2
Figure 2.1 Global incidence of malaria in the year 2017. ⁵	12
Figure 2.2 Life cycle of the malaria parasite taken from proteases as antimalarial.....	13
Figure 2.3 Catalytic mechanism of aspartic proteases taken from targeting the active site of aspartic proteases. ⁴³	15
Figure 2.4 Erythrocytic cycle of the red blood cell mediated by the malaria parasite. ⁶⁰	17
Figure 3.1 A Graphical representation of a two-dimensional potential energy surface ²⁷	27
Figure 3.2 Diagrammatic representation of the total potential energy function of a molecule, as mentioned above.....	29
Figure 4.1 A diagrammatic representation of plasmpesin VI-X, highlighting the role of each plasmepsin towards malarial infection and invasion.	52
Figure 4.2 Schematic representation of the parameters used to define the flap-structure motion: d1 the distance between the flap tip and flexible region hinge residue, the dihedral angle Φ and the TriC α angles, θ_1 and θ_2 . Asp1, Asp2, the flap tip and flexible region of PIX and PX	56
Figure 4.3 Flap and flexible loop region movements of apo Plm IX (magenta) and X (cyan) hroughout a 50 ns molecular dynamics simulation. The catalytic aspartate residues of Plm IX and X are highlighted in yellow, the flap tip in red and hinge residue in green.	57
Figure 4.4 C- α RMSF plot of <i>Plasmodium falciparum</i> Plm IX (purple), X (cyan), respectively	59
Figure 4.5 Graph showing the fluctuation in the distance, d1 as presented in Figure 2 of Plm IX (black) and Plm X (red).	61
Figure 4.6 : Graph of the dihedral angle throughout the 50 ns simulation, of Plm IX (blue) and X (red).	63
Figure 4.7 Graph showing the Tri C α angles θ_1 of Plm IX (pink) and X (cyan), respectively	64
Figure 4.8 Graph showing the Tri C α angles θ_2 of Plm IX (pink) and X (blue), accordingly	65
Figure 4.9 Radius of Gyration (ROG) of Plm IX (mustard yellow) and Plm X (bottle green)..	66
Figure 5.1 Schematic representation of inhibitor 49c blocking the various stages of egress and invasion of the <i>Plasmodium falciparum</i> parasite.	78

Figure 5.2 Root mean square fluctuation of the apo systems of PIX (blue), PIX in complex with 49c (red) and Pepstatin(green).....	82
Figure 5.3 Root mean square fluctuation of the apo systems of PX (pink), PX in complex with 49c (olive green) and Pepstatin(orange).	83
Figure 5.4 Graphical plots and visual representation of the Distance analysis between the flap tip and hinge residue throughout the 100ns MD simulation of apo PIX (blue), PIX-49c (red) and PIX-Pepstatin (green).....	86
Figure 5.5 Graphical plots and visual representation of the Distance analysis between the flap tip and hinge residue throughout the 100ns MD simulation of apo PX (pink), PX-49c (olive) and PX-Pepstatin (orange).	87
Figure 5.6 Radius of gyration (ROG) of apo PIX, PIX in complex with inhibitors 49c and Pepstatin on the left and apo PX, PX in complex with 49c and Pepstatin on the right	89
Figure 5.7 Molecular visualization of 49c at the catalytic sites (hydrophobic pockets) of plasmepsin IX. Accentuating the Inter-molecular interactions between 49c and catalytic site residues in PIX.....	91
Figure 5.8 Molecular visualization of 49c at the catalytic sites (hydrophobic pockets) of plasmepsin X. Highlighting the inter-molecular interactions between the catalytic site residues and 49c.	92
Figure 6.1 A PRED-based approach outline applied in the study	105
Figure 6.2 Pharmacophore model generation adapted from the PIX-49c complex. The orange circles spotlight the pharmacophoric moieties that were chosen for the model, based on the highest contributing residues, depicted in the binding affinity graph.....	106
Figure 6.3 Root Mean Square Fluctuation of PIX in complex with 49c, Compound 1, Compound 2 and Compound 3. The highlighted regions include the catalytic dyad Asp 53, Asp232 (magenta), the flap tip Gly97 (cyan) and hinge residue Glu305 (gold).	110
Figure 6.4 (a) Radius of gyration (ROG) of PIX in complex with 49c, Compound 1, Compound 2, and Compound 3. (b) Distance measured in Angstroms between the flap tip and hinge residue of PIX in complex with 49c, Compound 1, Compound 2 and Compound 3.....	112
Figure 6.5 MM/GBSA binding free energy profiles of PIX in complex with Compound 1. Highlighting the interacting amino acid residues of PIX and their energy contributions to the total binding free energy	113
Figure 6.6 MM/GBSA binding free energy profiles of PIX in complex with Compound 2. Highlighting the interacting amino acid residues of PIX and their energy contributions to the total binding free energy	114

Figure 6.7 MM/GBSA binding free energy profiles of PIX in complex with Compound 3. Highlighting the interacting amino acid residues of PIX and their energy contributions to the total binding free energy	114
Figure 6.8 MM/GBSA binding free energy profiles of PIX in complex with peptidomimetic inhibitor 49c. Highlighting the interacting amino acid residues of PIX and their energy contributions to the total binding free energy	115
Figure 7.1 General molecular scaffold and comparative characteristics of the three main classes of malaria protease inhibitors.	131
Figure 7.2 : Comparative sequence analysis of PIX and PX subsequent to binding of 49C and CWHM-117. Aspartic acid residues Asp80, Asp259, Asp53 and Asp232 remain conserved among all complexes (yellow). Highlighted residues show active site interactions.....	135
Figure 7.3 Stability of PIX and PX before and after the binding of CWHM-117. Conspicuous fluctuations are observed in the PIX systems at 70-80ns and in the PX systems between 45-65ns. This is validated by the structural modifications observed in the RMSF graphs.....	137
Figure 7.4 Stabilizing hydrogen bonds between plasmepsin complexes. Complexes on the right depict hydrogen bond donors and acceptors upon binding with CWHM-117.....	138
Figure 7.5 Binding free energy profile of PIX-CWHM-117 system. A) 3D predicted structure of PIX bound with CWHM-117. B) CWHM-PIX interaction network of binding site residues. C) Total contributing energies of binding site residues	141
Figure 7.6 Binding free energy profile of PX-CWHM-117 system. A) 3D predicted structure of PIX bound with CWHM-117. B) CWHM-PX interaction network of binding site residues. C) Total contributing energies of binding site residues	142

TABLE OF TABLES

Table 4.1 Residues used to calculate the distance (d1), dihedral (Φ), TriC α θ_1 and θ_2	58
Table 4.2 RMSF Values of Apo Plm IX and X for the Binding Site Residues, Tip and Hinge Residue, and Total Average Fluctuation.....	60
Table 4.3 Measurement of the distance by which the flap-structure moves, measured in Angstroms.	62
Table 5.1 RMSF values of apo Plm IX and X, as well as each Plm in complex with 49c and Pepstatin, highlighting the fluctuation observed in the binding site residues, tip, hinge residue, and total average fluctuation.....	83
Table 5.2 Measurement of the distance by which the flap-structure moves, detected in all PIX and PX systems measured in Angstroms	85
Table 5.3 MM/GBSA binding free energy profiles of 49c and Pepstatin interactions with PIX and PX.....	90
Table 6.1 MM/GBSA binding free energy profiles of 49c, Compound 1, Compound 2 and Compound 3 interactions with PIX and PX.....	116
Table 7.1 Currently available anti-malaria drugs.....	129
Table 7.2 Pharmacokinetic characteristics of CWHM-117 using SWISSADME	134
Table 7.3 Measurement of the distance by which the flap-structure moves, measured in Angstroms	138
Table 7.4 MM/GBSA-based binding free energy profile of CWHM-117	140

CHAPTER 1

1.1 Background and Rationale

Despite the global effort directed towards malarial intervention strategies which aims at the treatment and prevention of malaria, the prevalence of this parasitic disease remains relentless. Malaria continues to rampant the world on a global scale affecting the lives and health of millions of people worldwide.^{1,2} 90% of estimated malarial deaths are observed in Africa, as Africa carries the bulk of the global burden. The high mortality rate observed in Sub-Saharan Africa is reported in infected children under the age of 5 who are most susceptible to this vicious disease.^{3,4}

The transmission of malaria is mediated by the female *Anopheles* mosquito activated by the *Plasmodium* species of parasites.⁵ The four-parasite species that are responsible for the spread of malaria in humans include , *P. malariae* *P. falciparum*, *P. ovale* and *P. vivax* of which *P. falciparum*, and *P. vivax* are classified as the most relentless and *P. falciparum* as the most lethal of all species.^{6,7} Finding a therapeutic treatment against *P. falciparum* is the most viable route to impede and eradicate the spread of malaria, as *P. falciparum* is the highest contributor to the high mortality and morbidity rate associated with this invasive disease.⁸

Antimalarial drugs developed against *P. falciparum* has led to a meagre reduction in malaria-related deaths worldwide particularly, artemisinin-based combination therapies (ACTs)^{9,10}. The current failure of *P. falciparum* species to malarial therapeutics is due to developed drug resistance to frontline artemisinins and other affiliated drugs.¹¹ The life cycle of *P. falciparum* as presented in Figure 1.1 is quite complex and requires the expression of specialized proteins to facilitate survival within invertebrate and vertebrate host intra- and extracellular environments.^{12,13}

An attractive approach towards the effective development of antimalarial vaccines and drugs is to target the expression of specific proteins and most vulnerable phases of the parasites life cycle where it is exposed and defenceless.¹⁵ One specific group of therapeutic targets that have gained profound attention over the years due to the significant role they play in the survival of the malaria parasite is the aspartic protease class of enzymes.^{16,17,18}

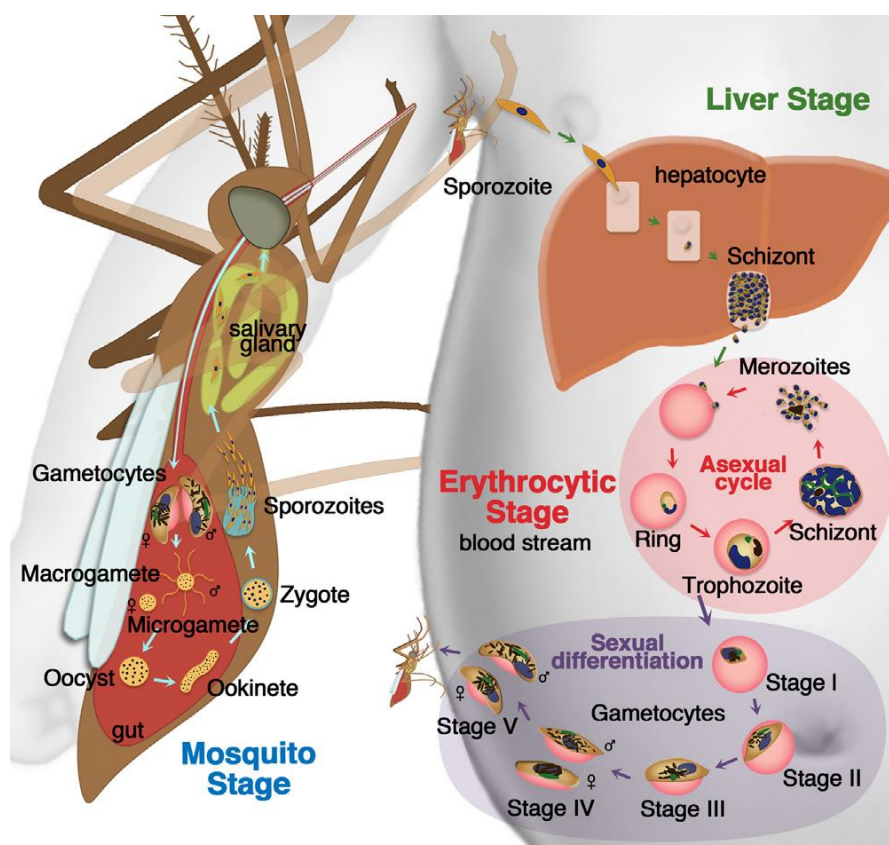


Figure 1.1 The complex life cycle of the *Plasmodium falciparum* parasite. Liver, erythrocytic, transmission, and mosquito stages¹⁴.

Aspartic proteases are a class of enzymes that play a connective role in number of life threatening diseases such as Alzheimer's disease (β -secretase), fungal infections (secreted aspartic proteases), and hypertension (renin) malaria (plasmepsins)^{19,20,21}. Major attention has been centred around the role undertaken by aspartic proteases in the progression of malaria. The aspartic protease family comprise of a diverse group of enzymes that play instrumental yet unique biological roles that aid in the survival and invasion of the malarial parasite. *P. falciparum* consists of 10 aspartic proteases family members (I-X), commonly known as Plasmepsins (Plm), of the 10 plasmepsins only PI (plasmepsin I), PII (plasmepsin II), PIII (plasmepsin III), PIV (plasmepsin IV) and PV (plasmepsin V) have been studied extensively. Some of the biological functions of plasmepsins include degradation of haemoglobin for the provision of nutrients to the parasite which is mediated by the four homologous plasmepsins PI, PII, PIII and PIV found in the digestive vacuole¹⁶. PV is sited at the parasitic endoplasmic reticulum, and facilitates the export of the dominant PEXEL (*Plasmodium* export element) flagged proteins into the cytoplasm as well as the surface of infected RBC's^{22,23}.

Plm VI, VII and VIII are released into the mosquito vector during sporozoite development, motility and midgut transversal of the parasite, although these enzymes play an imperative role within the vector they are not directly responsible for the clinical manifestations released by the parasite within the human host.^{24,25} Recent studies have disclosed the crucial role of plasmepsin IX (PIX) and plasmepsin X (PX) as egress mediators and parasite invasion machinery.²⁶ The egress and invasion stages of the parasites life cycle is the stage in which the parasite is most exposed and susceptible, therefore forming a new avenue towards the design of therapeutics tailored specifically to PIX and PX which may circumvent the development of drug resistance as well as terminate the spread of infection with minimal adverse effects to the host.^{27,28,18} A major challenge regarding this approach is the minimal literature available evidencing the functional and structural characteristics of PIX and PX, lack of tailored inhibitors specific to PIX and PX as well as evolving mutations of the disease. Therefore, developing a novel drug that meets all the requirements of inhibiting PIX and PX utilizing conventional techniques may be laborious and costly.

The construction of therapeutics that are target specific may eliminate factors such as drug resistance, as target specificity in combination with accurate surveillance of drug sensitivities of clinical infections is critical to gain an objective view on treatment strategies in addition to mass drug administration that aim to eliminate malaria.^{29,30,31} This has necessitated the development of new therapeutic strategies required for the effective treatment and prevention of malaria. Amidst the strides made in malaria treatment and prevention protocols, extended manufacturing and processing time involved in the drug discovery through to development remains a major challenge in the effective treatment of malaria.

Computational chemistry transcends into the drug discovery process, by reducing the drug development process by a fraction of the normal timescale^{32,33} therefore new therapeutic drugs are required for the effective treatment of malaria integrating the use of computational techniques which allows for the development of malaria treatment in the most cost-efficient manner. Computational techniques such as molecular modeling and docking^{34,35}, virtual screening³⁶, identification of pharmacophoric hot spots and molecular dynamic simulations³⁷ allow medicinal chemists and biological scientists to screen millions of compounds refining and selecting only the viable leads drugs which may then be validated experimentally.^{38,39} Employing the use of computational tools may alleviate the concept of “shooting in the dark” with experimental screening, by reducing the drug discovery time-line.

1.2 Aims and Objectives

The primary focus of this thesis was to generate the 3D predicted structures of plasmepsin VI-X, as well as investigate the structural and functional characteristics of the principal target proteins involved in the egress and invasion of the malaria parasite integrating the use of *in silico* techniques for the identification and development of potential inhibitors against malaria.

To achieve this, the following objectives were outlined:

1. To investigate the functional and structural characteristics of Plm VI-X, more explicitly PIX and PX due to the instrumental role each protein plays in the egress and invasion of the malaria parasite. This may be achieved by:
 - 1.1 Generating the 3D predicted structures of Plm VI-X, employing the use of the homology modeling protocol.
 - 1.2 Computing 50 ns Molecular dynamic simulations on PIX and PX and implementing successive post molecular dynamic analysis to validate the stability of the protein models.
 - 1.3 Utilizing the structural and sequence data to analyze the structural conformational changes and functional similarities of PIX and PX to other Plm family members to which the 3D crystal structures are known.
 - 1.4 Performing post dynamic analysis integrating a set of proposed parameters specific to aspartic proteases to analyse the flap dynamics of PIX and PX.
2. To investigate the proposed binding mode of experimental inhibitor 49c and Pepstatin to PIX and PX as well as perform thermodynamic analysis of each protein in complex with 49c and Pepstatin respectively. This may be accomplished by:
 - 2.1 Performing 100ns Molecular dynamic simulation on the following biological systems; apo PIX, apo PX, PIX-Pepstatin, PX-Pepstatin PIX-49c and PX-49c.
 - 2.2 Performing post molecular dynamic analysis utilizing a set of proposed parameters specific to aspartic proteases to evaluate the affect each ligand has on the flap dynamics of PIX and PX in respective models.
 - 2.3 Implementing per-residue energy decomposition analysis on all systems based on the Molecular Mechanics/ GB Area Surface Method (MM/GBSA) approach to identify the amino acid residues which form the highest contributions to the overall binding free energy.

3. To implement the *in silico* “Per-residue Energy Decomposition analysis Pharmacophore” screening protocol for the design of tailored inhibitors specific to protein target PIX. This may be achieved by:
 - 3.1 Running a 20ns molecular dynamic simulation of PIX in complex with 49c to generate a molecular dynamic ensemble.
 - 3.2 Generating a pharmacophore model based on the molecular interaction between PIX and 49C, highlighting moieties of 49c that are essential for binding as well as individual amino acid residues of PIX that extend high energy contributions to the total binding free energy.
 - 3.3 Identifying hit compounds with scaffolds that share high structural similarity to the pharmacophore model, by screening the ZINC database to identify compounds with potent pharmacokinetic properties.
 - 3.4 Docking of the top five hit compounds, followed by the identification of the top three hits with the most negative docking score to PIX.
 - 3.5 Computing a 100ns molecular dynamic simulation of PIX in complex with the respective top three hit compounds.
 - 3.6 Computationally validating the top three hits by performing per-residue energy decomposition analysis to decipher the binding free energy of each hit compound to PIX.
4. To investigate the proposed binding activity of hydantoin class of inhibitors, more specifically experimental inhibitor CWHM-117 against aspartic proteases PIX and PX.
 - 4.1 Investigating the pharmacokinetic properties of CWHM-117 that may contribute to its inhibitory activity and target as an aspartic protease inhibitor.
 - 4.2 Exploring the structural and functional binding landscape of PIX and PX in complex with CWHM-117.
 - 4.3 Perform thermodynamic calculations to determine the binding free energy summaries of PIX and PX each in complex with CWHM-117.
 - 4.4 Implementing per-residue energy decomposition analysis on each system based on the MM/GBSA approach to identify the amino acid residues which form the highest contributions to the overall binding free energy.

1.3 Novelty and significance

The global elimination net of malaria is expanding with rapid improvements being made in the targeted treatment and prevention of this parasitic disease. The contribution of malaria to the

overall high mortality rate observed in the year 2017 may be attributed to the emergence of drug resistant strains which has hampered the progress made and led to a swift decline in the predicted elimination of malaria.

An alternative route towards the effective treatment of malaria is the search for novel therapeutic targets. The newly discovered role of the proteins responsible for the egress and invasion of the malarial parasite of which there are no crystal structures, have been highlighted as prominent targets which may effectually lead to the complete eradication of malaria. This study focuses on the structural and functional conformational changes specific to the flap dynamics of PIX and PX, by constructing homology models of respective proteins and further integrating a proposed set of parameters to exhibit the “twisting motion” which are typical of aspartic proteases. This study highlights the sequential and structural characteristics of each of these enzymes to the aspartic protease family of enzymes thus elucidating the importance of designing inhibitors which are specific to each enzyme to alleviate factors such as drug resistance.

The scientific community has taken significant steps to develop effective inhibitors against the aspartic protease enzyme family however, there are still no available Food and Drug Administration (FDA) approved inhibitors against these enzymes. Therefore, in this study we provide a molecular perspective into the proposed binding landscape of two experimental inhibitors namely 49c and CWHM-117 in complex with PIX and PX, by characterizing the active sites, analysing the structural and functional composition, and highlighting the essential amino acid residues required for binding. Defining the binding landscape of each enzyme will offer a prospective avenue for the design of selective and distinct inhibitors with critical pharmaceutical characteristics that will aid in developing targeted and effective small molecule inhibitors

In this study we also employed a tailored pharmacophore approach specific to PIX for the design of inhibitors specific to this enzyme which is the first of its kind and which may assist scientists from different research domains in designing potential small molecule inhibitors against PIX and the Plm family of enzymes at large.

To this end, the work presented in this thesis remains fundamental for the advancement of research toward targeted drug design/delivery against Malaria

1.5 References

(1) Nkumama, I. N., Meara, W. P. O., and Osier, F. H. A. (2017) Changes in Malaria

- Epidemiology in Africa and New Challenges for Elimination. *Trends Parasitol.* 33, 128–140.
- (2) Biamonte, M. A., Wanner, J., and Le, K. G. (2013) Bioorganic & Medicinal Chemistry Letters Recent advances in malaria drug discovery. *Bioorg. Med. Chem. Lett.* 23, 2829–2843.
- (3) World Health Organization. (2017) World malaria report 2017. Geneva.
- (4) Jamison, D. T., Feachem, R. G., Makgoba, M. W., Bos, E. R., Baingana, F. K., Hofman, K. J., and Rogo, K. O. (2006) Disease and Mortality in Sub-Saharan Africa (Jamison, D. T., Feachem, R. G., Makgoba, M. W., Bos, E. R., Baingana, F. K., Hofman, K. J., and Rogo, K. O., Eds.) Second Edi. The World Bank, Washington,DC.
- (5) Ruiz, A. O., Postigo, M., Casanova, S. G., Cuadrado, D., Bautista, J. M., Rubio, J. M., Oroz, M. L., and Linares, M. (2018) *Plasmodium* species differentiation by non - expert on - line volunteers for remote malaria field diagnosis. *Malar. J.* 17, 1–10.
- (6) World Health Organization. (2016) Global technical strategy for malaria 2016–2030.
- (7) Autino, B., Noris, A., Russo, R., and Castelli, F. (2012) Epidemiology of Malaria in Endemic Areas. *Mediterr. J. Hematol. Infect. Dis.* 4, 1–9.
- (8) World Health Organization. (2018) World malaria report 2018. Geneva.
- (9) Ismail, H. M., Barton, V., Phanchana, M., Charoensutthivarakul, S., and Wong, M. H. L. (2016) Artemisinin activity-based probes identify multiple molecular targets within the asexual stage of the malaria parasites *Plasmodium falciparum* 3D7. *PNAS* 113, 2080–2085.
- (10) Mombo-ngoma, G., Remppis, J., Sievers, M., Manego, Z., Endamne, L., Kabwende, L., Veletzky, L., Nguyen, T. T., Groger, M., Lötsch, F., Mischlinger, J., Flohr, L., Kim, J., Cattaneo, C., Hutchinson, D., Duparc, S., Moehrle, J., Velavan, T. P., Lell, B., Ramharther, M., Adegnika, A. A., Mordmüller, B., and Kremsner, P. G. (2018) Efficacy and Safety of Fosmidomycin – Piperaquine as Nonartemisinin-Based Combination Therapy for Uncomplicated Falciparum Malaria : A Single-Arm , Age De-escalation Proof-of-Concept Study in Gabon. *J. Infect. Dis. Am.* 1–8.
- *(11) Haldar, K., Bhattacharjee, S., and Safeukui, I. (2018) Drug resistance in *Plasmodium*. *Nat. Publ. Gr.* 12, 1–15.
- (12) Cowman, A. F., Berry, D., and Baum, J. (2012) The cellular and molecular basis for malaria parasite invasion of the human red blood cell. *J. Cell Biol.* 198, 961–971.

- (13) Mohammed, S., Michael , Blackman, Patricia, D., He, C., Land, K. M., Lobo, C., and Mackey, Z. (2008) Drug Discovery and Development, in *Antiparasitic and Antibacterial Drug Discovery* (Selzer, P. M., Ed.), pp 177–250. Wiley-Vch Verlag GmbH & Co. KGaA, Weinheim.
- (14) Wirth, D. F., and Alonso, P. L. Biology in the Era of Eradication.
- (15) Florens, L., Washburn, M. P., Raine, J. D., Anthony, R. M., Grainger, M., Haynes, J. D., Moch, J. K., Muster, N., Sacci, J. B., Tabb, D. L., Witney, A. A., Wolters, D., Wu, Y., Gardner, M. J., Holder, A. A., Sinden, R. E., Yates, J. R., and Carucci, D. J. (2002) A proteomic view of the *Plasmodium falciparum* life cycle. *Nat. Publ. Gr.* 419, 520–526.
- (16) Moura, P. A., Dame, J. B., and Fidock, D. A. (2009) Role of *Plasmodium falciparum* Digestive Vacuole Plasmepsins in the Specificity and Antimalarial Mode of Action of Cysteine and Aspartic Protease Inhibitors. *Antimicrob. Agents Chemother.* 53, 4968–4978.
- (17) Coombs, G. H., Goldberg, D. E., Klemba, M., Berry, C., Kay, J., Mottram, J. C., and Mottram, J. C. (2001) Aspartic proteases of *Plasmodium falciparum* and other parasitic protozoa as drug targets. *Trends Parasitol.* 17, 532–537.
- (18) Dash, C., Kulkarni, A., Dunn, B., and Rao, M. (2003) Aspartic Peptidase Inhibitors : Implications in Drug Development Aspartic Peptidase Inhibitors : Implications in Drug Development. *Crit. Rev. Biochem. Mol. Biol.* 38, 89–119.
- (19) Mogire, R. M., Akala, H. M., Macharia, R. W., Juma, D. W., Cheruiyot, C., Andagalu, B., Brown, M. L., El-shemy, H. A., and Nyanjom, G. (2017) Target-similarity search using *Plasmodium falciparum* proteome identifies approved drugs with antimalarial activity and their possible targets. *PLoS One* 12, 1–24.
- (20) Ersmark, K., Samuelsson, B., and Hallberg, A. (2006) Plasmepsins as Potential Targets for NewAntimalarial Therapy. *Med. Res. Rev.* 26, 626–666.
- (21) Nguyen, J., Hamada, Y., Kimura, T., and Kiso, Y. (2008) Review Design of Potent Aspartic Protease Inhibitors to Treat Various Diseases. *Sci. J. Chem. Sci.* 314, 523–535.
- (22) Boddey, J. A., Carvalho, T. G., Hodder, A. N., Sargeant, T. J., Sleebs, B. E., Marapana, D., Lopaticki, S., Nebl, T., and Cowman, A. F. (2013) Role of plasmepsin V in Export of Diverse Protein Families from the *Plasmodium falciparum* Exportome. *Traffic* 14, 532–550.
- (23) McGillewie, L., Ramesh, M., and Soliman, M. E. (2017) Sequence , Structural Analysis

and Metrics to Define the Unique Dynamic Features of the Flap Regions Among Aspartic Proteases. *Protein J.* 12, 5–9.

(24) Nair, D. N., Singh, V., Angira, D., and Thiruvengadam, V. (2016) Proteomics & Bioinformatics Structural Investigation and In-silico Characterization of Plasmepsins from *Plasmodium falciparum*. *J. Proteomics Bioinforma.* 9, 181–195.

(25) Bedi, R. K., Patel, C., Mishra, V., Xiao, H., and Yada, R. Y. (2016) Understanding the structural basis of substrate recognition by *Plasmodium falciparum* plasmepsin V to aid in the design of potent inhibitors. *Sci. Rep.* 6, 1–17.

(26) Nasamu, A. S., Glushakova, S., Russo, I., Vaupel, B., Oksman, A., Kim, A. S., Fremont, D. H., Tolia, N., Beck, J. R., Meyers, M. J., Niles, J. C., Zimmerberg, J., and Goldberg, D. E. (2017) parasite egress and invasion. *Science* (80-.). 522, 518–522.

(27) Cai, H., Kuang, R., Gu, J., Wang, Y., Texas, S., Infectious, E., Antonio, S., and Antonio, S. (2011) Proteases in Malaria Parasites - A Phylogenomic Perspective. *Curr. Genomics* 12, 417–427.

(28) Berry, C., and Goldberg, D. E. (2001) Food Vacuole Plasmepsins, in *Handbook of Proteolytic Enzymes 3rd Edition*, pp 98–103. Cardiff.

(29) Rosenthal, P. J. (1998) Proteases of Malaria Parasites : New Targets for Chemotherapy. *Emerg. Infect. Dis.* 4, 49–57.

(30) Roy, K. K. (2018) International Journal of Antimicrobial Agents Targeting the active sites of malarial proteases for antimalarial drug discovery : approaches , progress and challenges. *Int. J. Antimicrob. Agents* 50, 287–302.

(31) Crunkhorn, S. (2017) Infectious disease: Blocking malaria parasite invasion and egress. *Nat. Publ. Gr.* 4, 2017.

(32) Adcock, S. A., and Mccammon, J. A. (2006) Molecular Dynamics : Survey of Methods for Simulating the Activity of Proteins. *Chem. Rev.* 106, 1589–1615.

(33) Kollman, P. (1993) Free Energy Calculations: Applications to Chemical and Biochemical Phenomena. *Chem. Rev.* 93, 2395–2417.

(34) Joseph-mccarthy, D., and Alvarez, J. C. (2003) Automated Generation of MCSS-Derived Pharmacophoric DOCK Site Points for Searching Multiconformation Databases. *PROTEINS*

Struct. Funct. Genet. 202, 189–202.

(35) Verdonk, M. L., Cole, J. C., Hartshorn, M. J., Murray, C. W., and Taylor, R. D. (2003) Improved Protein – Ligand Docking Using GOLD. *PROTEINS Struct. Funct. Genet.* 623, 609–623.

(36) Song, C. M., Lim, S. J., and Chuantong, J. (2009) Recent advances in computer-aided drug design. *Brief. Bioinforma.* 10, 579–591.

(37) Eweas, A. F., Maghrabi, I. A., and Namarneh, A. I. (2014) Advances in molecular modeling and docking as a tool for modern drug discovery. *Der Pharmaceutical Chem.* 6, 211–228.

(38) Anslyn, E. V, and Dougherty, D. A. (2006) Modern Physical Organic Chemistry (Murdzek, J., Ed.). University Science Books, California.

(39) Vanommeslaeghe, K., Guvench, O., and Mackarellel, D. A. (2015) Molecular Mechanics. *Curr. Protoc. Comput. drug Des.* 20, 3281–3292.

CHAPTER 2

2.1 Introduction

The Malaria parasite remains an eminent global health threat, predominantly in countries worldwide where the transmission of the parasite appears regularly in addition to where transmission is considered either restricted or eliminated.^{1,2,3} Globally there was a projected 219 million diagnosed cases of malaria in the year 2017 as presented in Figure 2.1, compared to the 239 million cases in the year 2010 and 217 million cases in the year 2016.^{4,5} Even though there was a substantial decrease of 20 million malaria cases, statistical data extracted between the timeframe 2015–2017 emphasize no significant improvement aimed at the global reduction in the number of diagnosed cases of malaria during this period.^{6,7,8} Although the impact of malaria is experienced on a global scale, 70% of estimated cases and deaths resulting from malaria emerged from 11 countries of which 10 were in sub-Saharan Africa.^{9,10} In these malaria burden regions, the most susceptible victims of this deadly disease are young children, expectant mothers, and non-immune visitors to malarious areas. Studies have shown that children under 5 years are the most exposed and defenseless to the malaria parasite as they accounted for over 60% towards the global mortality rate in 2017.⁵

Malaria is transmitted by the infected *Anopheles* mosquitoes and is triggered by parasitic protozoans of the *Plasmodium* genus.^{11,12} The complex life cycle of the *Plasmodium* parasite encompasses a asexual stage present within the human host and a sexual stage in the mosquito which acts a vector for the parasite.^{13,14} Among the diverse *Plasmodium* species responsible for the transmission of malaria, *P. falciparum* remains the most lethal strain responsible for the largest number of deaths, followed closely by the less deadly *P. vivax*.^{15,16}

Although there have been extensive efforts directed towards malaria control strategies, none remain formidable nor affordable in all frameworks. One of the major challenges affecting the elimination of malaria is the development of antimalarial drug resistance.^{17,18,19} Over the years drug resistance in malaria has played an instrumental role in the prevalence and relentlessness of epidemics witnessed in many parts of the world by aiding in the onset of malaria to new regions and re-emergence of malaria in regions where it once dormant or eradicated.^{13,20} This has brought us to a time of disparity as the malaria parasite has conveyed some level of resistance to almost all existing antimalarial therapeutics resulting in an expansive increase in the expense and complexity of achieving a parasitological cure. To address the incessant threat incurred by the development of drug resistance, the current objective of malaria eradication

tactics aims to develop malarial therapeutics that target parasite transmission to prevent the spread of pathogenic blood stages in humans.^{21,22} To meet this aim the identification of potential drug candidates is not only paramount but the availability of accurate drug protein targets remains a high priority for the design of novel antimalarial therapeutics. The availability of drug targets, may ensure the development of current therapeutic with enhanced drug target specificity and effectivity, potentially reducing developed drug resistance, lack of effectivity and potency.^{23,24}

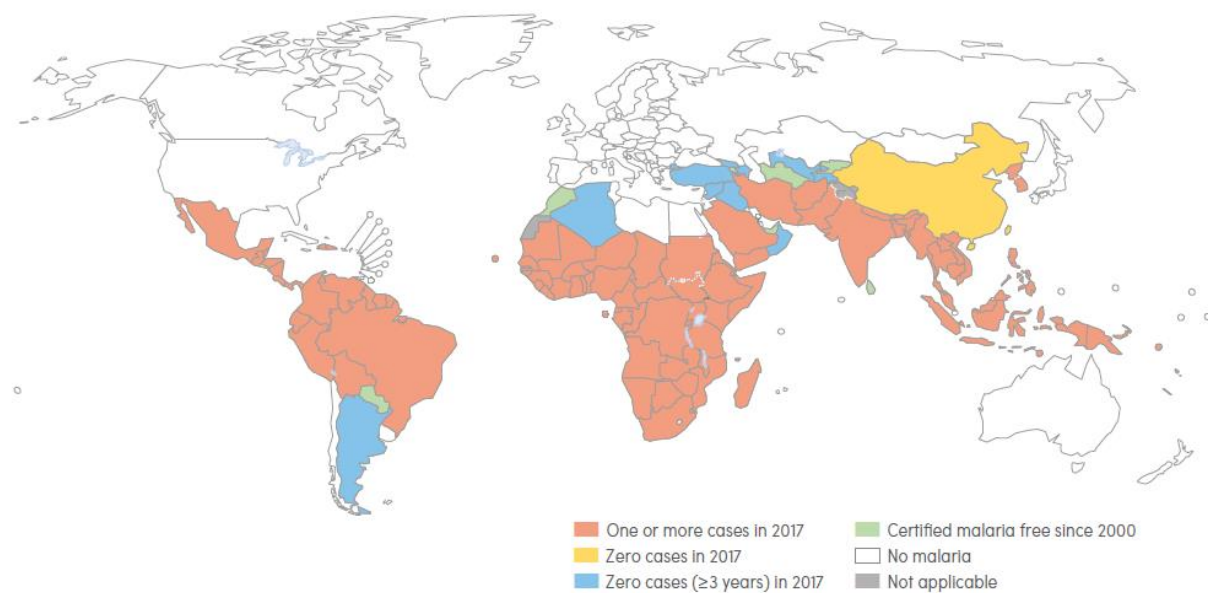


Figure 2.1 Global incidence of malaria in the year 2017.⁵

2.2 Life Cycle of the Malaria parasite

The intrinsic history of the malaria parasite involves recurrent infection of humans by the female *Anopheles* mosquitoes.^{25,26} When a human is infected by the malaria parasite initial growth begins when the parasite multiplies within the cells of the liver and spreads into the red blood cells within the human host. In the blood of an infected host, successive broods of parasites grow within the red cells and undergo egression, discharging daughter parasites “merozoites” that resume the cycle by undergoing invasion and entering surrounding red cells as presented in Figure 2.2.^{27,28}

The primary symptoms of malaria are caused by blood stage parasites. When specified forms of the blood stage parasite (gametocytes, present in female and male forms) are consumed by the female *Anopheles* mosquito during blood feeding, they undergo sexual reproduction in the gut of the mosquito and in doing so initiate a cycle of growth and reproduction within the mosquito.^{29,30} After a period of 10-18 days, the sporozoite form sporozoite relocates to the salivary glands of the mosquito vector.³¹ The mosquito vector partakes of a blood meal by feeding on a human, decoagulant saliva combined with the sporozoites and then injected into the human host, which migrates to the liver, commencing a new cycle.^{32,33} The infected mosquito acts as a vector mediating the disease from one human host to another easily causing the spread of the malaria virus.³⁴ Although the infected humans can transmit the parasite to the mosquito, the mosquito vector remains unaffected by the parasite. Extensive studies highlight the role of numerous parasite proteins that remain essential for malaria progression, which include a cascade of proteolytic activities. The egress and invasion of the malarial parasite which implicates the dynamics of merozoite release and encompasses two of the most crucial steps involved in parasite intraerythrocytic inexorable tactics, still remain to be deciphered. These two events in particular are often referred to as “explosive” events.^{35,34}

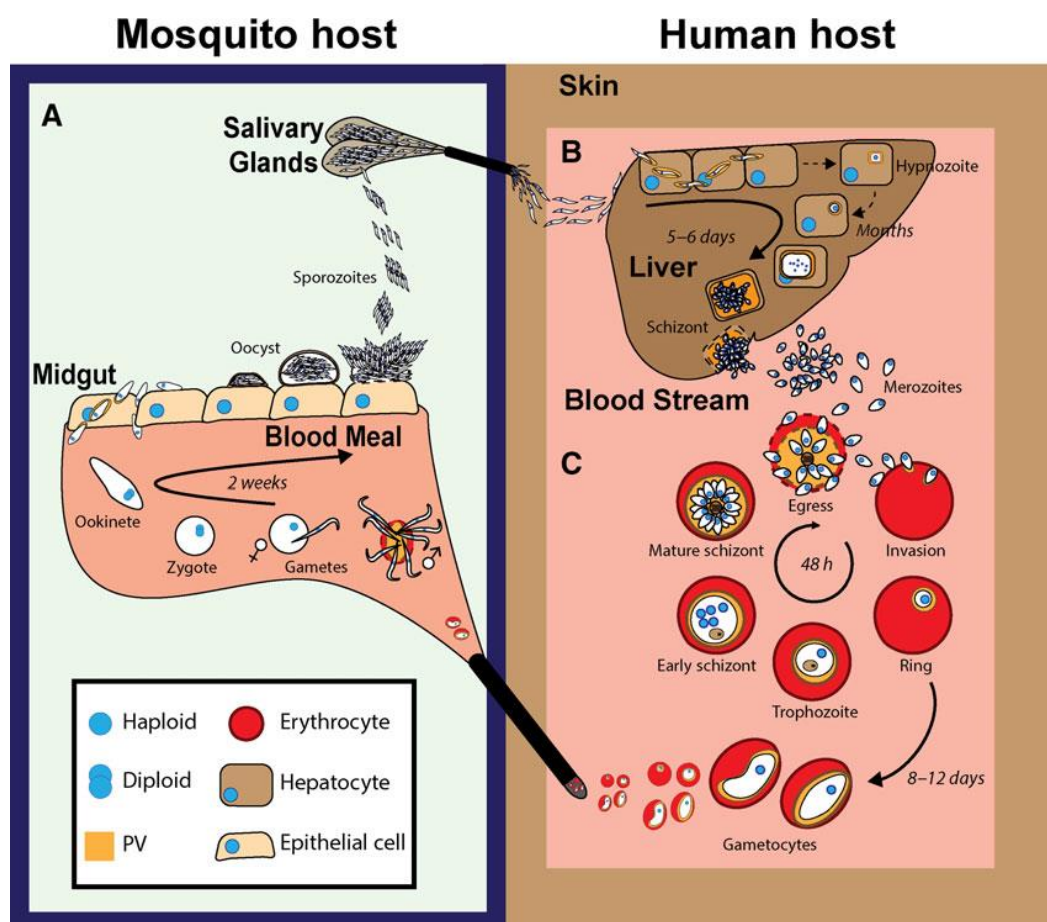


Figure 2.2 Life cycle of the malaria parasite taken from proteases as antimalarial (prepared by author).

2.3 Aspartic Proteases of *Plasmodium falciparum*

Aspartic proteases compose of a ubiquitous and unique group of catalytic and regulatory molecules that are ensnared in diverse biological roles within a large domain of living systems.^{36,37} Aspartic proteases play a connective role in number of life threatening diseases such as Alzheimer's disease (β -secretase), fungal infections (secreted aspartic proteases), and hypertension (renin) malaria (plasmepsins)³⁸. A combination of factors make these enzymes formidable targets for the development of novel anti-parasite therapies one of these factors include the existence of multiple proteases in all parasitic protozoa.³⁹

The family of aspartic proteases present within the *Plasmodium falciparum* parasite are commonly referred to as Plasmepsins and are believed to be excellent drug targets.⁴⁰ Extensive studies performed on protease inhibitors display three essential processes required for protease activity in the erythrocytic phase of intrusion by the parasite which include haemoglobin degradation, erythrocyte invasion and rupture.^{41,42} Therefore, dysregulation of these fundamental processes may alter the life expectation of the virulent parasite. The molecular mechanism adopted by the aspartic protease family of enzymes as can be observed in Figure 2.3 and entails the use a water molecule which acts as an active nucleophile that mediates an attack on the scissile peptide bond of a substrate.^{43,44} The reacting water molecule remains in a fixed pose and is stimulated by two aspartic acid residues. The most ostensible difference of aspartic proteases to cysteine and serine proteases counterparts, is the integration of the most applicable amino acid side-chains to render an attack on the nucleophile.^{45,46} The peptide bond resulting from the substrate assumes a fixed a position in close proximation to the juxtaposition of the catalytic water molecule embedded within the active site domain of the aspartic proteases which is stimulated by the interactions that remnant between the side-chain atoms and the corresponding backbone atoms of the substrate and enzyme.^{47,48} Therefore, inhibitors that possess hydroxyl moieties are integral as they target the active site domain of dual aspartic proteases rendering structural disposition to the crucial catalytic water molecule.⁴⁹

Crucial insight extracted from numerous studies performed on aspartic proteases have supplemented valuable insight pertaining to substrate-inhibitor specificities, functionalities and the crystal structures of these integral enzymes. The information extricated from conformational and structural analysis on these aspartic proteases has provided data regarding the design of inhibitors with high specificity and effectivity. Although a large stream of information exists on the molecular mechanism endorsed by the aspartic proteases which are

relatively similar however their sequential composition differs drastically, therefore highlighting the importance of designing specific inhibitors to the individual plasmepsins from *P. falciparum*.

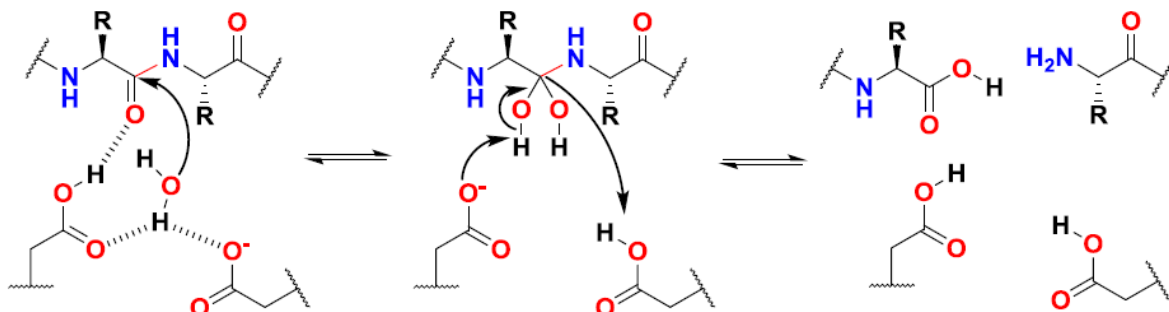


Figure 2.3 Catalytic mechanism of aspartic proteases taken from targeting the active site of aspartic proteases.⁴³

2.4 Plasmepsins of *Plasmodium falciparum*

The aspartic protease family of enzymes comprises of ten Plasmepsins encoded in the *P. falciparum* genome and are key activators of malaria pathogenesis, involved in an array of cellular functions. The Plm family of enzymes are divided into two distinct classes (1) Plm I-IV are all regarded as intronless with a single exon and are situated on adjacent positions on the fourteenth chromosome, often generated by tandem gene identical replication.^{50,51} (2) Plm's V-X form a sizeable clade which constitutes an ancestral type of Plasmepsins. Plm I, II, III, IV and the closely related histo-aspartic protease (HAP) (vacuolar Plm's) are expressed in the acidic food vacuole of parasite-infested red blood cells (RBC), where they are responsible for the breakdown of hemoglobin.^{52,30,53} PV is responsible for protein export whereas PIX and PX whose function was unknown has recently been found to play an instrumental role in the egress and invasion of the malarial parasite.^{22,23,54} Each Plm exhibits distinctive substrate specificity despite sharing of sequential similarity and preservation of the active site domain. Unique substrate specificity may be directly attributed to the sequence heterogeneity of amino acid residues outlining the active site domain, integrating intrinsic components such as the flap and flexible loop. Antimalarial discovery and development have extensively targeted aspartic proteases which play an essential role in the haemoglobin digestive pathway paving the way for new therapeutic targets that may render a final knockout to the malaria parasite. Therefore,

the two newly discovered enzymes namely PIX and PX whose role in the progression of malaria was unknown recently. Experimental studies have shown the instrumental role each of these enzymes play in the egress and the invasion of the malaria parasite and thus form reputable target for the effective treatment of malaria.^{55,35,27}

2.5 New targets for Malaria chemotherapy: Erythrocyte Rupture and Invasion

There is an urgent need for the discovery of new targets which currently limits potential antimalarial chemotherapy and the design of effective and innovative drugs.^{56,57} Two exemplary potential antimalarial targets for chemotherapy include mediators of egress and invasion of the malarial parasite. As the free merozoites seize entry into the erythrocytes, the parasite expounds making the erythrocytic life cycle of the malaria parasite directly responsible for the clinical symptoms of the parasite as depicted in Figure 2.4. PIX and PX, which are expressed at distinct stages in the parasite cycle, have recently emerged as propitious antimalarial targets as they fulfil central, hitherto unknown, functions. Recent studies have exhibited the crucial role adopted by PIX and PX which is required by the malaria parasite to initiate the rupture of infected erythrocytes by mature schizonts and the subsequent invasion of erythrocytes by free merozoites possibly to breach the erythrocyte cytoskeleton.^{35,11,58} Reduced expression of the latter family members PIX and PX, in the early ring-stage parasites using a new provisional knockdown technology resulted in reduced replication, further validating the critical role these enzymes play in the survival of the malaria. Further analysis of the progression of the cell cycle reveals the direct implication of PIX in RBC invasion and PX's dual role in egress and invasion of the parasite. The cellular localization of PIX facilitates the biogenesis of rhoptry secretory organelles that are directly involved in invasion which was discovered using subcellular localization, whereas PX localizes to exonemes and regulates the development of the imperative subtilisin-like serine protease (SUB1), which plays an instrumental role in the process of egression.⁵⁹

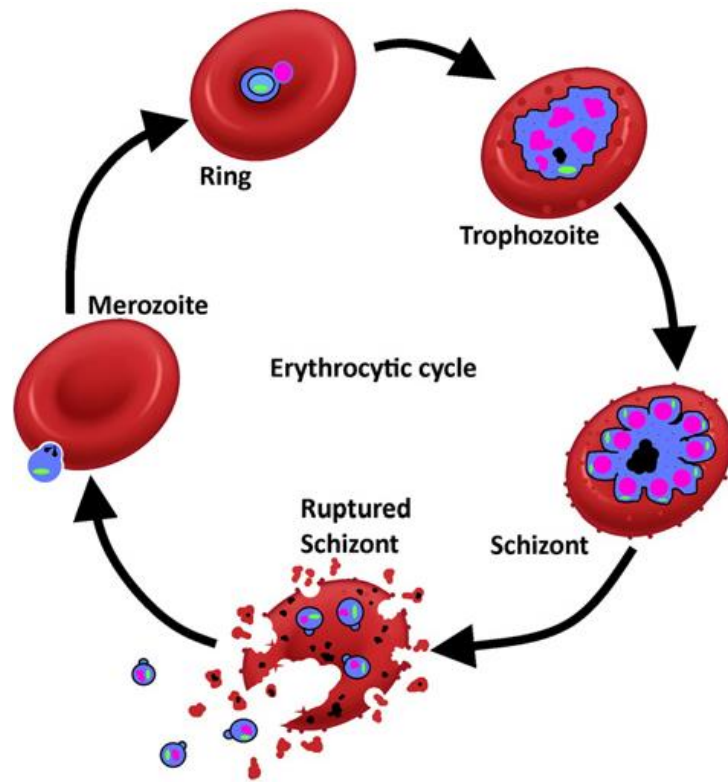


Figure 2.4 Erythrocytic cycle of the red blood cell mediated by the malaria parasite.⁶⁰

2.6 References

- (1) Maharaj, R., Raman, J., Morris, N., and Moonasar, D. (2019) Epidemiology of malaria in South Africa : From control to elimination. *South African Med. J.* 103, 1–11.
- (2) Torres, V. (2013) New Targets for Drug Discovery against Malaria. *PLoS One* 8, 1–14.
- (3) Burrows, J. N., Duparc, S., Gutteridge, W. E., Huijsduijnen, R. H. Van, Kaszubska, W., Macintyre, F., Mazzuri, S., Möhrle, J. J., and Wells, T. N. C. (2017) New developments in anti - malarial target candidate and product profiles. *Malar. J.* 16, 1–29.
- (4) World Health Organization. (2017) World malaria report 2017. Geneva.
- (5) World Health Organization. (2018) World malaria report 2018. Geneva.
- (6) Nkumama, I. N., Meara, W. P. O., and Osier, F. H. A. (2017) Changes in Malaria Epidemiology in Africa and New Challenges for Elimination. *Trends Parasitol.* 33, 128–140.
- (7) World Health Organization. (2016) Global technical strategy for malaria 2016–2030.
- (8) Bloland, P. B., and Williams, H. A. (2002) Malaria control during mass population movements and natural disasters. Washington,DC.
- (9) Autino, B., Noris, A., Russo, R., and Castelli, F. (2012) Epidemiology of Malaria in Endemic Areas. *Mediterr. J. Hematol. Infect. Dis.* 4, 1–9.
- (10) Mohammed, S., Michael , Blackman, Patricia, D., He, C., Land, K. M., Lobo, C., and Mackey, Z. (2008) Drug Discovery and Development, in *Antiparasitic and Antibacterial Drug Discovery* (Selzer, P. M., Ed.), pp 177–250. Wiley-Vch Verlag GmbH & Co. KGaA, Weinheim.
- (11) Nigussie, D., Beyene, T., Shah, N. A., and Belew, S. (2015) Malaria Control & Elimination New Targets in Malaria Parasite Chemotherapy: A Review. *Malar. Control Elimin. S1*, 1–6.
- (12) Ruiz, A. O., Postigo, M., Casanova, S. G., Cuadrado, D., Bautista, J. M., Rubio, J. M., Oroz, M. L., and Linares, M. (2018) *Plasmodium* species differentiation by non - expert on - line volunteers for remote malaria field diagnosis. *Malar. J.* 17, 1–10.
- (13) Halder, K., Bhattacharjee, S., and Safeukui, I. (2018) Drug resistance in *Plasmodium*. *Nat. Publ. Gr.* 12, 1–15.
- (14) Florens, L., Washburn, M. P., Raine, J. D., Anthony, R. M., Grainger, M., Haynes, J. D.,

Moch, J. K., Muster, N., Sacchi, J. B., Tabb, D. L., Witney, A. A., Wolters, D., Wu, Y., Gardner, M. J., Holder, A. A., Sinden, R. E., Yates, J. R., and Carucci, D. J. (2002) A proteomic view of the *Plasmodium falciparum* life cycle. *Nat. Publ. Gr.* 419, 520–526.

(15) McGeorge, R. D. (2014) *Plasmodium falciparum* Plasmepsins IX and X: Structure-function analysis and the discovery of new lead antimalarial drugs. Griffith University.

(16) Gnémé, A., Guelbéogo, W. M., Riehle, M. M., Tiono, A. B., Diarra, A., Kabré, G. B., Sagnon, N., and Vernick, K. D. (2013) *Plasmodium* species occurrence, temporal distribution and interaction in a child-aged population in rural Burkina Faso. *Malarial J.* 12, 1–9.

