Plasmodium falciparum Plasmepsins IX and X: Structure- Function Analysis and the Discovery of New Lead Antimalarial drugs

Plasmodium falciparum Plasmepsins IX and X: Structure-function analysis and the discovery of new lead antimalarial drugs

Rachael Dawn McGeorge
Bachelor of Forensic Science
Bachelor of Criminology and Criminal Justice
Bachelor of Science (Hons 1)

School of Biomolecular and Physical Sciences
Griffith Sciences
Griffith University

Submitted in fulfilment of the requirements of
the degree of Doctor of Philosophy

April, 2014

Abstract

Malaria is a deadly parasitic infection that poses an enormous threat to global health. While drugs are available, all of our current antimalarials are being gradually rendered ineffective by spreading drug resistance. Reports of increasing tolerance to artemisinin combination therapies, our most potent antimalarial treatments, are particularly concerning. To combat this problem there is an urgent need to identify new and unique targets within the malaria parasite against which novel chemotherapeutics can be developed. Studies investigating the antimalarial activity of HIV protease inhibitors (HIV PIs) have shown that these drugs can inhibit the growth of Plasmodium *in vitro, in vivo* and *ex vivo* at clinically relevant concentrations. While the anti-parasitic action of these drugs is not fully understood, it is believed these agents inhibit the growth of parasites by targeting an essential malarial aspartic protease or plasmepsin (PM) and that these enzymes may represent new targets for drug development. The aim of this thesis was to investigate the *Plasmodium falciparum* aspartic proteases PM IX and X as potential new targets for antimalarial development.

To determine if *Pf*PM IX is a viable drug target and to substantially contribute to existing knowledge on this enzyme, I sought to determine its function and location in asexual parasites. While the function of *Pf*PM IX remains unknown, expression and localisation data generated during this work, together with data suggesting that *Pf*PM IX is important to parasite growth, support the hypothesis that it is a feasible drug target. Data showing that *Pf*PM IX is transported outside of the parasite into the infected red cell suggest this enzyme has one or more essential function/s distinct from other malaria parasite aspartic proteases. Expression data also indicated that *Pf*PM IX is required throughout the asexual life cycle and that drugs targeting this enzyme will be active throughout asexual development. Additionally, transgenic parasites over-expressing *Pf*PM IX have been shown to have decreased sensitivity to selected HIV PIs. Although sufficient quantities of active recombinant enzyme could not be produced for in-depth analysis, inhibitor assessment and X-ray crystal structure determination, my work in this area has been of benefit to others who are now pursuing further optimisation and expression trials. Given the complexity of *Pf*PM IX and the lack of suitable templates for *in silico* modelling tools to accurately predict the structure of this enzyme, it is clear that recombinant enzyme will be required to accurately determine the structure of this *Pf*PM. The availability of recombinant enzyme will also permit further characterisation and inhibitor studies to be performed.

Data examining the role of *Pf*PM X in asexual parasites also supports the hypothesis that this enzyme is a good drug target for chemotherapeutic development. Similar to *Pf*PM IX, *Pf*PM X appears to perform a unique function within malaria parasites. Unlike *Pf*PM IX, *Pf*PM X is not exported into the host red cell. However, it also does not locate to the digestive vacuole or the endoplasmic reticulum, suggesting that it has a distinct function from the remaining *Pf*PMs expressed in asexual parasite stages. Data describing that GFP-tagged and purified *Pf*PM X is sensitive to HIV PIs, together with data demonstrating that parasites over-expressing *Pf*PM X are less sensitive to selected HIV PIs, suggests this enzyme may be a target of these inhibitors. *In silico* modelling are also supportive of this hypothesis, with top binding HIV PIs interacting with the model's catalytic aspartic residues. As the vast majority of *Pf*PM X transcription occurs during later schizont-stages, this enzyme may have a role in egress and/or invasion. The slow growth, as compared to wild-type parents, of parasites transfected with an antisense construct of *Pf*PM X which transcribed less *PfPM X*, together with the inability of this work to generate a population of parasites with a *PfPM X* targeted gene disruption, provides further evidence that this protein plays a vital role in asexual parasite stages.

Whilst the current body of research was not successful in identifying the function of *Pf*PM IX and *Pf*PM X, our data supports the hypothesis that these enzymes play an important role in parasite growth and development. The tools and resources developed during this project will continue to be useful for further studies investigating the role of these and other PMs in malaria parasites. They may also aid in the identification of more potent specific inhibitors.

Statement of Originality

This work has not previously been submitted for a degree or diploma in any university. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.

_____

Rachael Dawn McGeorge

Table of Contents

Abstract ..... **i**

Statement of Originality ..... **iii**

Table of Contents ..... **iv**

List of Figures and Tables ..... **xi**

List of Appendices ..... **xv**

Ethics ..... **xv**

List of Abbreviations ..... **xvi**

Acknowledgements ..... **xviii**

Publications ..... **xx**

Presentations: ..... **xx**

Chapter 1: Introduction and Literature Review **2**

1.1 Malaria: 4

1.1.1 A Lethal Parasitic Infection: 4

1.1.2 *Plasmodium falciparum*: 5

1.1.2.1 Life Cycle: 5

1.1.2.2 Falciparum Malaria: 5

1.1.2.3 Global Distribution: 7

1.1.2.4 Disease Demographics: 7

1.1.2.5 Disease Interactions: 8

1.1.3 Vector Control: 9

1.1.4 Vaccine Development: 9

1.1.5 Chemotherapy: 10

1.1.5.1 Prevention: 10

1.1.5.2 Current Treatment Guidelines for Falciparum Malaria: 11

1.1.5.2.1 Quinoline Antimalarials: 11

1.1.5.2.2 Antifolates: 14

1.1.5.2.3 Artemisinin and derivatives: (Sesquiterpene lactones and metabolites) 15
1.2 Identification of new Plasmodium drug targets 16
1.2.1 Haemoglobin digestion by *P. falciparum* 16
1.2.2 *P. falciparum* Plasmepsins 17
1.2.3 *Pf*PM IX and X: 21
1.3 Methods to Determine Protein Function in *P. falciparum*: 23
1.4 Inhibitor Discovery 25
1.5 The Current Project: 26