(17) Pussard, E., and Verdier, F. (1994) Antimalarial 4-aminoquinolines: mode of action and pharmacokinetics. *Fundam. Clin. Pharmacol.* 8, 1–17.

(18) Biamonte, M. A., Wanner, J., and Le, K. G. (2013) Bioorganic & Medicinal Chemistry Letters Recent advances in malaria drug discovery. *Bioorg. Med. Chem. Lett.* 23, 2829–2843.

(19) Gregson, A., and Plowe, C. V. (2005) Mechanisms of Resistance of Malaria Parasites to Antifolates. *Pharmacol. Rev.* 57, 117–145.

(20) Duraisingh, M. T., and Cowman, A. F. (2005) Contribution of the *pfmdr1* gene to antimalarial drug-resistance 94, 181–190.

(21) Aneja, B., Kumar, B., Jairajpuri, A. M., and Abid, M. (2013) Structure guided drug-discovery approach towards identification of *Plasmodium* inhibitors. *R. Soc. Chem.* 00, 1–50.

(22) Liu, J., Gluzman, I. Y., Drew, M. E., and Goldberg, D. E. (2005) The Role of *Plasmodium falciparum* Food Vacuole Plasmepsins. *J. Biol. Chem.* 280, 1432–1437.

(23) Li, F., Bounkeua, V., Pettersen, K., and Vinetz, J. M. (2016) *Plasmodium falciparum* ookinete expression of plasmepsin VII and plasmepsin X. *Malar. J.* 15, 1–10.

(24) Ashley, E. A., and Ashley, E. A. (2018) Drugs in Development for Malaria. *Drugs* 78, 861–879.

(25) Rosenthal, P. J. (1998) Proteases of Malaria Parasites: New Targets for Chemotherapy. *Emerg. Infect. Dis.* 4, 49–57.

(26) Gamo, F., Sanz, L. M., Vidal, J., Cozar, C. De, Alvarez, E., Lavandera, J., Vanderwall, D. E., Green, D. V. S., Kumar, V., Hasan, S., Brown, J. R., Peishoff, C. E., Cardon, L. R., and Garcia-bustos, J. F. (2010) Thousands of chemical starting points for antimalarial lead

identification. *Nature* 465, 305–310.

(27) Crunkhorn, S. (2017) Infectious disease: Blocking malaria parasite invasion and egress. *Nat. Publ. Gr.* 4, 2017.

(28) Suemizu, H., Snounou, G., Morosan, S., and Naour, G. Le. (2015) ovale liver stages in humanized mice. *Nat. Commun.* 6, 1–9.

(29) Cowman, A. F., Berry, D., and Baum, J. (2012) The cellular and molecular basis for malaria parasite invasion of the human red blood cell. *J. Cell Biol.* 198, 961–971.

(30) Bedi, R. K., Patel, C., Mishra, V., Xiao, H., and Yada, R. Y. (2016) Understanding the structural basis of substrate recognition by *Plasmodium falciparum* plasmepsin V to aid in the design of potent inhibitors. *Sci. Rep.* 6, 1–17.

(31) Marti, M., Marti, M., Good, R. T., Rug, M., Knuepfer, E., and Cowman, A. F. (2014) Targeting Malaria Virulence and Remodeling Proteins to the Host Erythrocyte. *Science* (80-.). 1930.

(32) Westling, J., Yowell, C. A., Majer, P., Erickson, J. W., Dame, J. B., and Dunn, B. M. (1997) *Plasmodium falciparum* , *P . vivax* , and *P . malariae* : A Comparison of the Active Site Properties of Plasmepsins Cloned and Expressed from Three Different Species of the Malaria Parasite 193, 185–193.

(33) Dhangadamajhi, G., Kar, S. K., and Ranjit, M. (2010) The Survival Strategies of Malaria Parasite in the Red Blood Cell and Host Cell Polymorphisms 2010.

(34) Lew, V. L. (2011) Malaria : Surprising Mechanism of Merozoite Egress Revealed. *Curr. Biol.* 21, R314–R316.

(35) Pino, P., Caldelari, R., Mukherjee, B., Vahokoski, J., Klages, N., Maco, B., Collins, C. R., Blackman, M. J., Kursula, I., Heussler, V., Brochet, M., and Soldati-favre, D. (2017) A multistage antimalarial targets the plasmepsins IX and X essential for invasion and egress. *Science* (80-.). 528, 522–528.

(36) Asojo, O. A., Gulnik, S. V, Afonina, E., Yu, B., Ellman, J. A., Haque, T. S., Silva, A. M., and Hall, B.-H. (2003) Novel Uncomplexed and Complexed Structures of plasmepsin II , an Aspartic Protease from *Plasmodium falciparum*. *J. Mol. Biol.* 2836, 173–181.

(37) Moura, P. A., Dame, J. B., and Fidock, D. A. (2009) Role of *Plasmodium falciparum*

Digestive Vacuole Plasmepsins in the Specificity and Antimalarial Mode of Action of Cysteine and Aspartic Protease Inhibitors. *Antimicrob. Agents Chemother.* 53, 4968–4978.

(38) Mogire, R. M., Akala, H. M., Macharia, R. W., Juma, D. W., Cheruiyot, C., Andagalu, B., Brown, M. L., El-shemy, H. A., and Nyanjom, G. (2017) Target-similarity search using *Plasmodium falciparum* proteome identifies approved drugs with antimalarial activity and their possible targets. *PLoS One* 12, 1–24.

(39) McGillewie, L., Ramesh, M., and Soliman, M. E. (2017) Sequence , Structural Analysis and Metrics to Define the Unique Dynamic Features of the Flap Regions Among Aspartic Proteases. *Protein J.* 12, 5–9.

(40) Coombs, G. H., Goldberg, D. E., Klemba, M., Berry, C., Kay, J., Mottram, J. C., and Mottram, J. C. (2001) Aspartic proteases of *Plasmodium falciparum* and other parasitic protozoa as drug targets. *Trends Parasitol.* 17, 532–537.

(41) Mahanti, M., and Bhakat, S. (2016) Flap Dynamics in Aspartic Proteases : A Computational Perspective. *Chem. Biol. Drug Des.* 1–19.

(42) Dash, C., Kulkarni, A., Dunn, B., and Rao, M. (2003) Aspartic Peptidase Inhibitors : Implications in Drug Development Aspartic Peptidase Inhibitors : Implications in Drug Development. *Crit. Rev. Biochem. Mol. Biol.* 38, 89–119.

(43) Nguyen, J., Hamada, Y., Kimura, T., and Kiso, Y. (2008) Review Design of Potent Aspartic Protease Inhibitors to Treat Various Diseases. *Sci. J. Chem. Sci.* 314, 523–535.

(44) Sharma, A., Eapen, A., and Subbarao, S. K. (2005) Purification and Characterization of a Hemoglobin Degrading Aspartic Protease from the Malarial Parasite *Plasmodium vivax* 78, 71–78.

(45) Ruiter, W. J. The Catalytic Role of the Active Site Aspartic Acid in Serine Proteases imidazole group to transfer the proton from.

(46) Na, B., Lee, E., Lee, H., Cho, S., and Bae, Y. (2004) Aspartic proteases of *Plasmodium vivax* are highly conserved in wild isolates 42, 61–66.

(47) Parr, C. L., Tanaka, T., Xiao, H., and Yada, R. Y. (2008) The catalytic significance of the proposed active site residues in *Plasmodium falciparum* histoaspartic protease. *FEBS J.* 275, 1698–1707.

- (48) Meyers, M. J., Tortorella, M. D., Xu, J., Qin, L., He, Z., Lang, X., Zeng, W., Xu, W., Qin, L., Prinsen, M. J., Sverdrup, F. M., Eickho, C. S., Griggs, D. W., Oliva, J., Ruminski, P. G., Jacobsen, E. J., Campbell, M. A., Wood, D. C., Goldberg, D. E., Liu, X., Lu, Y., Lu, X., Tu, Z., Lu, X., Ding, K., and Chen, X. (2014) Evaluation of Aminohydantoins as a Novel Class of Antimalarial Agents. *ACS Med. Chem. Lett.* 5, 6–10.
- (49) Cai, H., Kuang, R., Gu, J., Wang, Y., Texas, S., Infectious, E., Antonio, S., and Antonio, S. (2011) Proteases in Malaria Parasites - A Phylogenomic Perspective. *Curr. Genomics* 12, 417–427.
- (50) Recacha, R., Jaudzems, K., Akopjana, I., Jirgensons, A., and Tars, K. (2016) Crystal structure of *Plasmodium falciparum* proplasmepsin IV : the plasticity of proplasmepsins research communications. *Res. Commun.* 659–666.
- (51) Nair, D. N., Singh, V., Angira, D., and Thiruvenkatam, V. (2016) Proteomics & Bioinformatics Structural Investigation and In-silico Characterization of Plasmepsins from *Plasmodium falciparum*. *J. Proteomics Bioinforma.* 9, 181–195.
- (52) Banerjee, R., Liu, J., Beatty, W., Pelosof, L., Klemba, M., and Goldberg, D. E. (2001) Four plasmepsins are active in the *Plasmodium falciparum* food vacuole , including a protease with an active-site histidine *Curr. Genomics* 13, 324-333.
- (53) Sonoiki, E., Nsanzabana, C., Legac, J., Sindhe, K. M. V, Derisi, J., and Rosenthal, P. J. (2017) crossm Sensitivity to the Antiretroviral Protease Inhibitor Lopinavir Associated with. *Antimicrob. Agents Chemother.* 61, 1–4.
- (54) Deu, E. (2017) Proteases as antimalarial targets : strategies for genetic , chemical , and therapeutic validation. *FEBS Journal* 284, 2604–2628.
- (55) Nasamu, A. S., Glushakova, S., Russo, I., Vaupel, B., Oksman, A., Kim, A. S., Fremont, D. H., Tolia, N., Beck, J. R., Meyers, M. J., Niles, J. C., Zimmerberg, J., and Goldberg, D. E. (2017) parasite egress and invasion. *Science* (80). 522, 518–522.
- (56) Pandey, K. C., Singh, N., Arastu-kapur, S., Bogyo, M., and Rosenthal, P. J. (2006) Falstatin , a Cysteine Protease Inhibitor of *Plasmodium falciparum* , Facilitates Erythrocyte Invasion. *PLoS Pathog.* 2.
- (57) Walker, D. M., Oghumu, S., Gupta, G., McGwire, B. S., Drew, M. E., and Satoskar, A. R. (2015) Mechanisms of cellular invasion by intracellular parasites. *Cell. Mol. Life Sci.* 71, 1245–

1263.

(58) Boddey, J. A., Carvalho, T. G., Hodder, A. N., Sargeant, T. J., Sleebs, B. E., Marapana, D., Lopaticki, S., Nebl, T., and Cowman, A. F. (2013) Role of plasmepsin V in Export of Diverse Protein Families from the *Plasmodium falciparum* Exportome. *Traffic* 14, 532–550.

(59) Alam, A., and Chauhan, V. S. (2012) Inhibitory Potential of Prodomain of *Plasmodium falciparum* Protease Serine Repeat Antigen 5 for Asexual Blood Stages of Parasite. *PLoS One* 7, 1–8.

(60) Chakraborty, A. (2016) Understanding the biology of the *Plasmodium falciparum* apicoplast ; an excellent target for antimalarial drug development. *Life Sci.* 158, 104–110.

CHAPTER 3

3.0. Molecular Modeling and Computational Approaches to Biomolecular Structure and Drug Design

3.1 An Introduction into Computational Chemistry

The application of Computational chemistry in molecular drug design protocols is rapidly emerging as a subfield of theoretical chemistry, where the principal emphasis is focused on finding solutions to chemical related problems by employing the use of computers. Computational chemistry is not directly involved in developing theoretical methods, but rather in obtaining results relevant to chemical problems.¹ However a strong interplay remains dominant between the traditional stance of theory derived data and computational chemistry. The integration of computational chemistry may enable newly emerging, as well as long plaguing problems, to be studied in the search for effective solutions.² However, this remains dependent on the accuracy required, and the intrinsic characteristics of the system at hand, it is now possible to obtain meaningful information on systems that comprise of up to several thousand particles. The only limitation associated with computational chemistry may be the selection of an appropriate theory or method for a given problem, and the ability to assess the quality of the obtained results.³ There are two standard methods emanating from computational chemistry, the first being to study the chemistry of molecules at an electronic level, known as quantum mechanics, the second being molecular dynamics, which neglects explicit electron treatment and focuses on classical laws of physics. This chapter outlines the range of computational and theoretical tools applied in this study.

3.2 Quantum mechanics

In 1900, German theoretical physicist, Max Plank, inadvertently gave rise to the field of quantum mechanics, by inadvertently discovering that energy is discharged in small packets (called quanta) and emitted in wavelengths^{4,5}. Quantum mechanics is an essential constituent of computational chemistry, enabling the prediction of observable chemical properties. The fundamental law of quantum mechanics aims to illustrate that microscopic systems can be detailed by wave functions that capsule and characterize all physical properties of a system⁶. Quantum mechanics is the branch of mechanics that focuses on the mathematical analysis of

the motion and interaction of subatomic particles⁷, principally dealing with the influence of electromagnetic forces on the movement of electrons⁸. To perceive the electronic behaviour in molecules and consequently of the structures and reaction of molecules, knowledge of quantum mechanics, in particular the Schrödinger equation⁹. The characterization of a system in terms of a wave function can then be extracted by solving the Schrodinger equation in quantum chemistry. Quantum mechanics can be integrated to provide a better understanding and predict large-scale phenomena, initiating fundamental calculation of electronic structure and interactions¹.

3.3 The Schrödinger equation

The Schrödinger equation is regarded as the fundamental core of physics, as it entails a descriptive analysis of quantum mechanical behaviour within a system¹⁰. It was initially introduced by Austrian physicist Erwin Schrödinger in 1926¹⁰. In mathematical physics, the Schrödinger's equation undertakes the same role as the Hamilton's laws of motion as one of the basic equations in non-relativistic quantum mechanics and non-relativistic classical mechanics respectively^{11,12}.

There are two types of Schrödinger's equation: the first is the time-dependent Schrödinger's equation, this being the most applied equation in computational chemistry¹³, and defines the Hamiltonian operator as the accumulated value of the potential and kinetic energy. The second type is the time-independent Schrödinger's equation¹⁴. The simplest form of Schrodinger equation is presented as follows¹⁵:

$$\mathbf{H}\psi = \mathbf{E}\psi \quad \text{Eq.1}$$

Where H denotes the molecular Hamiltonian, ψ a wave function that expounds the probability of the electron and nuclear within disclosed locations, and E depicts the energy of the system (3, 5). The molecular Hamiltonian is the sum of the kinetic (T) and potential (V) energy, which can be denoted as¹⁶:

$$\mathbf{H} = \mathbf{T} + \mathbf{V} \quad \text{Eq. 2}$$

Particles are referred to as point masses, under the assumption that relativistic effects are not considered. The sum of kinetic and potential energy operators make up the composition of the Hamiltonian, which can then be presented in detail as:

$$\mathbf{H} = -\frac{\hbar^2}{8\pi^2} \left(\sum_i \frac{1}{m_i} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) + \sum_i \sum_j \left(\frac{e_i e_j}{r_{ij}} \right) \right) \quad \text{Eq.3}$$

The 1st term in Eq. 3 is operative of the kinetic energy of electrons, the 2nd term is the kinetic energy of the nuclei operator, the potential energy of electron-nuclei attractions operator is presented by term 3, the 4th term is the operator for potential energy of electron-electron repulsions, and the last term is the operator for potential energy of nuclei-nuclei repulsions¹⁸. The Schrödinger equation is highly complex, thus proving to be in-executable when solving for molecular systems, as it may contain thousands of atoms. Although, the Schrodinger equation integrates a range of equations, and cannot be resolved for a molecular system other than H₂, implementation of the Born-Oppenheimer approximation is considered an applicable solution as it compensates for molecular rather than atomic structure⁵.

3.4 The Born-Oppenheimer approximation

The Born-Oppenheimer approximation is considered an imperative commodity in rendering a solution to the Schrödinger equation, where the coupling between the nuclei and electronic *motion* is often omitted¹⁹. This enables the nuclear parameters to be taken into consideration when solving the electronic part and for the resulting *potential energy surface* (PES) to be integrated in finding a solution for motion²⁰. The Born-Oppenheimer approximation enables the use of the Schrodinger equation for a specific molecular system to be distinguished into two equations, namely the electron and nuclear equations from which the total energy of the system can be established¹³. The energy of a molecule is a function of the electron coordinates, but depends on the parameters of the nuclear coordinates, which define the molecular geometry. As nuclei are fixed, so the nuclear kinetic energy operator is neglected, as observed in Eq.4, permitting the statically distribution of electron input within a molecule to be determined^{21,14}.

$$\mathbf{T}_{elec} = -\frac{\hbar^2}{8\pi^2 m} \sum_i^{electrons} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \quad \text{Eq.4}$$

Presented below is the Schrödinger equation for fixed nuclei electrons¹⁷ which is incorporated in computational chemistry softwares:

$$\mathbf{H}^{elec} \varphi^{elec}(\mathbf{r}, \mathbf{R}) = \mathbf{E}^{eff}(\mathbf{R}) \varphi^{elec}(\mathbf{r}, \mathbf{R}) \quad \text{Eq.5}$$

3.5 Potential energy surface

The potential energy surface (PES) of a molecular system may be classified in terms of a set of force field energy equations that are fundamentally based on classical Newtonian physics. PES is a mathematical function required to correlate the energy of a molecule as a function of its geometry²¹, thereby enabling a deeper visual insight of structural characterization, as derived from the latent relationship of potential energy versus a molecules geometry²². PES is utilized to decipher energy minima, as well as the states of transition that occur within chemical reactions, with respect to the position of the nuclei²². The Born-Oppenheimer approximation is invoked in molecular systems to generate the PES²³. The concept of potential energy surface arises from variations in the mass and magnitude of the velocity between electrons and nuclei, a phenomenon defined by the Born-Oppenheimer approximation²⁴. The phenomenon of the Born-Oppenheimer approximation stipulates the instantaneous variation in the position of the electrons with regard to the nuclei displacement, thereby permitting the depiction of potential energy surface, as the potential of atoms motion present within a molecule or atoms in collision with one another is often referred to as the adiabatic motion^{25,22,26}.

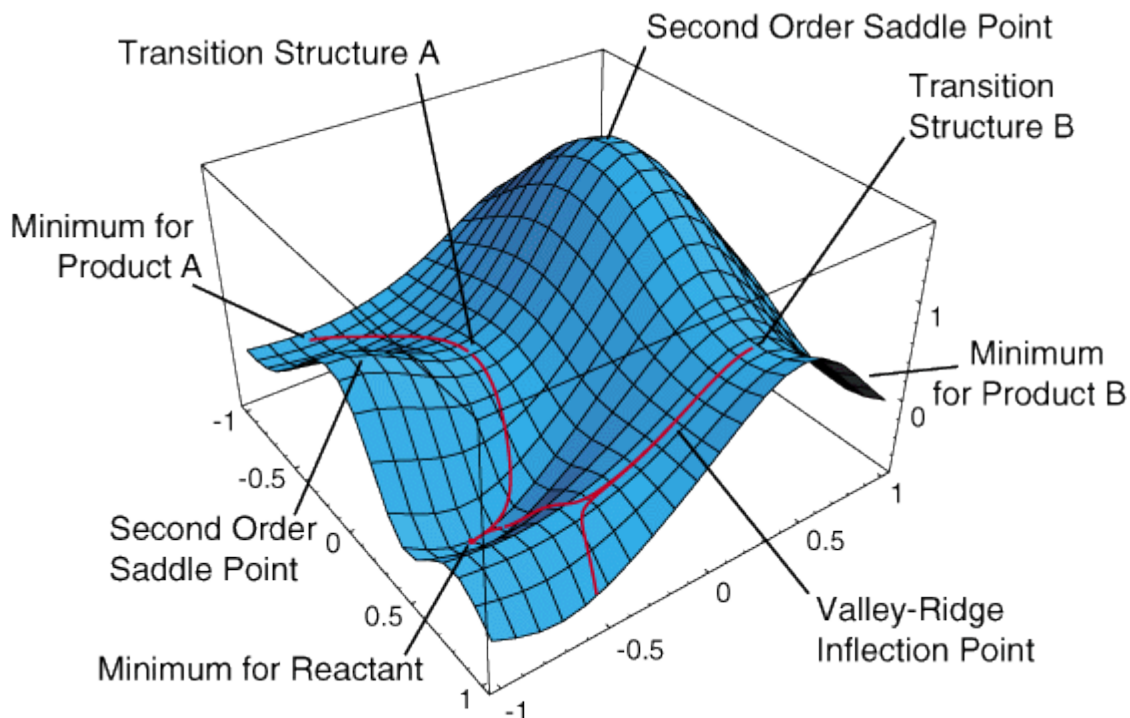


Figure 3.1 A Graphical representation of a two-dimensional potential energy surface²⁷.

The potential energy surface displays regions of unfavourable electronic interactions as high potential energy, as they depict molecular conformations or high-energy nuclear arrangements²⁷, with the regions of low energy being used as an indicator to display nuclear interactions presented as low energy molecular conformations and accompanied by favourable electronic interactions²⁸. Reactants and products are generally rendered as stable structures, resulting in relatively low potential energy that is often correlated with a reaction coordinate's minima on the PES. Figure 3.1 indicates that the minima represents the compounds structure during a state of equilibrium, whereas the first order saddle point represents the transition state or activated complex. The saddle point is referred to as an indicative factor of particular modifications in nuclear arrangements, which may result in decreased potential energy, while the others lead to higher energy.

3.6 Molecular mechanics

Molecular Mechanics (MM) may be defined as a set of models that utilize an empirical, algebraic, atomistic energy function for chemical systems^{29,30}. MM is based on the classical laws of physics, which is able to compute a molecule's geometrical and transitional state of equilibrium, as well as its relative energies present among conformers¹⁵. In this method, there is an assumption that the electrons will attain their optimal position once the position of the nuclei is defined and are hence not considered explicitly. The Born-Oppenheimer approximation forms the basis of the assumption, as it depicts nuclei as heavier than electrons, thus making their motion negligibly compared to that of electrons²¹.

MM is commonly applicable to larger systems, such as those in pharmaceutical and biological fields of research as they may exceed the use of more computer-intensive molecular orbital protocols³¹. The application to large systems can be attributed to suitable approximations (and its associated fast speed) and ability to determine the molecular conformation, or the atomic arrangement of a molecule, on the basis of its structural characterization and relative potential energies³². MM is an empirical method, is reliant on force field parameters that comprise of a range of parameters extracted from experimental data⁸. The total energy of a molecule as defined by molecular mechanics consists entirely of a range of interactions such as van der Waals and electrostatic contributions stemming from bond length and angles, torsions and non-bonded interactions, as presented Figure 3.2³³:

$$E_{tot} = E_{str} + E_{bend} + E_{tors} + E_{vdw} + E_{elec} \quad \text{Eq. 6}$$

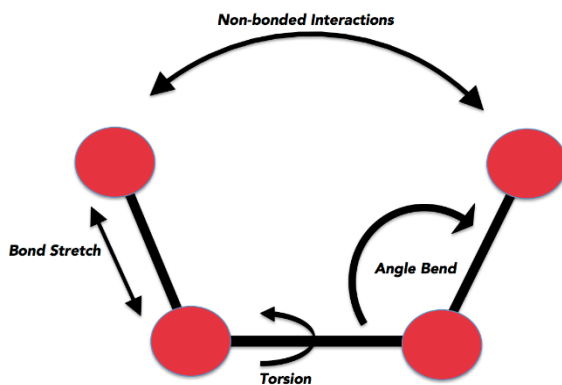


Figure 3.2 Diagrammatic representation of the total potential energy function of a molecule, as mentioned above (Prepared by Author).

Where; E_{tot} total energy

E_{str} bond-stretching energy

E_{bend} energy contribution from angle-bending

E_{tors} torsional energy contribution

E_{vdw} and E_{elec} van der Waals and electrostatic energy contributions respectively.

3.7 Force fields

A force field encompasses a set of functions and constants, defined as parameters, which can be applied to correlate the energy of the system in accordance with its particles³⁴. The parameters aim to define the reference points and force constants, providing a description of PES for various types of molecular systems with contrasting degrees of freedom that result from the inclusion of attractive or repulsive interactions between atoms. Force fields such as AMBER³⁵, CHARMM³⁶, NAMD³⁷ and GROMOS³⁸ display popularity, and are commonly used to set parameters that can be applied to the simulation of biomolecules. The parameters implemented in force fields are generated by ab initio method derivation or semi - empirical mechanical quantity calculations. These parameters can be from experimental data such as electron diffraction, X - rays and NMR and spectroscopy of neutrons^{39,34}. The forces acting within a molecule vary with regard to each system, with the administration of each force field needing to be adjusted accordingly. The different force fields are associated with a range of strengths and weaknesses, relative to the data applied, allowing a particular problem to be dealt with. The moderate low computational cost and accurate prediction protocols incorporated in the use of force fields highlights their use as an attractive option is molecular dynamics simulations and molecular mechanic calculations⁴⁰. There are a wide range of force fields that can be applied to a system, the force field of choice must be selected cautiously as each force field contains specific parameter that are designed for certain families of molecules⁴¹.

Therefore, it is crucial to choose the correct force field as it may increase the accuracy and applicability of theoretical structure function of your system of study. Force fields can only crudely approximate electrons interaction and hence cannot be used for bonding breaking nor formation calculations⁴². The AMBER force field⁴³ was employed in this study, by applying the General AMBER Force Field (GAFF) parameters accompanied by the standard AMBER⁴³ force field for the protein being introduced.

3.8 Molecular dynamics

Numerous studies have incorporated a vast range of molecular systems that include organic molecules in solution and biological macromolecules⁴⁴. Molecular dynamics^{45,46} and Monte Carlo⁴⁷ simulations are two techniques that provide useful insights into the structural, thermodynamic, and, for molecular dynamics, dynamical properties of systems in the condensed phase⁴⁸. Molecular dynamics (MD) is a computational technique used for simulating intricate molecular systems at an atomic level, as well as for computing the motions of individual molecules⁴⁹. The application of molecular dynamics is a sought after computational tool⁵⁰, as it enables the fluctuations that may occur in the motion of the system to be evaluated over a set period of time. MD enables the kinetic and thermodynamic properties of the molecular system to be determined⁴⁸. The low-energy deformation states derived from the inherent use of MD simulations can be incorporated to inspect the conformational space present in a large restricted system^{51,52}. If the initial geometry of a system is derived from experimental data stemming from X-ray or NMR structures, then MD techniques can be employed for sampling the conformational space⁵³. In order to set a MD system for simulation, the force, in combination with the energy of all particles within the system, must be calculated⁵⁴. MD integrates the use of Newton's equation of motion for atoms on an energy surface⁵⁵.

The main aim of using this specific computational application is to integrate the use of Newton's equations to decipher and gain a fundamental understanding of the energies and structural deviation that's may occur within a molecular network system. However, the following initial particle states are required:

1. The position and velocity of each particle within the molecular network system.
2. A formidable force field to distinguish the forces between atoms, e.g. AMBER or CHARMM
3. Boundary conditions must be specified

4. The classical equation of motion may then be solved

$$F_i = m_i \frac{d^2 r_i(t)}{dt^2} \quad \text{Eq. 7}$$

Where $r_i(t)$ represents the particle position vector, t represents the time-evolution, m depicts the mass of the particle and F_i is the interacting force on the particle.

The application of molecular dynamics, may be structured into four ongoing technical steps, that are continuous processed to generate a trajectory⁵⁶. The four technical steps include:

1. The fundamental requirements (states) of the biomolecular system are defined:
 - The co-ordinates of each atom within the MD system.
 - The bond characteristics present between each atom
 - The acceleration of each atom
2. The potential energy of each atom must be computed.
3. The equations of motion may be solved integrating the use of the potential energies extracted from step 2.
4. The cycle will proceed back to step 1 once the new “state” of the system is saved and co-ordinated of each atom is changed.

Once the trajectory has been completed, quantitative analysis will be performed on the system’s time- evolution.

3.9 Molecular Dynamics Post-Analysis

Molecular dynamic trajectories are generated directly from the production run of the simulation. The MD trajectories can be outlined as successive snapshots that are defined by the co-ordinates and velocity of each atom within the simulated system in addition to the time progression of the system in phase space^{57,58}.

The following requirements must be taken into consideration when choosing a software to perform molecular dynamic analysis

1. High qualitative visualization properties to enable accurate depiction of the trajectory’s video clips and generation of high-quality snapshots/images.
2. Software must be able to withhold and process large volumes of data efficiently.
3. Host a range of analytic tools easily accessible within one program

The techniques chosen for post-dynamic analysis must be directly dependent on the nature of the MD study. Although quantitative evaluation remains necessary, detailed analysis remains critical to support any visual systemization.

Therefore, post dynamic analysis of the projected trajectories remains essential for the following reasons:

1. To determine the energetic and conformational stability of the simulated system.
2. To determine and characterize the binding landscape of small molecules in addition calculate the thermodynamic energy fluctuations along assembled system's trajectory.
3. Present the dynamic conformational changes or variability that may occur within the biomolecular system throughout the MD trajectory.

3.10 System Stability of Simulated Systems

3.10.1 System Convergence

The term convergence is often used in molecular dynamic simulations to describe the structural stability of a protein system based on bond types and bond angle vibrations during the structural motion of a protein. In order to accurately analyse a MD trajectory a simulated system must depict a state of equilibrium which represents a state of final energetics and conformational plateau⁵⁹. This plateau may be directly linked to a protein-ligand system's energetically stable conformations being achieved.

3.10.2 Root Mean Square Deviation (RMSD)

The structural deviation of a complex may be determined by the spatial differentiation between two static structures of the same trajectory. The RMSD of a trajectory is defined as:

$$\text{RMSD} = \left(\frac{\sum_N (\mathbf{R}_i - \mathbf{R}_1^0)^2}{N} \right)^{\frac{1}{2}} \quad (\text{Eq 3.5})$$

Where: N depicts is the total number of atoms in a complex, $\mathbf{R}_i$ represents the vector position of the Ca atom of particle i is the conformational reference which is calculated after aligning the structure to an initial conformation (O) using the least square fitting protocol.

The average RMSD of a complex may be defined by taking the average structural deviation over the number of frames in each trajectory and can be calculated for the receptor, ligand and complex within a simulated system ⁶⁰.

3.10.3 Radius of Gyration (RoG)

The radius of gyration of system is measured as the root mean square distance of the atoms from their common centroid/center of gravity. This allows for the estimation of compactness of a protein complex along a trajectory. The RoG of a complex may be based on the following reaction:

$$r^2_{\text{gyr}} = \frac{(\sum_{i=1}^n w_i (r_i - r^-)^2)}{\sum_{i=1}^n w_i} \quad (\text{Eq 3.6})$$

Where: r_i is the position of the i th atom and r is the center weight of atom i .

The average RoG may be calculated by taking the average over the number of frames in a trajectory ⁶¹.

3.11 Thermodynamic calculations

The embodiment of thermodynamics in computational chemistry enables a deeper understanding of chemical reactions, provides a platform to calculate molecular properties and its derived entities, and predicts the chemical reactivity. The essential role of thermodynamic calculations is highlighted, due to its current contribution to the field of quantum mechanics. Thermodynamic calculations aid in distinguishing the energy surface associated with a specific chemical reaction. The use of thermodynamics can effectively justify minimization of energy, and its interconnection with the energy surface may therefore provide ample knowledge, based on the transition structure and reaction pathways⁶².

3.11.1 Binding free-energy calculations

Binding free energy calculations is an important end point method that provides essential information regarding the mechanism of binding between a ligand and enzyme, integrating both enthalpic and entropic contributions. Estimation of the binding free energy leads to development of various algorithms and approaches including free energy perturbation, thermodynamic integration, linear interaction energy and molecular docking calculations, to

mention a few^{63,64}. Free energy calculations have aided substantially in studies that include computational chemistry, thereby providing in-depth knowledge about drug design, protein structure determination⁶⁵ and protein-protein complexes^{66,67}. Two popular methods used to estimate the free binding energy with success of small ligands to biological macromolecules are the Molecular Mechanics/ Poisson-Boltzmann Surface Area (MMPB-SA) approach, followed by the Molecular Mechanics/Generalized Born Surface Area (MMGB-SA) approaches^{62,68}.

MM/GB-SA and MM/PB-SA rely on molecular simulations of the ligand-protein complex to compute rigorous statistical-mechanical binding free energy within a specified force field⁶⁹. Both approaches display favourable use, which can be attributed to their modular nature and lack of calculations that stem from training sets integrating continuum solvation models in combination with molecular mechanics calculations⁷⁰. Each approach displays avid accuracy and computational effort between the empirical scoring and stringent alchemical perturbation methods, and can be compared in order to reproduce and rationalize experimental data⁷¹. The MM/PB-SA and MM/GB-SA methods are utilised to determine free energy decomposition, which can be meticulously ranged into various groups, depending on the groups of atoms or types of interactions from which they originated^{66,72}.

The MM-PBSA employs a more rigorous algorithm than the MM/GB-SA, and simultaneously substitutes the MM/GB-SA model of electrostatics in water^{73,74}. However, with regard to calculations incorporating protein-drug interaction, including carbohydrates⁷⁵ and nucleic acids⁷⁶, the MM-GBSA method is favoured over the MM-PBSA⁷⁷. The use of binding free-energy calculations can also be utilized to enhance the results of virtual screening and the docking of therapeutics drugs⁷⁸. The binding free energy between the ligand and receptor highlighting the MM/GB-SA is given by⁷⁹:

$$G_{\text{bind}} = G_{\text{complex}} - G_{\text{receptor}} - G_{\text{ligand}} \quad \text{Eq. 9}$$

$$\Delta G_{\text{bind}} = \Delta E_{\text{mm}} + \Delta G_{\text{gsa}} - T\Delta S \quad \text{Eq.10}$$

$$\Delta E_{\text{mm}} = \Delta E_{\text{int}} + \Delta E_{\text{vdw}} + \Delta E_{\text{eel}} \quad \text{Eq. 11}$$

$$\Delta G_{\text{gsa}} = \Delta G_{\text{ebg}} + \Delta G_{\text{esurf}} \quad \text{Eq.12}$$

where ΔE_{MM} is the molecular mechanics energy of the system in a vacuum, ΔG_{GBSA} denotes the solvation free energy, $T\Delta S$ is the entropy, the sum of the bonded internal energy (ΔE_{Int}) is represented by ΔE_{MM} , non-bonded van der Waals (ΔE_{VDW}), electrostatic (ΔE_{EEL}) and ΔG_{GBSA} consists of polar contributions accounted for by the generalised born model (ΔG_{EGB}) and non-polar contributions (ΔG_{ESURF})^{80,81,82,83}.

The dynamic analysis of binding affinity aids in determining the approximate inhibitory activity of each inhibitor^{84,85}. During the MM-GBSA binding energy calculation, the correct binding conformation of each ligand can be determined prior to the binding energy estimation.⁸⁶ Herein, ligand-protein binding free energies were predicted using a MM/GB-SA approach.

3.12 Conformational Features of System

3.12.1 Root Mean Fluctuation (RMSF)

The root mean fluctuation (RMSF) of a protein measures residue's C α atom fluctuations based on the average protein structure along the system's trajectory. This extends to postulate the flexibility of regions of a protein based on the computed RMSF⁸⁷. To calculate the standardized RMSF, the following equation is applied:

$$sRMSF = \frac{(RMSF_i - \overline{RMSF})}{\sigma(RMSF)} \quad (Eq\ 3.8)$$

Where: $RMSF_i$ is the RMSF of the i^{th} residue, from which the average RMSF is subtracted. This is then divided by the RMSF's standard deviation to yield the resultant standardized RMSF.

The above method differs from RMSD and RoG as it is computed as the total residue fluctuation along the trajectory and is not analyzed at every frame in the trajectory.

3.13 Other Computer-Aided Drug Design Techniques Utilized in the Study

3.13.1 Molecular docking

Molecular docking is an essential component of computational chemistry, which involves determining the most optimal position of two molecules with respect to each other. Molecular

docking is utilized in structure-based drug design, and is often highlighted as one of the main contributing factors to the many problems arising in global optimization⁹⁰. The dynamic level of interaction displayed between ligand molecules and their receptors is often dependent upon the molecular recognition of the lock and key mechanism⁹¹.

Docking is often referred to as the positioning of a small molecule, such as an inhibitor or drug candidate, often referred to as a ligand, into the active site of macromolecules of known structural conformation⁴⁴. These macromolecules include proteins, such as nucleic acid, receptor or enzyme.⁷⁸. Being able to predict the binding of small molecules to target proteins plays a crucial role in structure-based drug design, as it enables the screening of virtual libraries⁹² of “drug like” molecules, thereby assisting in next generation drug development⁹³. The ligand-receptor binding energy is calculated as follow⁹⁴:

$$E_{binding} = E_{target} + E_{ligand} + E_{target-ligand} \quad \text{Eq. 8}$$

Numerous molecular docking programs are used for academic and commercial purposes⁹⁵, such as Dock⁹⁶, AutoDock GOLD⁹⁷, FlexX⁹⁸, -+ GLIDE⁹⁹, ICM⁹⁵, Surflex and others. While each program displays sufficient suitability for precise docking, the docking program Autodock generates two-orders of magnitude in comparison to other programs, while maintaining substantial accuracy in its binding mode prediction¹⁰⁰. The docking method used in this research study is the advanced version of AutoDock Vina¹⁰¹

The binding affinity generated by many docking softwares such as Autodock⁹⁶, neglect the presence of protons of the enzyme and inhibitors, thus the scores generated are often regarded as unreliable. MD takes into consideration protons and the solvent often water molecules. Thus MD results generates a more accurate binding affinity score as opposed to docking protocols⁴².

3.13.2 Homology modeling

In the midst of newly developing diseases, the crystal structure of a target protein is of utmost importance in the field of drug discovery and development¹⁰³. While various techniques that can be utilised to generate the structure of macromolecule, include the use of as X-ray crystallography, NMR spectroscopy and electron microscopy, they are often associated with multiple pitfalls. Homology modeling is gradually emerging as the sought after tool to be used to construct a 3D macromolecules structure due to its ease and accuracy in comparison to the other techniques¹⁰⁴. Homology modeling enables the construction of a proteins structure by using its sequence as a reference template, of which the X-ray crystal structure is known. An

accurate homology model depends on the existence, detection and quality of known template structures from the new protein maybe modelled. Although high-resolution structures are optimally extracted through X-ray crystallography, this approach is associated with a high cost increments, considerable experimental time and many trial runs¹⁰⁵. Furthermore, some biologically important macromolecules lack X-ray crystal structures or high-resolution 3D-structural properties, with reference to their protein sequence, with homology modeling being implemented to resolve this. An essential aim of drug discovery is to contrive bioactive molecules that are intended to target the disease condition, with minimal side effects, and hence are beneficial to the patient. The homology model of a 3D structure of a target protein may enable the full characterisation and exhibition of the mechanism of interaction on a structural and molecular level of a protein-ligand complex, elucidating the mode of action (MOA) of a drug molecule, and hence, greatly facilitate drug design.

3.13.3 Virtual Screening (VS)

Virtual screening (VS) plays a fundamental role in the drug discovery and development pipeline as the technique is defined by the assessment of extensive small molecule libraries in search of a new compound based on a biological target. The VS approach allows for the filtering of millions of small molecules to a manageable number of compounds that have the greatest chance as a lead drug. The method utilizes a wide variety of filters to identify biologically active alternatives to current inhibitors based on the “similar property principle”, which highlights that structures that share a high structural similarity with other molecules tend to have similar properties^{106,107,108}.

Pharmacophore based virtual screening (PBVS) has exhibited numerous benefits in computational hit identification and lead optimization. This approach uses pharmacophoric features based on a current inhibitor’s functional groups (hydrogen bond donors, hydrogen bond acceptors, cations, aromatics, hydrophobic areas). These pharmacophoric features are then established as the criteria when searching through existing small molecule libraries to identify a handful of compounds that may be validated as lead compounds.^{109–111}.

3.14 References

- (1) Breslow, R., and Tirrel, V. M. (2003) Beyond the Molecular Frontier: Challenges for Chemistry and Chemical Engineering. National Academy of Sciences, Yale.
- (2) Ekins, S., Mestres, J., and Testa, B. (2007) In silico pharmacology for drug discovery : methods for virtual ligand screening and profiling. *Br. J. Pharmacol.* 152, 9–20.
- (3) Grant, B. J., Gorfe, A. A., and Mccammon, J. A. (2009) Ras Conformational Switching : Simulating Nucleotide- Dependent Conformational Transitions with Accelerated Molecular Dynamics. *PLOS Comput. Biol.* 5, 1–10.
- (4) Atkins, P., and Friedman, R. (2005) Molecular Quantum Mechanics Fourth Edition. Oxford University Press, New York.
- (5) Jensen, F. (2007) Introduction to Computational chemistry. *Introd. to Comput. Chem.* John Wiley and Sons Ltd, West Sussex.
- (6) Sauer, J., and Sierka, M. (2000) Combining Quantum Mechanics and Interatomic Potential Functions in Ab Initio Studies of Extended Systems. *J. Comput. Chem.* 21, 1470–1493.
- (7) Amara, P., and Field, M. J. (2009) Combined Quantum Mechanical and Molecular Mechanical Potentials. *Oxford Univ. Press* 2, 1–9.
- (8) Bornemann, F. A., Nettesheim, P., Schu, C., and Introduction, I. (1996) Quantum-classical molecular dynamics as an approximation to full quantum dynamics “. *J. Chem. Phys.* 105, 1074–1083.
- (9) Tannor, J. D. (2010) introduction to quantum mechanics : a time dependent perspective. John Wiley and Sons, Houston.
- (10) Schrodinger, E. (1926) An undulatory theory of the molecules of atoms and molecules. *Phys. Rev.* 22.
- (11) Car, R. (2002) Introduction to Density-Functional Theory and ab-Initio Molecular Dynamics. *Quantum Struct. Act. relations* 21, 97–104.
- (12) Tomasi, J., Mennucci, B., and Cammi, R. (2005) Quantum Mechanical Continuum Solvation Models. *Chem. Rev.* 105, 2999–3093.
- (13) Shih, H.-W. (2009) Fundamentals of Computational Chemistry. Oxford University Press, West Sussex.

- (14) Atkins, P., and Paula, de J. (2006) ATKINS physical chemistry. W.H Freeman and Company, New York.
- (15) Lewars, E. (2004) Computational Chemistry: An introduction to the theory applications of molecular and quantum mechanics. Kluwers Academic Publishers, Moscow.
- (16) Quinn, R., and Schimdt, B. (2003) The theory behind Computational Chemistry. *PLOS Comput. Biol.* 4, 34–68.
- (17) Schroeder, D. V. (2016) The Schrodinger Equations, in *Techniques in Computational Chemistry*, pp 1–5.
- (18) Kruglyak, A. Y., and Whitman, D. . (1963) Calculations of Quantum chemistry. Russia.
- (19) Bechstedt, F. (2015) Born-Oppenheimer Approximation, in *Many-Body Approach to Electronic Excitations*, pp 3–12.
- (20) Bechstedt, F. (2015) Born-Oppenheimer Approximation. *Springer* 5, 3–12.
- (21) Cramer, C. J. (2004) Essentials of computational chemistry second edition: Theories and models. John Wiley and Sons Ltd, West Sussex.
- (22) Elsayy, K. M., Hodgson, M. K., and Caves, L. S. D. (2005) The physical determinants of the DNA conformational landscape : an analysis of the potential energy surface of single-strand dinucleotides in the conformational space of duplex DNA. *Nucleic Acids Res.* 33, 5749–5762.
- (23) Jaquet, R. (2002) Introduction to potential energy surfaces and graphical interpretation.
- (24) Lewars, E. G. (2011) The Concept of the Potential Energy Surface. *Comput. Chem.* 6, 9–13.
- (25) Hratchian, P. H., and Schlegel, H. B. (2005) Finding minima, transition states, and following reaction pathways on ab initio potential energy surfaces, in *Theory and Applications of Computational Chemistry: The first forty years*, pp 195–249.
- (26) Truhlar, D. G. (2001) Potential Energy Surfaces. Academic Press, New York.
- (27) Sherrill, C. D. (2009) Potential Energy Surfaces. Georgia.
- (28) Long, J. (2010) Use of potential energy surfaces (PES) in spectroscopy and reaction dynamics. University of Iceland.
- (29) Lin, H., and Truhlar, D. G. (2007) QM / MM : what have we learned , where are we , and

where do we go from here ? *Theor. Chem. Acc.* 117, 185–199.

(30) Burke, K., and History, I. I. A. B. (2012) Perspective on density functional theory. *J. Chem. Phys.* 150901, 1–9.

(31) Dejaegere, A., Lafont, V., Schaefer, M., and Stote, R. H. (2007) Protein – Protein Recognition and Interaction Hot Spots in an Antigen – Antibody Complex : Free Energy Decomposition Identifies “Efficient Amino Acids.” *Proteins Struct. Funct. Genet.* 434, 418–434.

(32) Kwan, E., Wiley, S. M., and Wiley, C. J. (2010) Introduction to Computational Chemistry. London.

(33) Vanommeslaeghe, K., Guvench, O., and Mackarellel, D. A. (2015) Molecular Mechanics. *Curr. Protoc. Comput. drug Des.* 20, 3281–3292.

(34) González, M. A. (2011) Force fields and molecular dynamics simulations. *EDP Sci.* 12, 169–200.

(35) Case, D. A., Cheatham, T. E., Darden, T. O. M., Gohlke, H., Luo, R. A. Y., Merz, K. M., Onufriev, A., Simmerling, C., Wang, B., and Woods, R. J. (2005) The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* 26, 1668–1688.

(36) Brooks, B. R., Iii, C. L. B., Mackerell, A. D., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A. R., and Feig, M. (2009) CHARMM : The Biomolecular Simulation Program. *J. Comput. Chem.* 30.

(37) Phillips, J. C., Braun, R., Wang, W. E. I., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R. D., and Poincare, H. (2005) Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* 26, 1781–1802.

(38) Bakowies, D., Baron, R., Christen, M., Hu, P. H., Geerke, D. P., Heinz, T. I. M. N., Kastenholtz, M. A., and Gunsteren, W. F. V. A. N. (2005) The GROMOS Software for Biomolecular Simulation :GROMOSO5. *J. Comput. Chem.* 26.

(39) Abildgaard, J. (1998) Quantum Chemical Models in Molecular Structure Elucidation. W.H Freeman and Company, Roskilde.

(40) Cusack, P. (2002) Molecular Force Fields. New Jersey.

(41) Mackerell, A. D. (2010) Protein Force Fields. Baltimore, USA.

- (42) Honarparvar, B., Govender, T., Maguire, G. E. M., Soliman, M. E. S., and Kruger, H. G. (2014) Integrated Approach to Structure-Based Enzymatic Drug Design : Molecular Modeling , Spectroscopy , and Experimental Bioactivity. *Chem. Rev.* 114, 493–537.
- (43) Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., Cheatham, T. E., Debolt, S., Ferguson, D., Seibel, G., and Kollman, P. (1995) AMBER , a package of computer programs for applying molecular mechanics , normal mode analysis , molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comput. Phys. Commun.* 91, 1–41.
- (44) Ramachandran, K. I., Deepa, G., and Namboori, K. (2008) Computational chemistry and molecular modeling. Springer-Verlag Berlin Heidelberg, Coimbatore.
- (45) Petruk, A. A., Defelipe, L. A., Marti, M. A., and Turjanski, A. G. (2013) Molecular Dynamics Simulations Provide Atomistic Insight into Hydrogen Exchange Mass Spectrometry Experiments. *J. Chem. Theory Comput.* 9, 658–669.
- (46) Nad, S. (2007) Molecular dynamics simulations and their application to four-stranded DNA. *Methods* 43, 278–290.
- (47) Neyts, E. C., Bogaerts, A., Carlo, M., Molecular, K., Monte, Á., and Long, C. Á. (2013) Combining molecular dynamics with Monte Carlo simulations : implementations and applications. *Theor. Chem. Acc.* 132, 1–12.
- (48) Petrenko, R. (2010) Molecular Dynamics, in *In:Encyclopedia of Life Sciences*, pp 1–13. John Wiley and Sons Ltd, Chichester.
- (49) Amadei, A., Linssen, A. B. M., and Berendsen, H. J. C. (1993) Essential Dynamics of Proteins. *Proteins Struct. Funct. Genet.* 425, 412–425.
- (50) Mhlongo, N. N., and Soliman, M. E. S. (2015) RSC Advances Single H5N1 in fl uenza A neuraminidase mutation develops resistance to oseltamivir due to distorted conformational and drug binding landscape : multiple molecular dynamics analyses †. *RSC Adv.* 5, 10849–10861.
- (51) Hehre, W. J. (2003) A Guide to Molecular Mechanics and Quantum Chemical Calculations. Wavefunction, Inc, Irvine.
- (52) Road, H. (1976) Theoretical Studies of Enzymic Reactions. *J. Mol. Biol.* 103, 227–249.
- (53) Dewar, M. J. S., Zoebisch, E. G., Healy, E. F., and Stewart, J. J. P. (1993) A quantum

mechanical molecular model. *J. Am. Chem. Soc.* 49, 5003.

(54) Rogers, D. W. (2003) Computational Chemistry Using the PC Third Edition. John Wiley and Sons, Inc, New Jersey.

(55) Adcock, S. A., and Mccammon, J. A. (2006) Molecular Dynamics : Survey of Methods for Simulating the Activity of Proteins. *Chem. Rev.* 106, 1589–1615.

(56) Jakobsson, E. (2001) Computational Biochemistry and Biophysics Edited by Oren M. Becker (Tel Aviv University), Alexander D. MacKerell, Jr. (University of Maryland), Benoît Roux (Cornell University), and Masa-katsu Watanabe (Wavefunction, Inc.). Marcel Dekker: New York and Ba. *J. Am. Chem. Soc.*

(57) Likhachev, I. V., Balabaev, N. K., and Galzitskaya, O. V. (2016) Available Instruments for Analyzing Molecular Dynamics Trajectories. *Open Biochem. J.* 10, 1–11.

(58) Jarosaw Meller. (2001) Molecular Dynamics. *Encycl. Life Sci.*

(59) Amadei, A., Ceruso, M. A., and Di Nola, A. (1999) On the convergence of the conformational coordinates basis set obtained by the Essential Dynamics analysis of proteins' molecular dynamics simulations. *Proteins Struct. Funct. Genet.* 36, 419–424.

(60) Kufareva, I., and Abagyan, R. (2012) Homology Modeling. *Methods Mol Biol.* 857, 231–257.

(61) Lobanov, M. Y., Bogatyreva, N. S., and Galzitskaya, O. V. (2008) Radius of Gyration as an Indicator of Protein Structure Compactness. *Mol. Biol.* 42, 701–706.

(62) Wang, J., Hou, T., and Xu, X. (2006) Recent Advances in Free Energy Calculations with a Combination of Molecular Mechanics and Continuum Models. *Curr. Protoc. Comput. drug Des.* 7, 287–306.

(63) Kamerlin, C. L. S., Haranczyk, M., and Warshel, A. (2009) Progress in Ab Initio QM / MM Free-Energy Simulations of Electrostatic Energies in Proteins : Accelerated QM / MM Studies of p K a , Redox Reactions and Solvation Free. *J. Phys. Chem.* 113, 1253–1272.

(64) Kollman, P. (1993) Free Energy Calculations: Applications to Chemical and Biochemical Phenomena. *Chem. Rev.* 93, 2395–2417.

(65) Homeyer, N., and Gohlke, H. (2012) Free Energy Calculations by the Molecular Mechanics Poisson À Boltzmann Surface Area Method. *Mol. Inf.* 31, 114–122.

- (66) Gohlke, H., Kiel, C., and Case, D. A. (2003) Insights into Protein – Protein Binding by Binding Free Energy Calculation and Free Energy Decomposition for the Ras – Raf and Ras – RalGDS Complexes. *J. Mol. Biol.* 2836, 891–913.
- (67) Hou, T., Chen, K., Mclaughlin, W. A., Lu, B., and Wang, W. (2006) Computational Analysis and Prediction of the Binding Motif and Protein Interacting Partners of the Abl SH3 Domain. *PLOS Comput. Biol.* 2.
- (68) Genheden, S., and Ryde, U. (2016) Expert Opinion on Drug Discovery The MM / PBSA and MM / GBSA methods to estimate ligand-binding affinities The MM / PBSA and MM / GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.* 0441.
- (69) Wang, J., Hou, T., Li, Y., and Wang, W. (2012) Assessing the performance of the MM/PBSA and MM/GBSA methods: I. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J. Chem. Inf. Model.* 51, 69–82.
- (70) Hou, T., Wang, J., Li, Y., and Wang, W. (2011) Assessing the Performance of the MM / PBSA and MM / GBSA Methods . 1 . The Accuracy of Binding Free Energy Calculations Based on Molecular Dynamics Simulations. *J. Chem. Inf. Model.* 51, 69–82.
- (71) Zoete, V., Meuwly, M., and Karplus, M. (2005) Study of the Insulin Dimerization : Binding Free Energy Calculations and Per-Residue Free Energy Decomposition. *Proteins Struct. Funct. Genet.* 93, 79–93.
- (72) Hou, Tingjun, Li Nan, W. W. (2013) Characterization of Domain–Peptide Interaction Interface: Prediction of SH3 Domain-Mediated Protein–Protein Interaction Network in Yeast by Generic Structure-Based Models. *NIH Public Access* 11, 2982–2995.
- (73) Case, D. A. (2001) Theory and Applications of the Generalized Born Solvation Model in Macromolecular Simulations. California.
- (74) Onufriev, A., Bashford, D., and Case, A. D. (2016) Modification of the Generalized Born Model Suitable for Macromolecules. *J. Phys. Chem.* 2000, 3712–3720.
- (75) Feig, M., and Iii, C. L. B. (2002) Evaluating CASP4 Predictions With Physical Energy Functions. *Proteins Struct. Funct. Genet.* 245, 232–245.
- (76) Wang, J., and Deng, Y. (2006) Absolute Binding Free Energy Calculations Using Molecular Dynamics Simulations with Restraining Potentials. *Biophys. J.* 91, 2798–2814.

- (77) Sun, H., Li, Y., Tian, S., and Hou, T. (2014) and MM / GBSA methodologies evaluated by various simulation protocols using PDBbind data set †. *Phys. Chem.* 16, 16719–16729.
- (78) Aalten, D. M. F. Van, Findlay, J. B. C., and Amadei, A. (1995) Essential dynamics of the cellular retinol-binding protein evidence for ligand-induced conformational changes. *Protein Eng.* 8, 1129–1135.
- (79) Decherchi, S., Masetti, M., Vyalov, I., and Rocchia, W. (2015) Implicit solvent methods for free energy estimation. *Eur. J. Med. Chem.* 16, 27–42.
- (80) Genheden, S., and Ryde, U. (2015) The MM / PBSA and MM / GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.* 10, 449–461.
- (81) Cappel, D., Hall, M. L., Lenselink, E. B., Qi, J., Bradner, J. E., Sherman, W., Cappel, D., Hall, M. L., Lenselink, E. B., Beuming, T., Qi, J., and Bradner, J. (2016) Relative Binding Free Energy Calculations Applied to Protein Homology Models. *J. Chem. Inf. Model.* 5, 1–15.
- (82) Huang, Y. M., Chen, W., Potter, M. J., and Chang, C. A. (2012) Insights from Free-Energy Calculations : Protein Conformational Equilibrium , Driving Forces , and Ligand-Binding Modes. *Biophys. J.* 103, 342–351.
- (83) Genheden, S., and Ryde, U. (2016) Expert Opinion on Drug Discovery The MM / PBSA and MM / GBSA methods to estimate ligand-binding affinities The MM / PBSA and MM / GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.* 0441, 449–461.
- (84) Chang, C. A., Chen, W., and Gilson, M. K. (2007) Ligand configurational entropy and protein binding. *J. Proceeding Natl. Acad. Sciences* 104, 1534–1539.
- (85) Wright, D. W., Hall, B. A., Kenway, O. A., Jha, S., and Coveney, P. V. (2014) Computing Clinically Relevant Binding Free Energies of HIV - 1 Protease Inhibitors. *Jounral Chem. theroy Comput.* 10, 1228–1241.
- (86) Srivastava, M., Singh, H., and Naik, K. P. (2010) Molecular modeling evaluation of the antimalarial activity of artemisinin analogues : molecular docking and rescoring using Prime / MM-GBSA Approach. *Curr. Res. J. Biol. Sci.* 83–102.
- (87) Bornot, A., Etchebest, C., and De Brevern, A. G. (2011) Predicting protein flexibility through the prediction of local structures. *Proteins Struct. Funct. Bioinforma.* 79, 839–852.

- (88) Kasahara, K., Fukuda, I., and Nakamura, H. (2014) A novel approach of dynamic cross correlation analysis on molecular dynamics simulations and its application to Ets1 dimer-DNA complex. *PLoS One* 9.
- (89) Tiberti, M., Invernizzi, G., and Papaleo, E. (2015) (Dis)similarity Index to Compare Correlated Motions in Molecular Simulations. *J. Chem. Theory Comput.* 11, 4404–4414.
- (90) Eweas, A. F., Maghrabi, I. A., and Namarneh, A. I. (2014) Advances in molecular modeling and docking as a tool for modern drug discovery. *Der Pharmaceutical Chem.* 6, 211–228.
- (91) Du, X., Li, Y., Xia, Y., Ai, S., Liang, J., Sang, P., and Ji, X. (2016) Insights into Protein – Ligand Interactions : Mechanisms , Models , and Methods. *Int. J. Mol. Sci.* 17, 1–34.
- (92) Talele, T. T., Khedkar, S. A., and Rigby, A. C. (2010) Successful Applications of Computer Aided Drug Discovery : Moving Drugs from Concept to the Clinic. *Curr. Top. Med. Chem.* 10, 127–141.
- (93) Young, D. C. (2001) CHEMISTRY COMPUTATIONAL CHEMISTRY A Practical Guide for Applying Techniques to Real-World Problems. John Wiley and Sons, Inc, Brisbane.
- (94) Nervall, M. (2007) Binding Free Energy Calculations on Ligand-Receptor Complexes Applied to Malarial Protease Inhibitors. Uppsala Universitet.
- (95) Cross, J. B., Thompson, D. C., Rai, B. K., Baber, J. C., and Fan, K. Y. (2009) Comparison of Several Molecular Docking Programs : Pose Prediction and Virtual Screening Accuracy. *J. Chem. Inf. Model.* 49, 1455–1474.
- (96) Ewing, T. J. A., Makino, S., Skillman, A. G., and Kuntz, I. D. (2001) DOCK 4 . 0 : Search strategies for automated molecular docking of flexible molecule databases. *J. Comput. Aided. Mol. Des.* 15, 411–428.
- (97) Verdonk, M. L., Cole, J. C., Hartshorn, M. J., Murray, C. W., and Taylor, R. D. (2003) Improved Protein – Ligand Docking Using GOLD. *Proteins Struct. Funct. Genet.* 623, 609–623.
- (98) Joseph-mccarthy, D., and Alvarez, J. C. (2003) Automated Generation of MCSS-Derived Pharmacophoric DOCK Site Points for Searching Multiconformation Databases. *Proteins Struct. Funct. Genet.* 202, 189–202.

- (99) Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., Repasky, M. P., Knoll, E. H., Shelley, M., Perry, J. K., Shaw, D. E., Francis, P., and Shenkin, P. S. (2004) Glide : A New Approach for Rapid , Accurate Docking and Scoring . 1 . Method and Assessment of Docking Accuracy. *J. Med. Chem.* 47, 1739–1749.
- (100) Vajda, S., and Kozakov, D. (2009) Convergence and combination of methods in protein – protein docking. *Curr. Opin. Struct. Biol.* 19, 164–170.
- (101) Trott Oleg, O. J. A. (2011) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function,efficient optimization and multithreading. *NIH Public Access* 31, 455–461.
- (102) Crane, E. A., and Gademann, K. (2016) Capturing Biological Activity in Natural Product Fragments by Chemical Synthesis. *Angew. Chemie Int. Ed.* 55, 2–23.
- (103) Ashburn, T. T., and Thor, K. B. (2004) Drug repositioning identifying and developing new uses for existing drugs. *Nat. Rev.* 3, 673–683.
- (104) Forrest, L. R., Tang, C. L., and Honig, B. (2006) On the Accuracy of Homology Modeling and Sequence Alignment Methods Applied to Membrane Proteins. *Biophys. J.* 91, 508–517.
- (105) Forrest, L. R., Tang, C. L., and Honig, B. (2006) On the Accuracy of Homology Modeling and Sequence Alignment Methods Applied to Membrane Proteins. *Biophys. J.* 91, 508–517.
- (106) Lionta, E., Spyrou, G., Vassilatis, D. K., and Cournia, Z. (2014) Structure-Based Virtual Screening for Drug Discovery : Principles , Applications and Recent Advances. *Curr. Top. Med. Chem.* 14, 1923–1938.
- (107) Vyas, V., Jain, A., Jain, A., and Gupta, A. (2008) Virtual screening: A fast tool for drug design. *Sci. Pharm.* 76, 333–360.
- (108) Anderson, A. . (2003) The process of Structure-Based Drug Design. *Chem. Biol.* 10, 787–797.
- (109) Kim, K., Kim, N. D., and Seong, B. (2010) Pharmacophore-based virtual screening : a review of recent applications. *Expert Opin. Drug Discov.* 5, 205–222.
- (110) Sliwoski, G., Kothiwale, S., Meiler, J., and Lowe, E. W. (2014) Computational Methods

in Drug Discovery. *Pharmacol. Rev.* 66, 334–395.

(111) Drie, J. H. (2007) Computer-aided drug design: The next 20 years. *J. Comput. Aided. Mol. Des.* 21, 591–601.

CHAPTER 4

Egress and Invasion Machinery of Malaria: An In-depth Look into The Structural and Functional Features of The Flap Dynamics of plasmepsin IX and X

Geraldene Munsamy^a, Pritika Ramharack^a and Mahmoud E. S. Soliman^{a*}

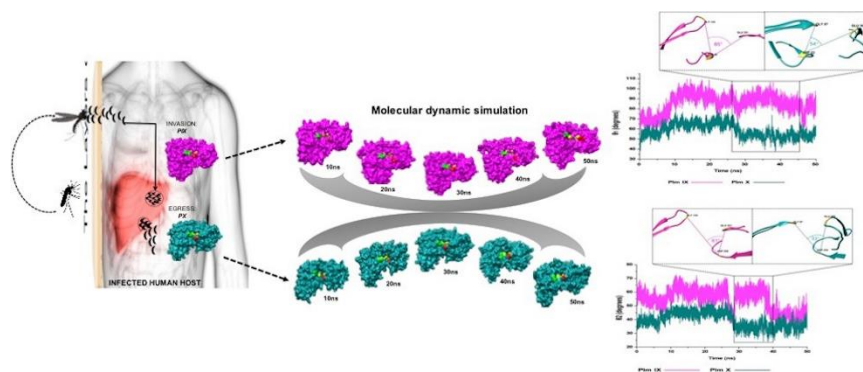
^a Molecular Bio-computation and Drug Design Laboratory, School of Health Sciences,
University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa

*Corresponding Author: Mahmoud E.S. Soliman

Email: soliman@ukzn.ac.za

Telephone: +27 (0) 31 260 8048, Fax: +27 (0) 31 260 7872

4.1 Graphical Abstract



4.2 Abstract

Plasmeprins, a family of aspartic proteases expressed by *Plasmodium falciparum* parasite, have been identified as key mediators in the onset of lethal malaria. Precedence has been placed on this family of enzymes due its essential role in the virulence of the parasite, thus highlighting its importance as novel drug targets. A previously published study by our group proposed a set of parameters used to define the flap motion of aspartic proteases. These parameters were used in the study of Plm I-V and focused on the flap flexibility as well as structural dynamics. Recent studies have highlighted the essential role played by Plm IX and X in egress and invasion of the malarial parasite. This study aims to close the gap on the latter family, investigating the flap dynamics of Plm's IX and X. Molecular dynamic simulations demonstrated an "open and close" mechanism at the region of the catalytic site. Further computation of the dihedral angles at the catalytic region revealed tractability at both the flap tip and flexible loop. This structural versatility enhances the interaction of variant ligand sizes, in comparison to other Plm family members. The results obtained from this study signify the essential role of structural flap dynamics and its resultant effect on the binding landscapes of Plm IX and X. We believe that this unique structural mechanism may be integrated in the design and development of effective antimalarial drugs.

Key Words: *P. falciparum*; Plasmeprin; Flap dynamics; Aspartic protease; Molecular dynamics, Plm IX, Plm X

4.3 Introduction

According to the World Health Organization (WHO), by 2016, there were an estimated 216 million reported cases of malaria worldwide¹. Most cases of malaria and deaths occur in sub-Saharan Africa^{2,3}, however, regions of South-East Asia, Eastern Mediterranean, Western Pacific, and the Americas are also at critical risk⁴. Since then, the infectious rate of malaria has almost doubled, with an estimated 500 million people being infected annually, with majority of these emerging cases being children under the age of five⁵.

In humans, the parasite is transmitted via the female *Anopheles* mosquito vector. Humans are infected by a range of *Plasmodium* species⁶, however, the most severe and common forms of malaria are caused by *Plasmodium falciparum* and *Plasmodium vivax*⁷. *P. falciparum*, predominating in Africa, is considered to be the most virulent species being responsible for 85% of human death⁸.

The complex life cycle of *P. falciparum* embodies multiple cycles of invasion, multiplication and egress in the human red blood cells^{9,10}. Evidently, targeting the egress/invasion machinery, at its most vulnerable stage of the life cycle of the parasite may not only reduce the severity of the malarial disease, but may jointly eradicate disease transmission¹¹. Even though the process of invasion occurs rapidly, it is the only time during the parasite's life cycle when it is directly exposed to the host immune system¹². There are several enzymes in *P. falciparum*, that have been implicated in hemoglobin proteolysis, parasite nutrition and development, these enzyme families include; cysteine proteases; aspartic proteases; metalloproteases and dipeptidyl aminopeptidases.^{13,14}

Aspartic proteases, or plasmepsins, have been identified as key mediators of cellular processes, including hemoglobin degradation for the export of *Plasmodium* proteins that essential for parasite growth/survival¹⁵ and particularly malarial egress and invasion¹⁶. *P. falciparum* possess a repertoire of 10 aspartic proteases (Plm I to X)¹⁷. Of the 10 Plms identified in *P. falciparum*, only Plm I,II, HAP, IV and V have been studied extensively¹⁸.

Plm VI-VIII are expressed in the vector during the parasite's intra-erythrocytic stages of sporozoite formation, motility as well midgut transversal¹⁹ (Figure 4.1). Studies have suggested that due to their expression during the sporogonial cycle in the mosquito, these Plm's may not be ideal drug target candidates²⁰. However, understanding the molecular mechanism by which the ookinete invades the mosquito vector can form a potential approach to developing malaria

transmission-blocking strategies, which may pave the way for future studies. Plm IX and X are expressed concurrently with Plm I–IV, but are not transported to the food vacuole, and have been identified as vital mediators in the progression of infection within the human host²². The essential roles undertaken by Plm IX and X further exemplifies their choice in our study in understanding the structural and functional anomalies associated with these enzymes. The ambiguity surrounding the functional and structural features of Plm IX and X may be attributed to the absence of tertiary structures of these neglected proteases. Gaining insight into the functional and structural features of Plm IX–X may aid in the design of effective therapeutics against malaria.

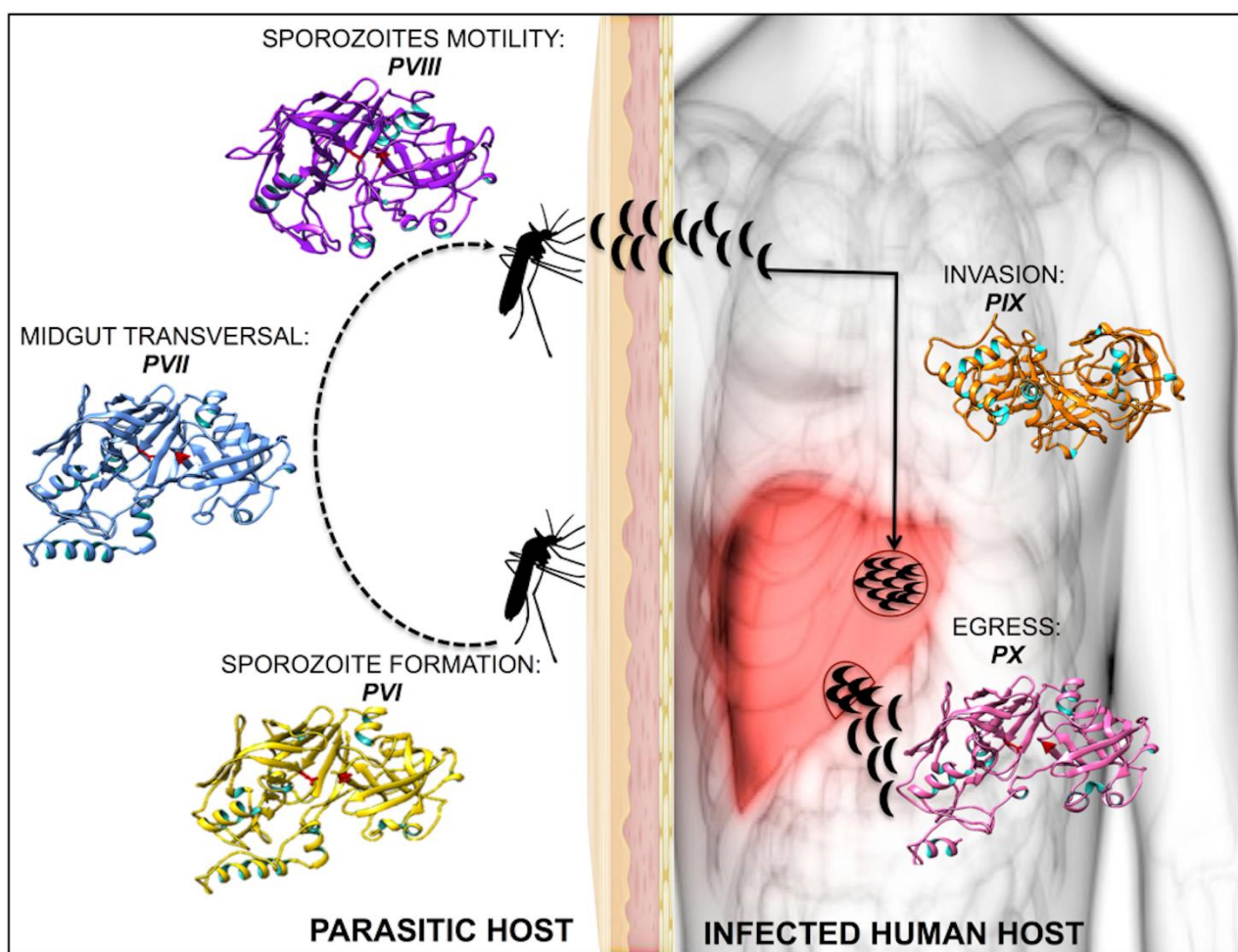


Figure 4.1 A diagrammatic representation of plasmepsin VI–X, highlighting the role of each Plm towards malarial infection and invasion (prepared by author).

The flap and flexible region of motion as well as the characteristic “twisting²³” motion during the opening and closing of Plm’s is essential in understanding the ligand binding landscape as well as conformational flexibility of Plm’s. Flaps covering the active site have a dual role in Plm function: (i) structural, as interactions formed between a ligand and the flaps stabilize the

ligand-protease complex, and (ii) kinetic, as flap closing induces ligand binding and flap opening induces ligand release²⁴.

Previous work reported by our research group proposed parameters to measure the flap dynamics of Plm I-V²³, which we have implemented in this study. In the absence of crystal structures of Plm IX-X, *in silico* models were generated for Plm IX-X and further refinement of the generated models were performed using molecular dynamic simulations.

Two recent studies by Paco Pino *et al* 2017²⁵ and Armiyaw Nasamua *et al* 2017²² have shown the expression of Plm IX and Plm X in late schizont/merozoites, and their potential involvement in egress-invasion. These studies have identified Plm IX and X as being essential for parasitic invasion and egress, and are evidently expressed in mature blood-stage schizonts and invasive merozoites and fulfil indispensable but unknown functions. A Conditional knockout study performed on Plm IX has revealed its role as a maturase during merozoite formation, and further substantiates its importance for RBC invasion²⁶. Plm IX and X have been identified as crucial mediators of invasion and egress of the malarial parasite. Therefore, qualifying as promising dual targets toward malaria eradication.

This study aims in generating a better understanding of the structural and mechanistic features of Plm IX and X, considering the importance surrounding the role of flap dynamics which is a characteristic and distinguishable feature of the family of aspartic proteases as well as utilise a range of computational techniques to provide comprehensive structural and functional data that will aid in the identification or design of inhibitors specific to Plm IX and X.

4.4 Computational Methodology

4.4.1 In silico Modelling

Crystal structures generated from experimental methods, such as X-ray crystallography and or NMR analysis, have aided researchers in the understanding of the form and function of protein targets over the past decade and are continuously being proven to be useful in the discovery of small molecule based protein inhibitors^{27,28}. However, limitations such obtaining a large quantity of absolute pure protein, as well as large molecular weight are challenges often associated with the crystallization of a protein, not to mention the substantial financial investments in equipment and human infrastructure that is required. A more cost efficient and sought after approach, for the determination of the tertiary structure of many proteins, not reported in the Protein Data Bank (PDB) is *in silico* modelling^{27,29}.

The amino acid sequence for *Plasmodium falciparum* aspartic proteases IX-X was accessed from NCBI³⁰, followed by the generation of the predictive 3-D structure using the Protein Homology/analogy Recognition Engine (PHYRE2)³¹ server, which selects templates for structure prediction based on secondary structural alignments. Plm's IX-X were modelled against the PDB structure of cathepsin D (4OBZ), the most closely related human aspartic protease¹⁹. The sequence similarity and Z-score of each model to reference template 4OBZ is presented in Table S4.1. A Ramachandran plot for the analyses of bond angles and torsional strain was generated using RAMPAGE³² (Figure S4.1). Results showed that all models generated had > 90.0 % of all residues in the favoured. With only a list of 12 outliers, none of which formed part of the active site of the protein. Each model was further validated using ProSA-web³³ presented in Figure S4.2.

4.4.2 System preparation

Plm IX and X were subjected to molecular dynamic simulations. The enzymes were prepared and visualised using Chimera. In total, two systems were subjected to continuous molecular dynamic simulations as described below in section "Molecular dynamics and post dynamics analyses." To verify and validate the results from the single continuous MD approach adopted herein, we ran multiple MD simulations (using three replicas) for one system and compared the results.

4.4.3 Molecular dynamics and post dynamic analysis

To investigate the flap dynamics of apo Plm IX and X, a continuous 50 ns MD approach was utilised in the present study. All-atom, explicit solvation unrestrained molecular dynamic simulations were performed using the GPU version of the SANDER³⁴ engine incorporated with the Amber 14 package which integrates the standard AMBER (FF14SB)³⁵ force field used to describe the protein systems.

Protein systems were modelled using the standard AMBER (FF14SB)³⁶ force field present in the Amber 14 package. To generate topologies of the protein systems and to ensure neutralization of the respective proteins system prior to the production run, the addition of hydrogen atoms and counterion were implemented by the LEaP module of Amber 14. The protein system was then immersed in orthorhombic TIP3P³⁵ water box to model the water molecules explicitly, at a distance of 10Å from all protein atoms. The particle mesh Ewald (PME) method³⁷ in the Amber 14 package was incorporated to compute the long-range

electrostatics interactions in the molecular dynamic simulation with a van der Waals cut of distance of 10Å. Systems were subjected to consecutive minimization steps, the initial partial minimization followed by full minimization.

An initial minimization of 2000 steps was carried out with an applied restraint potential of 500 kcal/mol. This was followed by a full minimization of 1000 steps carried out by conjugate gradient algorithm without any restraints. All systems were gradually heated from 0 to 300K for 50ps in the canonical ensembles (NVT), by the application of harmonic restraints of 10 kcal/mol Å and collision frequency of 1.0p. s⁻¹ to all solutes in the system. The simulation temperature was controlled using the Langevin thermostat. All systems were equilibrated at 300k in an NPT ensemble for 500ps with no restraints and a constant pressure (1 bar) was maintained using the Berendsen thermostat. The SHAKE algorithm was used to constrict the bonds of all hydrogen atoms. Continuous MD was performed on all systems for 50 ns in an NPT ensemble with a constant pressure of 1 bar, constant temperature of 300k and a pressure coupling constant of 2ps.

Trajectories were saved every 1ps and further analysed (for example, RMSD, RMSF, distance (d1) and dihedral angle Φ) using the CPPTRAJ³⁸ module incorporated Amber 14. As an additional validation of the continuous MD approach utilised in the current study, a multiple MD approach was performed for one system. In which three 50 ns MD runs with different initial velocities were performed and the average trajectory analysed. The trends for both approaches are similar, confirming that the reported approach is reliable and valid. The graphical interface of UCSF chimera³⁹ was used to visualise all enzyme structures, and data was analysed and plotted using the GUI of Microcal Origin data analysis software version 6.5²⁹.

4.4.4 Parameters used in this study

The present study utilised parameters previously proposed by our group to more accurately describe flap dynamics for aspartic proteases, which focused on Plms I-V. This study is a continuation of our work but primarily of Plm IX and X. These parameters are represented schematically in Figure 4.2.

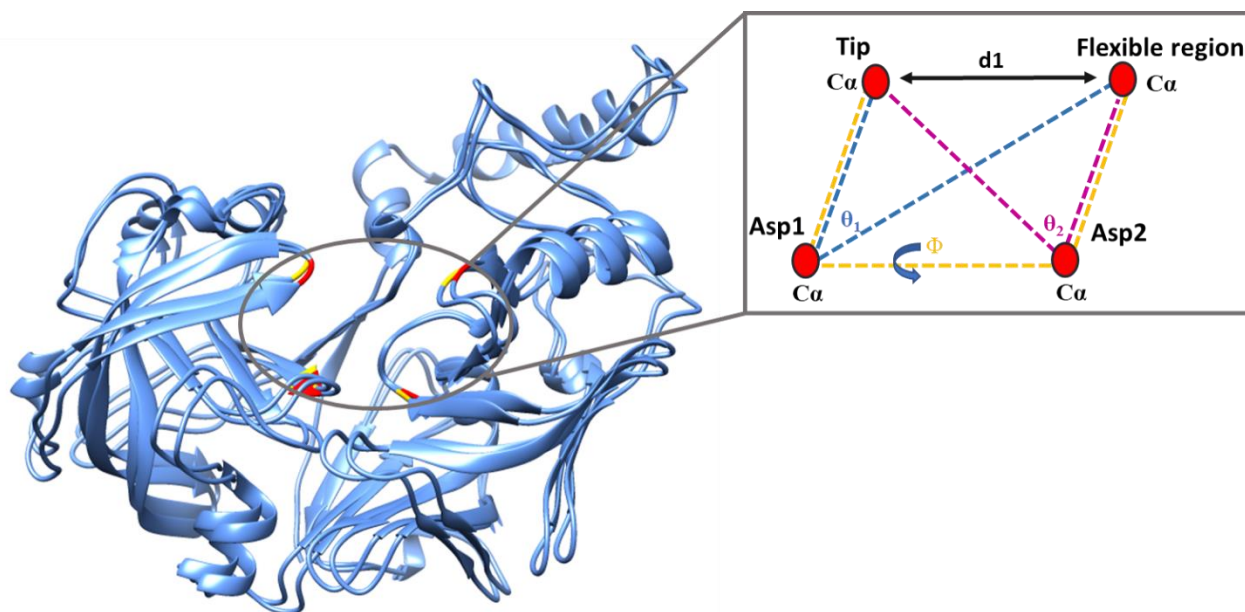


Figure 4.2 Schematic representation of the parameters used to define the flap-structure motion: $d1$ the distance between the flap tip and flexible region hinge residue, the dihedral angle Φ and the TriCa angles, θ_1 and θ_2 . Asp1, Asp2, the flap tip and flexible region of PIX and PX (prepared by author)

4.5 Results and discussion

4.5.1 Exclusive Structural Features of Plasmepsins

The active site of the *P. falciparum* Plm comprises of two catalytic aspartic acids, covered partially by a β -hairpin flap, a feature commonly shared with Plm IX and X⁴⁰. Numerous studies have shown that residues lining the active site such as the flap and flexible loop undergo large conformational changes instrumental for ligand binding. It has been reported that the flap and flexible domain of Plm I-V display a “twisting” dynamic motion which can be observed during the opening and closing of the active site region.

In this study, we could highlight this distinct feature shared in Plm IX and X from the visual analysis of the snapshots presented in Figure 4.3 of Plm IX and X. The apo tertiary structures transition between an open and semi-open conformation throughout the simulation. Present in all orthologs of Plm’s are the regions surrounding the catalytic active sites, while there is a distinctive structural difference between aspartic protease, predominantly in the loop regions at the top and bottom of the model. Plm I-X exhibit high structural similarity, as can be seen from the superimposed structures of Plm I, II, III, IV, VI, VII, IX and X (Figure S4.3), as opposed to their sequence composition and display an overall conservation of the active site (Figure S4.4). Although there is a deviation in the sequential composition of Plm’s, there is an overall conservation of the position of the catalytic aspartic residues and hydrophobic residues proximal to each binding cleft of Plm IX and X to other plasmepsins⁴¹.

Plm's are often associated with notable substrate specificity. This specificity is linked to the sequence heterogeneity of residues lining the active site, such as the flap and flexible loop domains. To date, there are no approved inhibitors against Plm. This paves the way towards the design of novel inhibitors that are more resilient against resistance. For Plm's, d1 accurately describes the opening and closing of the flap-structure, and the dihedral angle θ can be used to gain insight into the twisting motion of the flap-structure.

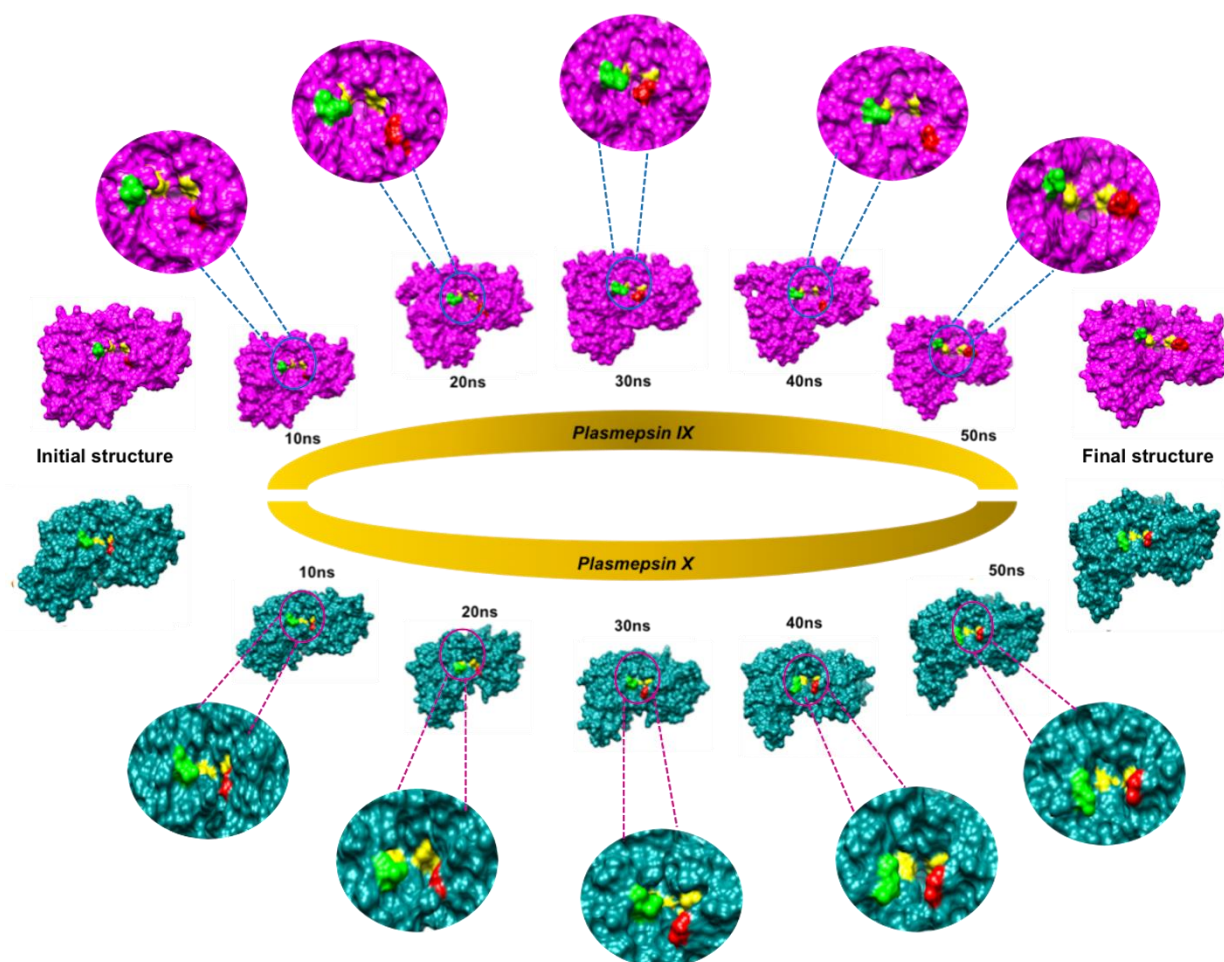


Figure 4.3 Flap and flexible loop region movements of apo Plm IX (magenta) and X (cyan) throughout a 50 ns molecular dynamics simulation. The catalytic aspartate residues of Plm IX and X are highlighted in yellow, the flap tip in red and hinge residue in green.

4.5.2 Mechanistic Flexibility of Plm IX and X

Root mean square deviation (RMSD) analysis projects information that can be used to measure the stability and disorder or stereochemical variability across a set of structural models, with the RMSD being complementary to the crystallographic B factor. Analysis of RMSD were performed to validate the stability of the apo systems of Plm IX and X (Figure S4.5), and to

ensure accuracy of successive post-dynamic analyses. The potential energy (PE) was also computed across the 50 ns MD simulation of Plm IX and X to further validate the convergence of each system. From the PE plots presented in Figure S4.6, the convergence of both systems was achieved within the 50 ns MD trajectory, which corresponds with the stability attained at 30 ns of the RMSD trajectory of Plm IX and X. The average PE for Plm X was substantially higher than that of Plm IX which were -158461.37 kcal/mol and -163213.16 kcal/mol respectively. The average PE correlate with the average RMSF values for both system, further accentuating a greater degree of stability in Plm IX as opposed to Plm X. This finding coincides with the “twisting motion” being more prominently exhibited by IX as opposed to X. The residues used to calculate the distance (d1), dihedral angle (Φ), Tri C α θ 1 and θ 2 can be found in Table 4.1.

Table 4.1 Residues used to calculate the distance (d1), dihedral (Φ), TriC α θ 1 and θ 2

	Plasmepsin IX	<i>Plasmepsin X</i>
Distance (d1)	Gly124-Glu331	Gly97-Glu305
Dihedral (Φ)	Gly124-Asp80-Asp259-Glu331	Gly97-Asp53-Asp232-Glu305
TriC α θ 1	Gly124-Asp80-Glu331	Gly97-Asp53-Glu305
TriC α θ 2	Glu331-Asp259-Gly124	Glu305-Asp232-Gly97

4.5.3 Plasmepsin Residue Fluctuations

Structural comparison of models Plm IX and X to well characterized aspartic proteases, observed no structural abnormalities. For quantifying the flexibility, Root Mean Square Fluctuation (RMSF) calculations were utilized.

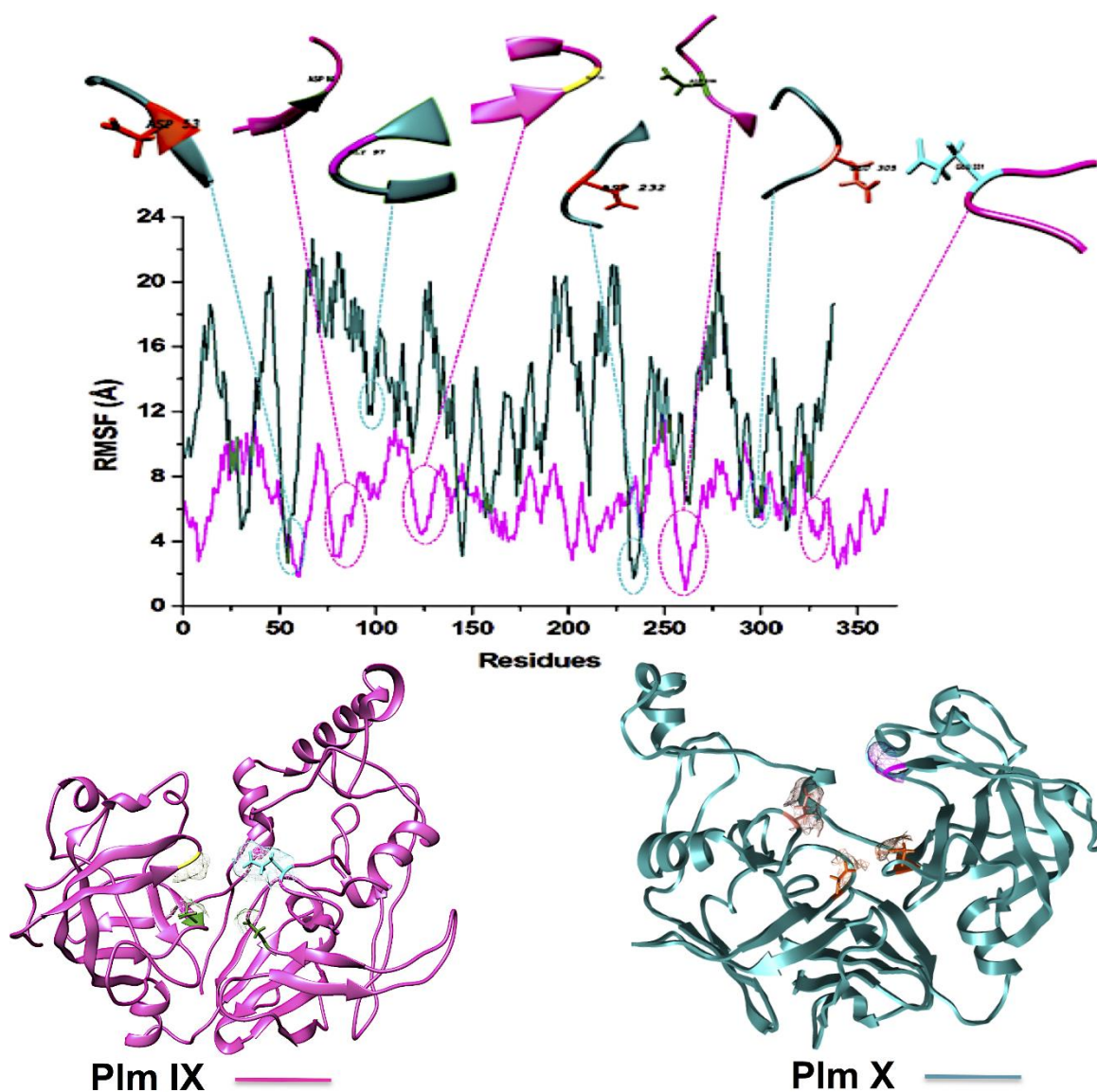


Figure 4.4 C- α RMSF plot of *Plasmodium falciparum* Plm IX (purple), X (cyan), respectively. Although both models displayed erratic behaviour in comparison to Plm I-V, Plm IX and X display a similar trend in the fluctuation observed in key residues. The average fluctuation of the of Plm IX and X were relatively higher than the fluctuation observed in the key residues, (Table 4.2). Displaying lower fluctuation of key residues in relation to the average fluctuation observed in both systems.

Table 4.2 RMSF Values of Apo Plm IX and X for the Binding Site Residues, Tip and Hinge Residue, and Total Average Fluctuation measured in Angstroms (Å)

	<i>PLASMEPSIN IX</i>	<i>PLASMEPSIN X</i>
ASP 1	0.6178	0.7249
ASP 2	0.6878	0.7213
FLAP TIP	3.6074	2.4348
HINGE	1.2962	1.4321
AVERAGE	6.44	12.63

It was interesting to note that the average values of fluctuation observed in the key residues of Plm IX and X (Table 4.2) follow a similar trend to those observed for Plm's I-V, displaying slightly greater flexibility in the aspartic residues, flap tip and hinge region as opposed to other Plm family members⁴². Plm IX and X displayed greater flexibility particularly in Asp 2 and the hinge residue as opposed to Asp 1 and the flap tip, which coincides with the similar observation made in Plm I-V. The increased flexibility interjected by Asp 2 and the hinge residue may be linked to the “twisting” characteristic of the opening and closing of the active site to enhance and stabilize ligand binding. Although there is a general trend of similarity shared by Plm I-V and Plm IX-X, the average fluctuation observed in Plm IX and X, was substantially higher than that observed in Plm I-V. This may be attributed to the use of homology models as opposed to crystal structures which display a greater level of stability.

In the absence of X-ray crystal structures, *in silico* tools are often implemented to assist structure-based drug design and eliminate occurrences of drug resistance. Some of the setbacks associated with the formation of protein crystals are the presence of highly flexible and disordered regions, and this bottleneck cannot be easily overcome even with the use of “predictive” tools.

Although Plm IX and X display a high level of structural similarity to Plm I-V, both Plm IX and X encompass highly flexible and disordered regions¹⁹ as can be observed in Figure 4.4 and Table 4.2. Even in the presence of these highly flexible and disordered regions, the key residues making up the active site of Plm IX and X display levels of flexibility relatable to other Plm's, with no drastic deviation in flexibility for these regions

Proteins destined for export contain an N-terminal signal sequence (R_xL_xE/Q/D) known as the *Plasmodium* export element (PEXEL)⁴³. *P. falciparum* Plm IX is one of the Plm that contains a PEXEL motif in its sequence (amino acid position 45-50), which marks the protein for export into the host red cell. This motif proves essential for protein export into the host red blood cell. When compared to Plm's I-VIII, *P. falciparum* Plm's V, IX and X also have comparatively large low complexity insert regions (LCR). Plm X contains 4 LCR's which is a greater quantity of LCR regions in comparison to other Plm's. The LCR contain non-random, limited alphabet amino acids. and are generally located in the solvent exposed hydrophilic loops of proteins. Although Plm IX present in all species of *Plasmodium* that contain LCR's, there is an increased number of these regions in the proteins encoded by species that infect humans⁴¹. All these structural discreteness attributes to the elevated levels of fluctuation observed in both Plm IX and X as observed in Figure 4.4.

4.5.4 Distance [d1] and Dihedral Angle [Φ] Analysis

Although the opening and semi-closed conformation observed in Plm's IX and X, is clearly visible in Figure 4.5. The distance (d1) between the flap tip and flexible loop referred to as the hinge residue is pivotal for accurately describing the opening and closing of the flap-structure, with the dihedral angle (Φ) being calculated to gain insight into the twisting motion of the flap structure.

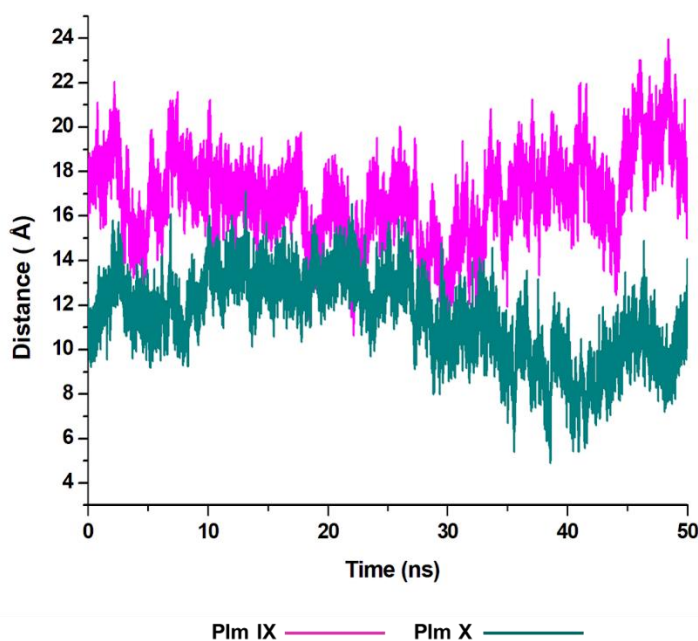


Figure 4.5 Graph showing the fluctuation in the distance, d1 as presented in Figure 4.2.

From the analysis of the distance (d1) between the flap tip and hinge residue of the flexible loop we could visually see the opening and closing conformations of Plm IX and X throughout the simulation.

With Plm IX exhibiting an open conformation from 10-16 ns, 35-50 ns, curled conformation at 18-34ns, and semi-closed conformation at 30-34 ns. Whereas, Plm X exhibiting an open formation at 12-14 ns, curled conformation at 23-28 ns and closed conformations at 5-10ns and 39-48 ns. At 30 ns we observe a shift in the dynamics of the flap tip and hinge residue of Plm X, as these regions move closer together we observe a significant decrease in the distance between the flap tip and hinge residue. The Plm X system maintains this distance and dihedral angle, with minimal fluctuation being observed while the system moves towards a more closed conformation. The meaningful change observed in the distance are presented in Table 4.3, with the opening and closing motion being more clearly prominent in Plm IX as compared to Plm X.

Table 4.3 Measurement of the distance by which the flap-structure moves, measured in Angstroms (Å).

	PLASMEPSIN IX	PLASMEPSIN X
AVERAGE	27.3671	20.5052
MINIMUM	13.5465	10.8648
MAXIMUM	21.27215	16.12234
Δ^A	13.8207	9.6404

^A Change between maximum and minimum distance

There is a large sense of ambiguity surrounding Plm IX and X, thus this study validates the flap-structure motion of an opening and closing conformation, of both enzymes as presented by other aspartic proteases. An ideal inhibitor may target the flap tip and hinge residues comprising of two glycine residues present in both enzymes and two negatively glutamate residues present in Plm IX and X respectively. Previous studies have highlighted the Gly-Gly omega bond in the glycine rich flap, which undergoes a cis-trans isomerization, at incremental time scales initiating the flap opening⁴⁴. The presence of these glycine residues makes the curling motion of the flap region possible. Targeting this specific domain may aid in the design

of potent structure-based inhibitors against plasmepsins such as allosteric inhibitors targeting flap pockets, hindering the opening and closing of the binding cavity.

From the analysis of the dihedral angles of Plm IX and X we can distinctly see the twisting motion observed from the graphs presented in Figure 4.6 respectively. The dihedral angle Φ , which is a change from negative/positive and or positive/negative, we observe the twisting motion more distinctly displayed in Plm IX in comparison to Plm X which coincides with the increase in d1. At open conformations, where the distance is at its maximum which is at 42-50 ns the angle changes from -10° to 20.25° , another twist is observed at 30-35ns the angle changes from -19° to 12.5° projecting a twisting motion of the flap region. In Plm X we observe a major twist observed between 23-28ns, the angle transcends from 9.5° to 35° , corresponding to the curled conformation of the flap tip and hinge residue. The dihedral angle of the twisting motion presented in Figure 4.6 is much more defined in Plm IX, as opposed to Plm X and this can also be correlated to the average distance in Plm IX, which is $21.27\text{\AA}$ as compared to that of Plm X, which is only $16.12\text{\AA}$.

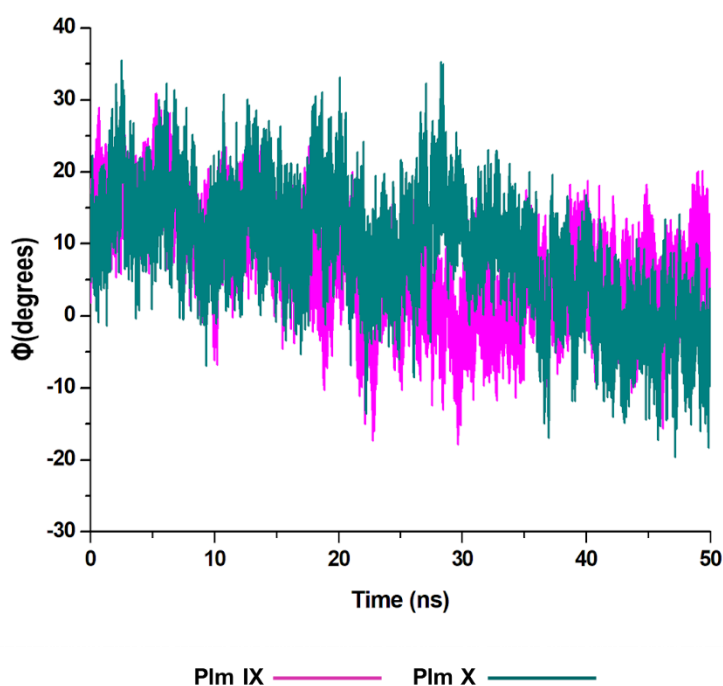


Figure 4.6 : Graph of the dihedral angle throughout the 50 ns simulation, of Plm IX (blue) and X (red).

4.5.5 TricCa angles θ_1 and θ_2

In this present study, we can see θ_1 and θ_2 follow similar trends to each other as observed in Figure 4.7 and 4.8. The trend observed in θ_1 and θ_2 corresponds to the fluctuations observed

in d1 for Plm IX, with the maximum θ_1 and θ_2 values coincide with opening of the flap structure, as the flap and flexible loop move away from each other exposing the active site. The minimum θ_1 and θ_2 values correlate with a more closed flap structure for Plm X as the flap folds inwards toward the active and the flexible loop recoils closing the active site.