Chapter 2: Characterisation of *Plasmodium falciparum* Plasmepsin IX: A potential new target for anti-malaria drug development 28
2.1 Introduction 30
2.2 Materials and Methods 31
2.2.1 Parasites 31
2.2.2 Expression and Localization of *Pf*PM IX 32
2.2.2.1 Real Time PCR 32
2.2.2.2 Antibody Generation and Purification 33
2.2.2.3 Immuno-blotting 34
2.2.2.4 Immuno-fluorescence 34
2.2.3 Examining the role and importance of *Pf*PM IX 35
2.2.3.1 Confirmation of transgenic parasites 35
2.2.3.2 Over-expression 35
2.2.3.3 Peptide Nucleic Acid Knock-down of *PfPM IX* 36
2.3 Results: 37
2.3.1 Localization and expression of *Pf*PM IX 37
2.3.2 Transfection of parasites with pHH1-PMIXcMycB and pHH1-PMIXGFPB generates transgenic parasites over-expressing *Pf*PM IX 41

2.3.3 *Pf*PM IX may be essential to intra-erythrocytic stage *P. falciparum* parasites 42

2.2.4 Peptide Nucleic Acid Knock-down of *PfPM IX* 43

2.4 Discussion and Conclusion 46

Chapter 3: Expression and Purification of *P. falciparum* Plasmepsin IX 52

3.1 Introduction: 54

3.2 Materials and Methods: 55

3.2.1 Expression Vectors 55

3.2.1.1 Semi-pro *Pf*PM IX (r*Pf*PM IX): 55

3.2.1.2 Truncated *Pf*PM IX: 55

3.2.2 Expression and purification of semi-pro *Pf*PM IX 56

3.2.2.1 Expression of insoluble semi-pro *Pf*PM IX 56

3.2.2.2 Purification of insoluble semi-pro *Pf*PM IX 56

3.2.2.2.1 Isolation of inclusion bodies 56

3.2.2.2.2 Solubilisation and refolding 57

3.2.3 Chromogenic activity assays and substrate cleavage 58

3.2.4 Crystallisation Trials with r*Pf*PM IX 59

3.2.5 Investigation of optimum pH for r*Pf*PM IX activation 59

3.2.6 Expression and purification trial of insoluble truncated r*Pf*PM IX (NCtrunc-r*Pf*PM IX) 59

3.3 Results: 59

3.3.1 Expression, refolding and purification of semi-pro *Pf*PM IX 59

3.3.2 Active semi-pro *Pf*PM IX cleaves the RS6 substrate 61

3.3.3 Optimal conditions for refolding and purifying semi-pro *Pf*PM IX 62

3.3.4 Semi-pro *Pf*PM IX is auto-activated at pH 2-3.5 67

3.3.5 NCtrunc-r*Pf*PM IX expression and purification trial 67

3.4 Discussion and Conclusion 70

Chapter 4: *In silico* modelling of Plasmepsin IX 72

4.1 Introduction: 74

4.1.1 *In silico* modelling and docking: 74

4.1.2 *Pf*PM IX and Orthologs: 76

4.2 Materials and Methods: 84

4.2.1 *Pf*PM IX homology model generation: 84

4.2.1.1 *Pf*PM IX model template and docking model selection: 84

4.2.1.2 Preparation and refinement of the docking model: 84

4.2.2 Validating the model: 85

4.2.2.1 Examining the model: 85

4.3 Results: 85

4.3.1 Choice of a template and selecting a *Pf*PM IX 'working' model 85

4.3.2 Refinement of the model required for docking studies. 86

4.3.3 Examination of the model: 87

4.4 Discussion and Conclusions: 90

Chapter 5: Characterisation of *Plasmodium falciparum* Plasmepsin X 94

5.1 Introduction: 96

5.2 Methods: 97

5.2.1 Parasites 97

5.2.2 Expression and Localization of *Pf*PM X 97

5.2.2.1 Real time PCR (qRT-PCR) 97

5.2.2.2 Antibody Generation 98

5.2.2.3 Immuno-blotting 98

5.2.2.4 Immuno-fluorescence 99

5.2.3 Examining the role and importance of *Pf*PM X 99

5.2.3.1 Confirmation of transgenic parasites 99

5.2.3.2 Over-expression of *Pf*PM X 100

5.2.3.3 Antisense knock-down of *Pf*PM X 100
5.2.4 Characterisation of *Pf*PM X 101
5.2.4.1 Recombinant enzyme 101
5.2.4.1.1 Expression of soluble r*Pf*PM X 101
5.2.4.1.2 GST purification 101
5.2.4.1.3 Ni-NTA purification 102
5.2.4.1.4 Enzyme assays and substrate cleavage. 102
5.2.4.2 Pull down of GFP tagged *Pf*PM X 103
5.3 Results: 104
5.3.1 *Pf*PM X expression and localisation 104
5.3.2 *Pf*PM X is over-transcribed in transgenic parasites 106
5.3.3 Is *Pf*PM X essential to *P. falciparum* growth in red blood cells? 108
5.3.3.1 Genetic disruption of *Pf*PM X 108
5.3.3.2 Knock-down of *PfPM X* transcription 109
5.3.4 *Pf*PM X characterisation 110
5.3.5.1 Recombinant enzyme 110
5.3.5.2 Pull down transgenic enzyme 112
5.4 Discussion and Conclusion: 113

Chapter 6: Generation of a *Pf*PM X *in silico* model and the identification of new inhibitors 120

6.1 Introduction: 122
6.1.1 *In silico* screening and inhibitor identification: 122
6.1.2 PM X and its orthologs: 123
6.2 Methods: 127
6.2.1 *Pf*PM X homology model generation: 127
6.2.1.1 *Pf*PM X template and model selection: 127
6.2.1.2 Preparation and Refinement of the working model: 127

6.2.2 Examining the model:..... 127
6.2.3 Preparation of the *Pf*PM X model for docking:..... 128
6.2.4 Preparation of Ligand libraries:..... 128
6.2.5 Inhibitor docking: 129
6.2.5.1 HIV PI library..... 129
6.2.5.2 GlaxoSmithKline..... 129
6.2.5.3 Sigma-Aldrich 129
6.2.5.4 NCI 129
6.2.5.5 Ambinter..... 129
6.2.6 Sensitivity of *P. falciparum* to docking hits *in vitro* 129
6.2.7 Enzyme Activity Assays..... 130
6.2.8 Pharmacophore hypothesis generation and screening..... 130
6.3 Results 131
6.3.1 Choice of a template and selecting a *Pf*PM X 'docking model' 131
6.3.2 Refinement of the *Pf*PM X 'docking model'..... 132
6.3.3 Examination of the model 132
6.3.4 Inhibitor Docking 133
6.3.5 Identification of lead inhibitors 136
6.3.5.1 GlaxoSmithKline library screen..... 136
6.3.5.2 Sigma-Aldrich library screen..... 137
6.3.5.3 NCI library screen 139
6.3.5.4 Ambinter library screen..... 141
6.3.6 Inhibition studies with lead docking "hits" 141
6.3.6.1 The effect of top docking "hits" on parasite growth *in vitro*..... 141
6.3.6.2 The effect of lead docking hits on enzyme activity..... 143
6.3.7 Pharmacophore model generation and library screening 145
6.4 Discussion and Conclusion: 146

Chapter 7: Discussion and Conclusion 152

Appendices 164

Appendix 1: Solutions 165

Appendix 2: Vector Maps 172

Appendix 3: Primers 175

Appendix 4: *Plasmodium falciparum* Plasmepsin IX and X sequences 176

A4.1 *Pf*PM IX 176

A4.1.1 *PfPM IX* genomic sequence 176

A4.1.2 *Pf*PM IX protein sequence 177

A4.1.3 r*Pf*PM IX protein sequence 178

A4.2 *Pf*PM X 178

A4.2.1 *PfPM X* genomic sequence 178

A4.2.2 *Pf*PM X and r*Pf*PM X protein sequence 179

Appendix 5: Crystal tray buffer conditions used in r*Pf*PM IX crystal screens.... 180

Appendix 6: NAMD Scripts for *in silico* modelling of *Pf*PM IX 184

A6.1: Step 1, initial minimisation of *Pf*PM IX in a water box 184

A6.2: Step 2, equilibrate temperature and pressure 187

A6.3: Step 3, 20 ps molecular dynamics simulation 192

A6.4: Step 4, 20 ns molecular dynamics simulation 198

Appendix 7: Top docking hits from *in silico* screens 204

Appendix 8: Phylogenetic tree 208

Bibliography 209

List of Figures and Tables

Chapter 1:

Figure 1.1: The *Plasmodium falciparum* life cycle. 6
Figure 1.2: Global Distribution of Malaria Risk.. 7
Figure 1.3: *Pf*PM IX and X transgenic parasites are less sensitive to HIV PIs than control parasites. 20
Figure 1.4: Schematic representations of *Pf*PM IX and X with the respective genes that encode them.. 22
Figure 1.5: Schematic mechanism of action of aspartic proteases 23

Chapter 2:

Table 2.1: Transfection of D10 parasites with pHH1-PM9cMycB generates transgenic parasites that over-transcribe *PfPM IX*. 42
Table 2.2: PNA treated cultures demonstrate reduced *PfPM IX* transcription.. 45
Table 2.3: qRT-PCR analysis demonstrates a reduction in *PfPM IX* transcription in PNA treated cultures. 46
Figure 2.1: Schematic representation of PF3D7_1430200 and *Pf*PM IX. 37
Figure 2.2: Stage specific transcription of *PfPM IX* as assessed by qRT-PCR analysis of cDNA from synchronised cultures, relative to the internal reference genes 38
Figure 2.3: *Pf*PM IX is expressed throughout the intra-erythrocytic cycle.. 39
Figure 2.4: *P. falciparum* PM IX is exported by intra-erythrocytic parasites into the host red cell.. 40
Figure 2.5: Transfection of D10 parasites with pHH1-PM9cMycB and pHH1-PM9GFPB generates *Pf*PM IX transgenic parasites. 41
Figure 2.6: PCR analysis of parasites containing a knock-out vector, pCD-PMIXKO, fails to show genomic integration 43
Figure 2.7: Treatment of *P. falciparum* with *Pf*PM IX targeted PNAs affects *in vitro* parasite growth.. 44
Figure 2.8: *Pf*PM IX targeted PNA treatment affects parasite morphology. 45

Chapter 3:

Table 3.1: Expression of semi-pro *Pf*PM IX. 60
Table 3.2: NCtrunc-r*Pf*PM IX expression and purification compared to Semi-pro *Pf*PM IX expression. 68