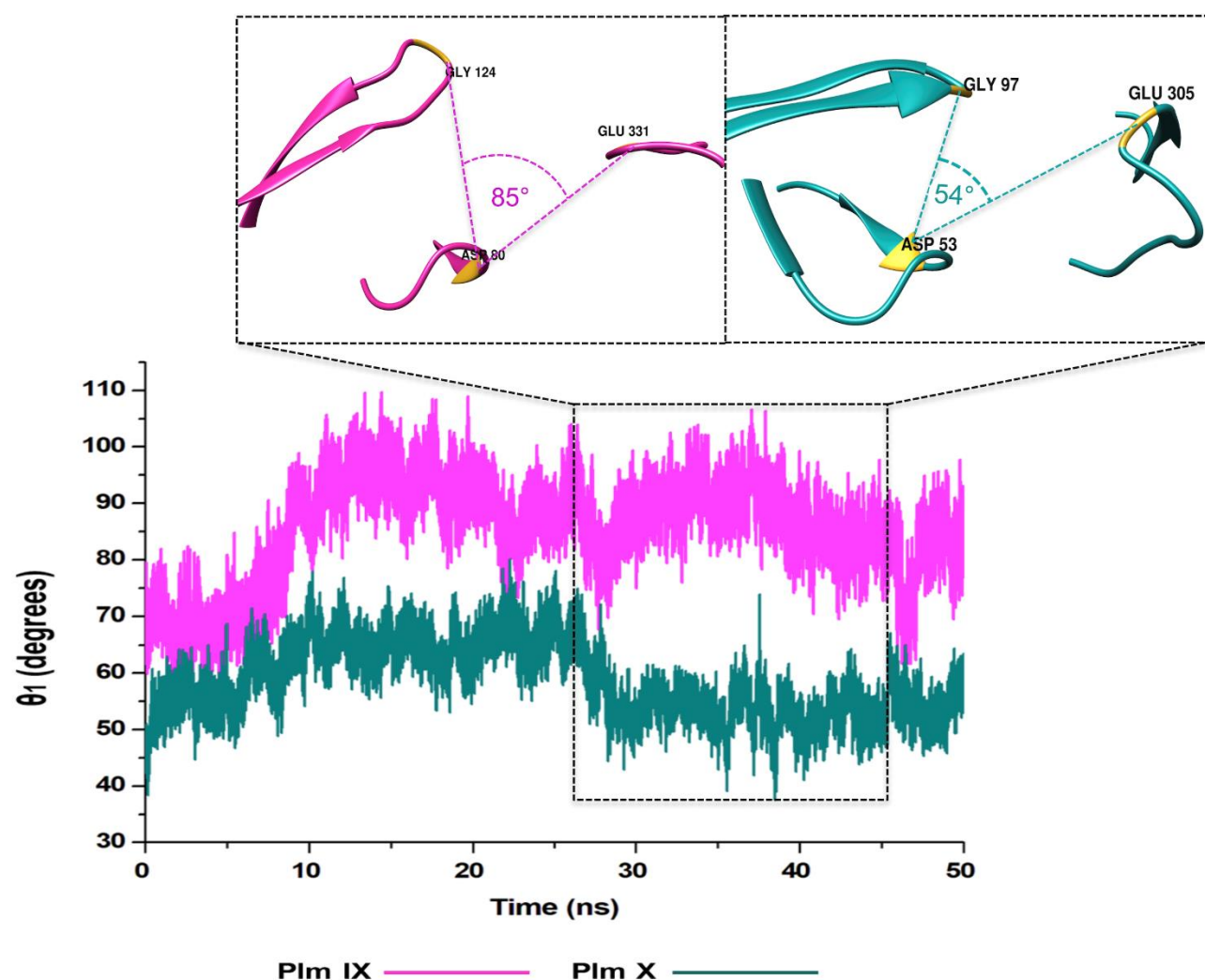


Figure 4.7 Graph showing the Tri C α angles θ_1 of Plm IX (pink) and X (cyan), respectively

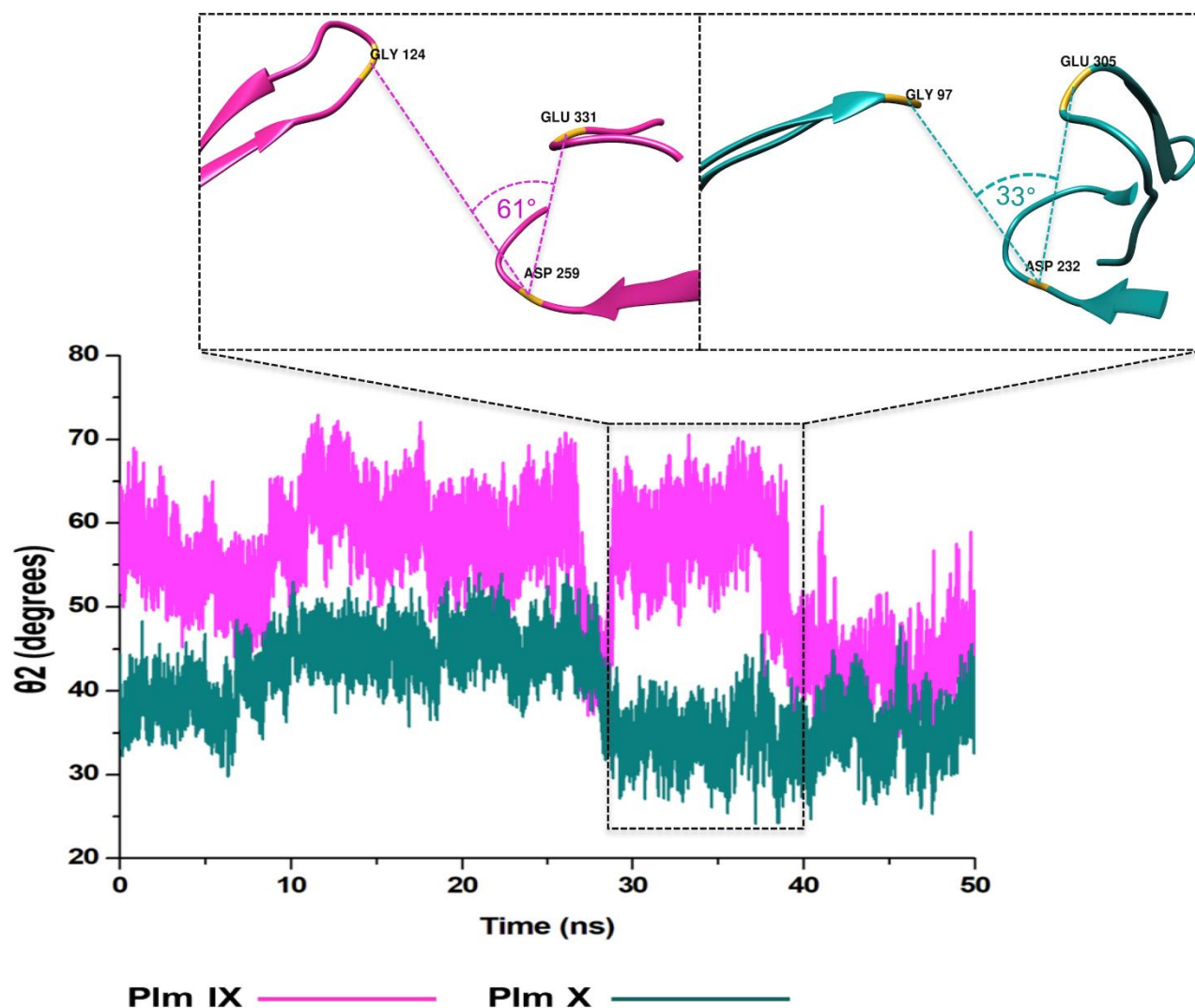


Figure 4.8 Graph showing the Tri C α angles θ_2 of Plm IX (pink) and X (blue), accordingly

Overall, as the flap-structure transitions into more open conformations (increase in d1), both θ_1 and θ_2 increase, whereas as the flap-structure moves toward a more closed conformation (decrease in d1), θ_1 and θ_2 decrease consequently. However, there is distinctive observation in Plm IX, there is an observed decrease in θ_2 as opposed to that of θ_1 between 40-50 ns, where the system maintains an open conformation. This decrease correlates to the conformational change which may be due to “closure-inducing” ligand binding. The selected residues making up θ_2 may actively aid in stabilization of the ligand once it enters the catalytic site in the open conformation. This unique further illustrates the structural versatility of Plm IX and its importance in drug design. Similarly, a very distinct observation can be made for Plm X as we observe an increase in both the θ_1 and θ_2 angles from 10-30ns, where the flap and hinge region remain in distal proximity, but in a compact conformation covering the aspartic catalytic

domain. At 30ns, there is an abrupt decline in the angles from 50° to 27°, where the flap tip and hinge region are near each other observed throughout the 50 ns simulation and reaches a state of stabilization.

4.5.6 Radius of Gyration (ROG)

The radius of gyration is indicative to the compactness of the tertiary structure of a protein that is, how folded or unfolded a protein is, and gives insight into the stability of biological molecules during the MD simulation⁴⁵. The Rg for Plm IX and X, were investigated in this study, displaying a similar trend as observed for that of d1 and 02. The Rg of Plm IX exhibited fluctuation, as the flap and flexible region moved between an open and closed conformation. The increase in the fluctuation observed from 40ns, of the Plm IX is where the system reaches a state of semi-stabilization in an open conformation. We can see from Figure 4.9 at 40 ns the flap tip and hinge residue move further apart, and maintains an open conformation at 50 ns, which coincides with an increase in the distance between the flap tip and flexible loop. The observations made for Plm X is like the observations made for d1 and 02, we see a decrease in the Rg at 30 ns where the flap and flexibles residues are at its closest distance to each other. After 30 ns the system experiences gradual fluctuation as the flap-structure remains in a semi-closed conformation.

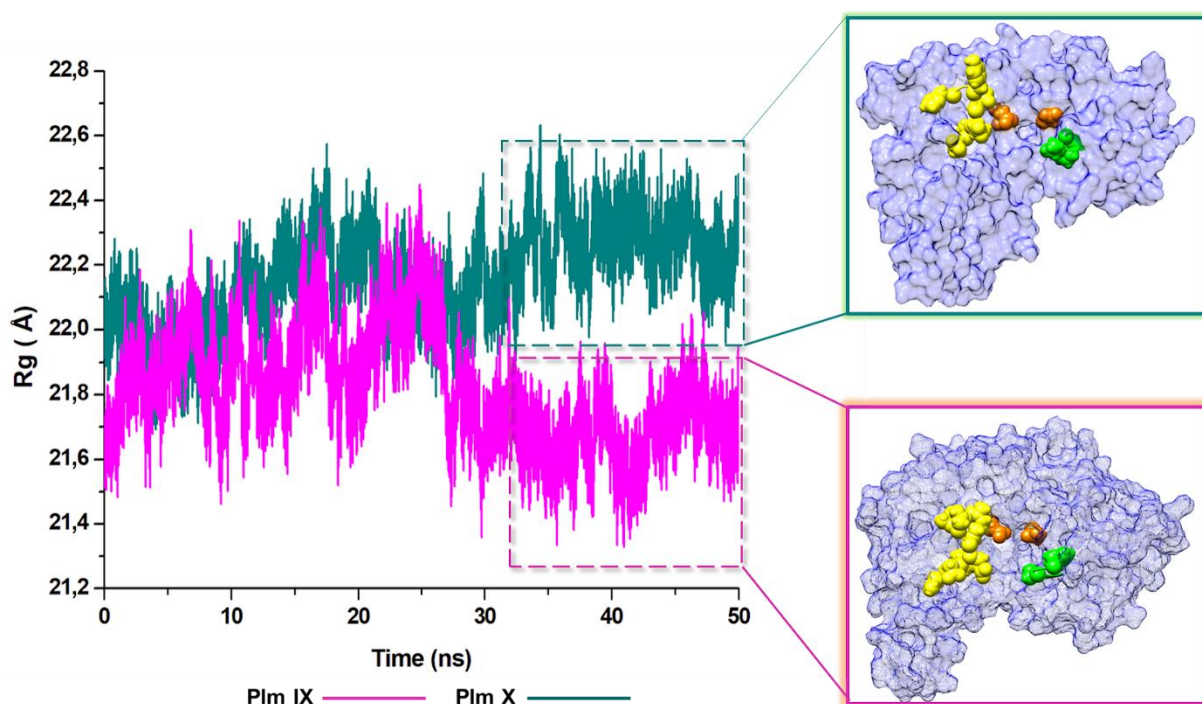


Figure 4.9 Radius of Gyration (ROG) of Plm IX (mustard yellow) and Plm X (bottle green).

4.6 Conclusions

Only a handful of proteases have been discovered and characterized prior to the completion of genome sequencing for *P. falciparum*. There are currently 6301 published crystal structures of aspartic proteases from experimental data (X-ray/NMR analysis)⁴⁶, amidst there are still no crystal structures of Plm IX and X. This study aimed to pave the way towards the design of potent inhibitors of Plm IX and X.

It is not coincidental that glycine mediates the role of the flap tip, observed in both Plm IX and X, as the amino acid glycine possesses enhanced conformational flexibility upon folding to form part of the α -helix⁴⁷. Flexibility and mobility of the flap tip is essential, as it lies perpendicular over the catalytic site, mediating the binding of inhibitors and or substrates. The presence of the glycine rich flap of Plm IX (Phe-Gly-Thr-Gly) and Plm X (Phe-Gly-Ser-Gly) are highly mobile and flexible and transitions between semi-open and open conformations in the absence of a ligand. The role undertaken by the flap tip regarding the opening and closing of the catalytic domain remains key in generating the twisting motion.

The hinge residue presents in the flexible loop of Plm IX and X, are represented by glutamate. Glutamate carries a hydrophilic acidic group with a strong negative charge. The presence of this negatively charged amino acid positioned towards the outer surface of a protein promotes water solubility and stability of a protein. This validates the reduced fluctuation observed in the hinge residue of Plm IX and X as compared to that of the flap tip. Information extracted from the Tri C α angles and θ_2 , we observe an increase in fluctuation during the open conformation of the models of Plm X, and a decrease in Plm IX. This decrease may aid in the effective binding a ligand during the open conformation, further implementing stability. This fluctuation may be stabilized primarily in the flap and flexible regions when bound to an inhibitor as observed in Plm II⁴⁸.

From the average distance of Plm IX and X being greater than 15 Å respectively, illustrates the flexibility and ability to accommodate inhibitors of variable size¹⁷. This flexibility is governed by the N-terminal flap and flexible loop. As Plm IX and X are structurally similar we do observe a trend in the transition between the open and closed conformation, however the transition between these two states of Plm IX and X are experienced at independent time intervals. Thus, displaying the activity of these aspartic proteases are not predictable in their

conformational flexibility and dynamic behaviour. Even though the overall RMSF fluctuation of Plm X is greater than Plm IX, there is greater fluctuation in the flap tip of Plm IX than Plm X of 3.6074 Å and 2.4348 Å accordingly which corresponds with the twisting motion which is more prominently exhibited by Plm IX than X.

The results obtained in the present study, have signified the essential role undertaken by the flap dynamics and provides new insight into the relevant conformational changes of the flexibility of Plm IX and X. Which will ultimately aid in the design and development of novel antimalarial drugs.

Conflicts of interest

There are no conflicts to declare

Acknowledgements

The authors acknowledge the National Research Foundation for their financial support and the Centre for High Performance Computing (<http://www.chpc.ac.za>) for their computational resources.

4.7 References

- 1 *World malaria report 2017*, Geneva, 2017.
- 2 D. T. Jamison, R. G. Feachem, M. W. Makgoba, E. R. Bos, F. K. Baingana, K. J. Hofman and K. O. Rogo, *Disease and Mortality in Sub-Saharan Africa*, The World Bank, Washington,DC, Second Edi., 2006.
- 3 I. N. Nkumama, W. P. O. Meara and F. H. A. Osier, Changes in Malaria Epidemiology in Africa and New Challenges for Elimination, *Trends Parasitol.*, 2017, **33**, 128–140.
- 4 *World malaria report*, 2017.
- 5 Who, Global Technical Strategy for Malaria 2016-2030, <http://www.tbfacts.org/tb-statistics/>.
- 6 V. Torres, New Targets for Drug Discovery against Malaria, *PLoS One*, , DOI:10.1371/journal.pone.0059968.
- 7 A. O. Ruiz, M. Postigo, S. G. Casanova, D. Cuadrado, J. M. Bautista, J. M. Rubio, M. L. Oroz and M. Linares, *Plasmodium* species differentiation by non - expert on - line volunteers for remote malaria field diagnosis, *Malar. J.*, 2018, **17**, 1–10.
- 8 A. Gnémé, W. M. Guelbéogo, M. M. Riehle, A. B. Tiono, A. Diarra, G. B. Kabré, N. Sagnon and K. D. Vernick, *Plasmodium* species occurrence , temporal distribution and interaction in a child-aged population in rural Burkina Faso, *Malarial J.*, 2013, **12**, 1–9.
- 9 S. Charpian, J. M. Przyborski, K. V. F. Strasse and J. M. Przyborski, Protein Transport Across the Parasitophorous Vacuole of *Plasmodium falciparum* : Into the Great Wide Open, *Traffic*, 2008, **9**, 157–165.
- 10 R. E. Farrow, J. Green, Z. Katsimitsoulia, W. R. Taylor, A. A. Holder and J. E. Molloy, Seminars in Cell & Developmental Biology The mechanism of erythrocyte invasion by the malarial parasite , *Plasmodium falciparum*, *Semin. Cell Dev. Biol.*, 2011, **22**, 953–960.
- 11 A. Adisa, M. Rug, N. Klonis, M. Foley, A. F. Cowman and L. Tilley, The Signal Sequence of Exported Protein-1 Directs the Green Fluorescent Protein to the Parasitophorous Vacuole of Transfected Malaria Parasites *, *J. Biol. Chem.*, 2003, **278**, 6532–6542.

- 12 D. M. Walker, S. Oghumu, G. Gupta, B. S. McGwire, M. E. Drew and A. R. Satoskar, Mechanisms of cellular invasion by intracellular parasites, *Cell. Mol. Life Sci.*, 2015, **71**, 1245–1263.
- 13 H. Cai, R. Kuang, J. Gu, Y. Wang, S. Texas, E. Infectious, S. Antonio and S. Antonio, Proteases in Malaria Parasites - A Phylogenomic Perspective, *Curr. Genomics*, 2011, **12**, 417–427.
- 14 S. Subramanian, M. Hardt, Y. Choe, R. K. Niles, E. B. Johansen, J. Legac, J. Gut, I. D. Kerr, C. S. Craik and P. J. Rosenthal, Hemoglobin Cleavage Site-Specificity of the *Plasmodium falciparum* Cysteine Proteases Falcipain-2 and Falcipain-3, *PLoS One*, , DOI:10.1371/journal.pone.0005156.
- 15 R. K. Bedi, C. Patel, V. Mishra, H. Xiao and R. Y. Yada, Understanding the structural basis of substrate recognition by *Plasmodium falciparum* plasmepsin V to aid in the design of potent inhibitors, *Nat. Publ. Gr.*, 2016, 1–17.
- 16 C. Berry, New Targets for Antimalarial Therapy: The Plasmepsins, Malaria Parasite Aspartic Proteinases, *Biochem. Educ.*, 2000, **4412**, 191–194.
- 17 K. K. Roy, International Journal of Antimicrobial Agents Targeting the active sites of malarial proteases for antimalarial drug discovery: approaches , progress and challenges, *Int. J. Antimicrob. Agents*, 2018, **50**, 287–302.
- 18 J. Liu, I. Y. Gluzman, M. E. Drew and D. E. Goldberg, The Role of *Plasmodium falciparum* Food Vacuole Plasmepsins *, *J. Biol. Chem.*, 2005, **280**, 1432–1437.
- 19 D. N. Nair, V. Singh, D. Angira and V. Thiruvengatam, Proteomics & Bioinformatics Structural Investigation and In-silico Characterization of Plasmepsins from *Plasmodium falciparum*, *J. Proteomics Bioinforma.*, 2016, **9**, 181–195.
- 20 F. Li, V. Bounkeua, K. Pettersen and J. M. Vinetz, *Plasmodium falciparum* ookinete expression of plasmepsin VII and plasmepsin X, *Malar. J.*, 2016, **15**, 1–10.
- 21 P. Bhaumik, A. Gustchina and A. Wlodawer, Structural studies of vacuolar plasmepsins, *Biochim. Biophys. Acta*, 2013, **1824**, 207–223.
- 22 A. S. Nasamu, S. Glushakova, I. Russo, B. Vaupel, A. Oksman, A. S. Kim, D. H. Fremont, N. Tolia, J. R. Beck, M. J. Meyers, J. C. Niles, J. Zimmerberg and D. E. Goldberg, parasite egress and invasion, *Science (80-.)*, 2017, **522**, 518–522.

- 23 W. Karubiu, S. Bhakat, L. McGillewie and M. E. S. Soliman, Molecular BioSystems Flap dynamics of plasmepsin proteases : insight dynamics †, *Mol. Biosyst.*, 2015, **11**, 1061–1066.
- 24 L. McGillewie, M. Ramesh and M. E. Soliman, Sequence , Structural Analysis and Metrics to Define the Unique Dynamic Features of the Flap Regions Among Aspartic Proteases, *Protein J.*, 2017, **12**, 5–9.
- 25 P. Pino, R. Caldelari, B. Mukherjee, J. Vahokoski, N. Klages, B. Maco, C. R. Collins, M. J. Blackman, I. Kursula, V. Heussler, M. Brochet and D. Soldati-favre, for invasion and egress, *Science (80-.)*, 2017, **528**, 522–528.
- 26 E. Deu, Proteases as antimalarial targets : strategies for genetic , chemical , and therapeutic validation, *FEBS jJournal*, 2017, **284**, 2604–2628.
- 27 H. Zheng, K. B. Handing, M. D. Zimmerman, S. C. Almo, W. Minor, B. Physics and J. P. Avenue, X-ray crystallography over the past decade for novel drug discovery – where are we heading next?, *Expert Opin. Drug Discov.*, 2016, **10**, 975–989.
- 28 A. Wlodawer, W. Minor, Z. Dauter and M. Jaskolski, Protein crystallography for asprinign crystallographers or how to avoid pitfalls and traps in macromolecular strucyure determination., *Fed. Eur. Biochem. Soc.*, 2014, **280**, 5705–5736.
- 29 A. Laura A. Deschenes and David A. Vanden BoutUniversity of Texas, Origin 6.0: Scientific Data Analysis and Graphing Software Origin Lab Corporation (formerly Microcal Software, Inc.), *J. Am. Chem. Soc.*, 2000, **122**, 9567–9568.
- 30 Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information (NCBI).
- 31 L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass and M. J. E. Sternberg, The Phyre2 web portal for protein modelling , prediction and analysis, *Nat. Protoc.*, 2017, **10**, 845–858.
- 32 W. Wang, M. Xia, J. Chen, F. Deng and R. Yuan, Data in Brief Data set for phylogenetic tree and RAMPAGE Ramachandran plot analysis of SODs in *Gossypium raimondii* and *G . arboreum*, *Data Br.*, 2016, **9**, 345–348.
- 33 M. Wiederstein and M. J. Sippl, ProSA-web : interactive web service for the recognition of errors in three-dimensional structures of proteins, *Nucleic Acids Res.*, 2007, **35**, 407–

- 410.
- 34 D. A. Pearlman, D. A. Case, J. W. Caldwell, W. S. Ross, T. E. Cheatham, S. Debolt, D. Ferguson, G. Seibel and P. Kollman, AMBER , a package of computer programs for applying molecular mechanics , normal mode analysis , molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules, *Comput. Phys. Commun.*, 1995, **91**, 1–41.
 - 35 D. A. Case, T. E. Cheatham, T. O. M. Darden, H. Gohlke, R. A. Y. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang and R. J. Woods, The Amber Biomolecular Simulation Programs, *J. Comput. Chem.*, 2005, **26**, 1668–1688.
 - 36 G. Explicit, S. Particle, M. Ewald, R. Salomon-ferrer, A. W. Goetz, D. Poole, S. Le Grand and R. C. Walker, Routine microsecond molecular dynamics simulations with Routine microsecond molecular dynamics simulations with AMBER on GPUs . 2 . Explicit Solvent Particle Mesh Ewald, *J. Chem. Theory Comput.*, 2013, 1–38.
 - 37 U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, A smooth particle mesh Ewald method A smooth particle mesh Ewald method, *J. Chem. Phys.*, 1995, **8577**, 8577–8593.
 - 38 D. R. Roe and T. E. Cheatham, PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data, *J. Chem. Theory Comput.*, 2013, **9**, 3084–3095.
 - 39 T. Z. Desantis, P. Hugenholtz, N. Larsen, M. Rojas, E. L. Brodie, K. Keller, T. Huber, D. Dalevi, P. Hu and G. L. Andersen, Greengenes , a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB, *Appl. Environ. Microbiol.*, 2006, **72**, 5069–5072.
 - 40 W. R. P. Scott and C. A. Schiffer, Curling of Flap Tips in HIV-1 Protease as a Mechanism for Substrate Entry and Tolerance of Drug Resistance, *Elsevier Sci.*, 2000, **8**, 1259–1265.
 - 41 R. D. McGeorge, Griffith University, 2014.
 - 42 L. McGillewie and M. E. Soliman, Flap flexibility amongst I, II, III, IV, and V: Sequence, structural, and molecular dynamic analyses, *Proteins Struct. Funct. Genet.*, 2015, 1693–1705.

- 43 R. E. D. Cells, M. Rug, M. Cyrklaff, A. Mikkonen, L. Lemgruber, S. Kuelzer, C. P. Sanchez, J. Thompson, E. Hanssen, M. O. Neill, C. Langer, M. Lanzer, F. Frischknecht, A. G. Maier and A. F. Cowman, Regular Article Export of virulence proteins by malaria-infected erythrocytes involves remodeling of host actin cytoskeleton, *BLOOD J. Hematol.*, 2016, **124**, 3459–3469.
- 44 M. Mahanti and S. Bhakat, Flap Dynamics in Aspartic Proteases : A Computational Perspective, *Chem. Biol. Drug Des.*, 2016, 1–19.
- 45 R. Lonsdale, J. N. Harvey, A. J. Mulholland and R. Lonsdale, A practical guide to modelling enzyme-catalysed reactions, *Cell Press Struct.*, 2012, **41**, 3025–3038.
- 46 H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne, The Protein Data Bank, 2000, **28**, 235–242.
- 47 M. C. Deller and B. Rupp, IYCr crystallization series Protein stability: a crystallographer ' s perspective IYCr crystallization series, *Struct. Biol. Commun.*, 2016, **72**, 72–95.
- 48 O. A. Asojo, S. V Gulnik, E. Afonina, B. Yu, J. A. Ellman, T. S. Haque, A. M. Silva and B.-H. Hall, Novel Uncomplexed and Complexed Structures of plasmepsin II , an Aspartic Protease from *Plasmodium falciparum*, *J. Mol. Biol.*, 2003, **2836**, 173–181.

CHAPTER 5

A Dual-Target of plasmepsin IX and X: Unveiling the atomistic superiority of a core chemical scaffold in Malaria therapy

Geraldene Munsamy^a, Clement Agoni^a and Mahmoud E. S. Soliman^{a*}

^a Molecular Bio-computation and Drug Design Laboratory, School of Health Sciences,
University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa

*Corresponding Author: Mahmoud E.S. Soliman

Email: soliman@ukzn.ac.za

Telephone: +27 (0) 31 260 8048, Fax: +27 (0) 31 260 7872

5.1 Abstract

PIX and PX, members of the prominent aspartic family of proteases whose function where hitherto unknown have only recently been established as key mediators of erythrocyte invasion and egress of the virulent malarial parasite. Inhibitor 49c, a potent antimalarial peptidomimetic inhibitor initially developed to target PII has recently been proven to exhibit potent inhibitory activity against PIX and PX. However, the molecular and structural dynamics supporting its inhibitory activity remain inconclusive. Hindering the motion of the flap and hinge region of an aspartic protease remains essential for disabling the catalytic activity of the enzyme. Integrating molecular dynamic simulations coupled with other advanced bio-computational tools, we reveal the enhanced inhibitory competence of 49c against PIX and PX relative to Pepstatin. Pepstatin, a known aspartic protease inhibitor which actively hinders the opening and closing of the flap tip and flexible loop and consequently limits access to the catalytic aspartic residues, however its administration has been related to elevated levels of toxicity. The inhibitor 49c demonstrated a higher binding free energy to PIX and PX relative to Pepstatin. A relatively less compact and structurally rigid 49c bound complexes resulted in the restriction of flap and hinge residues from cohesive movement, consequently hindering their “twisting motion” from transpiring. Findings unveil the atomistic perspectives into the enhanced inhibitory activity of 49c against PIX and PX inhibitor relative to Pepstatin. This study may hasten the search towards the discovery of novel antimalarial drugs amid the current wave of drug resistance.

5.2 Introduction

The infiltrating tactics of the malarial parasite remains a global threat in the effective treatment of the mosquito borne-parasitic infection. Alarming concerns have been linked to the unrelenting spread of malaria, predisposing half of the world's population to its infection. In the year 2016(WHO, 2015), malaria was reported as one of the highest contributors to the escalated morbidity and mortality rate observed(Roy, 2018).

Malaria is caused by five *Plasmodium* species, namely, *Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium ovale*, *Plasmodium knowlesi* and *Plasmodium Malariae*. Of these, *P. falciparum* is considered to be the most lethal strain, emanating in a deadly form of malaria in humans(Nkumama, Meara, & Osier, 2017; Li, Bounkeua, Pettersen, & Vinetz, 2016). Amongst a vast range of therapeutics targeting malaria, artemisinin has been established as one of the most potent antimalarial inhibitors(Harvey, Edrada-Ebel, & Quinn, 2011; Srivastava, Singh, & Naik, 2010; Newman & Cragg, 2016). However growing concerns have been linked to the development of drug resistance associated with artemisinin treatment. Thus, prompting the search for novel therapeutic targets, proposing a more alluring route towards the effective treatment and alleviation of malaria.

Aspartic proteases remain reputable targets due to their association and significant involvement in a growing number of human diseases(David, Wolfender, & Dias, 2014; Coombs et al., 2001; Asojo et al., 2003) with malaria high up on the growing list. The aspartyl protease family of enzymes are commonly referred to as Plasmpesin's (Plm). *P. falciparum* is host to a family of 10 Plm's, of these family members Plm I-IV are transported to the food vacuole, where they orchestrate the degradation of haemoglobin in a sequential manner with each Plm undertaking a distinctive task in the process in the haemoglobin degradation process(Banerjee et al., 2002; Parr, Tanaka, Xiao, & Yada, 2008). Plm V is accountable for the cleavage of proteins that are to be transmitted to the host cell(Moura, Dame, & Fidock, 2009), whereas Plm VI-VIII are expressed in the vector during the parasite's intra-erythrocytic stages of sporozoite formation, motility as well midgut transversal and do not play a direct role in transmission of the parasite within the human host(Nair, Singh, Angira, & Thiruvengatam, 2016). The latter family members PIX and PX are highlighted as mediators of egress and invasion of the parasite further exposing the parasite at its most vulnerable state accentuating these enzymes as formidable targets(Cai et al., 2011; Berry, 2000; Dash, Kulkarni, Dunn, & Rao, 2003).

PIX and PX are implicated in erythrocyte invasion and are expressed in mature blood-stage schizonts and aggressive merozoites, however their function hereto remains unknown. There has been countless efforts to find a more “permanent remedy” against the aspartyl protease family of enzymes as many of the proposed inhibitors were abandoned due to the limited potential of their therapeutic pathway(Sayer & Louis, 2010). Previous studies showed that pepstatin A, a hexa-peptide containing the unusual amino acid statine as a potent aspartic protease inhibitor which interrupts *Plasmodium* transmission to mosquitoes. However, the pharmacology action of pepstatin remains unsettled, a further concern is its association with adverse side effects upon administration(Argiulo et al., 2002;Munkhkargal et al., 2012). Dominant focus has now been placed on the aspartic protease family to circumvent the development of drug resistance. Fortunately, re-examining the once abandoned inhibitors unveiled an inhibitor that displayed potency against two newly discovered targets.

Experimental studies have highlighted the inhibitory activity of an aspartyl inhibitor referred to as compound 49c which consists of a hydroxyl-ethyl-amine–based scaffold against PIX and PX. Figure 5.1 highlights the stages of egress and invasion inhibited by 49c. 49c is defined as a peptidomimetic competitive inhibitor, that actively displays potency against *P. falciparum* in vitro and the rodent parasite *Plasmodium berghei* in vivo studies. The inhibitory facet of 49c is correlated to the presence of hydroxyl-ethyl-amine–based scaffold which inhibits the preexocytosis processing of several secreted rhoptry and microneme proteins by targeting the corresponding maturases PIX and PX(Pino et al., 2017). Studies have further revealed that inhibitor 49c blocks invasion and egress of *P. falciparum* without affecting intraerythrocytic growth(Mukherjee et al., 2018).

Computer-aided drug design and discovery ensembles have played an instrumental role in the development of potential inhibitory molecules for over three decades. Data extracted from computational integrated calculations, such as molecular dynamics simulations have previously revealed limitations and possible suggestions towards improvisations of small molecules in explicit drug discovery(Lonsdale, Harvey, Mulholland, & Lonsdale, 2012).

Previous studies have reported molecular docking studies of PIX and PX in complex with 49c(Mukherjee et al., 2018), these studies however do not elaborate on the structural mechanism of inhibition of these plasmepsins. Therefore, in this study we aim to uncover the structural architecture in addition to the binding mode of PIX and PX in complex 49c. Utilising

a range of computational techniques to provide comprehensive structural and functional data that will aid in the identification or design of inhibitors specific to Plm IX and X.

The proposed 3D structures of Plm IX and X are presented in our group's previous work(Munsamy, Ramharack, & Soliman, 2018). This study aims to give molecular insight into the binding mechanism of Plm IX and X in complex with 49c.

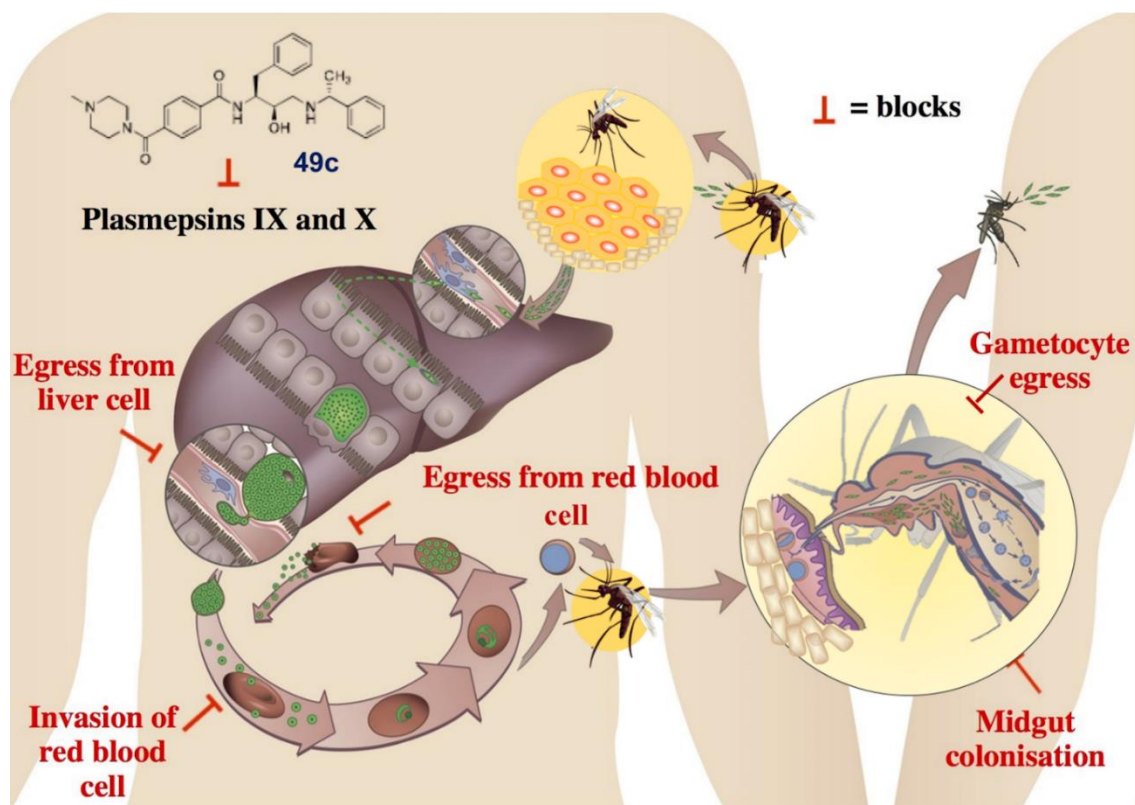


Figure 5.1 Schematic representation of inhibitor 49c blocking the various stages of egress and invasion of the *Plasmodium falciparum* parasite (prepared by author).

5.3 Computational Methods

5.3.1 Systems Preparation

In the absence of published crystal structures in-silico tools are often sought after for the generation of the 3D tertiary structure of a protein to aid in structure-based drug design. The proposed 3D structures of PIX and PX are presented in our group's previous work(Munsamy et al., 2018). UCSF Chimera(Pettersen et al., 2004) was used for visualisation of the 3D model.

The ligand 49c were generated with ChemDraw Ultra 9.0(Cousins, Kimberley R. (California State University, 2005). The 49c ligand was optimised incorporating the steepest descent method and MMFF94S force field in Avogadro(Hanwell et al., 2012). Softwares such as

Chimera as well as Gaussview 5.0(Hanwell et al., 2012) where implemented to ensure the correct angle bonds and hybridization were observable.

5.3.2 Molecular Docking

Molecular docking aims to provide the most energetically favourable binding pose as the ligand positions itself within the binding cavity of the target protein. Molecular docking was performed using AutoDock(Trott Oleg, 2011) a well-established docking program. Autodock Vina(Morris, Huey, & Lindstrom, 2009) was used to generate the calculations obtained from the docking scores. During the process of docking, Geister partial charges were added whilst Autodock atom types were defined using the Autodock Graphical user interface supplied by MGL tools(Honarparvar, Govender, Maguire, Soliman, & Kruger, 2014). The Lamarckian Genetic Algorithm, which is often associated with reliability and adequacy was applied to determine the docked conformations. Molecular visualization of the docked complexes was conducted using Chimera software. A previous study performed by Mukherjee *et al.*, (2018) highlighted the binding cleft of PIX and PX each in complex with 49c, even though the numbering of the active site residues do not correspond to those presented in our models there is a conformity amid the amino acid residues involved in the binding of 49c as can be visualised in **Figure S5.1**(Mukherjee et al., 2018). More detailed information regarding the parameters used for molecular docking as well as the pose selected for each ligand can be found in **Table S5.1**.

5.3.3 Molecular Dynamic Simulations

MD simulations were performed on six systems: Apo systems of PIX and PX, as well as PIX and PX each in complex with 49c, and pepstatin respectively. MD simulations were conducted using the GPU version of PMEMD incorporated in the Amber 14 package(Karubiu, Bhakat, McGillewie, & Soliman, 2015). Protein systems were modelled using the standard AMBER (FF14SB) force field present within the Amber package. The LEaP module of Amber 14 was employed to add hydrogen atoms and counter ions to aid in the stabilisation of the system. The Leap module of Amber 14 re-assigned the number of each amino acid from both PIX and PX starting from 1 in each system. Optimisation of ligands were performed by addition of partial atomic charges, which were calculated encompassing the restrained electrostatic potential (RESP)(Zhang & Lazim, 2017)(Kim, Blachly, Kaus, & Mccammon, 2015). Neutralisation was achieved by the addition of Na⁺ ions to all systems. Atomic partial charges comprising of

General Amber Force Field (GAFF) were prompted by the ANTECHAMBER module(Kim et al., 2015). Each system was solvated within the TIP3P waterbox consisting of a buffering distance of 8 Å amid water and the protein surface at box extremity. Long-range electrostatics interactions were performed adhering to the particle mesh Ewald (PME) method administered in the Amber 14 package with a van der Waals limitation of 10 Å distance.

Systems were subjected to consecutive partial minimization and full minimization steps. Initial energy minimization with a 500 kcal/mol Å² restraint potential related to the solute, were performed incorporating the steepest descent method for 1000 iterations. This was followed by 1000 iterations of conjugate gradient minimization. MD simulations were carried out for 10 ns during which the system was heated moderately between a range of 0-300K regulated with the aid of the Langevin thermostat(Paquet & Viktor, 2015). Systems were equilibrated at 300K under 1 atm pressure whilst retaining force constraints on the restrained solute for 500ps prior to production runs, followed by removal of restraints and maintenance of a constant pressure (1 bar) using a Berendsen barostat(Chetty & Soliman, 2015) All atoms covalently bound to a hydrogen atom were subjected to the SHAKE algorithm throughout the MD simulation(Explicit et al., 2013). From experimental studies, the preferred pH of the system was at 4.5 which is the optimum biological pH and was kept constant and validated in accordance with data projected from the H⁺⁺(Kim et al., 2015) tool which computes the pKa value of ionisable groups. All simulations were run at a 2-fs time step and the SPFP precision module. Trajectories were saved and analysed every 1 ps. The root mean square deviation (RMSD)(Zhang & Lazim, 2017), root mean square fluctuation (RMSF) various other post dynamic analysis was performed.

5.3.4 Binding Free Energy Calculations

Thermodynamic calculations assist in determining the binding free energy as an endpoint calculation which provide indispensable information regarding the interactions between a ligand and protein complex. The Molecular Mechanics/GB Surface Area method (MM/GBSA) method is computationally efficient and encompasses a versatile range of parameters for each energy term and is commonly utilized to calculate the binding free energies for macromolecules integrating continuum solvation models merging with molecular mechanics calculations(Wright, Hall, Kenway, Jha, & Coveney, 2014). The MM/GBSA(Ylilauri & Pentikäinen, 2013) method was employed to estimate the binding free energy of each of the systems,. Binding free energy was averaged over 20000 snapshots extracted from the 20ns trajectory.

The free binding energy (ΔG) computed by this method for each molecular species (complex, ligand and receptor) can be represented as:

$$\Delta G_{\text{bind}} = G_{\text{complex}} - G_{\text{receptor}} - G_{\text{ligand}} \quad (1)$$

$$\Delta G_{\text{bind}} = E_{\text{gas}} + G_{\text{sol}} - TS \quad (2)$$

$$E_{\text{gas}} = E_{\text{int}} + E_{\text{vdw}} + E_{\text{ele}} \quad (3)$$

$$G_{\text{sol}} = G_{\text{GB}} + G_{\text{SA}} \quad (4)$$

$$G_{\text{SA}} = \gamma \text{SASA} \quad (5)$$

The term E_{gas} denotes the gas-phase energy, which consist of the internal energy E_{int} ; Coulomb energy E_{ele} and the van der Waals energies E_{vdw} . The E_{gas} was directly estimated from the FF14SB force field terms. Solvation free energy, G_{sol} , was estimated from the energy contribution from the polar states, G_{GB} and non-polar states, G . The non-polar solvation energy, G_{SA} , was determined from the solvent accessible surface area (SASA), using a water probe radius of 1.4 Å, whereas the polar solvation, G_{GB} , contribution was estimated by solving the GB equation. S and T denote the total entropy of the solute and temperature respectively.

5.3.5 Dynamic cross correlation

By analysing the trajectories of the system, it is possible to calculate the dynamic correlation between all atoms within the molecule i.e. there organized movement (Arnold & Ornstein, 1997) (Mark & Gunsteren, 1995).

The time correlated motion (DCCM) of Ca atoms of each system during the 100ns simulation was calculated using the CPPTRAJ module from the AMBER 14 package (Mhlongo et al., 2015). Snapshots were taken at every 50 ps interval from the trajectory and a covariance matrix was made between residues i and j . The dynamic cross correlation matrices were assembled to represent cross correlated displacements of alpha atoms in trajectories throughout the respective systems. A cross-correlation coefficient was calculated for Ca atoms using eqn:

$$A_{ij} = \frac{\langle \Delta x_i \Delta x_j \rangle}{\sqrt{\langle \Delta x_i^2 \rangle} \sqrt{\langle \Delta x_j^2 \rangle}}$$

5.4 Results and Discussion

5.4.1 Assessment of structural stability of systems

Prior to MD analysis, the root mean square deviation (RMSD) of the Ca backbone and potential energy fluctuations of the trajectories were monitored during the 100ns molecular dynamic simulation to ensure stability within the systems and accuracy of successive post-dynamic analyses. Convergence and stabilization in the RMSD plots were observed for all systems as shown in Supplementary Information **Figure S5.2** and **S5.3**.

5.4.2 49c impedes residue fluctuation in Plm IX and X

Proteins are fascinating molecular machines capable of organizing themselves into well defined hierarchical structures through a huge number of conformational changes, in order to accomplish a wide range of cellular physiological functions (Deriu, Grasso, Tuszynski, & Gallo, 2016). In exploring this property of proteins, root mean square fluctuation plots for all systems were calculated to reveal the structural impact of 49c on the the flexibility of individual amino acids of PIX and PX relative to the inhibitory activity of the known aspartic protease inhibitor, Pepstatin. The RMSF plots of all six systems are presented in Figure 5.2 and

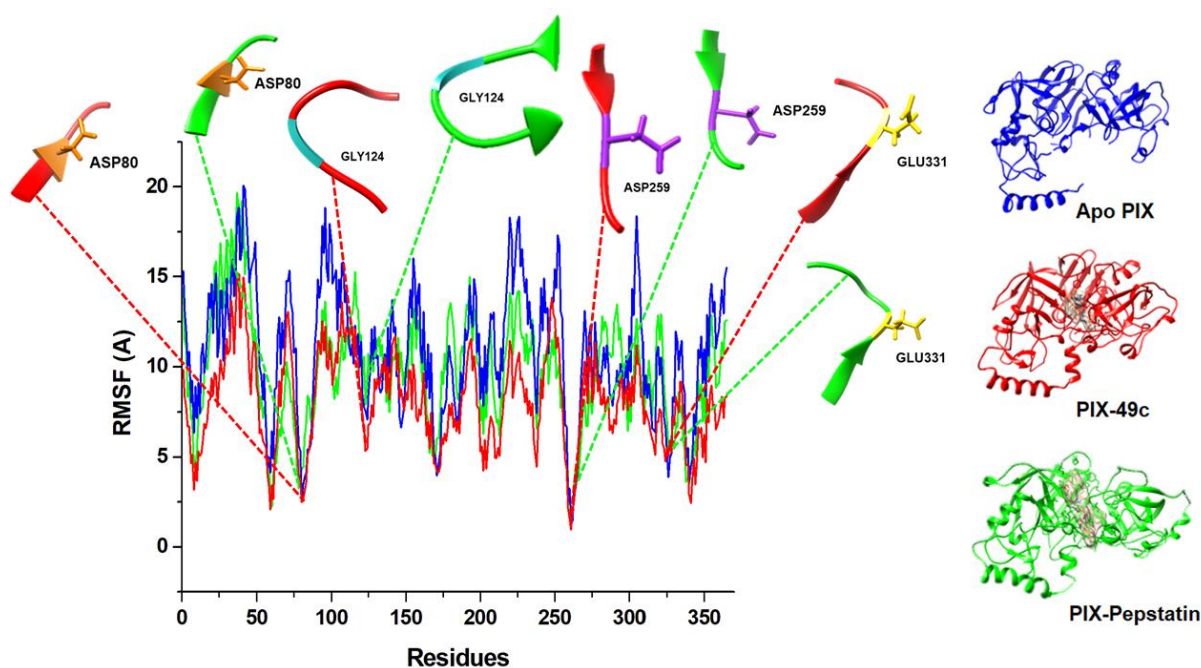


Figure 5.2 Root mean square fluctuation of the apo systems of PIX (blue), PIX in complex with 49c (red) and Pepstatin (green)

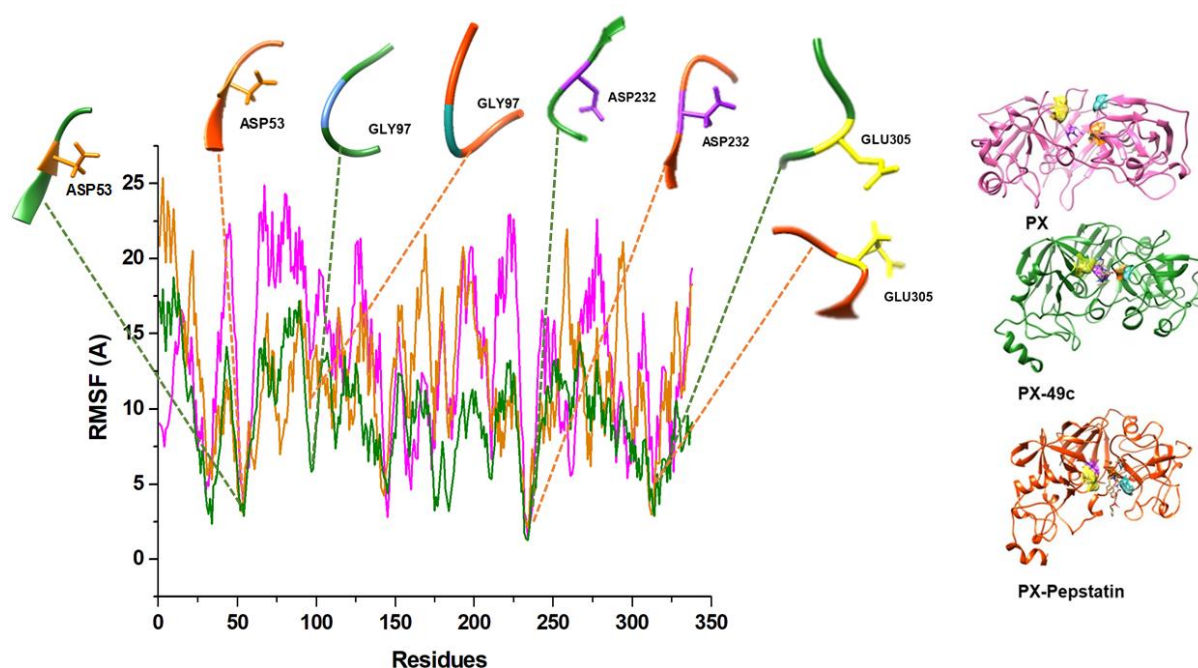


Figure 5.3 Root mean square fluctuation of the apo systems of PX (pink), PX in complex with 49c (olive green) and Pepstatin(orange).

Table 5.1 RMSF values of apo Plm IX and X, as well as each Plm in complex with 49c and Pepstatin, highlighting the fluctuation observed in the binding site residues, tip, hinge residue, and total average fluctuation measured in Angstroms (Å)

	<i>PIX</i>	<i>PX</i>	<i>PIX-49c</i>	<i>PX-49c</i>	<i>PIX-Pepstatin</i>	<i>PX-Pepstatin</i>
<i>Asp1</i>	0.1859	0.6881	0.5987	0.6828	0.5687	0.6189
<i>Asp2</i>	0.7572	0.7623	0.6576	0.6 715	0.6038	0.7695
<i>Flap Tip</i>	2.2294	3.8269	1.1161	1.4709	1.6181	1.6101
<i>Hinge</i>	1.9319	1.9875	1.3001	1.4011	1.6222	1.7999
<i>Average</i>	9.7447	13.12505	10.9075	11.8452	8.0502	9.3781

Unbound PIX and PX showcased the highest overall fluctuation, an indication of a more flexible system. The flexible nature of the unbound plasmepsins allows for their notable “twisting motion” to occur (Scott & Schiffer, 2000; Sores, Torres, Manuela, & Pascutti, 2016; Karubiu et al., 2015). This increase in fluctuation may be linked to the opening and closing motion exhibited by the flap tip and hinge region. However, upon binding of 49c and Pepstatin, observed fluctuation were relatively lower than the unbound systems. This relatively lower fluctuation suggests that the inhibitory activity of 49c and Pepstatin restricted the movement of individual residues, particularly the glycine flap tip and the glutamate hinge residues, subsequently preventing the “twisting” motion of each aspartic protease from transpiring. A comparison between RMSF plots of 49c and Pepstatin showed that the structural conformation of 49c within the active of each aspartic protease induced a decline in the fluctuation and flexibility of key catalytic residues suggestive of structural constraint in the motion of each aspartic protease.

5.4.3 Inhibitory impact of 49c on the flap motion of Plm IX and X

Considering the crucial relevance of flap motion in PIX and PX “twisting motion”, we employed distance calculations, $d1$, to unravel the structural influence of 49c on the flap motion of each aspartic protease. The distance analysis was performed by integrating the parameters employed by our group as mentioned in our previous publications (McGillwie, Ramesh, & Soliman, 2017; Karubiu et al., 2015; McGillwie & Soliman, 2015), the distance between the $C\alpha$ atom of the flap tip and the $C\alpha$ atom of the opposite hinge residue in the flexible loop were calculated. For Plasmepsin's, $d1$ accurately describes the opening and closing of the flap-structure.

Previous studies have highlighted the pivotal information that can be found by measuring the distance between the flap tip and hinge residue in analysing the “twisting motion”, which resembles the opening and closing of the aspartic active site (McGillwie et al., 2017). The greatest difference observed between the maximum and minimum distance among all systems is more pronounced in the apo systems of PIX and PX as can be observed in Table 5.2, this may be attributed to the opening and closing motion being more defined in the absence of

inhibitors . From the distance graphs presented of all six systems in Figure 5.4 and 5.5, the fluctuation observed in the apo systems of PIX and PX, it is clearly defined and further resembles the opening and closing of each aspartic protease.

Table 5.2 Measurement of the distance by which the flap-structure moves, detected in all PIX and PX systems measured in Angstroms (Å).

	<i>PIX</i>	<i>PX</i>	<i>PIX-49c</i>	<i>PX-49c</i>	<i>PIX_Pepstat</i> <i>in</i>	<i>PX_Pepstatin</i>
<i>Maximum</i>	22.423	21.1702	19.2517	20.7971	23.8411	21.4097
<i>Minimum</i>	9.4465	4.8958	9.554	11.5152	11.9507	8.2442
<i>Average</i>	16.2821	12.76035	15.9094	15.2199	19.75625	15.84595
<i>Δ^a</i>	12.9765	16.2744	9.6977	9.2819	11.8904	13.1655

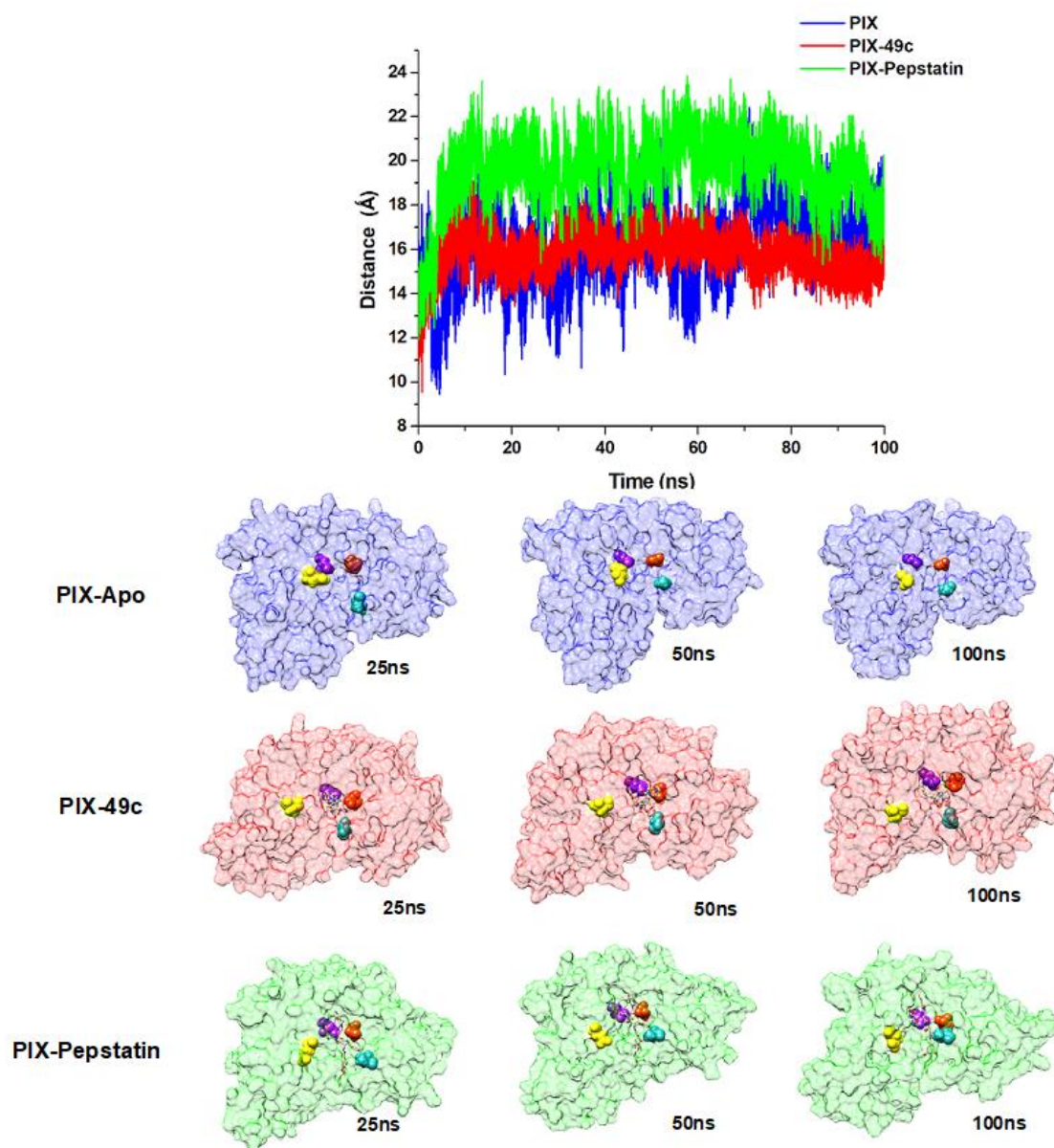


Figure 5.4 Graphical plots and visual representation of the Distance analysis between the flap tip and hinge residue throughout the 100ns MD simulation of apo PIX (blue), PIX-49c (red) and PIX-Pepstatin (green).