Figure 3.1: Expression and purification of semi-pro *Pf*PM IX.. 60
Figure 3.2: Concentration of dialysis material leads to loss of semi-pro PfPM IX 61
Figure 3.2: Purification of semi-pro PfPM IX by Gel Filtration and Cation Exchange.61
Figure 3.4: Activity test of Gel Filtration Fraction 28..... 62
Figure 3.5: Semi-pro *Pf*PM IX is active. 62
Figure 3.6: Activity and substrate cleavage by semi-pro *Pf*PM IX refolded in 50 mM sodium phosphate dibasic pH 7.0 buffer.). 63
Figure 3.7: Activity and substrate cleavage of semi-pro *Pf*PM IX refolded in 50 mM sodium phosphate dibasic pH 8.0 buffer.. 64
Figure 3.8: Activity and substrate cleavage of refolded semi-pro *Pf*PM IX dialysed in 50 mM bicine pH 8.0 buffer. 65
Figure 3.9: Increased yields and activity was seen when semi-pro *Pf*PM IX was refolded at pH 7.0. 66
Figure 3.10: Auto-activation of Semi-pro *Pf*PM IX. 67
Figure 3.11: Growth trial of *E. coli* containing the NCtrunc-r*Pf*PM IX expression construct.. 68
Figure 3.12: SDS-PAGE analysis of NCtrunc-r*Pf*PM IX expression trials.. 69

Chapter 4:

Table 4.1: Comparison of *P. falciparum* PM coding sequences (full length transcripts). 77
Table 4.2: Online server template choices for *Pf*PM IX structure prediction. 86
Table 4.3: Top BLAST results for deleted region of *Pf*PM IX. 90
Figure 4.1: Alignment of the 10 *P. falciparum* PMs. 81
Figure 4.2: Comparison of PM IX orthologs from six Plasmodium species..... 83
Figure 4.3: Homology model of *Pf*PM IX. 86
Figure 4.4: Ramachandran diagram of the Φ-Ψ dihedral angles of the *Pf*PM IX homology model.. 87
Figure 4.5: *Pf*PM IX in silico homology model shares similar active site architecture with aspartic protease structures. 88
Figure 4.6: Structural alignment of the *Pf*PM IX in silico homology model with 1AM5 and *Pf*PM II.. 89
Figure 4.7: Structure prediction removes a vital region of *Pf*PM IX. 90

Chapter 5:

Table 5.1: Parasites transfected with pHH1-PMXGFP have increased levels of *PfPM X* transcription. 108
Table 5.2: Transgenic parasites containing the *Pf*PM X antisense plasmid demonstrate reduced *PfPM X* transcription. 111
Figure 5.1: Schematic representation of PF3D7_0808200 and *Pf*PM X..... 104
Figure 5.2: *Pf*PM X is primarily transcribed in schizont stage asexual intra-erythrocytic *P. falciparum* parasites.. 105
Figure 5.3: Asexual intra-erythrocytic stage specific expression of *Pf*PM X by *P. falciparum*..... 105
Figure 5.4: *Pf*PM X is localised to the parasite cytoplasm..... 106
Figure 5.5: Transgenic cultures contain pHH1-PMXGFP. 107
Figure 5.6: PCR analysis of transgenic parasites containing pCD-PMXDKO. 109
Figure 5.7: Parasites transfected with pHHC-HX_AS_cMycB contain vector..... 110
Figure 5.8: Expression and purification of r*Pf*PM X. 111
Figure 5.9: r*Pf*PM X does not cleave the fluorogenic substrate M-2455.. 112
Figure 5.10: SDS-PAGE analysis of GFP tagged *Pf*PM X extracted from transgenic cultures.. 112
Figure 5.11: Western blot analysis of pull down material…….113
Figure 5.12: GFP tagged *Pf*PM X is active……..113

Chapter 6:

Table 6.1: Comparison of *Pf*PM X protein coding sequence to that of the remaining *Pf*PMs (full length transcripts). 123
Table 6.2: HIV PIs dock within the *Pf*PM X homology model.. 134
Table 6.3: Top docking “hits” from the GlaxoSmithKline library..... 136
Table 6.4: Top docking hits from a screen of the Sigma-Aldrich library.. 138
Table 6.5: *In silico* screening of the NCI library reveals top hits..... 139
Table 6.6: Top docking “hits” from an *in silico* screen of the Ambinter library..... 141
Table 6.7: *In vitro* sensitivities of *P. falciparum* to selected docking “hits”..... 142
Table 6.8: *In silico* pharmacophore screen of the Sigma-Aldrich library.. 146
Figure 6.1: Comparison of PM X orthologs from six Plasmodium species..... 126
Figure 6.2: Homology model of *Pf*PM X.. 131
Figure 6.3: Ramachandran diagram of the Φ-Ψ dihedral angles of the *Pf*PM X homology model.. 132

Figure 6.4: Structural alignment of the *Pf*PM X homology model with 1TZS and *Pf*PM II.. 133
Figure 6.5: HIV PIs dock within the *Pf*PM X homology model binding cleft..... 134
Figure 6.6: Structures of the HIV PIs. 135
Figure 6.7: Structures of top docking NCI hits.. 140
Figure 6.8: *In vitro* sensitivity of D10 and 3D7 to L-His and NFmoc.. 142
Figure 6.9: The *in vitro* sensitivities of transgenic cultures containing a C-terminally tagged *Pf*PM X to NFmoc and L-His. 143
Figure 6.10: Selected HIV PIs inhibit *Pf*PM X-GFP at 10 μM.. 144
Figure 6.11: Lead docking compounds do not inhibit *Pf*PM X-GFP. 144
Figure 6.12: Pharmacophore model for *Pf*PM X inhibition. 145
Figure 6.13: Structures of the top ranked pharmacophore search results. 147

Appendix 2:

Figure A2.1: pHH1-PMIXGFPB containing GFP tagged *PfPM IX* (McGeorge 2009). 172
Figure A2.2: pHH1-PMIXcMycB containing cMyc tagged *PfPM IX* (McGeorge 2009). 172
Figure A2.3: Double cross-over knock-out transfection vector for targeted gene disruption of *PfPM IX* (McGeorge 2009). 172
Figure A2.4: pHH1-PMXGFP containing GFP tagged *PfPM X* (Seidens 2010). 173
Figure A2.5: Double cross-over knock-out transfection vector for the targeted gene disruption of *PfPM X* (Seidens 2010). 173
Figure A2.6: Antisense transfection vector for the knock-down of *PfPM X* (Ljunggren 2006). 173
Figure A2.7: pGEX-6P-1 expression vector (GE Healthcare). 174

Appendix 3:

Table A3.1: Primers used and the respective sequences. 175

Appendix 5:

Table A5.1: Buffer conditions used in r*Pf*PM IX crystal screens. 180

Appendix 7:

Table A7.1: Top docking hits from the GlaxoSmithKline library.. 204

Table A7.2: Top docking hits from an *in silico* screen of the Ambinter library.. 206
Figure A7.1a: Structures of top binding GlaxoSmithKline inhibitors..... 204
Figure A7.1b: Structures of top binding GlaxoSmithKline inhibitors. 205
Figure A7.2: Structures of top binding Ambinter compounds 207

Appendix 8:

Figure A8.1: Phylogenetic tree of selected plasmepsins…….....208

List of Appendices

Appendix 1: Solutions

Appendix 2: Vector Maps

Appendix 3: Primers

Appendix 4: *Plasmodium falciparum* Plasmepsin IX and X sequences

Appendix 5: Crystal tray buffer conditions used in r*Pf*PM IX crystal screens.

Appendix 6: NAMD Scripts for *in silico* modelling of *Pf*PM IX

Appendix 7: Top docking hits from *in silico* screens

Ethics

All research was conducted in accordance with approved protocols and OGTR guidelines; DNIR486 v7298.

List of Abbreviations

Amp	ampicillin
ACT	Artemisinin Combination Therapy
APS	ammonium persulfate
ART	artemisinin
Asp	aspartic residue
BME	β-mercaptoethanol
BSA	bovine serum albumin
bp	base pair(s)
cDNA	complementary deoxyribonucleic acid
CHL	chloramphenicol
CRT	chloroquine resistance transporter
°C	degrees Celsius
DDT	dichloro-diphenyl-trichloroethane
DMSO	dimethyl sulfoxide
dNTP	deoxyribonucleic triphosphate
DNA	deoxyribonucleic acid
DV	digestive vacuole
EDTA	ethylene-diamine-tetra-acetic acid
ER	endoplasmic reticulum
ECL	enhanced chemiluminescence
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
GFP	green fluorescent protein
Gly	glycine residue
GST	glutathione s-transferase
HIV	human immune-deficiency virus
HIV PI	HIV protease inhibitor
HRP	horseradish peroxidase
hdhfr	human dihydrofolate reductase
His	histadine
hsp86	heat shock protein 86
IFA	immuno-fluorescence assay
IPTG	isopropylthio-B-D-galactoside
Kb	kilobase
kDa	kilodalton

KO	knock out
L-His	L-Histidine β-Napthylamide
LPV	lopinavir
NCI	national cancer institute
N-Fmoc	*N*-Fmoc-ethylenediamine hydrobromide
NFV	nelfinavir
O/N	overnight
PepA	pepstatin A
PCR	polymerase chain reaction
PbDT3'	*P.berghei dhfr* gene 3' UTR
PBS	phosphate buffered saline
PBST	phosphate buffered saline with tween 20
*Pf*EMP1	*P. falciparum* erythrocyte membrane protein 1
PM	Plasmepsin
RBCs	red blood cells
REX1	Ring exported protein 1
RFP	red fluorescent protein
rpm	revolutions per minute
RNA	ribonucleic acid
RT-PCR	reverse transcriptase polymerase chain reaction
RTV	ritonavir
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SQV	saquinavir
STE	Sodium Chloride-Tris-EDTA
TAE	tris-acetate EDTA (buffer)
TE	tris-EDTA (buffer)
TEMED	N,N,N,'N'- tetramethylethylenediamine
TM	transmembrane
TPV	tipranavir
UTR	untranslated region
WHO	World Health Organization

Acknowledgements

First and foremost I would like to thank Dr Tina Skinner-Adams. Her guidance and expertise were instrumental to my project and I certainly couldn't have done it without her. Tina was always there to give help and advice and her never ending patience is astounding. Enormous thanks goes to Assoc. Prof. Chris Brown. I am extremely grateful for his support and understanding with anything and everything. His patient explanations of anything chemistry and comments on 'bio babble' were always well humoured. Without his expertise nothing *in silico* in this project would have been possible. The last 4 years would not have been the same without Assoc. Prof. Don Gardiner. His ever-present humour meant there was rarely a dull moment in the laboratory. His bright ideas and in-depth knowledge were always invaluable.

Special thanks go to the members of the Malaria Biology Lab. It would not have been the same without Katharine's helpful advice, Aunty Karen's support, Chris' fatherly guidance, Fabio's entertaining stories and Liam's terrible 'dad jokes'. Thanks also to Liam for all of his help expressing and purifying *Pf*PM X.

Special mention must go to my malaria twin and plasmepsin pal, Alice Butterworth. She is awesome in general and has always encouraged me and pointed me in the right direction. I would have been lost without her constant support. Special mention must also go to the other students at QIMR-B, particularly Catherine and Leigh-Roy. They have made sure there is never a dull moment, especially with dinner/lunch conversation and family dining.

Thanks also goes to the members of the Andrews, McCarthy, McManus, Jones and Fredwards groups for brightening even the worst days in the lab.