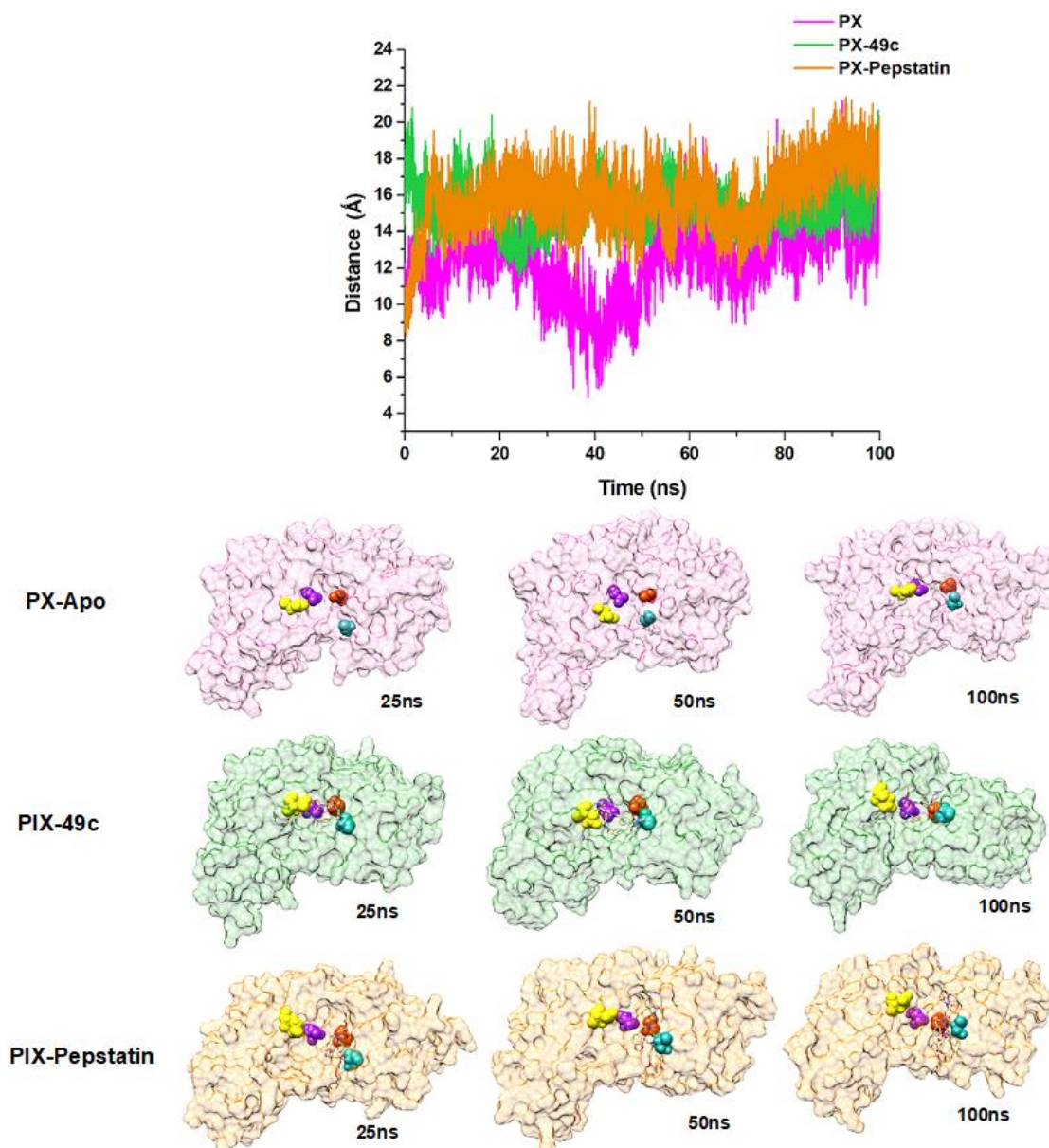


Figure 5.5 Graphical plots and visual representation of the Distance analysis between the flap tip and hinge residue throughout the 100ns MD simulation of apo PX (pink), PX-49c (olive) and PX-Pepstatin (orange).

The structural conformation of 49c and Pepstatin PIX and PX within the active site endorsed stability and actively hindered the opening and closing of the flap tip and flexible loop, limiting access to the catalytic aspartic residues. Overall, the difference between the maximum and minimum distance throughout the simulation display that systems in which inhibitors were present are shortened upon substrate binding indicative of the flap and hinge residues being drawn towards each other resulting in partial closing of this domain. From the visual observation there was an overall increase in the flexibility of the flap tip and hinge residue of the apo PIX and PX systems relative to the systems in complex with 49c and Pepstatin, which

exhibited little or no flexibility of the flap tip and hinge region over the simulation period. This finding coincides with the experimental data which illustrates inhibitory ability of 49c and Pepstatin as reported. However, relative to Pepstatin inhibited systems, the flap tip and hinge residue of the 49c complexed system showcased a greater decline in flexibility, suggesting a pronounced restriction of the “twisting motion” a typical motion of aspartic proteases.

5.4.4 Exploration of structural compactness of PIX and PX upon binding of 49c

The radius of gyration (ROG) of a protein is a measure of its compactness(Mhlongo & Soliman, 2015). If a protein is stably folded, it will likely maintain a relatively steady value of radius of gyration(Grasso et al., 2017). From the ROG of each system presented in Figure 5.6, it was observed that greater flexibility and loose structural conformation occurred in the in the apo systems of PIX and PX conferring with overall structural flexibility observed in the RMSF calculation for same systems. This prominent structural flexibility and decreased compactness observed throughout the simulation may be attributed to the typical “twisting motion” which is exhibited by aspartic proteases. Although, the binding of 49c and Pepstatin on PIX and PX induced a similar trend of minimal flexibility and structural compactness throughout the 100ns simulation with minimal fluctuation being observed during the beginning and at the end of the simulation. This observed minimal flexibility and structural compactness is indicative of a compact and stable system with restricted structural motions. 49c and Pepstatin prevented the flap and hinge residue from forming a cohesive movement and subsequently orchestrating the “twisting motion” from occurring as can be observed in the distance and RMSF calculations. Overall, 49c maintained prominence in its structural superiority on PIX and PX over Pepstatin as it maintained a steady level in the radius of gyration across each system.

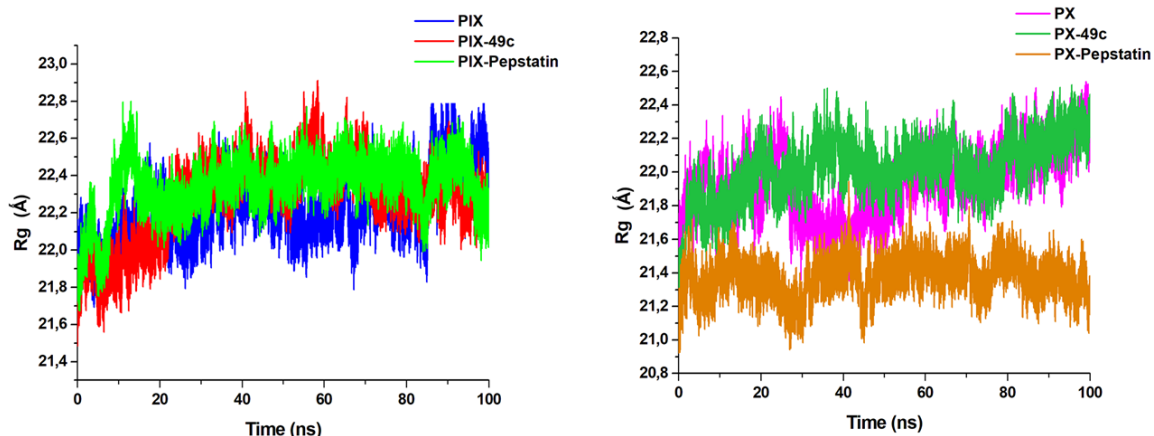


Figure 5.6 Radius of gyration (ROG) of apo PIX, PIX in complex with inhibitors 49c and Pepstatin on the left and apo PX, PX in complex with 49c and Pepstatin on the right

5.4.5 Analogous binding characteristics of 49c on PIX and PX

Using the MM/GBSA based approach, the binding free energy of 49c and Pepstatin on PIX and PX was estimated over the 100ns MD simulation as displayed in **Table 5.3**. Inhibitors 49c and Pepstatin revealed a binding free energy of -53kcal/mol and -47 kcal/mol respectively in complex with PIX. Whilst in aspartic protease PX, 49c and Pepstatin showed a binding free energy of -65.1kcal/mol and -38.5kcal/mol respectively. Taken together 49c exhibited a higher binding free energy and thermodynamic stability on PIX and PX across the simulation period relative to Pepstatin. This further suggests that favourable enthalpic contribution observed by the binding free energy is a consequence of the “tightening” of both PIX and PX around the active site induced by the structural conformation of inhibitor 49c. This further reaffirms the structural superiority and activity of 49c in complex with PIX and PX. 49c distinguished binding affinity correlates with literature as it has been reported to selectively and potently inhibit PIX and PX, aspartic proteases which remain crucial in egress and evasion of *P. falciparum* (Pino et al., 2017; Nasamu et al., 2017). The similarity observed in binding free energy of 49c in complex with PIX and PX may also suggest an analogous binding mode in the respective binding pockets of PIX and PX which further confers with reports by Mukherjee *et al.*, (2018) in which docking of 49c in modelled structures of PIX and PX showed comparable binding modes. The analogous binding modes of 49c may also contribute to its selective inhibitory activity against PIX and PX.

Table 5.2 MM/GBSA binding free energy profiles of 49c and Pepstatin interactions with PIX and PX

Complexes	ΔE_{vdw} (kcal/mol)	ΔE_{ele} (kcal/mol)	ΔG_{gas} (kcal/mol)	ΔG_{sol} (kcal/mol)	ΔG_{bind} (kcal/mol)
PIX_49c	-51.1 $\pm$ 0.1	-359.9 $\pm$ 0.4	- 410.1 $\pm$ 0.4	357.3 $\pm$ 0.3	-53.6 $\pm$ 0.1
PX_49c	-80.1 $\pm$ 0.1	-28.9 $\pm$ 0.2	- 109.1 $\pm$ 0.2	44.0 $\pm$ 0.2	-65.1 $\pm$ 0.1
PIX_pepstatin	-43.5 $\pm$ 0.4	-10.9 $\pm$ 0.1	- 112.5 $\pm$ 0.4	34.9 $\pm$ 0.1	-47.4 $\pm$ 0.4
PX_pepstatin	-35.2 $\pm$ 0.1	-326.4 $\pm$ 0.6	- 361.6 $\pm$ 0.6	323.1 $\pm$ 0.5	-38.5 $\pm$ 0.1

ΔE_{ele} = electrostatic energy; ΔE_{vdw} = van der Waals energy; ΔG_{bind} = total binding free energy; ΔG_{sol} = solvation free energy ΔG = gas phase free energy.

To further understand energy contributions of individual residues located at the binding site to overall total binding free energy, per residue decomposition analysis was performed. A depiction of the individual energy contributions and a map of the interactions between the inhibitor 49c and active site residues of PIX and PX are illustrated in Figure 5.7 and 5.8. This gave insights on binding pocket amino acid residues which are crucial to the binding 49c and Pepstatin. The primary interactions observed in all bound systems were dominated by van der Waals and hydrogen-bonding interactions, which are often associated with dominant negative signs for the enthalpy and entropy of the association (Crunkhorn, 2017) which correlates with the results obtained from this study.

As shown Figure 5.7 and 5.8 and Table 5.3. Ile338, Glu332, Thr262, Asp259, Gly124, Asp80 and Asp59 contributed the most towards total binding of 49c to PIX whilst in PX, the highest contributing residues to the overall binding energy included Glu305, Asp232, Thr235, Asp52 and Asp53. A critical exploration of specific interaction that existed between 49c and binding site residues revealed interaction in both PIX and PX inhibited system included conventional hydrogen bond interaction, pi-anion, and amide pi-stacked interactions as well as pi-alkyl interactions as shown in Figure 5.7 and 5.8 the conventional hydrogen bond interaction and pi-anion interaction formed between Asp259 and the hydroxy-ethylamine scaffold of 49c reaffirms its prominence in total energy contributions as shown in Figure 5.7 of the PIX-49c system. This interaction supplements a dynamic enthalpic contribution towards the binding free energy observed between PIX in complex with 49c. In a similar trend the conventional hydrogen bond interaction and pi-anion interactions occurring between Asp232 and Thr235 of

PX in complex with 49c also concurs with their prominence in total binding energy contribution as shown in **Figure 8**.

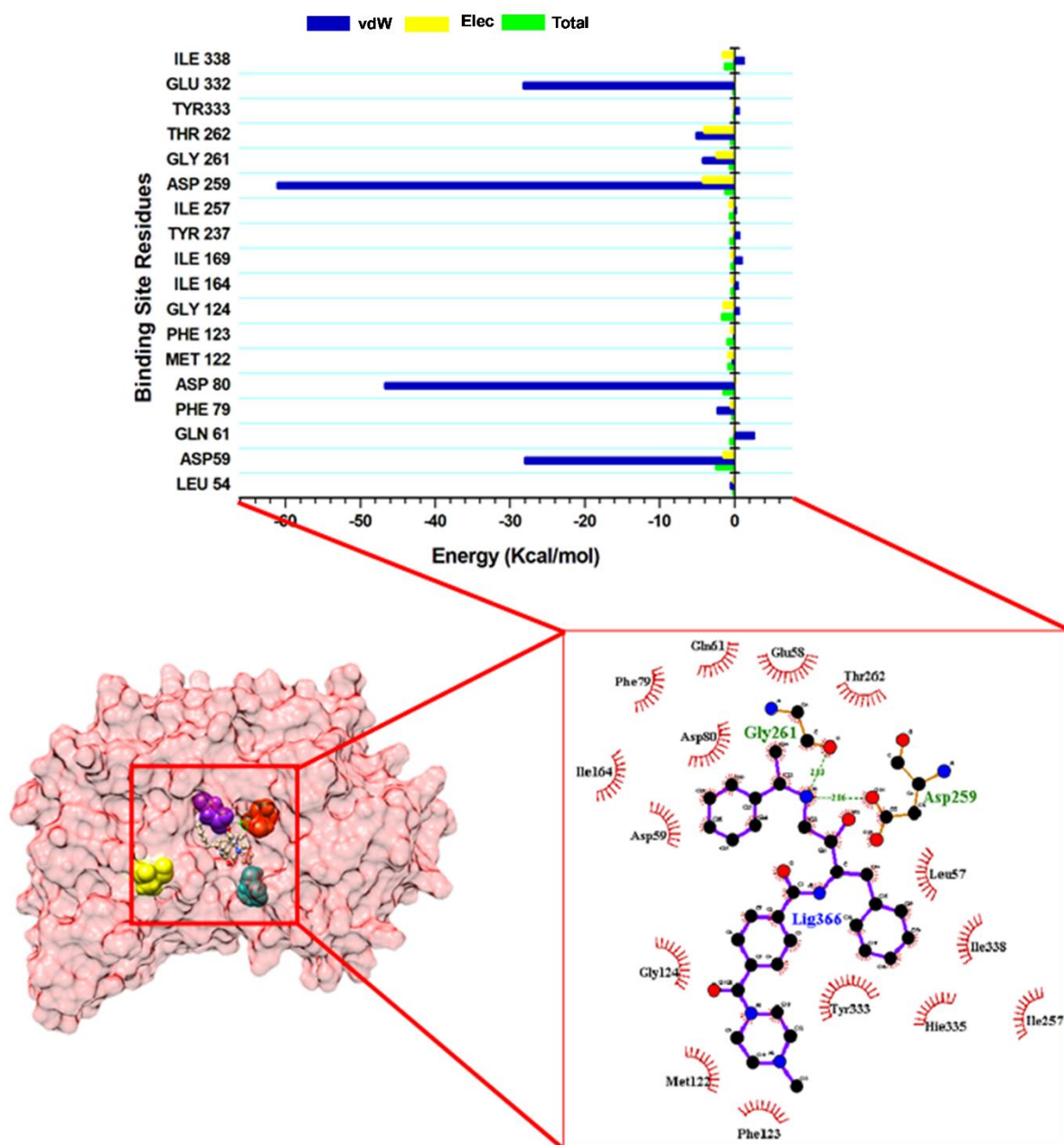


Figure 5.7 Molecular visualization of 49c at the catalytic sites (hydrophobic pockets) of Plm IX. Accentuating the Inter-molecular interactions between 49c and catalytic site residues in PIX.

Key residues that contributed towards the binding of Pepstatin in complex with PIX included Ile338, Thr262, Gly261, Thr237, Ile164, Glu56, Leu54, Leu57 and Ile78 while in PX Asp53, Arg31, Ile133, Phe96, Ile137, Phe139 and leu27 exhibited favourable binding interactions. As shown in **Figure S5.4** notable interactions in the Pepstatin inhibited proteins were also conventional hydrogen bond interaction, pi-anion, and amide pi-stacked interactions and pi-alkyl interactions.

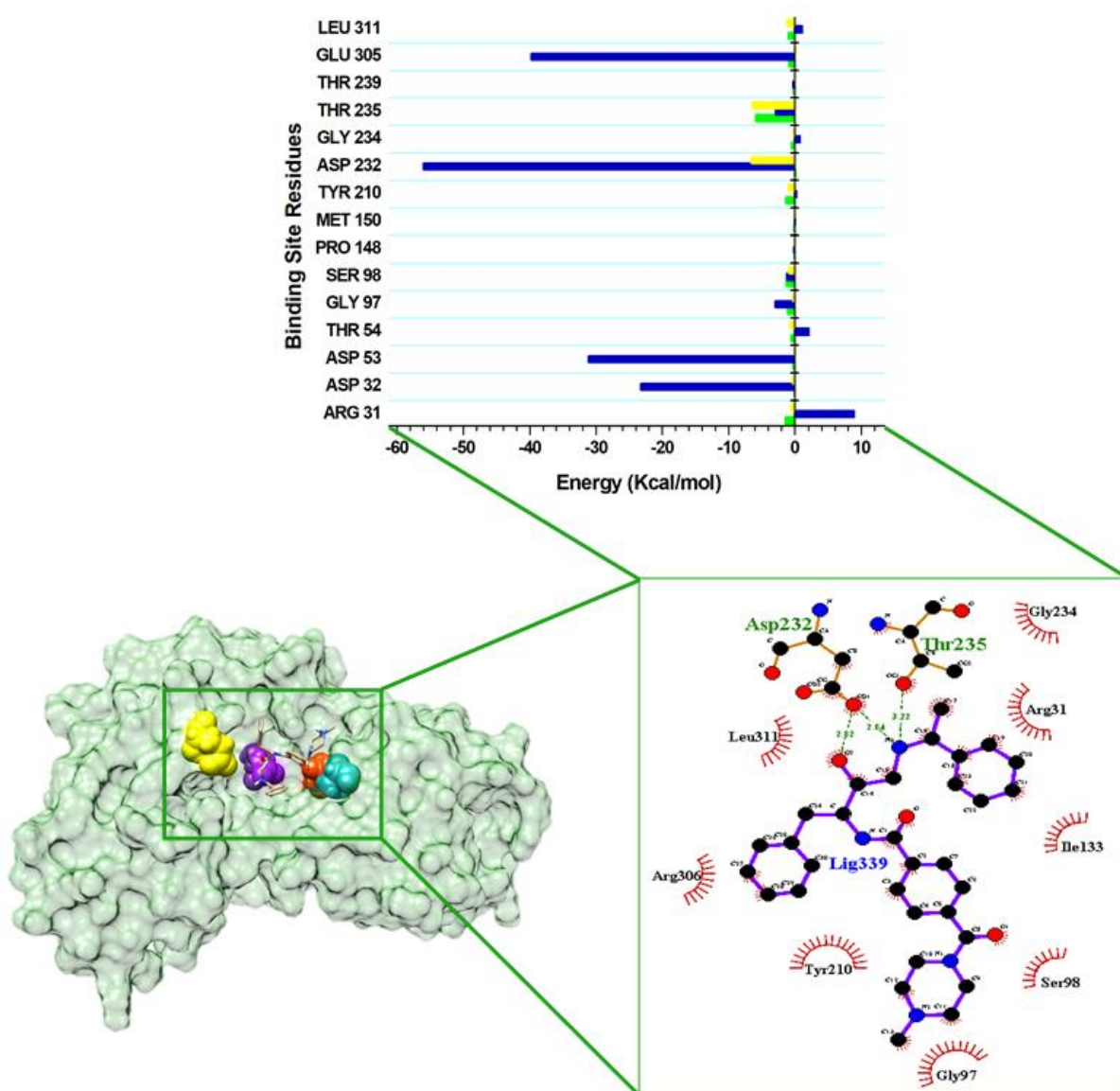


Figure 5.8 Molecular visualization of 49c at the catalytic sites (hydrophobic pockets) of PX. Highlighting the inter-molecular interactions between the catalytic site residues and 49c.

To further explore the structural dynamics that occurred upon binding of 49c to PIX and PX, dynamic cross correlation (DCC) analysis, a computational approach was employed to quantify

the correlation coefficients of motions between respective atoms based on C α atoms by analysing their backbone fluctuations and domain motions (Luo & Bruce, 2002). Positively correlated motions ranged from the colour yellow to deep red (+1). Anti-correlated motions which showed negative coloration ranged from cyan in colour to black (-1). As shown in **Figure S5.5** there was an overall anticorrelated movement of residues in PIX over the simulation period in the presence of 49c and Pepstatin rendering structural displacement within each aspartic protease. A closer look at residues between region “150 -350”, primarily made up of binding site residues revealed an anti-correlated motion with greater blue coloration in the 49c-PIX system relative to the same region in the Pepstatin in complex with PIX. This may be suggestive of notable inhibitory attributes of 49c in comparison to Pepstatin. This confers with the higher estimated binding energy observed in both PIX and PX in complex with 49c.

In an analogous manner, there was a notable overall anticorrelated motion of residues in PX upon binding of both 49c and Pepstatin particularly between region “200-300” which comprised predominantly of binding site residues. The relatively higher blue coloration in this region in 49c bound PX represents a higher anticorrelated motion of these residues in comparison with the apo and the PX-Pepstatin bound system. Although majority of the DCCM ‘s is made up of anti-correlated regions in the bound systems, there is also a larger presence of correlated regions in the bound systems as opposed to the apo systems which may imply enhanced conformational flexibility upon substrate binding which is further influenced by the mode in which the substrate binds on PIX and PX.

5.5 Conclusion

Malaria remains a global threat, with continuous efforts being made to find a more “permanent cure” to this invasive disease. Aspartic proteases have been identified as key mediators of malaria, recent studies have highlighted two reputable targets as being essential for the egress and invasion of this disease. Thus, this study aimed to pave the way towards the design of potent inhibitors of PIX and PX. The molecular dynamic analyses described in this study reveals the conformational evolutions of newly discovered inhibitor 49c in complex with PIX and PX. Analysis of structural dynamic revealed a relatively lower fluctuation of individual residues in all bound systems relative to the unbound systems, with 49c inhibited systems showing the lowest fluctuation. The apo systems of PIX and PX displayed greater structural versatility in comparison to systems in the presence of 49c, this is further correlated by the

DCCM calculations of the unbound systems in comparison to the bound. The “Twisting motion” resembling the opening and closing of PIX and PX could be explicated observed in the apo systems as opposed to those in the presence of inhibitors 49c and Pepstatin. Exhibiting stability conferred to each aspartic protease in the presence of an inhibitor which prevented the inherent movement of the flap and hinge domain. Although Pepstatin is associated with adverse side effects, it still displayed favourable binding free energy and effectually restrained the opening and closing of the flap and hinge residues. However, its molecular properties were superseded by 49c which displayed enhanced structural conformational, greater binding free energy in comparison to the counterpart drug, coinciding with experimental results and further highlighting the inhibitory potency of 49c. The information extended in this study maybe tailored for the development of novel inhibitors against the aspartic protease family of proteins implicated in a range of diseases

Conflicts of interest

There are no conflicts to declare

Acknowledgements

The authors acknowledge the National Research Foundation for their financial support and the Centre for High Performance Computing (<http://www.chpc.ac.za>) for their computational resources.

5.6 References

- Argiulo, M. G., Riatta, C. A., Aiteri, E. R., Sidorio, C. I., Astino, R. C., Ace, D. P., ... Molecolare, P. (2002). Lysosomal proteases as potential targets for the induction of apoptotic cell death in human neuroblastomas. *International Journal of Cancer*, 779(September 2001), 775–779. <https://doi.org/10.1002/ijc.10139>
- Arnold, G. E., & Ornstein, R. L. (1997). Molecular dynamics study of time-correlated protein domain motions and molecular flexibility: Cytochrome P450BM-3. *Biophysical Journal*, 73(3), 1147–1159. [https://doi.org/10.1016/S0006-3495\(97\)78147-5](https://doi.org/10.1016/S0006-3495(97)78147-5)
- Banerjee, R., Liu, J., Beatty, W., Pelosof, L., Klemba, M., & Goldberg, D. E. (2002). Four plasmepsins are active in the *Plasmodium falciparum* food vacuole, including a protease with an active-site histidine. *Proceedings of the National Academy of Sciences*, 99(2), 990–995. <https://doi.org/10.1073/pnas.022630099>
- Cai, H., Kuang, R., Gu, J., Wang, Y., Texas, S., Infectious, E., ... Antonio, S. (2011). Proteases in Malaria Parasites - A Phylogenomic Perspective. *Current Genomics*, 12, 417–427.
- Chetty, S., & Soliman, M. E. S. (2015). Possible allosteric binding site on Gyrase B , a key target for novel anti-TB drugs : homology modelling and binding site identification using molecular dynamics simulation and binding free energy calculations. *Medicinal Chemistry Research*, 24, 2055–2074. <https://doi.org/10.1007/s00044-014-1279-3>
- Cousins, Kimberley R. (California State University, S. B. (2005). ChemDraw Ultra 9.0. CambridgeSoft, 100 CambridgePark Drive, Cambridge, MA 02140. *Journal of the American Chemical Society*, 127(11), 4115–4116. Retrieved from <http://en.freownloadmanager.org/Windows-PC/ChemDraw-Ultra.html>
- Crunkhorn, S. (2017). Infectious disease: Blocking malaria parasite invasion and egress. *Nature Publishing Group*, 2017. <https://doi.org/10.1038/nrd.2017.253>
- David, B., Wolfender, J.-L., & Dias, D. A. (2014). The pharmaceutical industry and natural products: historical status and new trends. *Phytochemistry Reviews*, 14(2), 299–315. <https://doi.org/10.1007/s11101-014-9367-z>
- Deriu, M. A., Grasso, G., Tuszynski, J. A., & Gallo, D. (2016). Josephin Domain Structural Conformations Explored by Metadynamics in Essential Coordinates. *PLOS Computational Biology*, 1–14. <https://doi.org/10.1371/journal.pcbi.1004699>

- Explicit, G., Particle, S., Ewald, M., Salomon-ferrer, R., Goetz, A. W., Poole, D., ... Walker, R. C. (2013). Routine microsecond molecular dynamics simulations with Routine microsecond molecular dynamics simulations with AMBER on GPUs . 2 . Explicit Solvent Particle Mesh Ewald. *Journal of Chemical Theory and Computation*, 1–38. <https://doi.org/10.1021/ct400314y>
- Grasso, G., Tuszynski, J. A., Morbiducci, U., Licandro, G., Danani, A., & Deriu, M. A. (2017). Thermodynamic and kinetic stability of the Josephin Domain closed arrangement : evidences from replica exchange molecular dynamics. *Biology Direct*, 1–15. <https://doi.org/10.1186/s13062-016-0173-y>
- Hanwell, M. D., Curtis, D. E., Lonie, D. C., Vandermeersch, T., Zurek, E., & Hutchison, G. R. (2012). Avogadro : an advanced semantic chemical editor , visualization , and analysis platform, 1–17.
- Harvey, L. A., Edrada-Ebel, R., & Quinn, J. R. (2011). The re-emergence of natural products for drug discovery in the genomics era. *Biotechnology Advances*, 1–52.
- Honarparvar, B., Govender, T., Maguire, G. E. M., Soliman, M. E. S., & Kruger, H. G. (2014). Integrated Approach to Structure-Based Enzymatic Drug Design : Molecular Modeling , Spectroscopy , and Experimental Bioactivity. *Chemical Reviews*, 114, 493–537.
- Karubiu, W., Bhakat, S., McGillevie, L., & Soliman, M. E. S. (2015). Molecular BioSystems Flap dynamics of plasmepsin proteases : insight dynamics †. *Molecular BioSystems*, 11, 1061–1066. <https://doi.org/10.1039/C4MB00631C>
- Kim, M. O., Blachly, P. G., Kaus, J. W., & Mccammon, J. A. (2015). Protocols Utilizing Constant pH Molecular Dynamics to Compute pH- Dependent Binding Free Energies. *The Journl of Physical Chemistry*, 1–12.
- Lonsdale, R., Harvey, J. N., Mulholland, A. J., & Lonsdale, R. (2012). A practical guide to modelling enzyme-catalysed reactions. *Chemical Society Reviews*, 41, 3025–3038. <https://doi.org/10.1039/c2cs15297e>
- Luo, J., & Bruice, T. C. (2002). Ten-nanosecond molecular dynamics simulation of the motions of the horse liver alcohol dehydrogenase PhCH 2 O complex. *PNAS*, 99(26), 16597–16600.
- Mark, A. E., & Gunsteren, W. F. Van. (1995). Fluctuation and Cross-correlation Analysis of

- Protein Motions Observed in Nanosecond Molecular Dynamics Simulations. *Journal of Molecular Biology*, 252, 492–503.
- McGillewie, L., Ramesh, M., & Soliman, M. E. (2017). Sequence , Structural Analysis and Metrics to Define the Unique Dynamic Features of the Flap Regions Among Aspartic Proteases. *The Protein Journal*, 12(0), 5–9. <https://doi.org/10.1007/s10930-017-9735-9>
- Mhlongo, N. N., Ebrahim, M., Skelton, A. A., Kruger, H. G., Williams, I. H., & Soliman, M. E. S. (2015). Dynamics of the thumb-finger regions in a GH11 xylanase *Bacillus circulans*: comparison between the Michaelis and covalent intermediate. *RSC Advances*, 5(100), 82381–82394. <https://doi.org/10.1039/c5ra16836h>
- Mhlongo, N. N., & Soliman, M. E. S. (2015). RSC Advances Single H5N1 in fl uenza A neuraminidase mutation develops resistance to oseltamivir due to distorted conformational and drug binding landscape : multiple molecular dynamics analyses †. *RSC Advances*, 5, 10849–10861. <https://doi.org/10.1039/C4RA13494J>
- Morris, G. M., Huey, R., & Lindstrom, W. (2009). *Autodock :Raccoon*. California.
- Moura, P. A., Dame, J. B., & Fidock, D. A. (2009). Role of *Plasmodium falciparum* Digestive Vacuole Plasmepsins in the Specificity and Antimalarial Mode of Action of Cysteine and Aspartic Protease Inhibitors. *Antimicrobial Agents and Chemotherapy*, 53(12), 4968–4978. <https://doi.org/10.1128/AAC.00882-09>
- Mukherjee, B., Tessaro, F., Vahokoski, J., Kursula, I., Marq, J., Scapozza, L., & Soldati-favre, D. (2018). Modeling and resistant alleles explain the selectivity of antimalarial compound 49 c towards apicomplexan aspartyl proteases. *The EMBO Journal*, 1–19. <https://doi.org/10.15252/emj.201798047>
- Munsamy, G., Ramharack, P., & Soliman, M. E. S. (2018). Egress and invasion machinery of malaria: an in depth look into the structural and functional features of the fl ap dynamics of plasmepsin IX and X. *RSC Advances*, 8, 21829–21840. <https://doi.org/10.1039/C8RA04360D>
- Nair, D. N., Singh, V., Angira, D., & Thiruvenkatam, V. (2016). Proteomics & Bioinformatics Structural Investigation and In-silico Characterization of Plasmepsins from *Plasmodium falciparum*. *Journal of Proteomics and Bioinformatics*, 9(7), 181–195. <https://doi.org/10.4172/jpb.1000405>

- Nasamu, A. S., Glushakova, S., Russo, I., Vaupel, B., Oksman, A., Kim, A. S., ... Goldberg, D. E. (2017). parasite egress and invasion. *Science*, 522(October), 518–522.
- Nkumama, I. N., Meara, W. P. O., & Osier, F. H. A. (2017). Changes in Malaria Epidemiology in Africa and New Challenges for Elimination. *Trends in Parasitology*, 33(2), 128–140. <https://doi.org/10.1016/j.pt.2016.11.006>
- Paquet, E., & Viktor, H. L. (2015). Molecular dynamics, monte carlo simulations, and langevin dynamics: A computational review. *BioMed Research International*, 2015. <https://doi.org/10.1155/2015/183918>
- Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera — A Visualization System for Exploratory Research and Analysis. *Journal of Computer Chemistry*, 25, 1605–1612. <https://doi.org/10.1002/jcc.20084>
- Pino, P., Caldelari, R., Mukherjee, B., Vahokoski, J., Klages, N., Maco, B., ... Soldati-favre, D. (2017). for invasion and egress. *Science*, 528(October), 522–528.
- Roy, K. K. (2018). International Journal of Antimicrobial Agents Targeting the active sites of malarial proteases for antimalarial drug discovery : approaches , progress and challenges. *International Journal of Antimicrobial Agents*, 50(3), 287–302. <https://doi.org/10.1016/j.ijantimicag.2017.04.006>
- Sayer, J. M., & Louis, J. M. (2010). Interactions of different inhibitors with active site aspartyl residues of HIV-1 protease and possible relevance to pepsin. *Journal of Proteins*, 75(3), 556–568. <https://doi.org/10.1002/prot.22271>.
- Scott, W. R. P., & Schiffer, C. A. (2000). Curling of Flap Tips in HIV-1 Protease as a Mechanism for Substrate Entry and Tolerance of Drug Resistance. *Science Direct*, 8(00), 1259–1265.
- Trott Oleg, O. J. A. (2011). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function,efficient optimization and multithreading. *NIH Public Access*, 31(2), 455–461. <https://doi.org/10.1002/jcc.21334>.AutoDock
- Who. (2015). Global Technical Strategy for Malaria 2016-2030. Retrieved from <http://www.tbfacts.org/tb-statistics/>
- Wright, D. W., Hall, B. A., Kenway, O. A., Jha, S., & Coveney, P. V. (2014). Computing

Clinically Relevant Binding Free Energies of HIV - 1 Protease Inhibitors. *Journal of Chemical Theroy and Computation*, 10, 1228–1241.

Ylilauri, M., & Pentikäinen, O. T. (2013). MM/GBSA as a tool to understand the binding affinities of filamin-peptide interactions. *Journal of Chemical Information and Modeling*, 53(10), 2626–2633. <https://doi.org/10.1021/ci4002475>

Zhang, D., & Lazim, R. (2017). Application of conventional molecular dynamics simulation in evaluating the stability of apomyoglobin in urea solution. *Nature Publishing Group*, 7(October 2016), 1–12. <https://doi.org/10.1038/srep44651>

CHAPTER 6

Unveiling a new era in Malaria therapeutics: A tailored molecular approach towards the design of plasmepsin IX inhibitors

Geraldene Munsamy^a, and Mahmoud E. S. Soliman^{a*}

^a Molecular Bio-computation and Drug Design Laboratory, School of Health Sciences,
University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa

*Corresponding Author: Mahmoud E.S. Soliman

Email: soliman@ukzn.ac.za

Telephone: +27 (0) 31 260 8048, Fax: +27 (0) 31 260 7872

6.1 Abstract

The invasive strategies employed by the malarial parasite renders malaria a global health threat, further impeding the effective treatment of the mosquito borne-parasitic disease. Although there have been countless efforts directed towards the development of effective therapeutics, factors such as emerging strains of drug resistance, enhanced toxicity and poor pharmacokinetic properties of current therapeutics has hampered the drug discovery process relenting in the spread of this parasitic disease. A promising target of the most lethal strain of the *Plasmodium* species that plays a predicted role in erythrocyte invasion of the virulent malarial parasite is aspartic protease IX commonly referred to PIX. The integration of computer aided-drug design platforms has revolutionized the 21st century and has opened avenues to render a final “knock out” in the elimination and eradication of this parasitic disease. Hitherto, this is the first attempt directed towards the design of therapeutics tailored explicitly to PIX. A potent peptidomimetic inhibitor referred to as 49c which is a known inhibitor of PII, has recently been proven to exhibit further potent inhibitory activity against PIX. *In-silico* structural and physicochemical inspection of 49c displayed poor pharmacokinetic properties thus paving the way for development of tailored inhibitors with desirable therapeutic properties against PIX. In this study we implement the pharmacophore model approach in combination with per-residue energy decomposition analysis to serve as a powerful cornerstone, that may assist medicinal experts in the composition of multifunctional therapeutics that may dispose factors such as cross-resistance and toxicity, with enhanced pharmacokinetic properties.

Keywords: *Plasmodium falciparum*; PIX; Twisting motion; Pharmacophore

6.2 Introduction

Malaria is categorized as one of the highest contributors to the unremitted global health crisis. The rapid elevation of morbidity and mortality rates escalating from malarial infections, remain predominant in endemic regions such as sub-Saharan Africa as they carry 80% of the global malaria burden (World Health Organization, 2016) (World Health Organization, 2017). Over one million deaths have resulted primarily from the infection of the *P. falciparum* parasite, which is acclaimed to be the most virulent species of *Plasmodium* (Deu, 2017) (Liu, Gluzman, Drew, & Goldberg, 2005) (Nair, Singh, Angira, & Thiruvengatam, 2016) (Asojo et al., 2003). Despite the effort directed towards malaria treatment and elimination of emerging strains of multidrug resistance of artemisinin and partner drugs, the eradication of malaria still remains ambiguous (Halder, Bhattacharjee, & Safeukui, 2018). The regression in malaria treatment has led to the decline in the sustainability of the ongoing global effort to reduce the impending burden of malaria (Rosenthal, 1998) (Nigussie, Beyene, Shah, & Belew, 2015). Amid emergent strains of drug resistance against malaria, there is still an ever-growing search for reputable targets to administer a final “knock-out” in the permanent treatment of malaria. Efficacious antimalarial therapeutics for specific and selective targets remains critical to malaria control and eradication.

The last decade has observed an expansive research interest in the development of potent inhibitors against the aspartic protease family of enzymes (Dash, Kulkarni, Dunn, & Rao, 2003) (Ersmark, Samuelsson, & Hallberg, 2006). Aspartic proteases are a small family of enzymes which have received interest due to their significant role in human diseases some examples include; HIV-1 protease in HIV, cathepsin D in metastasis of breast cancer, β -Secretase in Alzheimer's Disease, and more specifically Plasmeprin's in malaria pathogenesis (Cai et al., 2011) (Coombs et al., 2001) (Banerjee et al., 2002). The essential role of aspartic proteases are exemplified in the role they play in parasite survival, cell invasion, nutrition, growth and processing of precursor proteins (Cowman, Berry, & Baum, 2012) (Berry, 2000).

P. falciparum are host to 10 aspartic proteases (I-X) commonly referred as Plasmeprin's (Plms). Of the 10 Plms identified in *P. falciparum*, Plm I-IV are delivered to the food vacuole, where they play an instrumental role in the degradation of haemoglobin in a consecutive manner with each Plm undertaking a specific yet independent role in haemoglobin degradation (Parr, Tanaka, Xiao, & Yada, 2008). Plm V is responsible for the export of effector proteins mediated to the host cell (Moura, Dame, & Fidock, 2009), whereas Plm VI-VIII are

expressed in the vector during the parasite's intra-erythrocytic stages of motility, formation of sporozoites and midgut transversal and hence are not directly involved in the transmission of the malarial parasite within the human host(Nair et al., 2016). The hindmost family members PIX and PX are accentuated as key mediators of egress and invasion of the parasite relaying exposure to the parasite at its most vulnerable state(McGeorge, 2014)(Nasamu et al., 2017)(Cells et al., 2016)(Nair et al., 2016).

PIX and PX are consecutively expressed with Plm I–IV, although they are restricted entrance to the food vacuole, PIX and PX share prohibitive structural similarity to the PI-PIV and thus may be the targets of digestive vacuole Plm inhibitors that have antimalarial activity. PIX and PX intercedes the progression of infection within the human host assimilating the role they play particularly in egress and invasion of the malaria parasite. A provisional knockout study implemented on PIX has disclosed its role as a maturase during the formation of merozoites, and the crucial role it plays in erythrocyte invasion making it a formidable target in the design of antimalarial inhibitors(Crunkhorn, 2017). A unique structural domain present within aspartic proteases forming the flap and flexible hinge region perpendicular to the aspartic catalytic site remain essential for exhibiting the characteristic “twisting motion” during the opening and closing movement of Plm's. This region is crucial in understanding the ligand binding landscape as well as conformational flexibility of Plm's. Previous studies have highlighted the how the enclosure on an inhibitor within the active site restricts the characteristic “twisting motion” of aspartic proteases from transpiring(Chang, Chen, & Gilson, 2007)(Mao, 2011)(Petrenko, 2010).

Inhibitor 49c, which is a hydroxy-ethyl amine peptidomimetic competitive aspartic protease inhibitor, demonstrated efficacy in experimental infection models. 49c was initially developed to target PII but studies have shown its potency against latter family members PIX and PX. This core scaffold of 49c displayed potency against the preexocytosis processing of several secreted rhoptry and microneme proteins by directly targeting the resultant maturases PIX and PX respectively (Pino et al., 2017). A previous study by our group has highlighted the structural binding conformation of PIX and PX in complex with 49c, with both systems displaying reduced flexibility and limitation in the motion of opening and closing of the flap and hinge domain of PIX and PX.

The SwissADME online tool was integrated in this study to compute the physicochemical descriptors as well as predict the pharmacokinetic properties and drug-like nature of inhibitor

49c to PIX(Galindo-murillo et al., 2016)(Martin & Park, 2005). SwissADME exploits the ‘Brain Or Intestinal Estimated permeation (BOILED-Egg)’ method which computes the lipophilicity and polarity of small molecules(Daina, Michielin, & Zoete, 2017). From the results displayed in table S5.1, 49c exhibits poor solubility, as well as a violation of Lipinski’s rule of five, unravelling a loophole in the search of novel antimalarial therapeutics with desirable pharmacokinetic properties.

The quest for potential curative compounds, sought after in the drug discovery and development pipeline remain a time staking and laborious task and is currently proving to unsuccessful in target discovery(Saddala & Adi, 2018). Alternatively, implementing *in silico* tools in the generation of potential therapeutics has proven to be a powerful and indispensable tool in discovering small molecular inhibitors and the visualisation of these small molecules as they interact with drug targets on a molecular level.

The integration of the pharmacophore model protocol serves as a robust cornerstone, that may assist medicinal experts in the composition of multifunctional inhibitors that adhere to sought after pharmacokinetics characteristics essential for drug like molecules as well as eliminate factors in drug design such as cross-resistance and toxicity and may further enhance patient adherence(Kaalia, Kumar, Srinivasan, & Ghosh, 2011)(Machaba, 2017). A combined ligand and structure-based drug design approach provides a synergistic advantage over either method performed independently. The pharmacophore model approach presented in Figure 6.1 was implemented in our study to form an assistive measure in the design of small molecules against an array of aspartic proteases, more specifically to PIX to which there are currently no available FDA-approved inhibitors. This approach has the capabilities to eliminate collective drug design challenges, particularly cross-resistance and toxicity(Favourite N, Ramesh, & Mahmoud E.S, 2016).

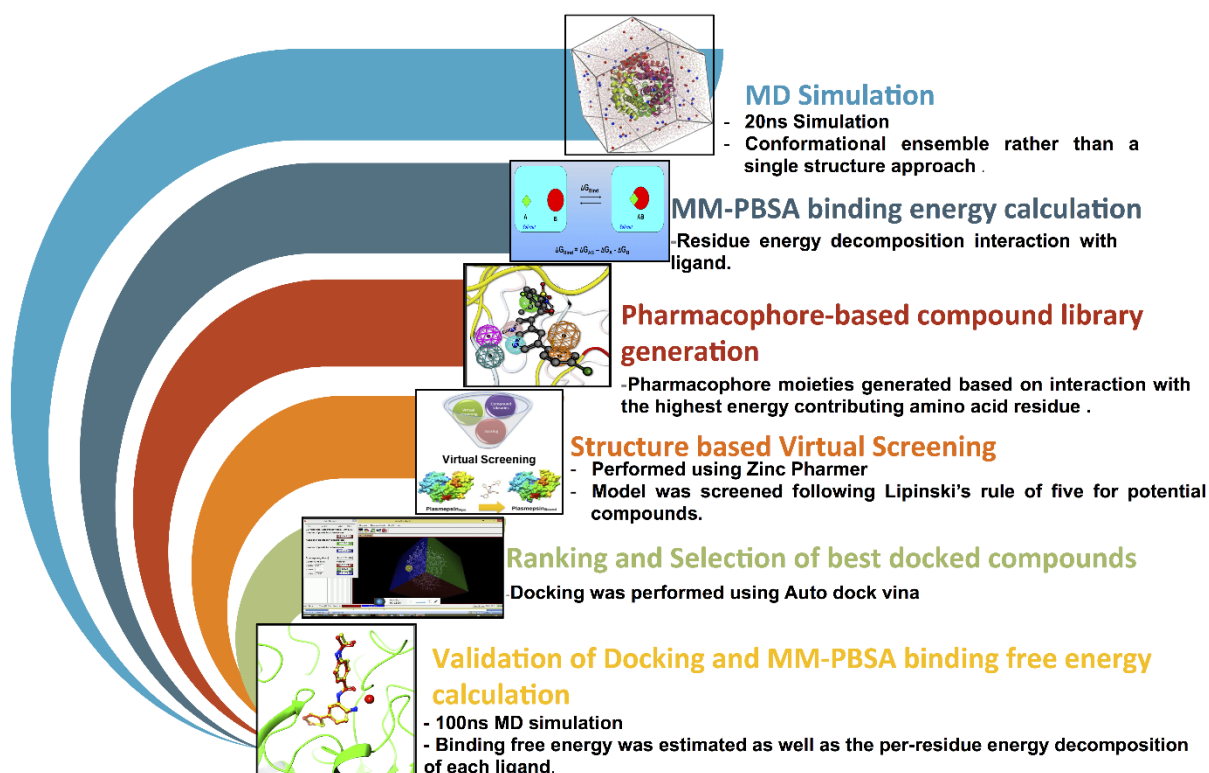


Figure 6.1 A PRED-based approach outline applied in the study

6.3 Computational Methodology

6.3.1 Pharmacophore model generation and Screening

In the absence of experimental models, homology modeling is often sought after in determining a proteins 3D structure. In this study the homology models of PIX and PX were generated using the PHYRE2 server, more in-depth information regarding the generation of these models are explicated in our previous publication (Munsamy, Ramharack, & Soliman, 2018). A 50 ns MD simulation of PIX in complex with inhibitor 49c were simulated to generate the bound conformation of 49c within the active site of PIX as well as to decipher which amino acids of PIX projected a substantial contribution towards the total binding of 49c to PIX. The CPPTRAJ module within AMBER 14 package were integrated to perform the per-residue energy decomposition analysis of 49c in complex with PIX. The pharmacophoric moieties of 49c that displayed the highest contribution towards the total binding energy as observed in Figure 6.2 were selected to construct the pharmacophore model that was used to screen for other inhibitors using the ZINC PHARMER online platform. The model was screened using an explicit selection criteria (molecular weight of <500 Da, rotatable bonds <6, hydrogen bond donors <5, and hydrogen bond acceptors <10), to screen the ZINC database(Lionta, Spyrou, Vassilatis, & Cournia, 2014; Irwin & Shoichet, 2006). The Lipinski's rule of five and toxicity (ADMET)

properties were further analysis to refine the number of compounds screened and to nondrug-like hits(Lipinski, Lombardo, Dominy, & Feeney, 2001).

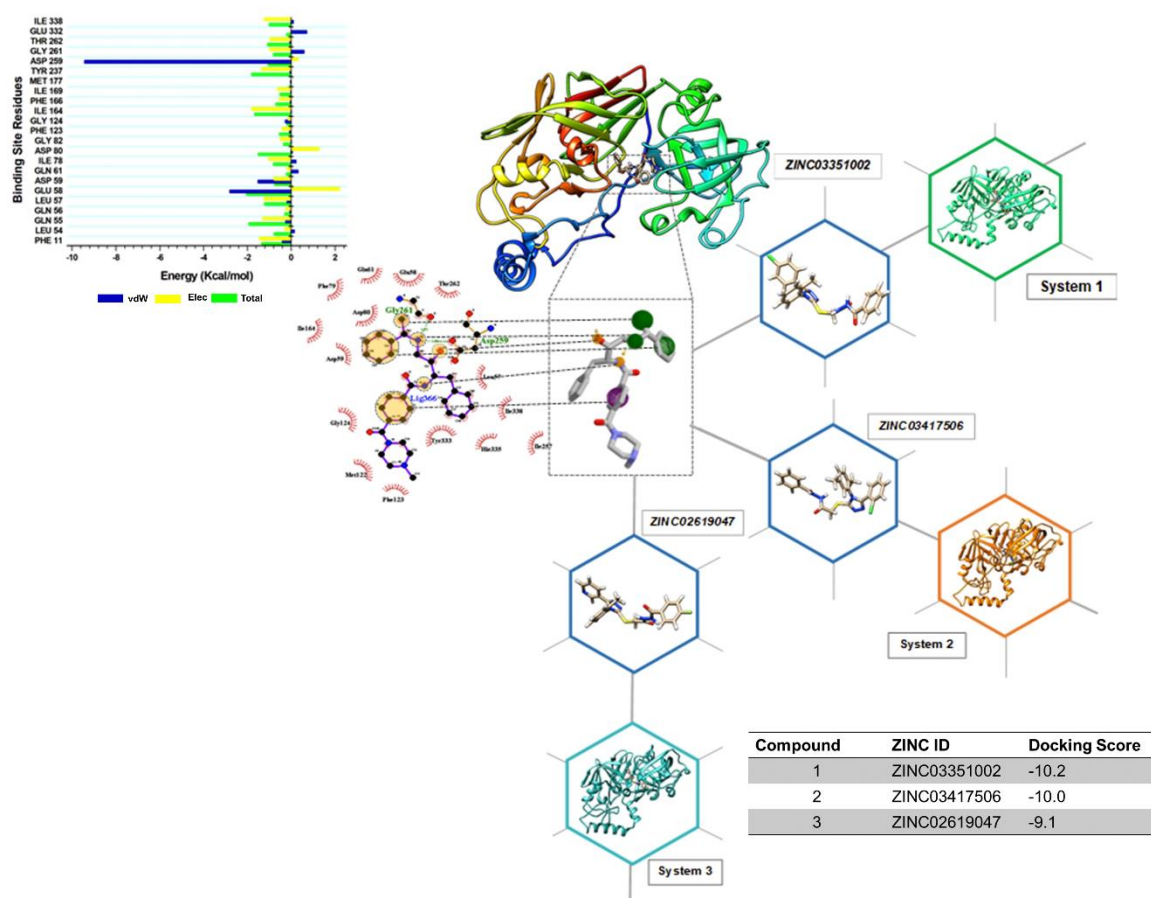


Figure 6.2 Pharmacophore model generation adapted from the PIX-49c complex. The orange circles spotlight the pharmacophoric moieties that were chosen for the model, based on the highest contributing residues, depicted in the binding affinity graph.