Not everyone is as lucky as I was to get Denise and Kate as reviewers. They were awesome and gave me a lot of helpful advice.

Thank you to Tatjana Seidens and Elin Ljunggren for their previous work investigating PM X. A big part of my work built on what had been done by them.

Thanks must go to Ben Dunn for allowing me to become a part of his lab at the University of Florida. Thanks also for providing recombinant proteins. Awesome

friends such as Jeyko, Ashleigh, Meray, Adwoa, Svetlana and Alina made my time so far from home enjoyable and unforgettable.

Thank you to Dr Bela Stantic for his help with *de novo* modelling of *Pf*PM IX.

This project wouldn't have been possible without the funding I received throughout my PhD. Thank you to the Australian Government for my Australian Postgraduate Award, the Queensland Government for my Smart Futures PhD Scholarship and QIMR Berghofer Medical Research Institute for my Higher Degrees Committee Top-up Scholarship.

A gratuitous amount of thanks goes to my friends and family who have supported me throughout the many ups and downs of the last 4 years and endured my babbles about my parasites. They have always encouraged me to do great things and have been there for me when things have been a bit of a mess. Special thanks to my family who have dealt with my crazyness and supported me even when they had no idea what I was talking about. "You're a schizont!" "Santrap mattress!" I would be nowhere without them.

Publications

Rachael D. McGeorge, Alice S. Butterworth, Christopher L. Brown, Donald L. Gardiner, Katherine. T. Andrews, Ben M. Dunn and Tina S. Skinner-Adams. "HIV Protease Inhibitors target *Plasmodium falciparum* Plasmepsin IX: A potential new target for anti-malaria drug development". (Under revision, International Journal for Parasitology)

Rachael D. McGeorge, Alice S. Butterworth, Tatjana Seidens, Elin Ljunggren, Christopher L. Brown, Donald L. Gardiner and Tina S. Skinner-Adams. "*Plasmodium falciparum* Plasmepsin X: A potential new target for anti-malaria drug development". (In preparation)

Presentations:

2009 "Examining the antimalarial action of the HIV protease inhibitors". Australian Society for Medical Research Student Conference, Brisbane. Poster presentation.

2009 "Examining the antimalarial action of the HIV protease inhibitors". Australian Society for Parasitology Annual Conference, Sydney. Poster presentation.

2009 "Examining the antimalarial action of the HIV protease inhibitors". Queensland Institute of Medical Research Student Retreat, Mt Tamborine. Poster presentation.

2010 "Inhibitors of the Malarial Aspartic Protease PM IX – New directions in Malarial Chemotherapeutics?". Australian Society for Medical Research Student Conference, Brisbane. Poster presentation.

2010 "Inhibitors of the Malarial Aspartic Protease PM IX- New Directions in Malarial Chemotherapeutics?". XIIth International Congress of Parasitology, Melbourne. Poster Presentation.

2011 "Plasmepsins IX and X: New Antimalarial Drug Targets?". Australian Society for Medical Research Student Conference, Brisbane. Poster Presentation.

2011 "Plasmepsins IX and X: New Antimalarial Drug Targets?". Australian Society for Parasitology Annual Conference, Cairns. Oral Presentation.

2012 “Plasmepsins IX and X and the Search for New Antimalarials”. Northern Australia Malaria Symposium, Brisbane. Oral Presentation.

2012 “Models and Drugs: A Useful Combination?”. Queensland Institute of Medical Research Student Symposium, Brisbane. Oral Presentation.

2012 “Plasmepsins IX and X: New Anti-malaria Drug Targets?”. Australian Society for Medical Research Student Conference, Brisbane. Poster Presentation.

2012 “*Plasmodium falciparum* Plasmepsin X”. ESKITIS Student Seminar Series Griffith University, Brisbane. Oral Presentation.

2012 “Plasmepsin IX and X and the Search for New Antimalarials”. Australian Society for Parasitology Annual Conference, Launceston. Oral Presentation.

2012 “Plasmepsin IX and X: Targets for the Anti-Malaria Activity of Selected Antiretroviral Protease Inhibitors and the Search for Specific Inhibitors”. XVIII International Congress of Tropical Medicine and Malaria and XLVIII Congress of the Brazilian Society for Tropical Medicine, Rio de Janeiro, Brazil. Poster Presentation.

2012 “Towards a new generation of anti-malarial drugs with anti-disease and anti-transmission activity: Investigating PM 9 and 10.” South East Queensland Structural Biology Symposium, Brisbane. Oral Presentation.

2013 “*P. falciparum* aspartic proteases: Characterisation and inhibitor identification.” ESKITIS Student Symposium, Griffith University, Brisbane. Oral Presentation.

2013 “Towards a new generation of anti-malarial drugs with anti-disease and anti-transmission activity.” QIMR Berghofer Medical Research Institute Student Symposium, Brisbane. Oral Presentation.

Chapter 1: Introduction and Literature Review

1.1 Malaria:

1.1.1 A Lethal Parasitic Infection:

Malaria is a significant threat to public health and is considered man's most lethal parasitic infection. It threatens ~3.4 billion of the world's population (WHO, 2013) and causes the death of up to 1.17 million people each year (Murray, Rosenfeld *et al.*, 2012, Lozano, Naghavi *et al.*, 2012, WHO, 2013). The name malaria came from the Latin *mala aria* meaning "bad air", as this disease was originally thought to be caused by inhaling contaminated air from marshes and swamps (reviewed in Buonsenso and Cataldi, 2010). It wasn't until 1880 that the first malaria parasite was described by the French physician Charles Louis Alphonse Laveran (reviewed in Buonsenso and Cataldi, 2010). Malaria has plagued mankind for centuries and despite extensive control efforts and scientific advancements, it still remains one of the leading causes of death in third world countries (WHO, 2013, Murray, Vos *et al.*, 2012, Lozano, Naghavi *et al.*, 2012). Many of the deaths caused by malaria are of children, with this disease being responsible for a significant proportion of the childhood mortality in endemic countries (Hay, Guerra *et al.*, 2004, WHO, 2013, Lozano, Naghavi *et al.*, 2012). It has been estimated by the World Health Organisation (WHO) that a child dies of malaria every 60 seconds (WHO, 2013). In addition to high rates of mortality, malaria causes significant morbidity. Countries with high rates of disease also often suffer from high economic burdens and decreased economic growth rates (WHO, 2013).