6.3.2 Molecular docking

Optimized conformations and binding affinities of the 29 hits generated from pharmacophore screening were subjected to molecular docking via UCSF Chimera software and Auto Dock vina(Morris et al., 2010). All 29 hits were individually docked to PIX (grid box of spacing of 0.375 Å and x, y, z dimensions of 40 x 56 x 46). Molecular dynamic simulations were run on the top three complexes with the most negative binding energy (kcal.mol⁻¹).

6.4 System preparation

6.4.1 Molecular dynamics

The integration of Molecular dynamic (MD) simulations in the study of biological systems enables the exploration of the physical motion of atoms and molecules that cannot be easily

accessed by any other means(Tamar, 2002). The insight extracted from performing this sort of simulations provide an intricate perspective into the dynamical evolution within the biological systems such as conformational changes and molecule association(Anslyn & Dougherty, 2006). The MD simulations of all systems were performed using the GPU version of the PMEMD engine present in the 14 AMBER package(Case et al., 2005).

The atomic partial charge of each compound were generated using ANTECHAMBER utilizing General Amber Force Field (GAFF) protocol(Wang, Wolf, Caldwell, Kollman, & Case, 2004). Each system was implicitly solvated within an orthorhombic box of TIP3P water molecules within 10Å of any box edge, performed by the Leap module of the AMBER 14 package. Neutralisation of each system were further implemented via the addition of Na⁺ and Cl⁻ counter ions integrated in the Leap module(Galindo-murillo et al., 2016).

An initial minimization of each system was performed for a period of 2000 steps in the presence of an applied restraint potential of 500 kcal/mol. This was followed by a full minimization of 1000 steps carried out by conjugate gradient algorithm with no restraints.

Each system was gradually heated during the MD simulation from 0K to 300K for 500ps, ensuring that all systems maintained a fixed number of atoms with a fixed volume. The solutes within the system were imposed with a potential harmonic restraint of 10kcal/mol and collision frequency of 1ps. Ensuing heating, equilibration of each system was performed for 500ps at a constant temperature of 300K. The number of atoms and pressure within in each system for each production simulation were kept constant to mimic an isobaric-isothermal (NPT) ensemble, with the systems pressure being maintained at 1 bar using the Berendsen barostat(Berendsen et al., 1984).

Each system was MD simulated for 100ns, in each simulation the SHAKE algorithm was employed to constrain the hydrogen bond atoms. The pH of our systems were set to 4.5 which correlated with experimental studies as it is the optimum biological pH and was kept constant and corroborates with data projected from the H++ tool which computes the pKa value of ionisable groups. The step size of each simulation was 2fs and integrating an SPFP precision model. The simulations coincided with isobaric-isothermal ensemble (NPT), with randomized seeding, constant pressure of 1 bar, a pressure-coupling constant of 2ps, a temperature of 300K and Langevin thermostat(Davidchack et al., 2009) with collision frequency of 1ps.

6.4.2 Post-Dynamic Analysis

The coordinates of the free enzyme and bound complexes were saved after every 1ps and the trajectories were analysed using the CPPTRAJ module employed in AMBER 14 suit. The CPPTRAJ module present within the AMBER 14 package were employed to perform post dynamic analysis on each system. The Root Mean Square Deviation (RMSD) and thermodynamic energy of each system were investigated.

6.4.3 Per-Residue Decomposition Analysis & Binding Free Energy Computation

The relative binding free energy (ΔG_{bind}) of PIX to PIX, were computed using molecular mechanics integrated with the Poisson-Boltzmann or generalized Born and surface area continuum solvation (MM/PBSA and MM/GBSA) approach(Wang, Hou, Li, & Wang, 2012). MM/GB-SA and MM/PB-SA rely on molecular simulations of the ligand-protein complex to compute rigorous statistical-mechanical binding free energy within a specified force field(Du et al., 2016). These methods have been expounded in the referred papers(Erik, 2010)(Wright, Hall, Kenway, Jha, & Coveney, 2014)(Ylilauri & Pentikäinen, 2013).

The molecular dynamic simulation yielded a trajectory depicted by 50,000 snapshots, which were averaged to generate ΔG_{bind} . The free binding energy (ΔG) computed by this method for each molecular species (complex, ligand and receptor) can be represented as:

$$\Delta G_{\text{bind}} = G_{\text{complex}} - G_{\text{receptor}} - G_{\text{ligand}} \quad (1)$$

$$\Delta G_{\text{bind}} = E_{\text{gas}} + G_{\text{sol}} - TS \quad (2)$$

$$E_{\text{gas}} = E_{\text{int}} + E_{\text{vdw}} + E_{\text{ele}} \quad (3)$$

$$G_{\text{sol}} = G_{\text{GB}} + G_{\text{SA}} \quad (4)$$

$$G_{\text{SA}} = \gamma \text{SASA} \quad (5)$$

where:

E_{ele} Electrostatic potential energy from Coulomb forces

E_{gas} Gas-phase energy (based on FF14SB force field terms)

E_{int} Internal energy

E_{vdW} van der Waals energy

G_{sol} Solvation free energy

G_{GB} Polar solvation energy

G_{SA} non-polar solvation energy

S Total entropy of solute

SASA Solvent accessible surface area (water probe radius of 1.4 Å)

T Total entropy of temperature

A per-residue free energy atomistic decomposition was employed to estimate the contribution of each residue to the total ΔG_{bind} at the binding site, for significant residues of the PIX using the AMBER14 MM/GBSA method.

6.5 Molecular dynamic analysis and Post MD Analysis

6.5.1 Screened compounds facilitate the restriction of PIX structural flexibility

Root Mean Square Deviation (RMSD) shows the deviation from the minimized predicted structure while Root Mean Square Fluctuation (RMSF) shows the deviation from the mean structure over a dynamic ensemble. In this sense, RMSF gives a picture of which parts of the protein are moving while RMSD gives an overall picture of how much each part of the protein has changed over the course of the simulation. Analysis of RMSD were performed to validate the stability of all systems. PIX in complex with Compound 1, Compound 2 and 49c all exhibited RMSD values lower than 3 Å, however PIX in complex with Compound 3 projected an RMSD value of 3.28 Å (Figure S6.2). However, post MD analysis was performed on all systems to decipher factors that have attributed to the PIX-Compound 3 system not reaching stability.

The RMSF for all systems are presented in Figure 6.3. Intuitively, binding between a protein to a drug molecule reflects the restriction of the intrinsic flexibility of the binding region in a protein and in its binding partner, resulting in a decline in the conformational entropy. Inadvertently the PIX-49c system exhibits the highest fluctuation in comparison to the other systems, however we observe a significant decrease in the fluctuation observed in residues Asp80, Asp259, Gly124 and Glu331 which are the main constituting residues essentially for ligand binding and stabilization which remains conserved in all systems.

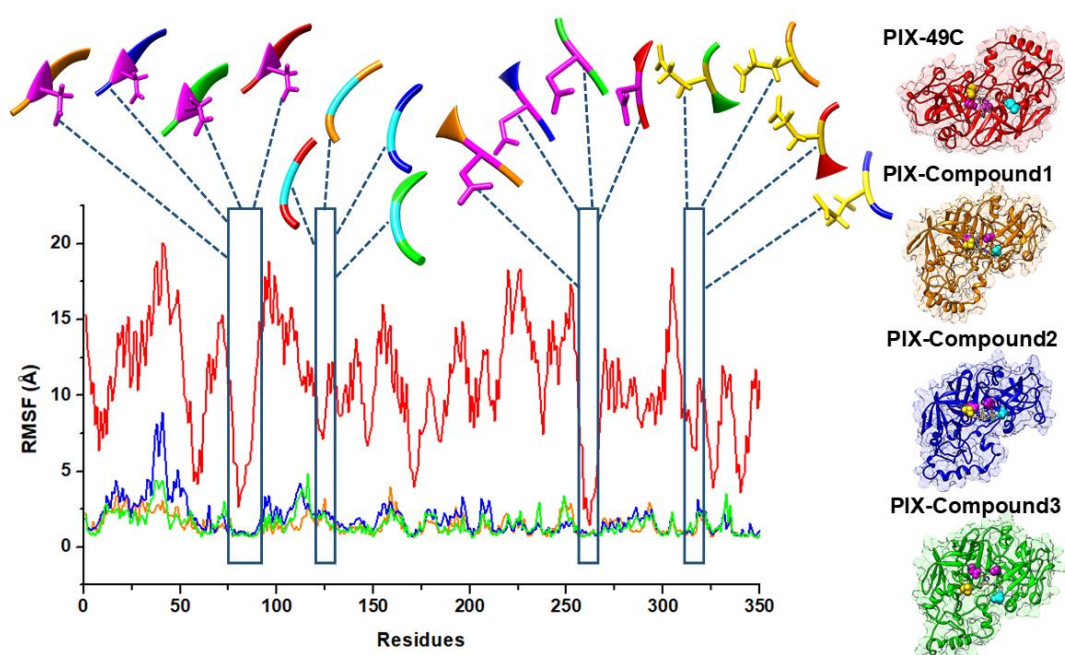


Figure 6.3 Root Mean Square Fluctuation of PIX in complex with 49c, Compound 1, Compound 2 and Compound 3. The highlighted regions include the catalytic dyad Asp 53, Asp232 (magenta), the flap tip Gly97 (cyan) and hinge residue Glu305 (gold).

The fluctuation observed in the PIX-49c system may resonate from the structural transition as the loop recoils within the binding cavity in relation to the catalytic dyad (Asp80, Asp259) during the binding of 49c which forms noncovalent interaction with the catalytic domain and assumes a lateral shift deviating from the principal structure. The instantaneous binding of PIX to Compound 1 and Compound 2, restricts the rigorous asymmetrical motion of opening and closing of the flap and hinge region, exhibiting minimal fluctuation throughout the 100ns simulation.

The reduced flexibility of PIX in complex with Compound 1, Compound 2 and Compound 3 is further corroborated by the DCCM plots presented in Figure S6.3, which displays an overall anti-correlated motion in all PIX systems in complex with Compound 1, Compound 2 and Compound 3, whereas the PIX system in complex with 49c displayed anti-correlated motion as well as correlated motion.

6.5.2 Distance analysis between the flap tip and hinge residue

Considering the crucial relevance of flap dynamics of aspartic proteases which exhibit a “twisting motion” related to the catalytic activity of aspartic proteases. We employed distance calculations, to resolve the structural influence of each ligand on the flap motion of PIX. The

distance analysis was performed assimilating the parameters employed by our group as mentioned in our previous publications(Mcgillewie & Soliman, 2015)(Mcgillewie, Ramesh, & Soliman, 2017)(Karubiu, Bhakat, McGillewie, & Soliman, 2015)(Arodola & Soliman, 2016).

PIX in complex with Compound1, Compound 2 and 49c adopt a compact conformation throughout the 100ns simulation, whereas PIX in complex with Compound 3, displays greater mobility as PIX dispositions between an opening and closing motion in complex with Compound 3 as can be observed in Figure 6.4b.

The distinct increase in the distance between the flap tip and hinge region of PIX in complex with 49c during the first 20ns of the simulation may resemble the opening of the catalytic domain facilitating conformational stability upon binding of 49c which maintains that stability during the remainder of the simulation. The binding of Compound 1 and Compound 2 to PIX renders immediate restriction of the opening and closing of the flap and hinge region during binding to PIX with minimal fluctuation in the distance between the flap tip and hinge region during the 100ns simulation. This contoured stability correlates with the presence of the adjacent amine moieties present in Compound 1 and Compound 2 which form energetically favourable noncovalent interactions with key catalytic residues preventing the “twisting motion” from transpiring. Compound 1 and Compound induce a similar conformational orientation within the active site and this may be attributed to their high structural similarity differing only by the presence of a single methyl group.

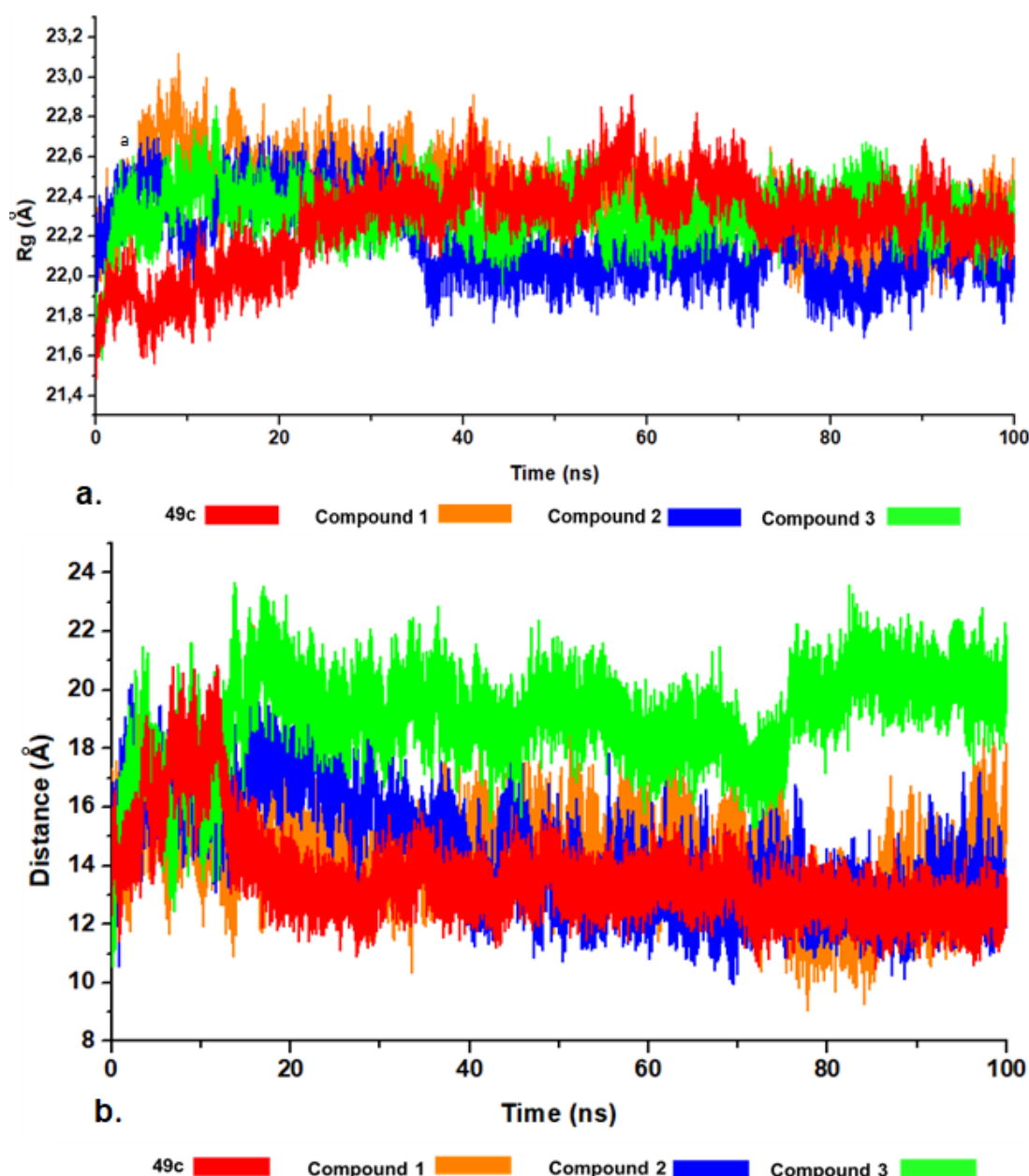


Figure 6.4 (a) Radius of gyration (ROG) of PIX in complex with 49c, Compound 1, Compound 2, and Compound 3. (b) Distance measured in Angstroms between the flap tip and hinge residue of PIX in complex with 49c, Compound 1, Compound 2 and Compound 3.

6.5.3 Radius of Gyration

The radius of gyration (ROG) is indicative of the compactness of the tertiary structure of a protein that is a measure of the spatial spread of the protein mass or how folded or unfolded a protein is and conveys insight into the stability of biological molecules during the MD simulation. From the ROG analysis, all systems exhibit compact conformations with minimal fluctuations, however, PIX in complex with Compound 3 still exhibits the least compact system. The ROG analysis of all systems presented in Figure 6.4a, concur with the distance analysis, as we were able to observe a significant change in the ROG during the first ~35 ns in

PIX in complex with Compound 1, Compound 2 and 49c all of which are suggestive of ligand binding , which includes the formation of hydrogen bonds and electrostatic interactions which play a critical role in in the binding of each ligand within the catalytic domain.

6.5.4 Binding free energy calculations and Per-residue energy decomposition analysis

The high relative binding free energy observed in all systems are resultant of the noncovalent interactions observed between each ligand within the active site which is depicted by the restriction of the mobility of the interacting residues in the presence of each ligand. As can be depicted in Figures 6.5, 6.6, 6.7 and 6.8.

PIX in complex with 49c exhibits the highest binding free energy in comparison to PIX in complex with Compound 1, Compound 2 and Compound 3, however a major contributing factor to the high binding free energy correlated to 49c stems from the high electrostatic energy observed between PIX and 49c within the catalytic domain.

The high relative binding free energy observed in all systems is linked to the increased number of noncovalent interactions observed between each ligand within the active site which is depicted by the restriction of the mobility of the interacting residues in the presence of each ligand. As can be depicted in Figure 6.5, 6.6, 6.7 and 6.8.

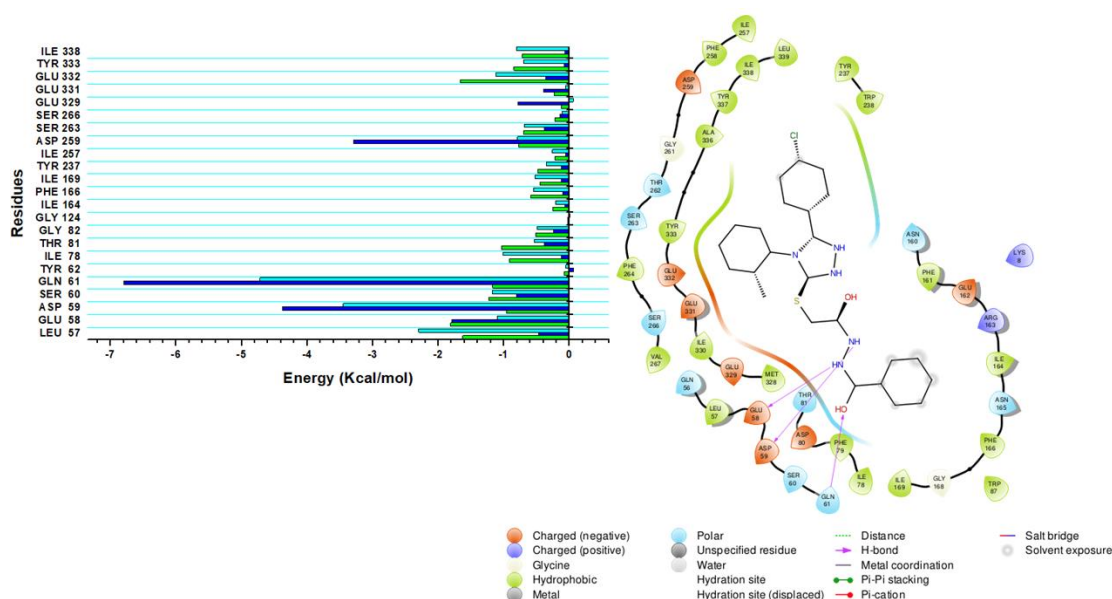


Figure 6.5 MM/GBSA binding free energy profiles of PIX in complex with Compound 1. Highlighting the interacting amino acid residues of PIX and their energy contributions to the total binding free energy

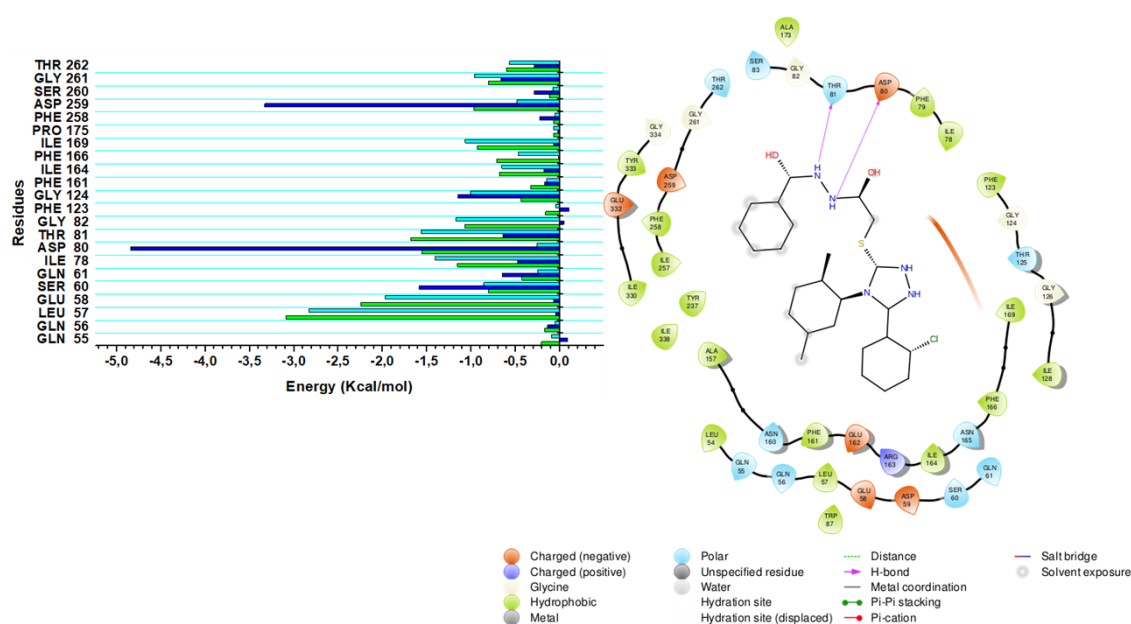


Figure 6.6 MM/GBSA binding free energy profiles of PIX in complex with Compound 2. Highlighting the interacting amino acid residues of PIX and their energy contributions to the total binding free energy

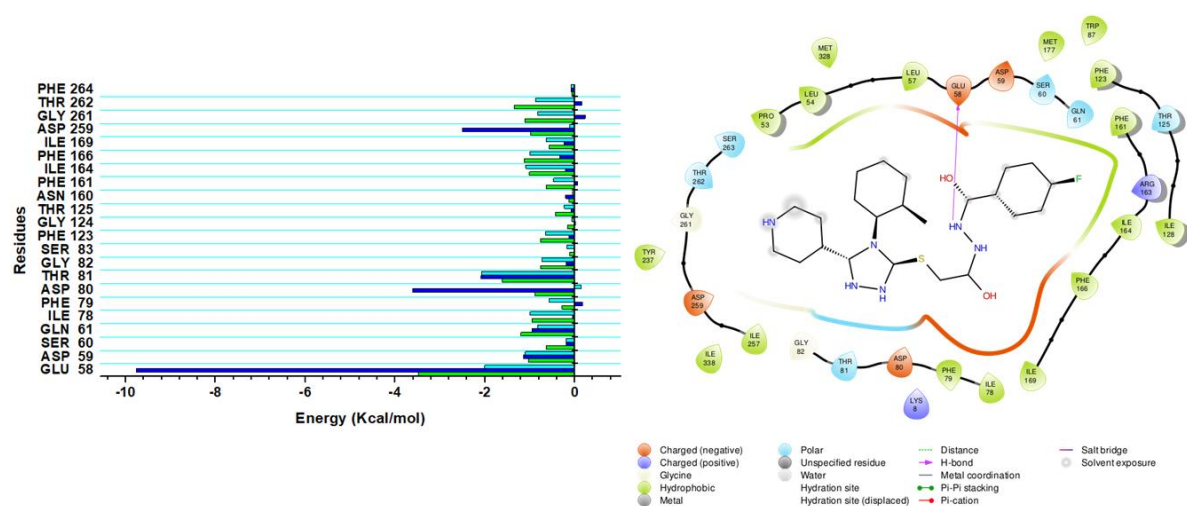


Figure 6.7 MM/GBSA binding free energy profiles of PIX in complex with Compound 3. Highlighting the interacting amino acid residues of PIX and their energy contributions to the total binding free energy

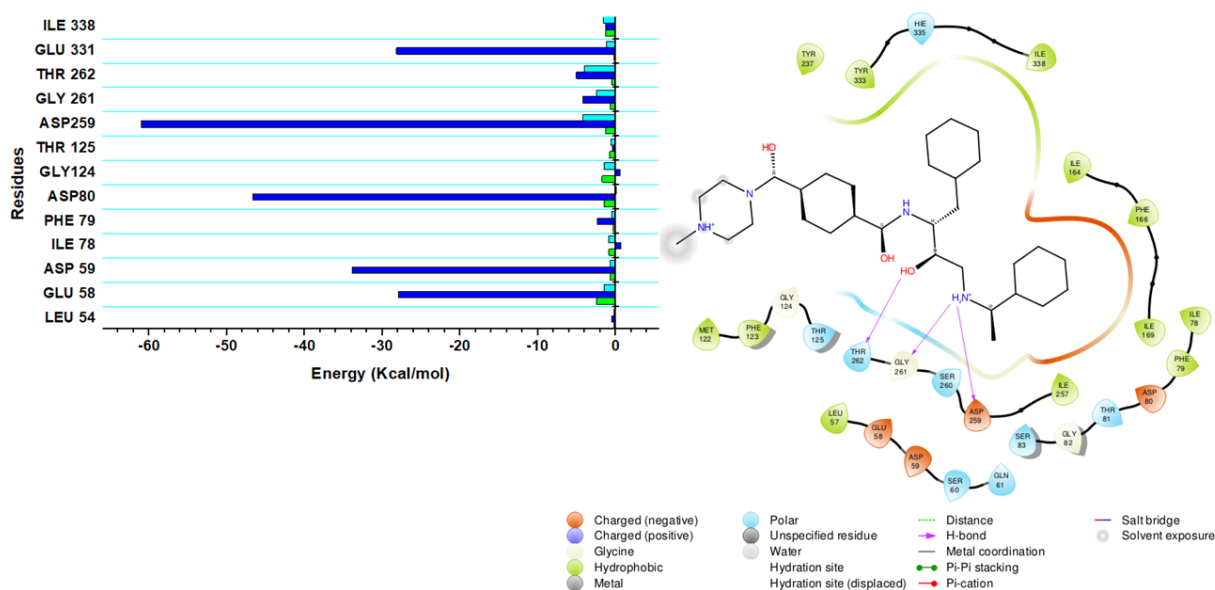


Figure 6.8 MM/GBSA binding free energy profiles of PIX in complex with peptidomimetic inhibitor 49c. Highlighting the interacting amino acid residues of PIX and their energy contributions to the total binding free energy

As presented in **Table 6.1**, PIX in complex with 49c exhibits the highest binding free energy of -53.64 Kcal/mol in comparison to PIX in complex with Compound 1, Compound 2 and Compound 3, which displayed binding free energies of -47.97 Kcal/mol, -43.48 Kcal/mol and -34.93 Kcal/mol respectively. A major contributing factor to the high binding free energy correlated to 49c stems from the high electrostatic energy of -359.90 Kcal/mol observed between PIX and 49c within the catalytic domain.

Table 6.1 MM/GBSA binding free energy profiles of 49c, Compound 1, Compound 2 and Compound 3 interactions with PIX and PX

Complexes	ΔE_{vdw} (Kcal/mol)	ΔE_{ele} (Kcal/mol)	ΔG_{gas} (Kcal/mol)	ΔG_{sol} (Kcal/mol)	ΔG_{bind} (Kcal/mol)
Compound1_PIX	-49.52±0.1	-24.99±0.2	-69.51±0.2	28.54±0.1	-47.97±0.1
Compound 2_PIX	-55.08±0.1	-18.89±0.2	-70.97±0.1	27.49±0.1	-43.48±0.1
Compound 3_PIX	-51.61±0.2	-26.29±0.1	-77.90±0.2	42.98±0.1	-34.93±0.1
49c_PIX	-51.04±0.1	-359.90±0.4	-410.95±0.4	357.31±0.4	-53.64±0.1

ΔE_{ele} = electrostatic energy; ΔE_{vdw} = van der Waals energy; ΔG_{bind} = total binding free energy; ΔG_{sol} = solvation free energy ΔG = gas phase free energy.

The high electrostatic energy associated with 49c can be associated to the proximity of the two hydrogen bond acceptors to each other which creates a broad expanse of negative electrostatic potential between them. From the 2D molecular representation of PIX in complex with 49c in Figure 6.8, Asp 259 and Gly 261 form hydrogen bonds interactions with the amine (NH₂) moiety of the potent hydroxyl-ethyl-amine scaffold of 49c. The high electrostatic interaction exhibited by 49c is constituted by the negative potential between Asp 259 and Gly 261 as they interact with the same chemical moiety. Although the presence of hydrogen bonds and electrostatic interactions are present between PIX in complex with Compound 1, Compound 2 and Compound 3, no two amino acid residues interact with the same moiety with each Compound accounting for the difference in the electrostatic energy contribution.

From visualisation analysis of all four ligands in complex with PIX, Compound 1 and Compound 2 display greater hydrogen bond interaction with PIX as can be observed in Figures 6.5 and 6.6, this may be attributed to the presence of benzohydrazide moiety present in each compound reflecting more than a single amine group presenting a greater number of hydrogen bond interactions.

Hydrogen bonds are key contributors to the specificity of receptor–ligand interactions. This distance is generally considered to be from 2.7 to 3.3 Å. The angle the bond forms is also indicative of the strength of the hydrogen bond. The closer the hydrogen bond is to correct geometry, the stronger the bond (D & R.C, 2007). In all the systems we observe hydrogen bond distances < 3.3 Å, however, Compound 1 and Compound 2 display the shortest hydrogen bond interactions in comparison to systems in complex with Compound 3 and 49c. The bond distance is not the only factor that affects the conformation of a ligand within the active site, the charge distribution during ligand binding remains essential, which is extended in all systems.

The lowest binding free energy is observed between PIX in complex with Compound 3 which may be linked to a range of polar interaction present within this system, emanating from the presence of fluorine only observed in Compound 3. Studies have shown hydrophobic interactions are more frequent in high-efficiency ligands(Ferreira De Freitas & Schapira, 2017), which remain present within PIX-Compound 3 system but not to the extent to which it is observed in the other three systems.

From the per-residue energy decomposition analysis of all systems, there is a concurrence among the amino acid residues that play a pivotal role in the binding of each ligand within the catalytic domain of PIX. These residues include, Glu58; Asp59; Gln61; Asp 80; Thr81; Asp259; Glu331 and Glu333. The interaction between Asp259 and Glu331 of PIX are observed between all ligands in each system and are further accentuated as the highest contributing amino acids towards the total binding energy in all systems. Studies have shown the crucial role of Asp 259 and Glu331 also identified as the second aspartic residue of the catalytic dyad and the hinge residue respectively, as they make up two of the four essential residues required for distortion of the “Twisting motion” exhibited by aspartic proteases.

6.6 Conclusion

Malaria remains a global threat amid continuous efforts made towards the prevention and eradication of this parasitic virus. The integration of in silico techniques in the field of drug design and development, enables a deeper understanding underlying molecular interactions integral in ligand-protein binding. The insight extracted from these molecular interactions may further supplement the interpretation of experimental results at an atomic-level. The use of computers in drug discovery aids in delivering new drug candidates at a faster rate whilst integrating cost efficiency. The techniques implemented in this study were selected for their robustness, efficacy, as well as for interpretation to enable efficient translation to medicinal chemistry through molecular design for lead optimization and potential drug development. This study proposes two new compounds with enhanced pharmacokinetic properties and adherence to “drug-likeness”. The lead compounds, namely Compound 1 and Compound 2, have displayed stability in complex with the target enzyme, prohibiting the typical pivoting motion of PIX from transpiring and stabilizing catalytic aspartic dyad however further experimental evaluation is necessary for efficacy and toxicity validation.

Conflicts of interest

There are no conflicts to declare

Acknowledgements

The authors acknowledge the National Research Foundation for their financial support and the Centre for High Performance Computing (<http://www.chpc.ac.za>) for their computational resources.

6.7 References

- Anslyn, E. V, & Dougherty, D. A. (2006). *Modern Physical Organic Chemistry*. (J. Murdzek, Ed.). California: University Science Books.
- Arodola, O. A., & Soliman, M. E. S. (2016). Molecular Dynamics Simulations of Ligand-Induced Flap Conformational Changes in Cathepsin-D — A Comparative Study. *Journal of Celular Biochemistry*, 15(February), 1–15. <https://doi.org/10.1002/jcb.25564>
- Asojo, O. A., Gulnik, S. V, Afonina, E., Yu, B., Ellman, J. A., Haque, T. S., ... Hall, B.-H. (2003). Novel Uncomplexed and Complexed Structures of plasmepsin II , an Aspartic Protease from *Plasmodium falciparum*. *Journal of Molecular Biology*, 2836(3), 173–181. [https://doi.org/10.1016/S0022-2836\(03\)00036-6](https://doi.org/10.1016/S0022-2836(03)00036-6)
- Banerjee, R., Liu, J., Beatty, W., Pelosof, L., Klemba, M., & Goldberg, D. E. (2002). Four plasmepsins are active in the *Plasmodium falciparum* food vacuole, including a protease with an active-site histidine. *Proceedings of the National Academy of Sciences*, 99(2), 990–995. <https://doi.org/10.1073/pnas.022630099>
- Berendsen, H. J. C., Postma, J. P. M., Gunsteren, W. F. Van, Dinola, A., Haak, J. R., Berendsen, H. J. C., ... Haak, J. R. (1984). Molecular dynamics with coupling to an external bath Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, 81(8), 3685–3689. <https://doi.org/10.1063/1.448118>
- Berry, C. (2000). New Targets for Antimalarial Therapy: The Plasmepsins, Malaria Parasite Aspartic Proteinases. *Biochemical Education*, 4412(97), 191–194.
- Cai, H., Kuang, R., Gu, J., Wang, Y., Texas, S., Infectious, E., ... Antonio, S. (2011). Proteases in Malaria Parasites - A Phylogenomic Perspective. *Current Genomics*, 12, 417–427.
- Case, D. A., Cheatham, T. E., Darden, T. O. M., Gohlke, H., Luo, R. A. Y., Merz, K. M., ... Woods, R. J. (2005). The Amber Biomolecular Simulation Programs. *Journal of Computational Chemistry*, 26, 1668–1688. <https://doi.org/10.1002/jcc.20290>
- Cells, R. E. D., Rug, M., Cyrklaff, M., Mikkonen, A., Lemgruber, L., Kuelzer, S., ... Cowman, A. F. (2016). Regular Article Export of virulence proteins by malaria-infected erythrocytes involves remodeling of host actin cytoskeleton. *BLOOD Journal of Hematology*, 124(23), 3459–3469. <https://doi.org/10.1182/blood-2014-06-583054>.The
- Chang, C. A., Chen, W., & Gilson, M. K. (2007). Ligand configurational entropy and protein

- binding. *Proceeding of National Academy of Sciences*, 104(5), 1534–1539.
- Coombs, G. H., Goldberg, D. E., Klemba, M., Berry, C., Kay, J., Mottram, J. C., & Mottram, J. C. (2001). Aspartic proteases of *Plasmodium falciparum* and other parasitic protozoa as drug targets. *Trends in Parasitology*, 17(11), 532–537.
- Cowman, A. F., Berry, D., & Baum, J. (2012). The cellular and molecular basis for malaria parasite invasion of the human red blood cell. *Journal of Cell Biology*, 198(6), 961–971. <https://doi.org/10.1083/jcb.201206112>
- Crunkhorn, S. (2017). Infectious disease: Blocking malaria parasite invasion and egress. *Nature Publishing Group*, 2017. <https://doi.org/10.1038/nrd.2017.253>
- D, M., & R.C, W. (2007). Computer-Assited Drug Design. In J. B. Taylor & T. David J (Eds.), *Comprehensive Medicinal Chemistry II* (2nd ed.). New York: Elsevier Science.
- Daina, A., Michielin, O., & Zoete, V. (2017). SwissADME : a free web tool to evaluate pharmacokinetics , drug- likeness and medicinal chemistry friendliness of small molecules. *Nature Publishing Group*, (March), 1–13. <https://doi.org/10.1038/srep42717>
- Dash, C., Kulkarni, A., Dunn, B., & Rao, M. (2003). Aspartic Peptidase Inhibitors : Implications in Drug Development Aspartic Peptidase Inhibitors : Implications in Drug Development. *Critical Reviews in Biochemistry and Molecular Biology*, 38(2), 89–119. <https://doi.org/10.1080/713609213>
- Davidchack, R. L., Handel, R., Tretyakov, M. V, Davidchack, R. L., Handel, R., & Tretyakov, M. V. (2009). Langevin thermostat for rigid body dynamics Langevin thermostat for rigid body dynamics. *The Journal of Chemical Physics*, 234101(130). <https://doi.org/10.1063/1.3149788>
- Deu, E. (2017). Proteases as antimalarial targets : strategies for genetic , chemical , and therapeutic validation. *The FEBS jJournal*, 284, 2604–2628. <https://doi.org/10.1111/febs.14130>
- Du, X., Li, Y., Xia, Y., Ai, S., Liang, J., Sang, P., & Ji, X. (2016). Insights into Protein – Ligand Interactions : Mechanisms , Models , and Methods. *International Journal of Molecular Sciences*, 17(i), 1–34. <https://doi.org/10.3390/ijms17020144>
- Erik, J. (2010). *Electrostatics in proteins and protein-ligand complexes* . Dublin.

- Ersmark, K., Samuelsson, B., & Hallberg, A. (2006). Plasmepsins as Potential Targets for NewAntimalarial Therapy, 26(5), 626–666. <https://doi.org/10.1002/med.20082>
- Favourite N, C., Ramesh, M., & Mahmoud E.S, S. (2016). Per-residue energy decomposition pharmacophore model to enhance virtual screening in drug discovery : a study for identification of reverse transcriptase inhibitors as potential anti-HIV agents. *Drug Design, Development and Therapy*, 10, 1365–1377.
- Ferreira De Freitas, R., & Schapira, M. (2017). A systematic analysis of atomic protein-ligand interactions in the PDB. *MedChemComm*, 8(10), 1970–1981. <https://doi.org/10.1039/c7md00381a>
- Galindo-murillo, R., Robertson, J. C., Zgarbova, M., Jir, S., Otyepka, M., Jurec, P., & Cheatham, T. E. (2016). Assessing the Current State of Amber Force Field Modifications for. *Journal of Chemical Theory and Computation*, 12, 4114–4127. <https://doi.org/10.1021/acs.jctc.6b00186>
- Haldar, K., Bhattacharjee, S., & Safeukui, I. (2018). Drug resistance in *Plasmodium*. *Nature Publishing Group*, 1–15. <https://doi.org/10.1038/nrmicro.2017.161>
- Irwin, J. J., & Shoichet, B. K. (2006). ZINC- A free database of commercially available compounds for virtual screening. *Journal of Chemical Information and Modeling*, 45(1), 177–182.
- Kaalia, R., Kumar, A., Srinivasan, A., & Ghosh, I. (2011). An Ab Initio Method for Designing Multi-Target Specific Pharmacophores using Complementary Interaction Field of Aspartic Proteases. *Molecular Informatics*, 1–15. <https://doi.org/10.1002/minf.201400157>
- Karubiu, W., Bhakat, S., McGillemie, L., & Soliman, M. E. S. (2015). Molecular BioSystems Flap dynamics of plasmepsin proteases : insight dynamics †. *Molecular BioSystems*, 11, 1061–1066. <https://doi.org/10.1039/C4MB00631C>
- Lionta, E., Spyrou, G., Vassilatis, D. K., & Cournia, Z. (2014). Structure-Based Virtual Screening for Drug Discovery : Principles , Applications and Recent Advances. *Current Topics in Medicinal Chemistry*, 14, 1923–1938.
- Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and

- development settings. *Advanced Drug Delivery*, 46, 3–26.
- Liu, J., Gluzman, I. Y., Drew, M. E., & Goldberg, D. E. (2005). The Role of *Plasmodium falciparum* Food Vacuole Plasmepsins *. *The Journal of Biological Chemistry*, 280(2), 1432–1437. <https://doi.org/10.1074/jbc.M409740200>
- Machaba, K. E. (2017). Medicinal Chemistry. *Future Medicinal Chemistry*, 9, 1055–1071.
- Mao, Y. (2011). Dynamical Basis for Drug Resistance of HIV-1 Protease. *BMC Structural Biology*, 11(31), 1–9.
- Martin, Y. C., & Park, A. (2005). A Bioavailability Score. *Journal of Medical Chemistry*, 48, 3164–3170. <https://doi.org/10.1021/jm0492002>
- McGeorge, R. D. (2014). *Plasmodium falciparum Plasmepsins IX and X: Structure- function analysis and the discovery of new lead antimalarial drugs*. Griffith University.
- McGillewie, L., Ramesh, M., & Soliman, M. E. (2017). Sequence , Structural Analysis and Metrics to Define the Unique Dynamic Features of the Flap Regions Among Aspartic Proteases. *The Protein Journal*, 12(0), 5–9. <https://doi.org/10.1007/s10930-017-9735-9>
- McGillewie, L., & Soliman, M. E. (2015). Flap flexibility amongst I, II, III, IV, and V: Sequence, structural, and molecular dynamic analyses. *Proteins: Structure, Function and Genetics*, (April), 1693–1705. <https://doi.org/10.1002/prot.24855>
- Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2010). AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *Journal of Computational Chemistry*, 30(16), 2785–2791. <https://doi.org/10.1002/jcc.21256>
- Moura, P. A., Dame, J. B., & Fidock, D. A. (2009). Role of *Plasmodium falciparum* Digestive Vacuole Plasmepsins in the Specificity and Antimalarial Mode of Action of Cysteine and Aspartic Protease Inhibitors. *Antimicrobial Agents and Chemotherapy*, 53(12), 4968–4978. <https://doi.org/10.1128/AAC.00882-09>
- Munsamy, G., Ramharack, P., & Soliman, M. E. S. (2018). RSC Advances depth look into the structural and functional features of the fl ap dynamics of plasmepsin IX and. *RSC Advances*, 8, 21829–21840. <https://doi.org/10.1039/C8RA04360D>
- Nair, D. N., Singh, V., Angira, D., & Thiruvengatam, V. (2016). Proteomics & Bioinformatics

- Structural Investigation and In-silico Characterization of Plasmepsins from *Plasmodium falciparum*. *Journal of Proteomics and Bioinformatics*, 9(7), 181–195. <https://doi.org/10.4172/jpb.1000405>
- Nasamu, A. S., Glushakova, S., Russo, I., Vaupel, B., Oksman, A., Kim, A. S., ... Goldberg, D. E. (2017). parasite egress and invasion. *Science*, 522(October), 518–522.
- Nigussie, D., Beyene, T., Shah, N. A., & Belew, S. (2015). Malaria Control & Elimination New Targets in Malaria Parasite Chemotherapy : A Review. *Malaria Control and Elimination*, S1(S1007), 1–6. <https://doi.org/10.4172/2470-6965/1000S1-007>
- Parr, C. L., Tanaka, T., Xiao, H., & Yada, R. Y. (2008). The catalytic significance of the proposed active site residues in *Plasmodium falciparum* histozoic protease. *FEBS Journal*, 275(8), 1698–1707. <https://doi.org/10.1111/j.1742-4658.2008.06325.x>
- Petrenko, R. (2010). Molecular Dynamics. *In: Encyclopedia of Life Sciences*. <https://doi.org/10.1002/9780470015902.a0003048.pub2>
- Pino, P., Caldelari, R., Mukherjee, B., Vahokoski, J., Klages, N., Maco, B., ... Soldati-favre, D. (2017). for invasion and egress. *Science*, 528(October), 522–528.
- Rosenthal, P. J. (1998). Proteases of Malaria Parasites : New Targets for Chemotherapy. *Emerging Infectious Diseases*, 4(1), 49–57.
- Saddala, M. S., & Adi, P. J. (2018). Discovery of small molecules through pharmacophore modeling , docking and molecular dynamics simulation against *Plasmodium vivax*. *Heliyon*, 4(February), e00612. <https://doi.org/10.1016/j.heliyon.2018.e00612>
- Tamar, S. (2002). *Molecular Modelin and Simulation: An Interdisciplinary Guide* (2nd ed.). New York.
- Wang, J., Hou, T., Li, Y., & Wang, W. (2012). Assessing the performance of the MM/PBSA and MM/GBSA medthods: I. The accuracy of binding free energy calculations based on molecular dynamics simulations. *Journal of Chemical Information and Modelling*, 51(1), 69–82. <https://doi.org/10.1021/ci100275a.Assessing>
- Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., & Case, D. A. (2004). Development and Testing of a General Amber Force Field. *Journal of Computational Chemistry*, 25, 1157–1174.

World Health Organization. (2016). *Global technical strategy for malaria 2016–2030*. Retrieved from http://apps.who.int/iris/bitstream/handle/10665/176712/9789241564991_eng.pdf?sequence=1

World Health Organization. (2017). *World malaria report 2017*. Retrieved from <http://www.who.int/malaria/publications/world-malaria-report-2017/report/en/>

Wright, D. W., Hall, B. A., Kenway, O. A., Jha, S., & Coveney, P. V. (2014). Computing Clinically Relevant Binding Free Energies of HIV - 1 Protease Inhibitors. *Journal of Chemical Theory and Computation*, 10, 1228–1241.

Ylilauri, M., & Pentikäinen, O. T. (2013). MM/GBSA as a tool to understand the binding affinities of filamin-peptide interactions. *Journal of Chemical Information and Modeling*, 53(10), 2626–2633. <https://doi.org/10.1021/ci4002475>

CHAPTER 7

Fundamental class of inhibitors activates “fireman’s grip”: An enhanced binding analysis in search of effective malarial therapeutics.

Geraldene Munsamy^a, Clement Agoni^a, Pritika Ramharack^a, Mahmoud E. S. Soliman^{a*}

^aMolecular Bio-computation and Drug Design Laboratory, School of Health Sciences,
University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa

*Corresponding Author: Mahmoud E.S. Soliman

Email: soliman@ukzn.ac.za

Telephone: +27 (0) 31 260 8048, Fax: +27 (0) 31 260 7872

7.1 Abstract

The past decade has witnessed unprecedented efforts to control the invasive tactics of the malarial parasite, including focused research in the design of inhibitors against selective targets of *Plasmodium falciparum*. Regrettably, the infectious disease still represents a major health burden, particularly in Africa. Of the parasitic targets, prominent members of the aspartic protease family, PIX and PX, have gained attention in drug discovery studies due to its essential role in erythrocyte invasion. As these enzymes have only recently been identified as key mediators of egress and invasion of the malarial parasite, an effective inhibitor is yet to be brought to light. In our previous study, a hydroxylethylamine inhibitor was reported to demonstrate structural inhibitory characteristics towards PIX and PX however, the drug exhibited poor pharmacokinetic properties. This augmented our search for a potent and effective inhibitor by the utilization of *in silico* tools. In this study, we effectuated a comparative structural analysis on an aminohydantoin derived inhibitor, CWHM-117, based on its nanomolar inhibitory activity against *Plasmodium* aspartic proteases. Pharmacokinetic analysis revealed enhanced solubility, lipophilicity and oral bioavailability. Active site sequence analysis demonstrated aspartic acid residues; Asp80, Asp259, Asp53 and Asp232, remained conserved upon binding, whilst creating a network of hydrogen bonds known as a “fireman’s grip”. As a characteristic protease inhibitor, CWHM-117 triggered restriction of the flap and hinge residues by restraining cohesive movement, consequently hindering an activating “twisting motion” from occurring. From the investigations carried out in this study, CWHM-117 has established structural inhibitory superiority, thus advancing the understanding of aminohydantoins in malaria therapy.

Keywords: Pharmacokinetic profiling; Aminohydantoin; Malarial therapy; PIX; PX; Flap dynamics; Aspartic proteases

7.2 Background

Malaria is a deadly protozoan infectious disease, that infects 200 to 500 million people resulting in an annual mortality of 1 to 2 million persons[1]. The World malaria report 2018 estimated 219 million cases of malaria in the year 2017. In Africa, the 10 countries that experience the highest Malarial burden observed an estimated increase of 3.5 million more cases of malaria incidence in the year 2017[2]. Although the last two decades have seen a surge in antimalarial drug development focused on the advancement of current therapeutics take on a leading role[3]. The availability of several effective antimalarial drugs such as chloroquine, sulfadoxine and artemisinin[3], have resulted in the onset of increased drug resistance of the malaria parasite, which now necessitates an urgent need for development and design of potent antimalarial compounds that will render a final knock against recently discovered novel targets[4,5].

Malaria transmission is mediated by the female *Anopheles* mosquito and is infected to humans by six *Plasmodium* species. Of the six species *Plasmodium falciparum* is the most lethal species[6] and remains endemic in Africa. Plasmepsins are a group of aspartic proteases that are essential for the survival and growth of the malaria parasite[6–8]. The *Plasmodium falciparum* species constitute more than 10 genes possibly encoding plasmepsins or related enzymes that have been identified. The four highly homologous *P. falciparum* Plasmepsins, Plm I, II, IV and histo-aspartic protease (HAP) play a vital role in hemoglobin degradation and have been considered as potential targets for antimalarial drug development[9–13]. However, these enzymes have been studied extensively over the decades. To avoid redundancy, this study focuses on PIX and PX released in the mature blood-stage schizonts of the parasite hitherto functions have been unknown. PIX and PX have recently been discovered to play an imperative role in the egress and invasion of the malaria parasite within the human host[14].

The only known inhibitors against PII and V is pepstatin A, other aspartic protease inhibitors targeting HIV proteases include lopinavir and ritonavir. These inhibitors have been associated with poor pharmacokinetic properties such as large molecular weights, polymorphic drug properties and poor peptidomimetic characteristics and thus make relatively poor starting points for the development of effective antimalarial drugs[15]. Recent optimization of the antimalarial activity of a hydroxylethylamine inhibitor known as 49c was reported. Although inhibitor 49c displayed potent inhibitory activity against PIX and PX, the pharmacokinetic analysis of 49c projected poor solubility and high molecular weight, all of which are undesirable properties of this hydroxylethylamine inhibitor[16,17].

Aminohydantoin is a class of antimalarial inhibitors which display inhibitory activity against aspartic protease family of enzymes[18]. Structure activity relationship analyses performed on the aminohydantoin class of antimalarial compounds unveil its potent nanomolar inhibitory activity against the *Plasmodium* aspartic proteases Plm II and Plm IV and likely inhibitory activity against other *Plasmodium* aspartic acid proteases. Therefore, the integration of this potent chemotype as a starting point for the development of effective inhibitors against the aspartic protease family of enzymes may lead to a permanent cure, with limited drug resistance in malaria.

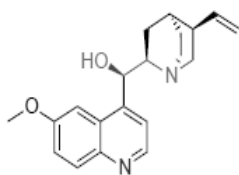
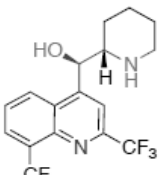
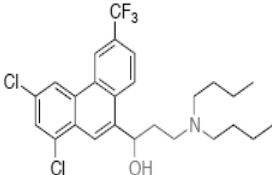
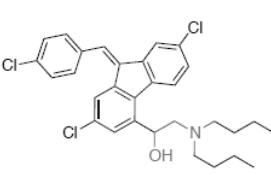
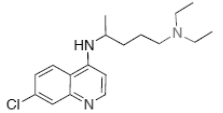
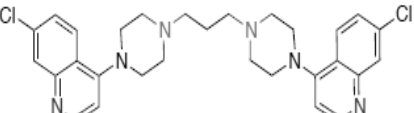
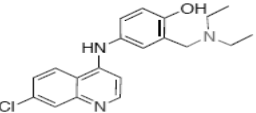
The use of atomistic computer simulations may provide unique insights into the microscopic aspects of binding of PIX and PX to aminohydantoin inhibitor CWHM-117, providing insight into the intermolecular interactions showing promise in advancing the understanding the interaction of this class of inhibitors at a molecular level. In this study we also integrate the use of *in silico* tools to analyse the pharmacokinetic profile of aminohydantoin inhibitor CWHM-117 and map out the binding landscape for the development of potential antimalarial therapeutics.

7.3 Characterizing Existing Malarial Aspartic Protease Inhibitors

Given the threat of drug resistance, there is an acute need for new classes of antimalarial agents that act via a unique mechanism of action relative to currently used drugs[15]. Whilst still maintaining appropriate ADME (absorption, distribution, metabolism and excretion properties with minimal drug-induced toxicity levels). In the pharmaceutical industry this has become a major hurdle in the drug design and discovery process as it is more desirable that compounds with poor pharmacokinetic properties be removed early in the discovery phase rather than during the costlier drug development phases[19]. Current malarial therapeutic drugs are listed in **Table 7.1**, although these drugs have proved to be effective in the treatment of malaria they are often associated with the emergence of drug resistance and adverse side effects. A new approach in Malaria treatment is the combination of current drugs to limit the emergence of resistance, however, this technique is not infallible. Despite the urgent need, most antimalarial drugs in late stage clinical development do not target new mechanisms of action. Finding a permanent therapeutic cure for Malaria lies in the need to identify antimalarial drugs with novel chemotype and mechanism of action. The aspartic protease family of enzymes proves to be

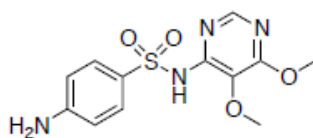
good starting point for drug discovery due to the role they play in a host of diseases and more especially in Malaria.[20] The function of the latter aspartic proteases namely, IX and X have been identified as key targets which may lead to the complete eradication of this viral parasite. The discovery of three chemotype scaffolds which pose as potential starting point for the design of aspartic protease inhibitors include, hydroxyethylamines, Aminohydantoin and Spiropiperidines. Aminohydantoin propose to be the most formidable drug scaffold for the development of aspartic protease inhibitors due to its oral bioavailability and demonstration of its druglike nature chemotype and low inhibitory concentration against aspartic protease family members PII and PV.