The causative agents of malaria are protozoan parasites of the genus *Plasmodium*. Of the many species of *Plasmodium*, there are five which infect humans. These are *P. falciparum, P. vivax, P. knowlesi, P. ovale* (2 sub-species) and *P. malariae*. *P. falciparum* is medically the most important as it gives rise to the most severe form of malaria and is responsible for the majority of the mortality associated with this disease (Mackintosh, Beeson *et al.*, 2004, Svenson, MacLean *et al.*, 1995, Hay, Guerra *et al.*, 2004, Mueller, Zimmerman *et al.*, 2007, Singh, Kim Sung *et al.*, 2004, White, 2008, Sutherland, Tanomsing *et al.*, 2010). While *P. vivax* and *P. knowlesi* can also cause severe disease and death, *P. ovale* and *P. malariae* typically result in less severe forms of malaria (Mueller, Zimmerman *et al.*, 2007, Svenson, MacLean *et al.*, 1995, Singh, Kim Sung *et al.*, 2004, White, 2008). The research described in the current project focuses on *P. falciparum* and the discussion of the disease caused by this parasite will be the focus of this review.

1.1.2 *Plasmodium falciparum*:

1.1.2.1 Life Cycle:

While malaria parasites can be transmitted by blood transfusion or can be passed from mother to child (congenital malaria) a human host usually becomes infected when it is bitten by a female *Anopheles* mosquito (as reviewed in Frederich, Dogne *et al.*, 2002).

The *anopheline* mosquito injects sporozoites into the human blood stream during a blood meal (Figure 1.1 step i). In less than an hour these sporozoites migrate to the host liver and enter hepatocytes where they begin to divide asexually into merozoites (Figure 1.1 step ii). Hepatic parasites mature into schizonts that rupture to release thousands of new merozoites into the bloodstream. These parasites invade the host red blood cells and develop into asexual ring-stage parasites (Figure 1.1 step iii). These then mature into trophozoites before segmenting into schizonts. Each schizont contains many daughter merozoites. At the end of this 48 hour life cycle, infected erythrocytes lyse releasing more merozoites into the blood stream to infect other red blood cells to continue this cycle of asexual development (discussed in Frederich, Dogne *et al.*, 2002). In a process which is not completely understood, some of these parasites develop into gametocytes. When the human host is bitten by the *anopheline* mosquito, the gametocytes are taken up and finish the sexual stage of the cycle within the mosquito vector (Figure 1.1 step iv). The sexual stage of the parasite life cycle cannot be completed without the mosquito.

1.1.2.2 Falciparum Malaria:

The severe pathology of Falciparum malaria is believed to be associated with the ability of the mature asexual intra-erythrocytic parasites to sequester within the microvasculature of the host. Sequestration, or cytoadherence, is the process by which the parasitised red blood cells adhere to host endothelial cells, causing blockages which restrict blood flow (Mackintosh, Beeson *et al.*, 2004, MacPherson, Warrell *et al.*, 1985, Beeson, Amin *et al.*, 2002, Silamut, Phu *et al.*, 1999). The most severe pathology associated with this cytoadherence is cerebral malaria, a form of severe disease characterized by changes in mental status and coma (WHO2000). Other factors, including the ability of this parasite to invade red blood cells of all ages, are also likely to be associated with the virulence of Falciparum malaria (Fried and Duffy, 1998, Fried, Nosten *et al.*, 1998). Efforts to treat *P. falciparum* infection or protect humans from infection by the development of vaccines have been hindered by the high degree of