Table 7.1 Currently available anti-malaria drugs

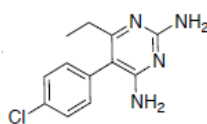
<u>Aryl-amino Alcohols</u>			
These drugs are associated with adverse cardiac effects. The high price of this class of drugs has also limited its use[21].			
			
Quinine	Mefloquine	Halofantrine	Lumefantrine
<u>4-Aminoquinolines</u>			
The extensive spread of parasite resistance has severely limited its use. These drugs were first recognized in the second world war. Their initial use was associated with toxicity in humans[22]			
			
Chloroquine	Piperaquine	Amodiaquine	

Antifolates

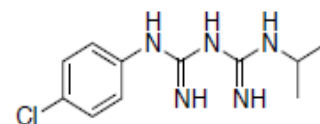
A combination of these drugs are currently the first line treatment for Malaria in many parts of Africa[23]



Sulfadoxine



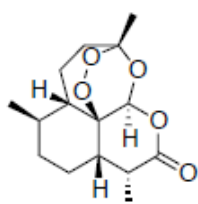
Pyrimethamine



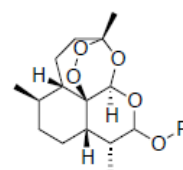
Proguanil

Peroxides

Historically these drugs were used in China as treatment for Malaria. Antimalarial action is mediated by free radicals and involves covalent linkage to the parasite membrane, protein and heme[24].



Artemisinin

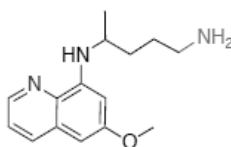


Dihydroartemisinin R=H/ Artemether R=Me

Artesunate R= COCH₂CH₂CO₂⁻

8-Aminoquinolines

These drugs contain a radical cure targeting the killing of hypnozoites. These drugs target the parasite mitochondria[25]



Primaquine

7.4 Assessing the paradigm of Aminohydantoin in malaria therapy

Structure activity relationship analysis performed on the aminohydantoin class of antimalarial compounds unveils their potent nanomolar inhibitory activity against the *Plasmodium* aspartic proteases Plm II and PIV and likely inhibitory activity against other *Plasmodium* aspartic acid proteases. Experimental studies performed on aminohydantoin compound CWHM-117 reveal its antimalarial properties associated low molecular weight, modest lipophilicity, oral bioavailability, and *in vivo* antimalarial activity in mice. Therefore, highlighting this chemotype as an indisputable starting point for the development of effective inhibitors against the aspartic protease family of enzymes.

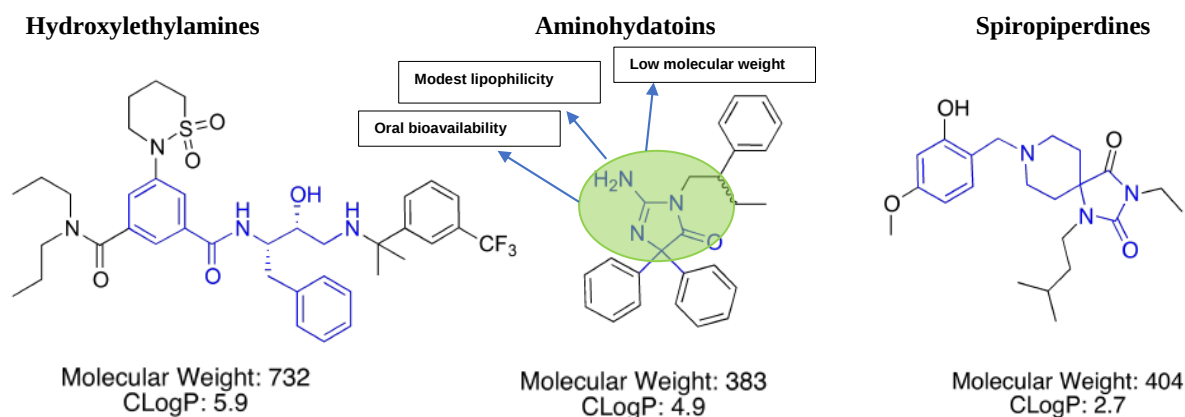


Figure 7.1 General molecular scaffold and comparative characteristics of the three main classes of malaria protease inhibitors.

7.5 Plm and CWHM-117 acquisition and preparation

A protein's 3D structure is often required for functional and structural characterization of a protein, which is related to its biological activity[26]. However, the fraction of experimentally known 3D models is currently less than 1% due to the inherently time-consuming and complicated nature of structure determination techniques. In this study the homology model of PIX was generated using the PHYRE2 server, a comprehensive protocol used in the generation of the PIX and PX models are explicated in our previous publication [27].

The ligand CWHM-117 was generated using the ChemDraw Ultra 9.0. program. The ligand was then optimized integrating the steepest descent method of the MMFF94S force field within the Avogadro software. The Chimera and GaussView 5.0 software's were employed to ensure the correct angle bonds and hybridization in were correct in the optimized structure of the ligand.

7.5.1 Molecular Docking

The Molecular docking software utilized in this study included Raccoon [28], Autodock Graphical user interface supplied by MGL tools [29] and AutoDockVina [30] with default docking parameters. Prior to docking, Gasteiger charges were added to CWHM-117 and the non-polar hydrogen atoms were merged to carbon atoms. The CHWM-117 compound was then docked into the binding pocket of PIX and PX (by defining the grid box with spacing of 1 Å and size of 46 × 60 × 66 for PIX and 56 × 55 × 68 for PX respectively pointing in x, y and z

direction). The best-docked pose for each compound was then subjected to molecular dynamic simulations.

7.5.2 Molecular Dynamic (MD) Simulations

Molecular dynamic (MD) simulations provide insight into the dynamical evolution of biological systems by exploring the physical movements of atoms and molecules[31]. The GPU version of the PMEMD engine provided with the AMBER package [32] were integrated to run the MD simulation on all systems.

The atomic partial charges for the ligands were generated using the ANTECHAMBER program which implements the General Amber Force Field (GAFF)[33]. The Leap module of AMBER 14[34] was used to implicitly solvate each system in an orthorhombic box of TIP3P water molecules such that all atoms were within 10Å of any box edge. Leap was further implemented for the addition of Na⁺ or Cl⁻ counter ions to ensure neutralization of all systems prior to the production run[35].

An initial minimization of 2000 steps was carried out with an applied restraint potential of 500 kcal/mol. This was subsequently followed by a full minimization of 1000 steps carried out by conjugate gradient algorithm in the absence of all restraints.

Each system underwent gradual heating ranging from 0K to 300K for 50ps, to ensure that all systems maintained a fixed volume and number of atoms. The solutes within all systems underwent a potential harmonic restraint of 10kcal/mol and collision frequency of 1ps. Each system was then equilibrated using an equilibration step estimate of 500ps at a constant temperature of 300K. The of number of atoms and pressure within each system were kept constant, mimicking an isobaric-isothermal ensemble (NPT). The systems pressure was maintained at 1 bar using the Berendsen barostat.

The total time for the MD simulation of each system was 100ns. The SHAKE algorithm was employed in the simulation of each system to constrict the bonds of hydrogen atoms. The step size of each simulation was 2fs integrating the SPFP precision model. The simulations coincided with isobaric-isothermal ensemble (NPT), with randomized seeding, constant pressure of 1 bar, a pressure-coupling constant of 2ps, a temperature of 300K and Langevin thermostat with collision frequency of 1ps.

7.6 Post-Dynamic Analysis

The coordinates of the free enzyme and bound complexes were then saved after every 1ps and the trajectories were analysed using the CPPTRAJ module employed in AMBER 14 suit. The Root Mean Square Deviation (RMSD) and thermodynamic energy of each system was then investigated.

7.6.1 Binding Free Energy Calculations

The Molecular Mechanics/GB Surface Area method (MM/GBSA) [36] was implemented to estimate the binding free energy of each of the systems,. Binding free energy was averaged over 20000 snapshots extracted from the 100ns trajectory.

The free binding energy (ΔG) computed by this method for each molecular species (complex, ligand and receptor) can be represented as:

$$\Delta G_{\text{bind}} = G_{\text{complex}} - G_{\text{receptor}} - G_{\text{ligand}} \quad (1)$$

$$\Delta G_{\text{bind}} = E_{\text{gas}} + G_{\text{sol}} - TS \quad (2)$$

$$E_{\text{gas}} = E_{\text{int}} + E_{\text{vdw}} + E_{\text{ele}} \quad (3)$$

$$G_{\text{sol}} = G_{\text{GB}} + G_{\text{SA}} \quad (4)$$

$$G_{\text{SA}} = \gamma \text{SASA} \quad (5)$$

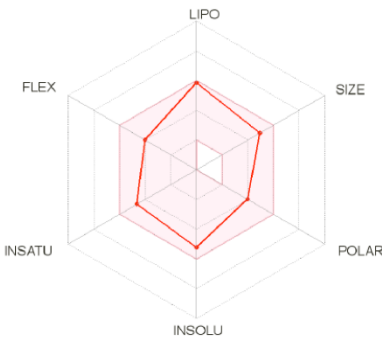
The term E_{gas} represents the gas-phase energy, comprising of the internal energy E_{int} ; Coulomb energy E_{ele} and the van der Waals energies E_{vdw} . The E_{gas} was directly estimated from the FF14SB force field terms. Solvation free energy, G_{sol} , was estimated from the energy contribution from the polar states, G_{GB} and non-polar states, G . The non-polar solvation energy, G_{SA} , was estimated from the solvent accessible surface area (SASA), using a water probe radius of 1.4 Å, whereas the polar solvation, G_{GB} , contribution was estimated by solving the GB equation. S and T denote the total entropy of the solute and temperature respectively.

7.7 Results and Discussion

7.7.1 Pharmacokinetic Profiling of CWHM-117 demonstrates favourable drug-likeness

In silico tools to investigate absorption, distribution, metabolism, excretion, and pharmacokinetics (ADME-PK) properties of new chemical entities are an integral part of the current industrial drug discovery paradigm[37]. In our study we performed pharmacokinetic drug profile analysis on CWHM-117 using the SWISSADME online tool[35] which is presented in **Table 7.2**.

Table 7.2 Pharmacokinetic characteristics of CWHM-117 using SWISSADME

CWHM-117								
Molecular Formula	Molecular Weight (g/mol)	Lipophilicity (iLOGP)	Water Soluble	GIT Absorption	BBB Permeability	Bioavailability Score	Synthetic Accessibility	Druglikeness (Lipinski)
C ₂₄ H ₂₉ N ₃ O ₃	407.51	3.60	Yes	High	Yes	0.55	3.59	Yes
“Boiled-Egg” Method Summary 								

7.7.2 Comparative bioinformatic analysis of Plasmepsins demonstrate greater interactions with Aminohydantoins

Amino acid sequence analysis has been extensively used to understand the functional relationship and molecular evolution of proteins [38]. In this study, we aligned the full amino acid sequences of PIX and PX to identify conserved amino acids when bound to 49C and CWHM-117. The binding landscape and final energy of 49c bound to PIX and PX was discussed elaborately in our previous publication and will therefore not be re-iterated in this study. The residues of the 49C-complexes are, however, demonstrated for comparative purposes. **Figure 7.2** exhibits the binding site residues of 49C and CWHM-117 bound to PIX and PX.

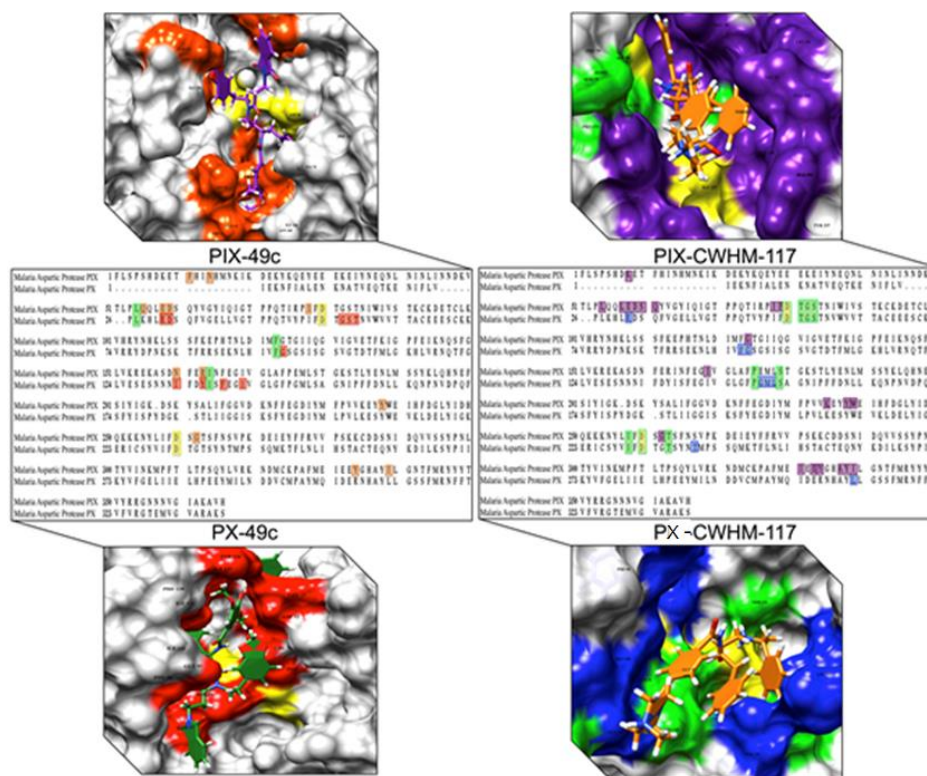


Figure 7.2 : Comparative sequence analysis of PIX and PX subsequent to binding of 49C and CWHM-117. Aspartic acid residues Asp80, Asp259, Asp53 and Asp232 remain conserved among all complexes (yellow). Highlighted residues show active site interactions

It was interesting to note that although binding of the compounds occurred at varying spatial landscapes, the key aspartic residues, being Asp80 and Asp259 for PIX and Asp53 and Asp232 for PX, remained conserved in all complexes. Whilst comparing the amino acids of PIX, it was noted that Leu54, Glu58, Asp59, Ile78, Tyr237, Gly261, Tyr333 and Ile338 remained conserved upon binding of both 49C and CWHM-117. PX illustrated only five-conserved residue interactions (Arg31, Gly55, Ser56, Phe96 and Gly97) in both drug complexes. It is evident from the overall analysis that CWHM-117 interacts with a greater number of amino acid residues (30 and 17 interactions for PIX and PX, respectively) in comparison with 49C (17 and 15 interactions for PIX and PX, respectively). This comparative investigation further augments the need for a comprehensive structural analysis on the effects of CWHM-117 on PIX and PX in malaria therapy. In the subsequent discussion, the structural integrity, conformational characteristics and binding landscapes of PIX and PX, upon binding of CWHM-117, will be critically explored.

7.7.3 Sensitivity of PIX and PX to CWHM-117 result in transformed conformational ensembles

The average RMSD of the bound complexes PIX-CWHM-117 and PX-CWHM-117 is 2.79 Å and 2.25 Å respectively. The RMSD of the apo enzymes PIX and PX is 2.83 Å and 3.76 Å,

thus indicating a slight trend of increased flexibility of the apo enzymes. This may have been possibly due to the absence of bound ligands. The root mean square fluctuation (RMSF) profiles of the bound and apo systems for PIX and PX are presented in **Figure 7.3**. In the bound systems of PIX and PX the secondary structural elements present within the flap domain of each aspartic protease maintain a state of stability throughout the 100ns simulation. Molecular dynamics simulations of apo- as well as complexed PIX and PX structures show that the catalytic domain remains relatively stable in nature with respect to the core structure. The core of the complexed structures displays overall structural stability and accommodates the substrate in the active site unlike the apo systems of PIX and PX, in which the flap transitions between open and closed conformations throughout the 100ns simulation. Experimental studies performed to show the oral efficacy of CWHM-117 to PIX and PX exhibited a lower half maximal effective concentration to PX indicating greater hypersensitivity of PX to aminohydantoins as compared to PIX[14]. This hypersensitivity transcends into the observed RMSF of PX in complex with CWHM-117 which displays a substantially lower RMSF as compared to the PIX-CWHM-117 system.

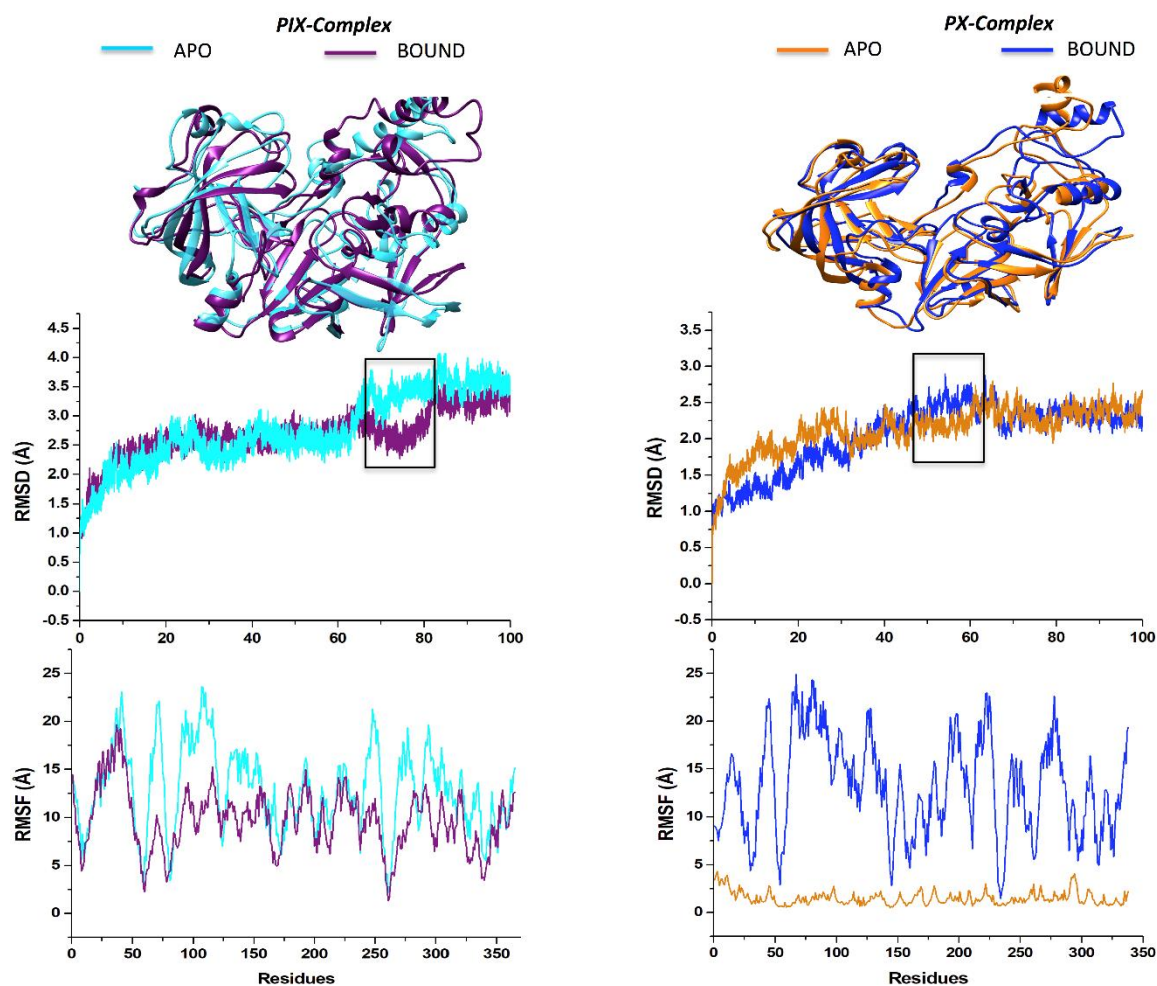


Figure 7.3 Stability of PIX and PX before and after the binding of CWHM-117. Conspicuous fluctuations are observed in the PIX systems at 70-80ns and in the PX systems between 45-65ns. This is validated by the structural modifications observed in the RMSF graphs

The Radius of Gyration (RoG) further indicates a more compact structure of the bound systems as opposed to the apo enzymes as can be observed in **Figure S7.1**. These observations are indicative of conformational changes associated with ligand binding and an opening closing of aspartic proteases PIX and PX, which maintain a close structure upon ligand binding. PX is characterized as having overall lower levels of flexibility and a more compact structure than family member PIX.

Figure 7.4 depicts the hydrogen bonds present in each Plm system, with the bound systems displaying a greater number of hydrogen bonds as compared to the apo systems. This may be attributed to a conserved network of hydrogen bonds, termed the “fireman’s grip”, which stabilizes the catalytic site structure and a β -hairpin turn, also known as the “flap” covers the binding cleft with the ability to interact with substrates and inhibitors.

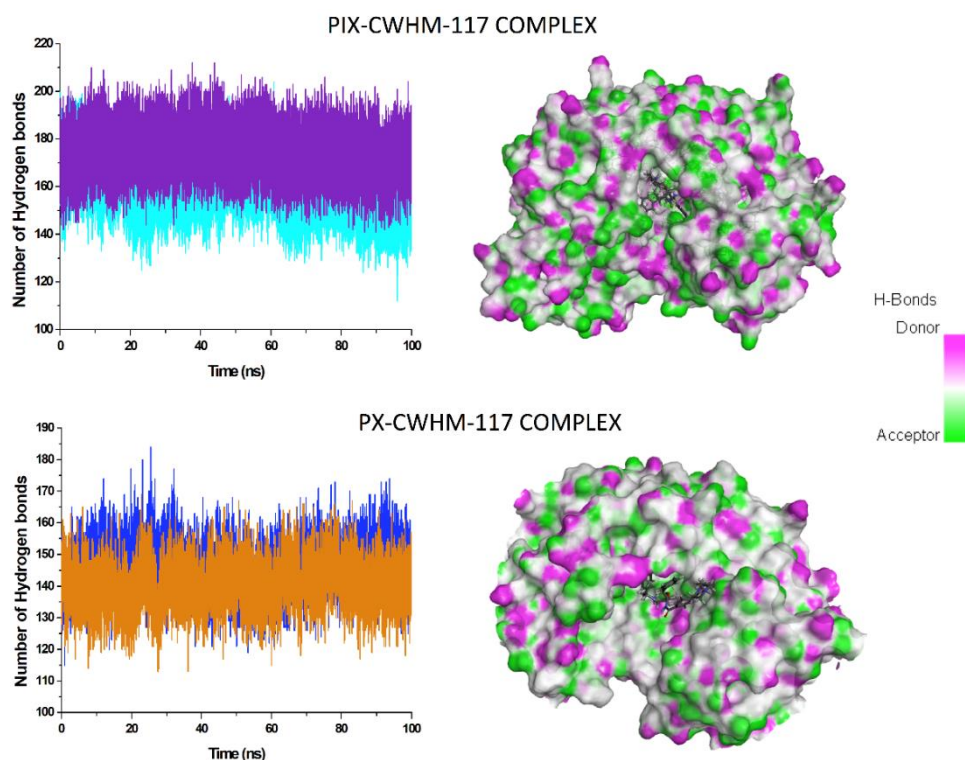


Figure 7.4 Stabilizing hydrogen bonds between Plm complexes. Complexes on the right depict hydrogen bond donors and acceptors upon binding with CWHM-117

It is evident from these results that there is greater stability in the PX-CWHM-177 complex, which correlates well with the RMSD and RMSF results. To further understand the dynamics of the Plm systems, the characteristic aspartic residues were assessed using distance analysis as presented in **Table 7.3**.

Table 7.3 Measurement of the distance by which the flap-structure moves, measured in Angstroms

	PIX-Apo	PX-Apo	PIX-117CHWM	PX-117CHWM
Maximum	22.42	21.17	20.87	21.81
Minimum	9.45	4.89	12.57	10.89
Average	16.21	12.57	15.98	16.32
Δ^a	12.97	16.28	8.3	10.92

^a change between maximum and minimum distance

Aspartic proteases are distinguished by the twisting motion of the flap and tip domain which is positioned parallel to the active site of the enzyme. The distance between the flap tip and hinge region of each system throughout the 100ns simulation is presented in **Figure S7.2**. The flap region in aspartic proteases is a unique structural feature to this class of enzymes and found to

have a profound impact on protein overall structure, function, and dynamics. **Table 7.3** shows the change in this region during the analysed trajectory. More importantly, our results characterize the range of flap conformation in an uninhibited form, showing distances that span 9.45-22.42 Å, and 4.89-21.17 Å in PIX and PX respectively, which indicative of a broad range of conformations in this uninhibited state exhibiting the opening and closing of the flap domain[39].

7.7.4 Enhanced binding prowess of CWHM-117 to PIX over PX

Using snapshots from the generated trajectory of the 100ns simulation of the bound complexes, MM/GBSA based free energy calculations for all simulated complexes were performed. This allowed for a quantitative assessment of binding of CWHM-117 to both PIX and PX as presented in **Table 7.4**. The binding free energy of CWHM-117 is higher when bound to PIX with a total binding free energy of -58.19kcal/mol relative to when bound to PX with a binding free energy of -31.66 kcal/mol. The superior binding free energy of CWHM-117 towards PIX suggests a possible superior inhibitory potency relative to PX upon further experimental validation. Other components of the binding free energy calculations reveal deep insights into the collective forces supporting the binding of CWHM-117. Prominent intermolecular interactions towards binding including van der Waals and electrostatic forces are also presented in Table with estimated values of -49.96kcal/mol and -393.51kcal/mol respectively. The favourable electrostatic interactions are neutralized by the polar solvation effect resulting in a much lower electrostatic contribution. As such, in both complexes a large van der Waals interactions contribute to the overall total binding of CWHM-117. The higher van der Waal interaction recorded in the CWHM-117-PIX complex therefore possibly contributed to the higher total binding estimated.

Table 7.4 MM/GBSA-based binding free energy profile of CWHM-117

Systems	Energy components (kcal/mol)				
	ΔE_{vdw}	ΔE_{ele}	ΔG_{gas}	ΔG_{sol}	ΔG_{bind}
PIX-C117	-49.96±0.20	-393.51±1.53	-443.47±0.56	385.28±0.56	-58.19±0.16
PX-C117	-31.19± 0.98	-280.20±0.61	-311.39±1.52	279.73±1.29	-31.66±0.38

ΔE_{ele} = electrostatic energy; ΔE_{vdw} = van der Waals energy; ΔG_{bind} = total binding free energy; ΔG_{sol} = solvation free energy ΔG = gas phase free energy

The energy contributions of individual residues within a 5Å radius of the CWHM-117 binding site play a very essential role into the overall binding of CWHM-117 towards both PIX and PX. As such, we calculated their individual energies using the MM/GBSA based approach. This allowed for the identification of very crucial residues to the inhibitory activity of CWHM-117 towards PIX and PX. **Figure 7.5 and 7.6** showcases the per-residue total energy contribution of binding site residues in both PIX and PX. Particularly residues with energies -1.0kcal/mol or less were considered very essential considering their prominence. Overall, the major interacting residues in the CWHM-117-PIX complex included; Met177(-1.03), Tyr237(-1.357), Ile257(-1.147), Asp259(-8.357)), Thr262 (-1.33) and Ile338(-1.353) whiles Pro148(-1.638), Met150(-1.945), Asp232(-4.312) and Thr235(-1.491) were the prominent energy contributing residues in the CWHM-117-PX complex. Asp259 was the highest energy contributing residue in PIX while Asp232 contributed the most in PX. Identifying these crucial residues and their respective interactions provides a bases for further structure-based drug design of new compounds which could possess enhanced inhibitory potency and minimal toxicity against PIX and PX. This could also allow for the construction of pharmacophore model for screening new of new hits from large databases. As shown in **Figure 7.5**, specific hydrophobic interaction that culminated into the overall superior binding of CWHM-117 towards PIX included notable conventional hydrogen bond, Pi-cation and Pi-anion interaction between residues such as Met177, Asp80, Asp59, Glu332 and Ile338. Comparatively, there were more hydrophobic interactions occurring in the CWHM-117-PIX complex which could have accounted for its higher total binding free energy relative the CWHM-117-PX complexes which exhibited few interaction as shown in **Figure 7.6** Notable hydrophobic intermolecular

interactions in the CWHM-117-PX complex included one conventional hydrogen bond with Asp232, two carbon hydrogen bonds with Gly149, a Pi-Pi T-shaped with Tyr210 and Pi alkyl interactions with Met150, Ile230 and Leu311.

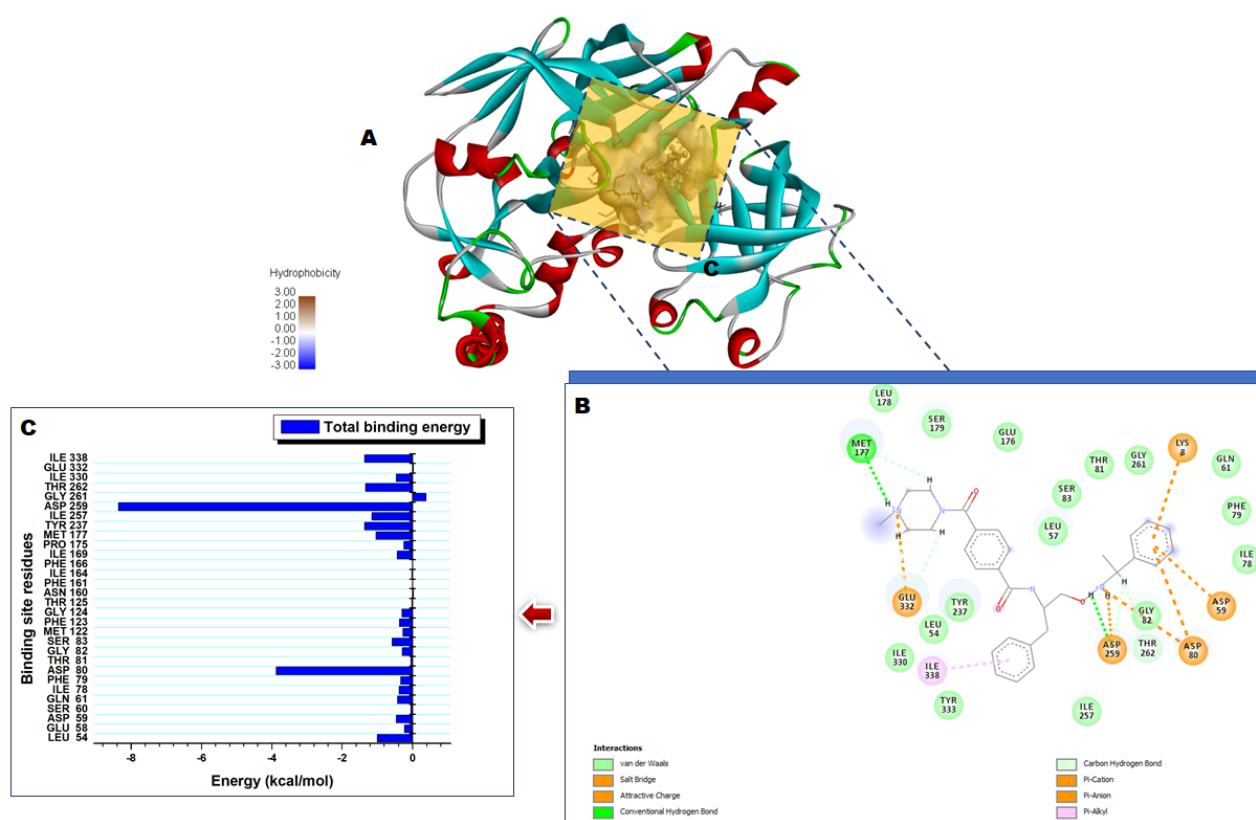


Figure 7.5 Binding free energy profile of PIX-CWHM-117 system. A) 3D predicted structure of PIX bound with CWHM-117. B) CWHM-PIX interaction network of binding site residues. C) Total contributing energies of binding site residues.

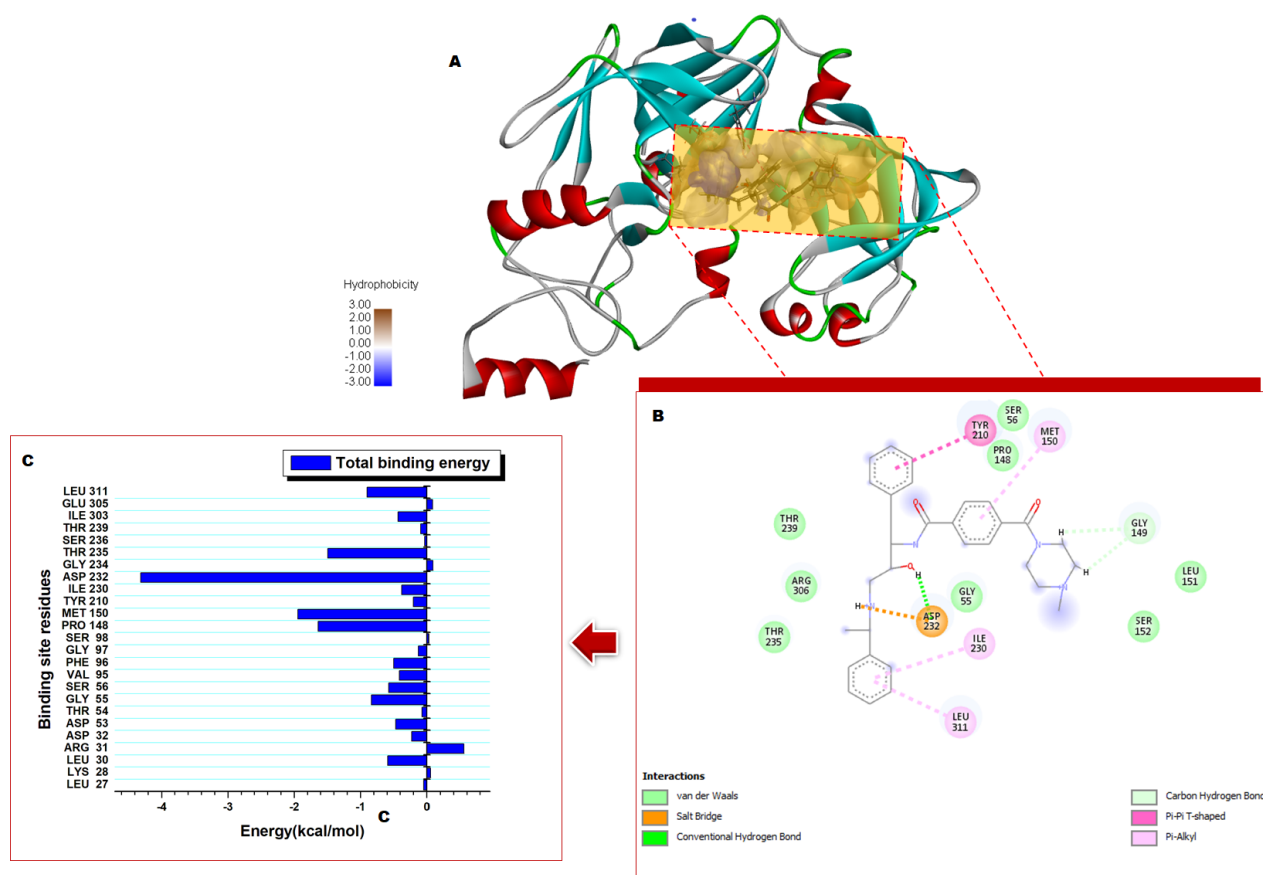


Figure 7.6 Binding free energy profile of PX-CWHM-117 system. A) 3D predicted structure of PIX bound with CWHM-117. B) CWHM-PX interaction network of binding site residues. C) Total contributing energies of binding site residues

7.5 Conclusion

In this study we aimed to provide structural and functional analysis of the aminohydantoin class of drugs as potential inhibitors against the enzymes responsible for egress and invasion of the malarial parasite within the human host. The pharmacokinetic profiling of the aminohydantoin class of inhibitors pose as ideal drug candidates due to their associated low molecular weight, modest lipophilicity, oral bioavailability, and *in vivo* antimalarial activity in mice. From the *in silico* analysis performed the overall binding landscape of PIX and PX to CWHM-117 displayed a stable trend in the RMSF, RoG, hydrogen bond and distance analysis which depict greater stability in the bound systems as opposed to the apo systems. Aminohydantoin derived inhibitor CWHM-117 displayed greater restriction of the prominent “twisting motion” of PX which remains essential for the catalytic activity of the enzyme mediated by the network of hydrogen bonds also known as the “fireman grip” which stabilizes the catalytic site and forms crucial hydrogen bond interactions in the flap and hinge region of the enzyme. To this end, designing novel inhibitors, which can interact with the flap tips as well as catalytic aspartic residues, will be important in retaining the closed conformation of

PIX and PX. This can be achieved by integrating the optimization of functionalized small molecules derived from the aminohydantoin class of drugs which will provide a new avenue for the discovery of antimalarial agents.

Competing Interests

Authors declare there are no competing interest

Funding

The authors would like to acknowledge the National Research Foundation for their financial support (grant number:115451)

Acknowledgements

The Centre for High Performance Computing for providing computational resources (<http://www.chpc.ac.za>)

7.6 References

- [1] Ashley EA, Ashley EA. Drugs in Development for Malaria. *Drugs* [Internet]. 2018;78:861–879. Available from: <https://doi.org/10.1007/s40265-018-0911-9>.
- [2] World Health Organization. World malaria report 2018. Geneva; 2018.
- [3] Biamonte MA, Wanner J, Le KG. Bioorganic & Medicinal Chemistry Letters Recent advances in malaria drug discovery. *Bioorg. Med. Chem. Lett.* [Internet]. 2013;23:2829–2843. Available from: <http://dx.doi.org/10.1016/j.bmcl.2013.03.067>.
- [4] Bedi RK, Patel C, Mishra V, et al. Understanding the structural basis of substrate recognition by *Plasmodium falciparum* plasmepsin V to aid in the design of potent inhibitors. *Sci. Rep.* [Internet]. 2016;1–17. Available from: <http://dx.doi.org/10.1038/srep31420>.
- [5] World Health Organization. World malaria report 2017. Geneva; 2017.
- [6] Cai H, Kuang R, Gu J, et al. Proteases in Malaria Parasites - A Phylogenomic Perspective. *Curr. Genomics*. 2011;12:417–427.
- [7] Farrow RE, Green J, Katsimitsoulia Z, et al. Seminars in Cell & Developmental Biology The mechanism of erythrocyte invasion by the malarial parasite , *Plasmodium falciparum*. *Semin. Cell Dev. Biol.* [Internet]. 2011;22:953–960. Available from: <http://dx.doi.org/10.1016/j.semcdb.2011.09.022>.
- [8] Nkumama IN, Meara WPO, Osier FHA. Changes in Malaria Epidemiology in Africa and New Challenges for Elimination. *Trends Parasitol.* [Internet]. 2017;33:128–140. Available from: <http://dx.doi.org/10.1016/j.pt.2016.11.006>.
- [9] Karubiu W, Bhakat S, McGillevie L, et al. Molecular BioSystems Flap dynamics of plasmepsin proteases : insight dynamics †. *Mol. Biosyst.* [Internet]. 2015;11:1061–1066. Available from: <http://dx.doi.org/10.1039/C4MB00631C>.
- [10] Mahanti M, Bhakat S. Flap Dynamics in Aspartic Proteases : A Computational Perspective. *Chem. Biol. Drug Des.* 2016;1–19.
- [11] Berry C. New Targets for Antimalarial Therapy: The Plasmepsins, Malaria Parasite

- Aspartic Proteinases. *Biochem. Educ.* 2000;4412:191–194.
- [12] Bhaumik P, Gustchina A, Wlodawer A. Structural studies of vacuolar plasmepsins. *Biochim. Biophys. Acta.* 2013;1824:207–223.
- [13] McGillem L, Soliman ME. Flap flexibility amongst I, II, III, IV, and V: Sequence, structural, and molecular dynamic analyses. *Proteins Struct. Funct. Genet.* 2015;1693–1705.
- [14] Nasamu AS, Glushakova S, Russo I, et al. parasite egress and invasion. *Science* (80-.). 2017;522:518–522.
- [15] Meyers MJ, Tortorella MD, Xu J, et al. Evaluation of Aminohydantoin as a Novel Class of Antimalarial Agents. *ACS Med. Chem. Lett.* 2014;5:6–10.
- [16] Munsamy G, Agoni C, Soliman MES. A dual target of plasmepsin IX and X : Unveiling the atomistic superiority of a core chemical scaffold in malaria therapy. *J. Cell. Biochem.* 2018;1–12.
- [17] Singh AK, Rathore S, Tang Y, et al. Hydroxyethylamine Based Phthalimides as New Class of plasmepsin Hits : Design , Synthesis and Antimalarial Evaluation. *PLoS One.* 2015;28:1–20.
- [18] Aneja. B, Kumar. B, Jairajpuri. AM, et al. Structure guided drug-discovery approach towards identification of *Plasmodium* inhibitors. *R. Soc. Chem.* 2013;00:1–50.
- [19] Lin J, Diana C, Morais SMF De, et al. The Role of Absorption , Distribution , Metabolism , Excretion and Toxicity in Drug Discovery. *Curr. Top. Med. Chem.* 2003;3:1125–1154.
- [20] Pratap S, Badhauria RS, Gurjar H. Docking studies of Aminohydantoin derivatives as antimalarial agents. 2018;8:322–326.
- [21] Duncan. GW, Colwell. TW, Scott. RC, et al. Arylamino alcohols antimalarials. A new method for incorporating the side chain. *J. Org. Chem.* 1985;21:9–11.
- [22] Pussard E, Verdier F. Antimalarial 4-aminoquinolines : mode of action and pharmacokinetics. *Fundam. Clin. Pharmacol.* 1994;8:1–17.
- [23] Gregson A, Plowe C V. Mechanisms of Resistance of Malaria Parasites to Antifolates. *Pharmacol. Rev.* 2005;57:117–145.

- [24] Chem M, Slack RD, Jacobine M, et al. MedChemComm Antimalarial peroxides : advances in drug discovery and design. *Med. Chem. Commun.* 2012;3:281–297.
- [25] Grewal RS. Pharmacology of 8-aminoquinolines. *Bull. World Health Organ.* 1981;59:397–406.
- [26] Webb B, Sali A. Comparative Protein Structure Modeling Using Modeller. *Model. Struct. from Seq.* New York: John Wiley and Sons; 2014. p. 1–32.
- [27] Munsamy G, Ramharack P, Soliman MES. Egress and invasion machinery of malaria: an in depth look into the structural and functional features of the fl ap dynamics of plasmepsin IX and X. *RSC Adv.* 2018;8:21829–21840. Available from: <http://dx.doi.org/10.1039/C8RA04360D>.
- [28] Ramirez D, Caballero J. Is it reliable to use common molecular docking methods for comparing the binding affinities of enantiomer pairs for their protein target? *Int. J. Mol. Sci.* 2016;17:1–15.
- [29] Sanner MF. The Python interpreter as a framework for integrating scientific computing software-components. *Scripps Res. Inst.* 2008;26:1–12.
- [30] Trott O, Olson AJ. AutoDock Vina. *J. Comput. Chem.* 2010;31:445–461.
- [31] Manuscript A. Molecular dynamics study of the ribosomal A-site. 2009;112:15227–15243.
- [32] Nair PC, Miners JO. Molecular dynamics simulations: from structure function relationships to drug discovery. *Silico Pharmacol.* 2014;2:1–4.
- [33] Wang J, Wolf RM, Caldwell JW, et al. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* 2004;25:1157–1174.
- [34] Pearlman DA, Case DA, Caldwell JW, et al. AMBER , a package of computer programs for applying molecular mechanics , normal mode analysis , molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comput. Phys. Commun.* 1995;91:1–41.
- [35] Galindo-murillo R, Robertson JC, Zgarbova M, et al. Assessing the Current State of Amber Force Field Modifications for. *J. Chem. Theory Comput.* 2016;12:4114–4127.
- [36] Ylilauri M, Pentikäinen OT. MM/GBSA as a tool to understand the binding affinities of

- filamin-peptide interactions. *J. Chem. Inf. Model.* 2013;53:2626–2633.
- [37] Lombardo F, Desai P V, Arimoto R, et al. In Silico Absorption , Distribution , Metabolism , Excretion , and Pharmacokinetics (ADME-PK): Utility and Best Practices . An Industry Perspective from the International Consortium for Innovation through Quality in Pharmaceutical Development Miniperspe. *J. Med. Chem.* 2017;60:9097–9113.
- [38] Prakash B, Murthy MRN. Source and target enzyme signature in serine protease inhibitor active site sequences. *J. Biosci.* 1997;22:555–565.
- [39] Galiano L, Bonora M, Fanucci GE, et al. Interflap Distances in HIV-1 Protease Determined by Pulsed EPR Measurements. *J. Am. Chem. Soc.* 2007;129:11004–11005.

CHAPTER 8

8.1 Conclusion

The development of drug resistance against current malaria therapeutics, has initiated a global effort directed towards the complete eradication of malaria. One of the major challenges affecting the design of potential therapeutics is the emergence of new drug targets that play a crucial role in the relentless tactics of the malaria parasite.

The present study aimed at gaining a more intricate understanding detailing the structural and molecular characteristics of the flap region of two newly discovered targets that play an instrumental role in the regulation of egress and invasion of the malaria parasite. In the absence of the crystal structures of Plasmpesin IX and X, this study integrated the use of homology modeling to generate a 3D model of each enzyme. To gain further insight regarding the structural dynamics the implementation of previously proposed parameters specific to aspartic proteases were integrated to reveal the structural similarity of PIX and PX to other prominent aspartic proteases and simultaneously highlighted their sequential uniqueness that must be taken into consideration when designing potential inhibitors to prevent the development of drug resistance.

The binding landscape of PIX and PX were further investigated using two experimental inhibitors each derived from the two major anti-malaria class of inhibitors. The effectivity of each inhibitor was scrutinized based on “per-residue energy decomposition analysis” and overall binding free energy contributions. These finding revealed the most integral chemical scaffolds of each inhibitor which effectively restricted the prominent “twisting motion” of the flap domain of PIX and PX. This further affirms the inhibitory capabilities of each inhibitor as reduced flexibility of the flap domain is closely linked to the complete reduction of biological active of aspartic proteases. A pharmacophore approach was further implemented for the design of tailored inhibitors specific to PIX, which revealed three potential inhibitors with desired pharmacokinetic properties and enhanced structural versatility.

Overall, this study presents indispensable insights into the design and development of malaria inhibitors specifically PIX and PX through the integration of molecular modeling and CADD.

8.2 Future perspectives

Quantification of the structural dynamics and functional motion of the flap dynamics portrayed in this study may be integrated in the design of potential Plm inhibitors. Further refinement of the core chemical scaffolds of the proposed experimental inhibitors must be implemented to ensure derived inhibitors adhere to “Lipinski’s rule of five” while maintaining appropriate pharmacokinetic properties. Furthermore, the potential inhibitors presented in this study provide promising protein-ligand interactions and binding energies and may therefore be employed as lead compounds. However, to further validate the effectivity of these proposed inhibitors the following approaches may be used:

1. QM or hybrid QM/MM computational studies may be implemented to supplement a better understanding regarding the dynamic motions of the flap domain and active site cleft providing a more detailed account of the molecular interactions and bond formations with respect to the distribution of motion.
2. Prospective biological testing of the lead compounds derived from in silico studies to verify their effectivity and inhibitory profiles.
3. Explore the potential effectivity of the proposed lead compounds to other aspartic proteases and members of the Plm family.

SUPPLEMENTARY INFORMATION

Table S4.1. The Z-score and sequence identity of 4OBZ template used for Plm IX-X

<i>PLASMEPSIN</i>	<i>TEMPLATE</i>	<i>SEQUENCE IDENTITY</i>	<i>Z-SCORE</i>
IX	4OBZ	35%	-6.16
X	4OBZ	35%	-6.22

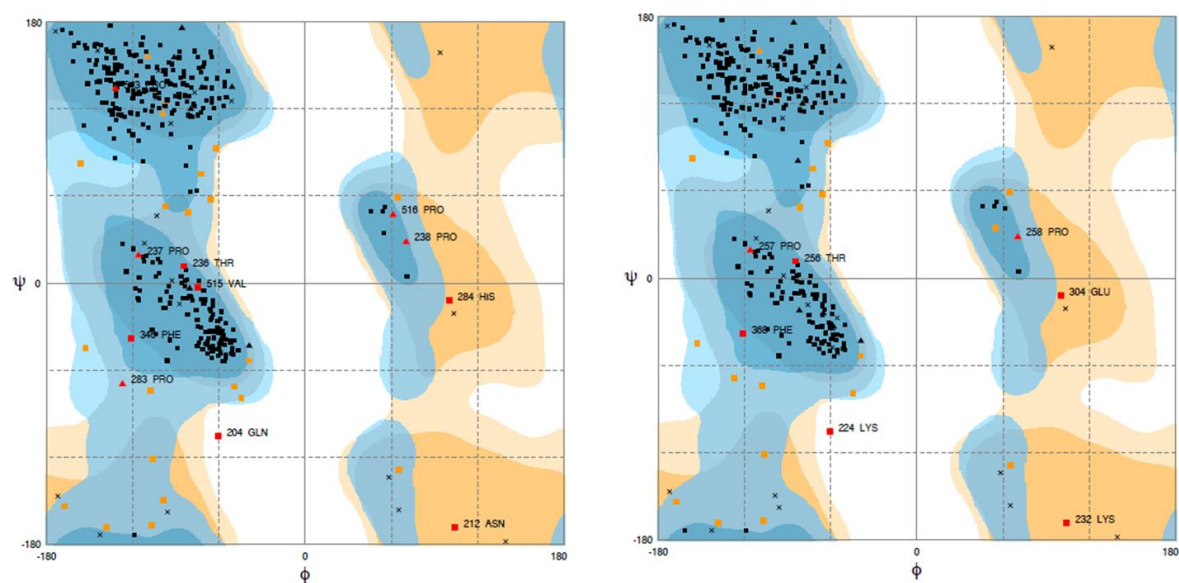


Figure 4.1. Ramachandran plots for models Plm IX (left) and Plm X (right).

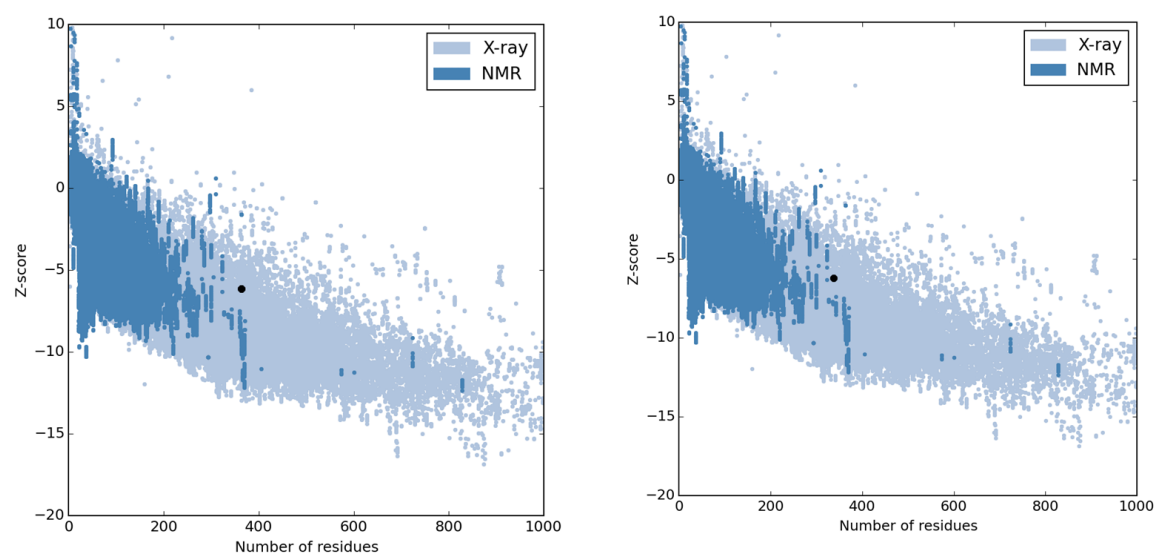


Figure S4.2. Overall model quality generated using ProSA-web of Plm IX (left) and Plm X (right).

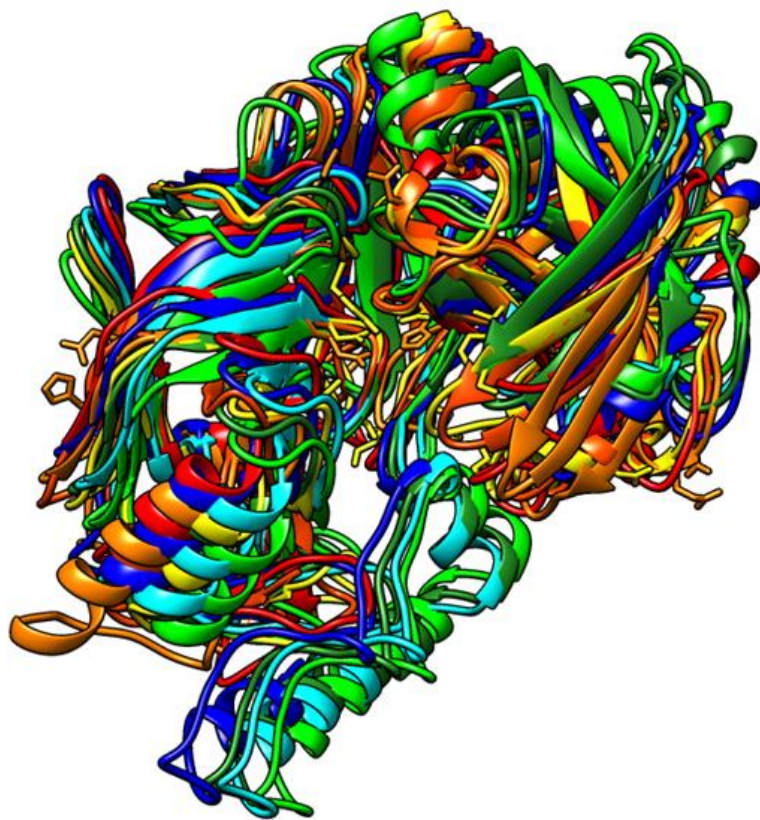


Figure S4.3. Superimposed image of Plasmepsins I (PDB: 3QRV; red), II (PDB:1LF4, orange red), III (PDB: 3FNS, orange), IV (1LS5, green), VI (dark green), VII (cyan), IX (sea green) and X (blue).

```

PM I      NAGDSYIIND.VA.VMYYTGAAGGDNKKK.FAFIPDTGSANLWVTSAGQNTIG.....KTENLFTSGSKK..TYEK
PM II     SSENDNIRIVD.FQNIMPTGDAEYSDNQCP.FTFILDTGSANLWVPSYKCTTAGC.....LTENLFTSGSKK..TYEK
HAP       SEFDONVIRKD.LANVLSFGDAKLSDNGCK.KNFLHTAGSSVWVPSIKSTSES.....KSENHWYSSSKK..TYEK
PM IV     SENDSIRDD.VANLMPTGSGGCTGNTKCP.KMFPIDTGSANLWVPSYVMDSIGC.....STENLFTSGSKK..TYEK
PM V      DLYKYKLTGDIIDRYATYTLDDIDGKPSCK.IGLILDTGSSSSLSFGCNGCK..DGIHMKKFTYLNLY..NN.TSSI
PM VI     LSYIQNDL.N.FM.SQFIADIGVGNPPGV.KVVPDTGSSSLAIPSTKCIKGGC.....ASIKKFNPMKSK..TYEK
PM VII    IKLKKQDIPS.KKLSTYYGTVQ.GRQSENMMNVLPDTGSSQVYTLNDTCNLSLC.....NNIN.S.YR.KSF
PM VIII   IGISTIVKGGTINQFICHENICNPPT.KVLPDTGSSSLNTPSKMCFTRAC.....YNNKRYTHKI..SENY
PM IX     NDKVTLFQQG.LSDSQIVYTIQGTFFPT.INFIPDTGSSINIVYTKCKDRTC.....LKVHRRNKKLS.SFKY
PM X      KNIFLVPK.N.LSDSQIVGALLVGTFFPT.VTFIPDTGSSIVVVTACENKSC.....KKVHRYPMKSK..PMK

PM I      .DGT.....VMMNVVSGT.VSGPPSKLIVTIAN.LSFPYKFIKVTDTN.....GFPATYTLG
PM II     .DGT.....VMMNVVSGT.VSGPPSKLIVTIGN.LSFPYKFIKVIDTN.....GFPATYTLG
HAP       .DDTF.....VKLTSGAGT.ISGIPSKLIVTIGK.LSFPYKFIKMTIV.....GFPATYTLG
PM IV     .DGT.....VLSVSGT.VRSPYKFIKVIDTD.LSFPYKFIKVIDTD.....DLPIYSGS
PM V      LYCNKSNCP-(11)-.....YLQSGCEGSGQYTFPSDIIVTLPSTNNKNNKISPEKLMGCH.....MNRSLFLHQ
PM VI     NLKSNNII-(24)-KKTPTTLFVGTCT.SSLQSGYTDVTLKGLYKKNQCEIGLAING.....KSLNLPYDL
PM VII    VYKYDKKG.-...-LPSV.IIPFLSGKIVAFRANVLKECIYTFPLIK.....NSHHR
PM VIII   KL.VKKDF.-...-VILFPTGR.IHIAVTTDHL.GDEKVKNGRFGASMI.....SDFPDM
PM IX     YRPHTN.-...-LIMPGTCT.SGVIQVNTFKIGP.FEIKNQSPGLVERKKASDNKS..NVFRI
PM X      SFIRKN.-...-LHIVFGSGS.SGSGVGTDTFPLCK.HLVNRQTFGLVRSK.SNNKKNCCNIPDTI

PM I      QFDGIVGL.....GKKLS..IGSVDPVVVRKNONKIQAVTFYLPVDDKHK..YLTIGGIDNR..YTG..
PM II     TFDGILGL.....GKKLS..IGSVDPVIVVRKNONKIRNALTFYLPVNDKHT..SPLTIGGIDNR..YTG..
HAP       DVGQVFG.....GKKLS..ISIDIVIVVRKTONKINQAVTSIYLPENKHK..YLTIGGIDNR..YTG..
PM IV     TFDGILGL.....GKKLS..ISIDIVIVVRKNONKIDNALTFYLPVNDKHV..SPLTIGGIDNR..YTG..
PM V      QATGVLPALTEKPKCVFTF.....VQLPKHTPSEKPIYKICVSRH.....GSLIIGGYPDYPLSNQK-
PM VI     QFDGIVGL.....GFSR.....LERNISFV.....NRS.SAIFGKANKYTVGK..
PM VII    .DGIIGL.....GFGNGDSIKKGIKLLDTKDDKILTKKNTKNQFGYLLSOKKSYITLGGIDNKLENTPDN-
PM VIII   QFDGLFG.....GISDK..K..KKQLIYDSIFKILEKNMAITYKNVDDDD.SAIFGQYKKN..NS..
PM IX     NYEGTVGL.....APPMLT..TC.KSTLYRNMSYKLGHNKSIYIGKD.SKY.SAIFGQYKKN..YTG..
PM X      SFGIVGL.....SFGMLT..AG.NIFFDNLKNPNDVDPGSPYISPY.DGK.STLIGGSKSK..YTG..

PM I      .....QLTTEKLNHDL.FHQVDGLDFGMLTVKKA.....TA.VDSGTSSTIA..XFLNKF
PM II     .....PLTTEKLNHDL.FHQTLDAEYVGNMLKKA.....NCVDSGTSSTIA..VDFLNKML
HAP       .....PLTTEKLNHDL.FHQVGLDVRVFGMSSEKKA.....NVVLSGATEVITVTHFFNQFV
PM IV     .....PLTTEKLNHDL.FHQDGLDFGKYYVMQKA.....NAVDSGTSSTIA..XFLNKF
PM V      (44)-EDIVNQAITK..Y..YTIKIYGLDLYGTINMDKKE.....LMDLVDSGTSSTIA..XFLNKF
PM VI     .....SIRFFVISLI.YWRNLDIQLSEKMLFDCN.....SKKCRAAITGSLIG..STFIQPLI
PM VII    ....-EIMVMTFVSTMGNTIQINGIRKKEVNNHFFRNKKEK-(11)-.GGKNSIDTCTYLLYAKNTMENYL
PM VIII   ....-EIMVMTFVSTMGNTIQINGIRKKEVNNHFFRNKKEK-(11)-.GGKNSIDTCTYLLYAKNTMENYL
PM IX     ....-EIMVMTFVSTMGNTIQINGIRKKEVNNHFFRNKKEK-(11)-.GGKNSIDTCTYLLYAKNTMENYL
PM X      ....-EIMVMTFVSTMGNTIQINGIRKKEVNNHFFRNKKEK-(11)-.GGKNSIDTCTYLLYAKNTMENYL

PM I      NGLDVVKIPFLPLYIT-....-TCN.NPKLSTIPRSA.....TNVYTKRNYYLQQIPFG..ISLCN
PM II     QNLDVVKIPFLPLYIT-....-LCN.NSKLSTIPRSA.....NGNYTKRNYYLQQIPFG..ISLCN
HAP       NSASVFKVFPFLPLYIT-....-TCG.NTKLSTIPRSA.....NKVYTKRNYYLQQIPFG..ISLCN
PM IV     NDMNVKIPFLPLYIT-....-TCG.NDKLSTIPRSA.....NKVYTKRNYYLQQIPFG..ISLCN
PM V      YYLDILCINDMTNITE-(44)-CWKSLNHLNMYITLS.....NNYKMIKNSSYLYKKKSPW....CK
PM VI     NKNL.LNRD@SNKES.....LTIISVLEN.....VRGKNITLDFMNDYIIEGQTMNTLNCM
PM VII    KDLKIN.NCDKKN.....LHILIQIKSDRIKTIEGSAIINIVINDYVIRYVKKNNTKKCI
PM VIII   E.....LTPKDYI.....VNSRID.....VLRSF.....CN
PM IX     RVVPSKKCDSDNIDQV-....-VSSYKTYVIN.....KMPSTINSQVLVRKNM....CK
PM X      NLINSTACTRQNYKDI-....-LKSYIKYVFG.....ELIIRHNRVWILNDQV....CM

PM I      VSILVGLNKN...TFILGD.FPMKKYFTVVDY...DNHTVGFALAKKL
PM II     LNIIGLFFVP...TFILGD.FPMKKYFTVVDY...DNHTVGFALAKKL
HAP       LNIIGLFFVP...TFILGD.FPMKKYFTVVDY...DNHTVGFALAKKL
PM IV     LYLLVGLNKN...TFILGD.FPMKKYFTVVDY...DNHTVGFALAKKL
PM V      GLEKQVNNKP.....LLGL.TFKNKQVIFDL...QQNQIARINS.CFENLT-(67)
PM VI     FGMLPLVPPFNG.PIPFPM.SVISEYIIPDN...DNKLIILINNNMF
PM VII    LGIQDQSQSHRDNDVGGTLQ.VALKAYYIPK...DNKLIILINNNMF
PM VIII   FAFMIRISSSNG.YLYILQ.IPLQKYAIREK...DNKLIILINNNMF
PM IX     FAFMIRISSSNG.YLYILQ.IPLQKYAIREK...DNKLIILINNNMF
PM X      FAFMIRISSSNG.YLYILQ.IPLQKYAIREK...DNKLIILINNNMF

```

Figure S4.4. Sequence alignment of Plm I-X from *Plasmodium falciparum*.

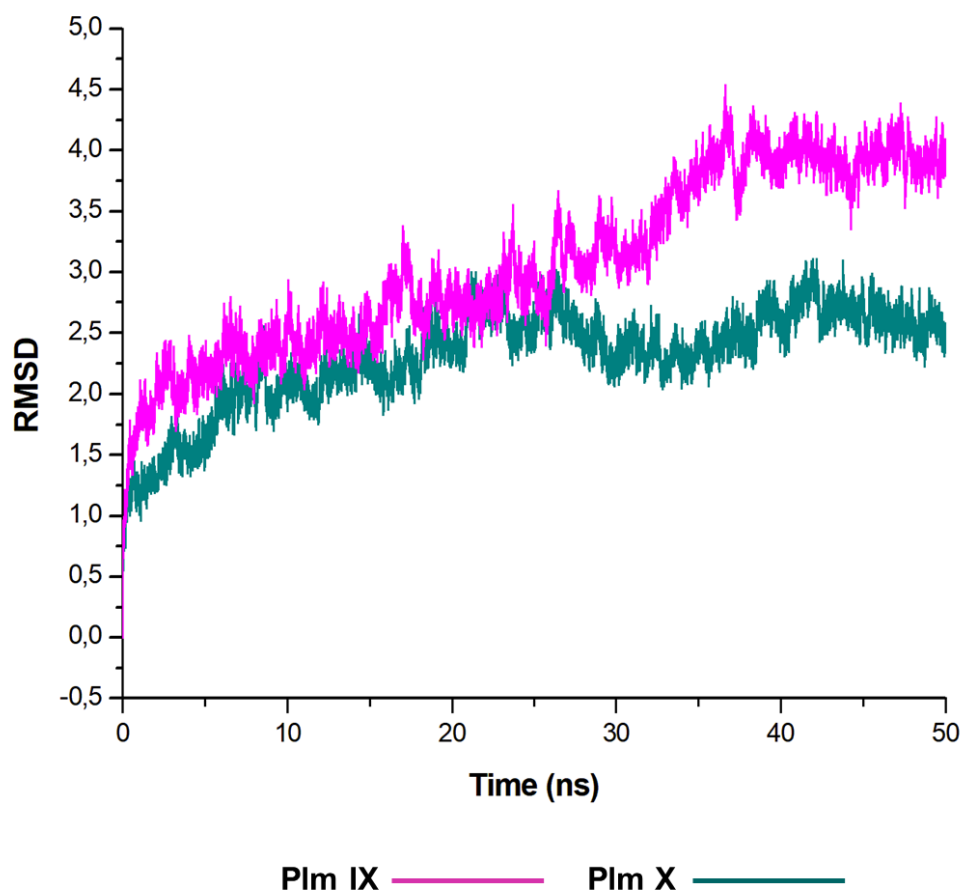


Figure S4.5. RMSD plot of PIX and PX, displaying system stability after 30ns.

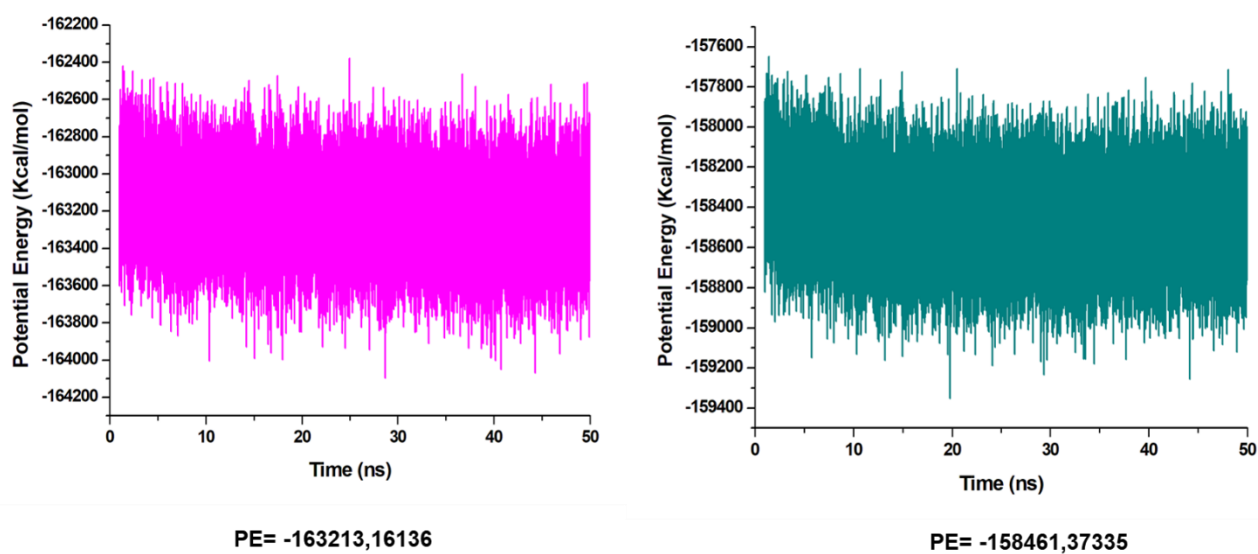


Figure S4.6. Potential energy (PE) plots of Plm IX (left) and Plm X (right) with interaction cutoff of 12 Å. Mean values are presented below the respective plots.

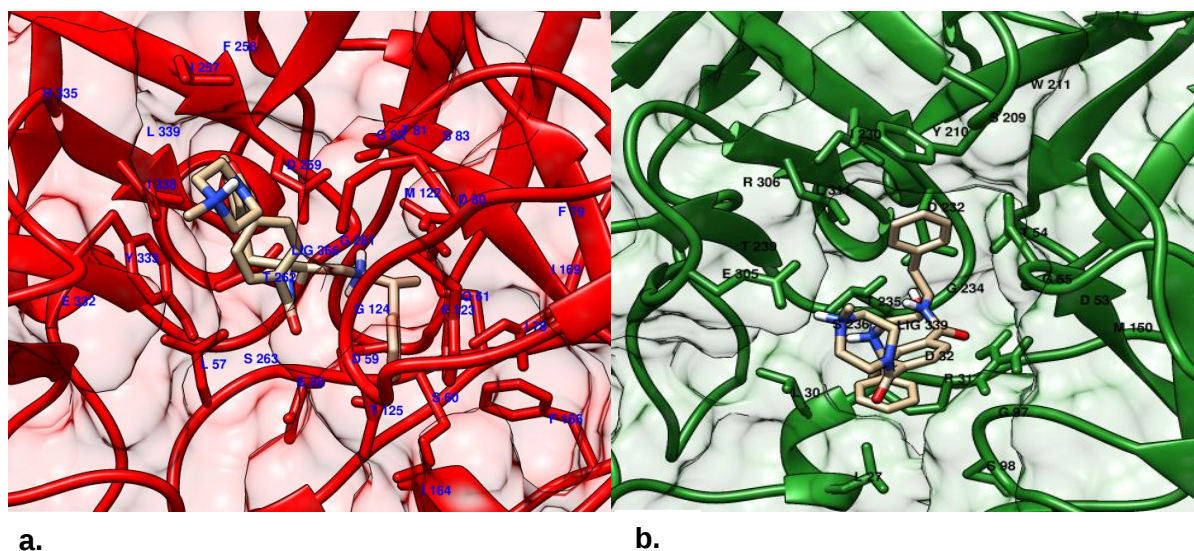
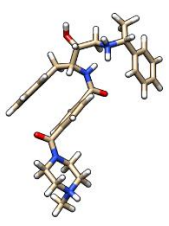
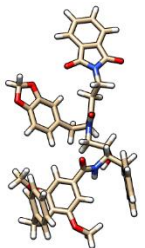
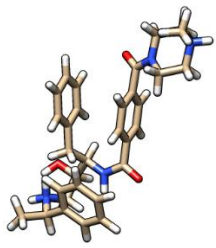
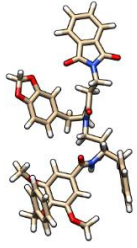


Figure S5.1. Binding mode of PIX (a) and PX(b) in complex with 49c respectively, with the interacting residues depicted

Table S5.1. Parameters used for docking of 49c and Pepstatin to PIX and PX

plasmepsin IX		plasmepsin X	
49c	Pepstatin	49c	Pepstatin
 Binding Affinity = -8.2	 Binding Affinity= -7.6	 Binding Affinity= -7.9	 Binding Affinity= -6.4
Docking Dimensions X-68 Y-40 Z-44	Docking Dimensions X-66 Y-50 Z-65	Docking Dimensions X-52 Y-52 Z-40	Docking Dimensions X-72 Y-54 Z-59

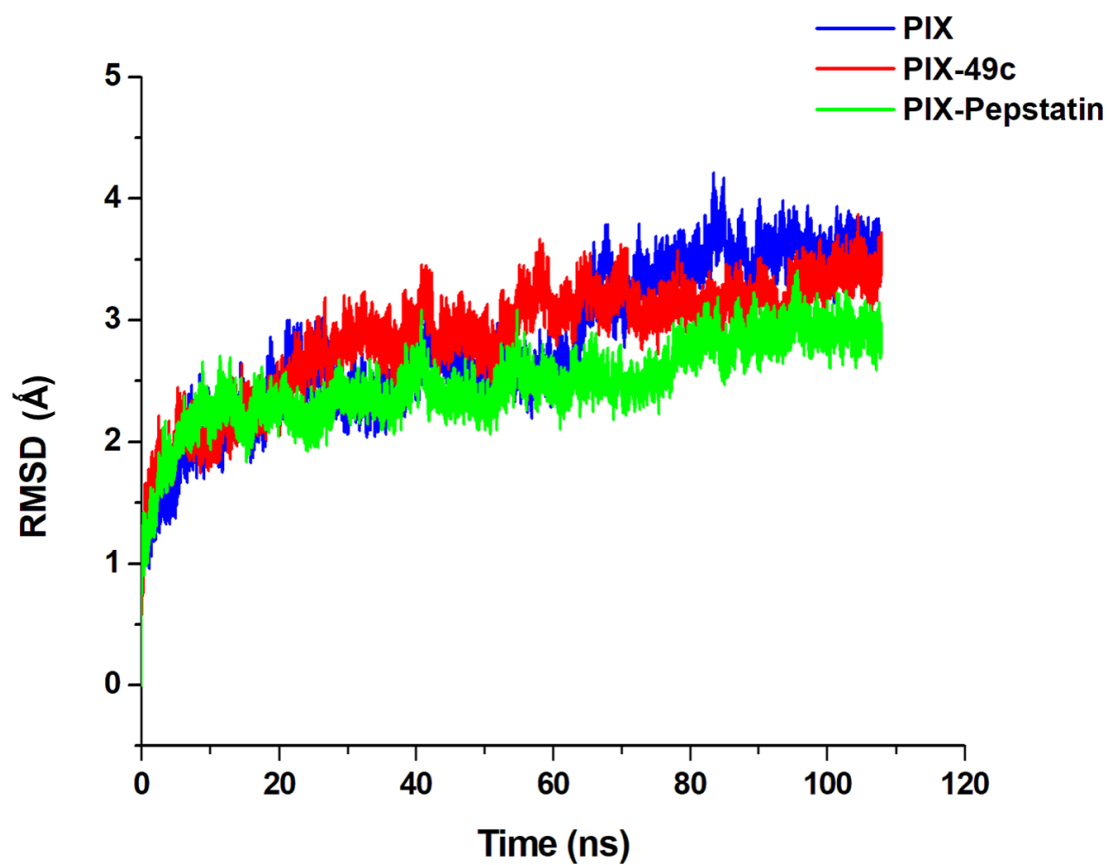


Figure S5.2. RMSD plot of PIX (blue), PIX-49c (red) and PIX-Pepstatin (green).

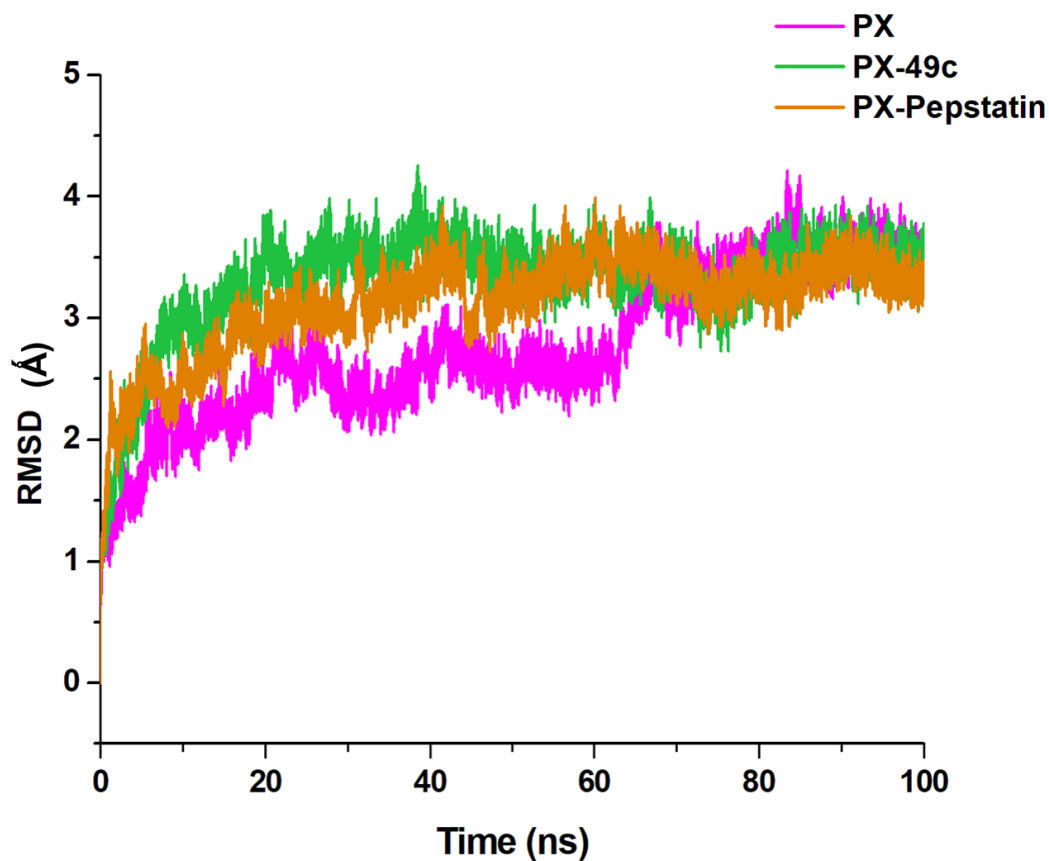


Figure S5.3. RMSD plot of PX (pink), PX-49c (olive) and PX-Pepstatin (orange).

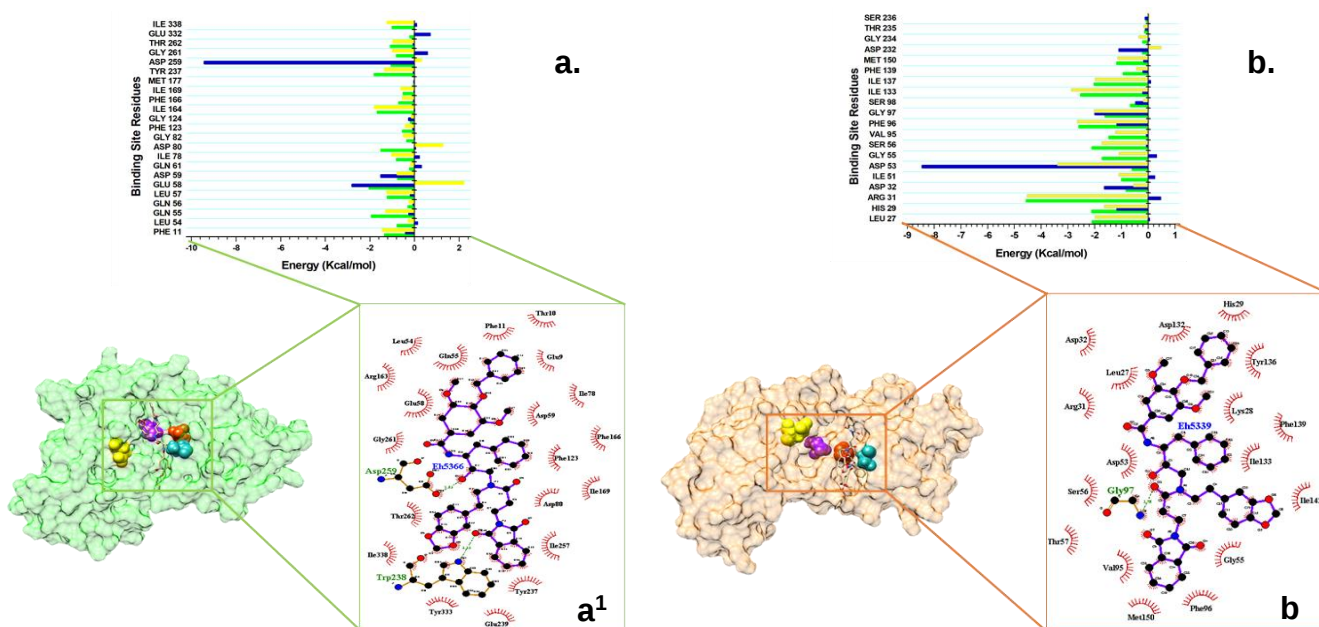


Figure S5.4. Molecular visualization of Pepstatin at the catalytic sites (hydrophobic pockets) of **[a]** PIX and **[b]** PX. Inter-molecular interactions between Pepstatin and catalytic site residues in PX and PIX are shown in **a¹** and **b¹** respectively.

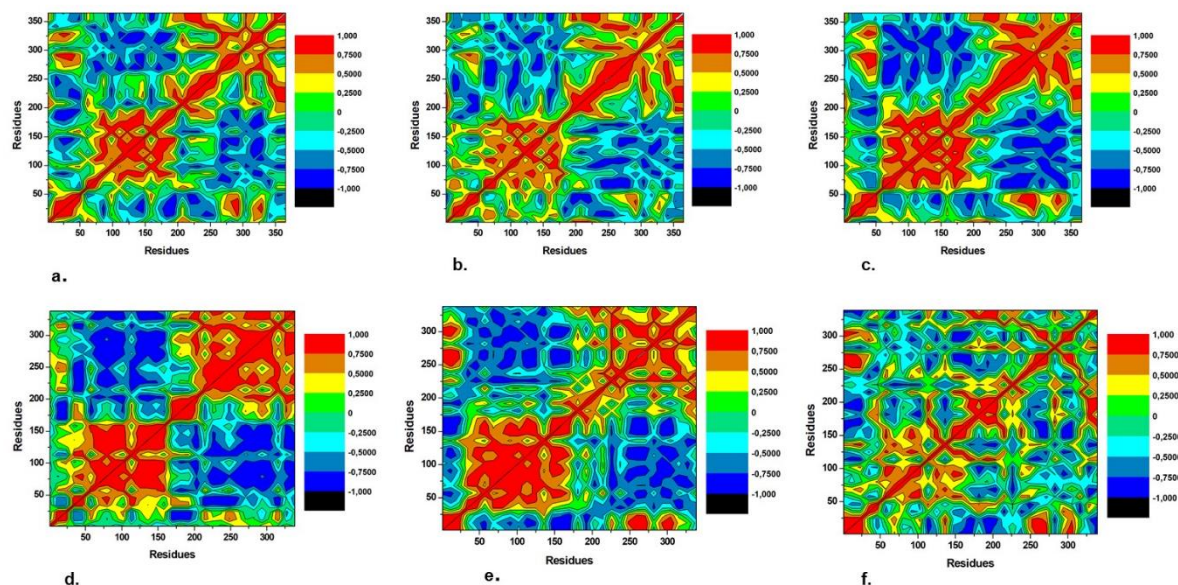


Figure S5.5. DCC map during simulation time taking in account C α residues of unbound PIX (a), PIX-49c complex (b), PIX-Pepstatin complex (c), unbound PX (d), PX-49c complex (e) and PX-Pepstatin complex (f).

Table S6.1. Physicochemical properties and pharmacokinetics properties of 49c

49c								
<i>Molecular Formula</i>	<i>Molecular Weight (g/mol)</i>	<i>Lipophilicity (iLOGP)</i>	<i>Water Soluble</i>	<i>GIT Absorption</i>	<i>BBB Permeability</i>	<i>Bioavailability Score</i>	<i>Synthetic Accessibility</i>	<i>Druglikeness (Lipinski)</i>
C13H14N4O4	514.66	3.95	Low	High	No	0.55	4.27	No

"Boiled-Egg" Method Summary					

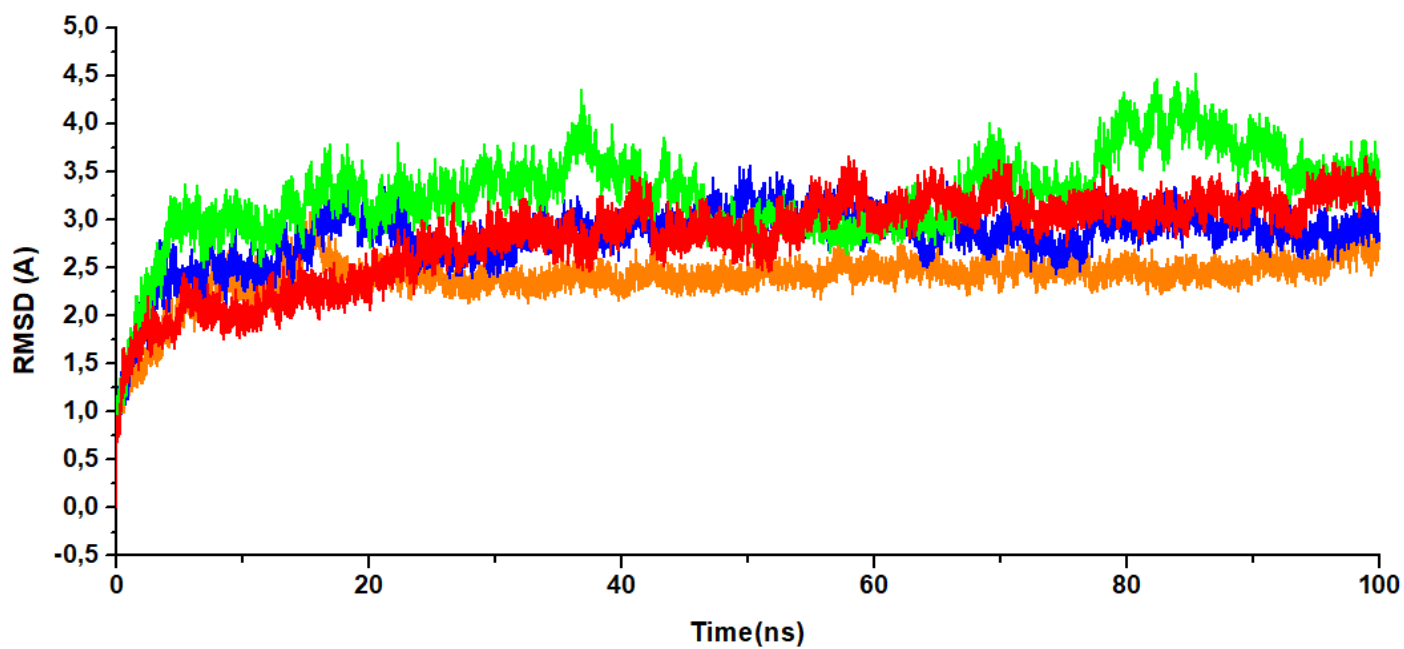


Figure S6.2. Root Mean Square Fluctuation of PIX in complex with 49c (red), Compound 1 (orange), Compound 2 (blue) and Compound 3 (green).

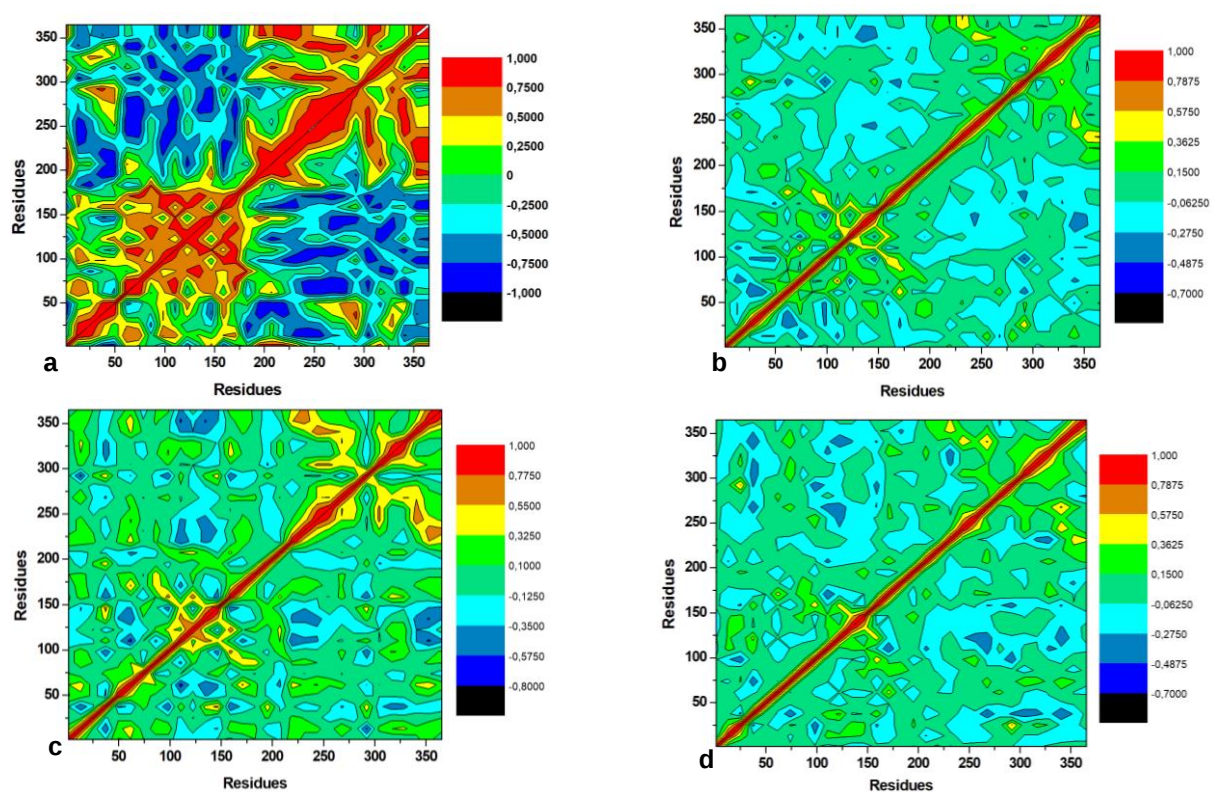


Figure S6.3. DCCM plots of PIX in complex with 49c (a), Compound 1 (b), Compound 2 (c) and Compound 3 (d).

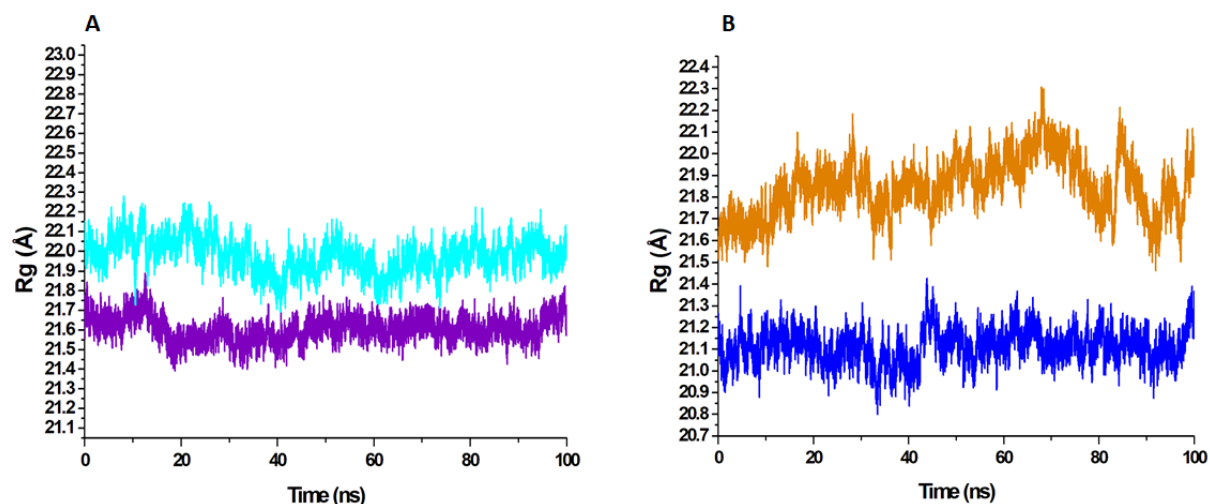


Figure S7.1. Radius of gyration of each system for the 100ns MD simulation A) PIX-apo (cyan), PIX-117-CWHM (purple). B) PX-apo (orange) and PX-117-CWHM (dark blue)

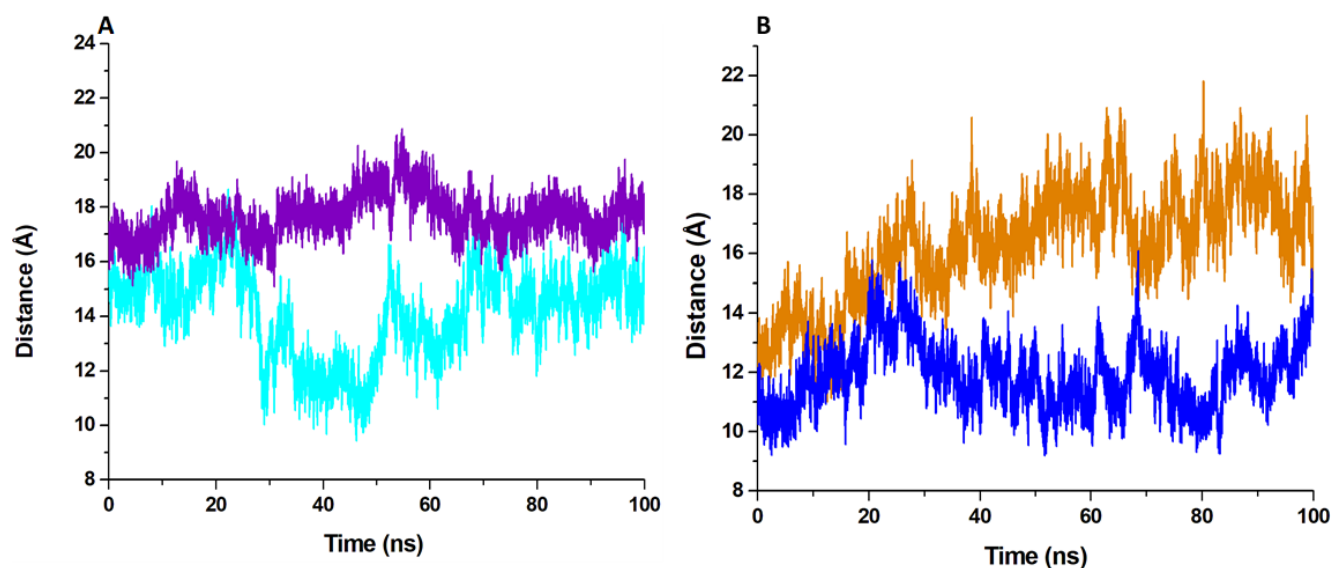


Figure S7.2. Distance measured between flap tip and hinge residue for the 100ns MD simulation measured in Angstroms (Å). A) PIX-apo (cyan), PIX-117-CWHM (purple). B) PX-apo (orange) and PX-117-CWHM (dark blue)

APPENDICES

APPENDIX A



RSC Advances

PAPER



Cite this: RSC Adv., 2018, 8, 21829

Egress and invasion machinery of malaria: an in-depth look into the structural and functional features of the flap dynamics of plasmepsin IX and X†

Geraldene Munsamy, Pritika Ramharack and Mahmoud E. S. Soliman *

Plasmepsins, a family of aspartic proteases expressed by *Plasmodium falciparum* parasite, have been identified as key mediators in the onset of lethal malaria. Precedence has been placed on this family of enzymes due their essential role in the virulence of the parasite, thus highlighting their importance as novel drug targets. A previously published study by our group proposed a set of parameters used to define the flap motion of aspartic proteases. These parameters were used in the study of Plm I–V and focused on the flap flexibility as well as structural dynamics. Recent studies have highlighted the essential role played by Plm IX and X in egress and invasion of the malarial parasite. This study aims to close the gap on the latter family, investigating the flap dynamics of Plms IX and X. Molecular dynamics simulations demonstrated an ‘open and close’ mechanism at the region of the catalytic site. Further computation of the dihedral angles at the catalytic region revealed tractability at both the flap tip and flexible loop. This structural versatility enhances the interaction of variant ligand sizes, in comparison to other Plm family members. The results obtained from this study signify the essential role of structural flap dynamics and its resultant effect on the binding landscapes of Plm IX and X. We believe that this unique structural mechanism may be integrated in the design and development of effective anti-malarial drugs.

Received 22nd May 2018
Accepted 7th June 2018

DOI: 10.1039/c8ra04360d
rsc.li/rsc-advances

Introduction

According to the World Health Organization (WHO), by 2016, there were an estimated 216 million reported cases of malaria worldwide.¹ Most cases of malaria and deaths occur in Sub-Saharan Africa,^{2,3} however, regions of South-East Asia, Eastern Mediterranean, Western Pacific, and the Americas are also at critical risk.⁴ Since then, the infection rate of malaria has almost doubled, with an estimated 500 million people being infected annually, with the majority of these emerging cases being children under the age of five.⁵

In humans, the parasite is transmitted via the female *Anopheles* mosquito vector. Humans are infected by a range of *Plasmodium* species,⁶ however, the most severe and common forms of malaria are caused by *Plasmodium falciparum* and *Plasmodium vivax*.⁷ *P. falciparum*, predominating in Africa, is considered to be the most virulent species being responsible for 85% of human death.⁸

The complex life cycle of *P. falciparum* embodies multiple cycles of invasion, multiplication and egress in the human red

blood cells.^{9,10} Evidently, targeting the egress/invasion machinery, at its most vulnerable stage of the life cycle of the parasite may not only reduce the severity of the malarial disease, but may jointly eradicate disease transmission.¹¹ Even though the process of invasion occurs rapidly, it is the only time during the parasite's life cycle when it is directly exposed to the host immune system.¹² There are several enzymes in *P. falciparum*, that have been implicated in hemoglobin proteolysis, parasite nutrition and development, these enzyme families include; cysteine proteases; aspartic proteases; metalloproteases and dipeptidyl aminopeptidases.^{13,14}

Aspartic proteases, or plasmepsins, have been identified as key mediators of cellular processes, including hemoglobin degradation for the export of *Plasmodium* proteins that essential for parasite growth/survival¹⁵ and particularly malarial egress and invasion.¹⁶ *P. falciparum* possess a repertoire of 10 aspartic proteases (Plm I to X).¹⁷ Of the 10 Plms identified in *P. falciparum*, only Plm I, II, HAP, IV and V have been studied extensively.¹⁸

Plm VI–VIII are expressed in the vector during the parasite's intra-erythrocytic stages of sporozoite formation, motility as well midgut transverse¹⁹ (Fig. 1). Studies have suggested that due to their expression during the sporogonic cycle in the mosquito, these Plm's may not be ideal drug target candidates.²⁰ However, understanding the molecular mechanism by

Molecular Bio-computation and Drug Design Laboratory, School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa. Email: soliman@ukzn.ac.za; Fax: +27 (0) 31 260 7872; Tel: +27 (0) 31 260 8048

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c8ra04360d



A dual target of Plasmepsin IX and X: Unveiling the atomistic superiority of a core chemical scaffold in malaria therapy

Geraldene Munsamy | Clement Agoni | Mahmoud E. S. Soliman

Department of Pharmaceutical Chemistry, Molecular Bio-Computation and Drug Design Laboratory, School of Health Sciences, Westville Campus, University of KwaZulu-Natal, Durban, South Africa

Correspondence
Mahmoud E. S. Soliman, Molecular Bio-Computation and Drug Design Laboratory, School of Health Sciences, University of KwaZulu-Natal, Westville Campus, Durban 4001, South Africa.
Email: soliman@ukzn.ac.za

Funding information
National Research Foundation, Grant/Award Number: 115451; Centre for High Performance Computing

Abstract

Plasmepsin IX and X, members of the prominent aspartic family of proteases whose function were hitherto unknown have only recently been established as key mediators of erythrocyte invasion and egress of the virulent malarial parasite. Inhibitor 49c, a potent antimalarial peptidomimetic inhibitor initially developed to target Plasmepsin II has lately been proven to exhibit potent inhibitory activity against Plasmepsin IX and X. However, the molecular and structural dynamics supporting its inhibitory activity remain inconclusive. Hindering the motion of the flap and hinge region of an aspartic protease remains essential for disabling the catalytic activity of the enzyme. Integrating molecular dynamic simulations coupled with other advanced biocomputational tools, we reveal the enhanced structural mechanistic competence of 49c in complex with Plasmepsin IX and X relative to Pepstatin. Pepstatin, a known aspartic protease inhibitor which actively hinders the opening and closing of the flap tip and flexible loop and consequently limits access to the catalytic aspartic residues, however, its administration has been related to elevated levels of toxicity. Thermodynamic calculations reveal a higher relative binding free energy associated with Plasmepsin IX and X in complex with 49c as opposed to Pepstatin. A relatively compact and structurally rigid 49c bound complexes sequel into the restriction of the flap and hinge residues by restraining cohesive movement, consequently hindering their “twisting motion” from transpiring. Findings unveil an atomistic perspective into the structural superiority of 49c in complex with Plasmepsin IX and X.

KEYWORDS

49c, molecular dynamics simulation, Plasmepsin X, Plasmepsin IX, twisting motion

1 | INTRODUCTION

The infiltrating tactics of the malarial parasite remains a global threat in the effective treatment of the mosquito borne-parasitic infection. Alarming concerns have been linked to the unremitting spread of malaria, predisposing half of the world's population to its infection. In the year

2016,¹ malaria was reported as one of the highest contributors to the escalated morbidity and mortality rate observed.² Malaria is caused by five *Plasmodium* species, namely, *Plasmodium falciparum*, *Plasmodium vivax*, *Plasmodium ovale*, *Plasmodium knowlesi* and *Plasmodium Malariae*. Of these, *P. falciparum* is considered to be the most lethal strain, emanating in a

Appendix C

The Protein Journal

Unveiling a new era in Malaria therapeutics: A tailored molecular approach towards the design of Plasmepsin IX inhibitors –Manuscript Draft–

Manuscript Number:	JOPC-D-18-00182R2	
Full Title:	Unveiling a new era in Malaria therapeutics: A tailored molecular approach towards the design of Plasmepsin IX inhibitors	
Article Type:	Original Research	
Keywords:	Keywords: Plasmodium falciparum; Plasmepsin IX; Pharmacophore; Per-residue energy decomposition	
Corresponding Author:	Mahmoud E. S. Soliman University of KwaZulu-Natal - Westville Campus Durban, SOUTH AFRICA	
Corresponding Author Secondary Information:		
Corresponding Author's Institution:	University of KwaZulu-Natal - Westville Campus	
Corresponding Author's Secondary Institution:		
First Author:	Geraldene Munsamy	
First Author Secondary Information:		
Order of Authors:	Geraldene Munsamy Mahmoud Soliman	
Order of Authors Secondary Information:		
Funding Information:	National Research Foundation (115451)	Miss Geraldene Munsamy
Abstract:	The invasive strategies employed by the malarial parasite renders malaria a global health threat, further impeding the effective treatment of the mosquito borne-parasitic disease. Although there have been countless efforts directed towards the development of effective therapeutics, factors such as emerging strains of drug resistance, enhanced toxicity and poor pharmacokinetic properties of current therapeutics has hampered the drug discovery process relenting in the spread of this parasitic disease. A promising target of the most lethal strain of the Plasmodium species that plays a predicted role in erythrocyte invasion of the virulent malarial parasite is aspartic protease IX commonly referred to Plasmepsin IX. The integration of computer aided-drug design platforms has revolutionized the 21st century and has opened avenues to render a final "knock out" in the elimination and eradication of this parasitic disease. Hitherto, this is the first attempt directed towards the design of therapeutics tailored explicitly to Plasmepsin IX. A potent peptidomimetic inhibitor referred to as 49c which is a known inhibitor of Plasmepsin II, has recently been proven to exhibit further potent inhibitory activity against Plasmepsin IX. In-silico structural and physicochemical inspection of 49c displayed poor pharmacokinetic properties thus paving the way for the development of tailored inhibitors with desirable therapeutic properties against Plasmepsin IX. In this study we implement the pharmacophore model approach in combination with per-residue energy decomposition analysis to serve as a powerful cornerstone, that may assist medicinal experts in the composition of multifunctional therapeutics that may dispose of factors such as cross-resistance and toxicity, with enhanced pharmacokinetic properties.	

Appendix D

RSC Advances



RSC Advances

**Fundamental class of inhibitors activates “fireman’s grip”:
An enhanced binding analysis in search of effective malarial
therapeutics**

Journal:	RSC Advances
Manuscript ID	RA-ART-03-2019-002290
Article Type:	Paper
Date Submitted by the Author:	25-Mar-2019
Complete List of Authors:	Munsamy, Geraldene; University of KwaZulu-Natal College of Health Sciences Agoni, Clement; University of KwaZulu-Natal College of Health Sciences Ramharack, Pritika; University of KwaZulu-Natal, Pharmaceutical Sciences Soliman, Mahmoud; University of KwaZulu-Natal College of Health Sciences
Subject area & keyword:	Computational - Biological < Biological

SCHOLARONE™
Manuscripts