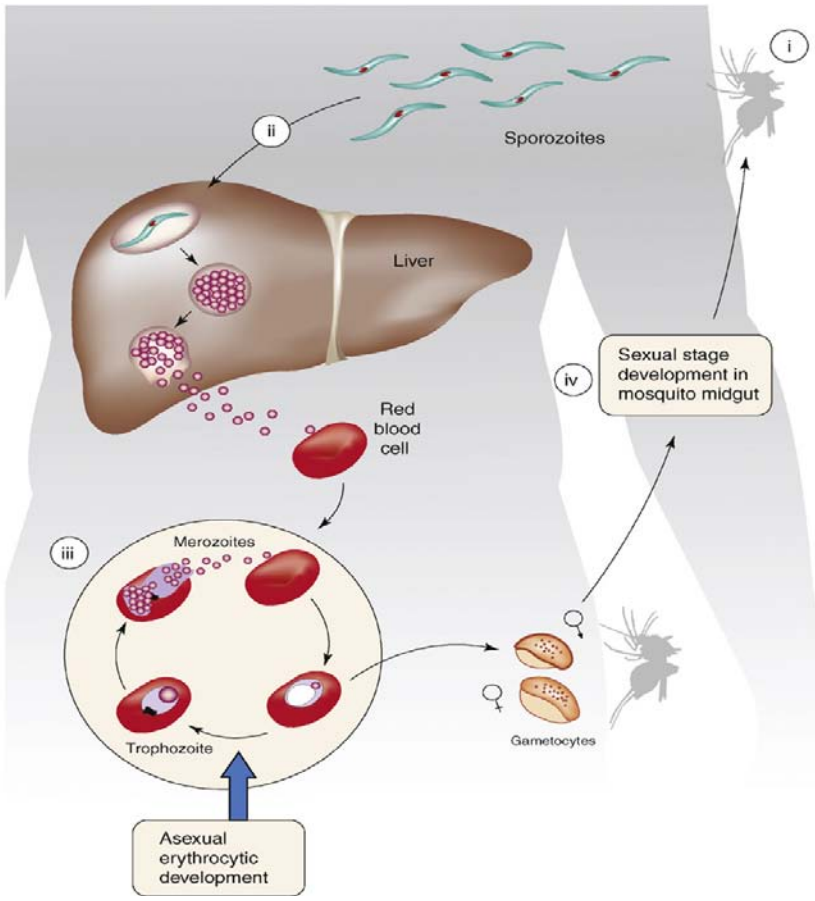


Figure 1.1: **The *Plasmodium falciparum* life cycle.** (Skinner-Adams, Stack *et al.*, 2010) The complex life cycle, steps i to iv, involves both a mosquito and human host. Within the human body, *P. falciparum* has multiple distinct developmental stages in a number of tissues with the intra-erythrocytic stages being responsible for the clinical symptoms of malaria.

antigenic variation in this parasite's exported membrane proteins, such as *Pf*EMP1 (the primary parasite ligand involved in cytoadherence) (as reviewed in Flick and Chen, 2004), and the by the ability of parasites to develop resistance to chemotherapeutics.

Falciparum malaria often causes nonspecific and viral-like symptoms that can lead to a delay in diagnosis and treatment (reviewed in Murphy and Oldfield, 1996). In non-immune patients, the incubation time for a *P. falciparum* infection is typically 11 days (reviewed in Murphy and Oldfield, 1996), however, this can be prolonged if the patient is taking prophylaxis. Fever, chills and headaches are the most common symptoms encountered, however, abdominal pain, nausea, myalgia and dizziness can also occur (Genton and D'Acremont, 2001). Complications associated with severe Falciparum

malaria include acute renal failure, cerebral malaria (discussed above), severe anaemia, acidosis, hypoglycaemia, coma and death (WHO 1990, WHO 2000).

1.1.2.3 Global Distribution:

Malaria is primarily a disease of the tropics/sub-tropics (Figure 1.2) with the greatest burdens (90% of malaria related deaths) occurring in Africa (WHO, 2013). Approximately 50% of the global population are located in these endemic countries (Hay, Guerra *et al.*, 2004, WHO, 2013).

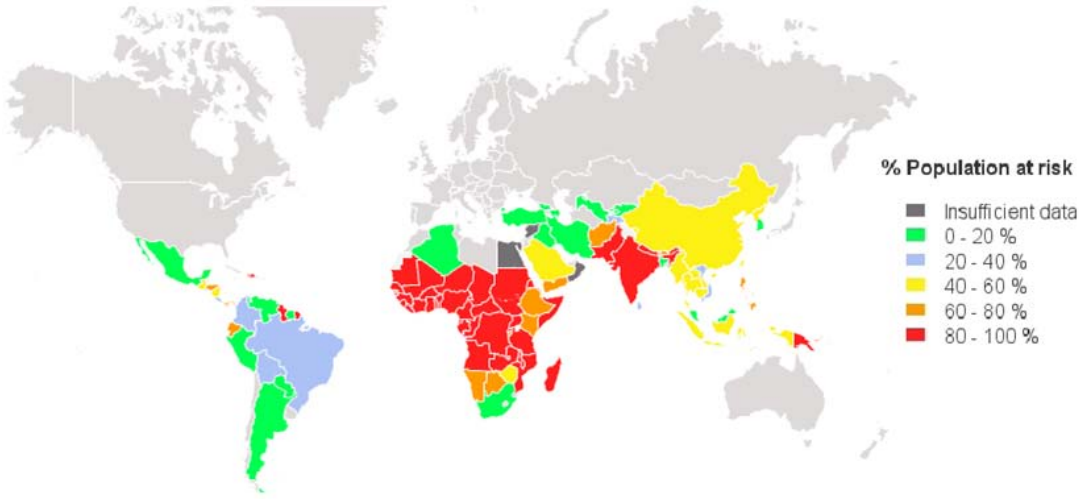


Figure 1.2: **Global Distribution of Malaria Risk.** Endemic regions of malaria are in the tropics and subtropics, with the majority of populations at high risk of malaria residing in Africa. Map generated with the Global Malaria Mapper (WHO Global Malaria Programme, 2014).

1.1.2.4 Disease Demographics:

Those populations most at risk of severe Falciparum malaria are pregnant women, infection-naive travellers, children under five and those co-infected with other diseases (WHO, 2012). Malaria is a particular problem in children as it takes time for individuals in endemic regions to develop immune tolerance. It is the same for travellers who go to endemic regions who have not previously been exposed to infection. Individuals at high risk who have not received antimalarial prophylaxis, have a pre-existing medical condition and/or have a delay in receiving treatment may develop severe disease and die (Schwartz, Sadetzki *et al.*, 2001, Bruneel, Hocqueloux *et al.*, 2003, Blumberg, Lee *et al.*, 1996).

Malaria is a particular problem if contracted by a pregnant woman (particularly if the woman is non-immune). This disease can increase the chance of miscarriage and is a major cause of death in women living in endemic regions (WHO, 2013). Pregnant women often develop associated complications such as severe anaemia and placental malaria (Nosten, Rogerson *et al.*, 2004, ter Kuile, Parise *et al.*, 2004). While the precise reasons behind a pregnant woman's increased risk of malaria and severe disease are not yet understood, it is believed to be due, at least in part, to a modified immune response (Rogerson, 2003, Nosten, Rogerson *et al.*, 2004). Cytoadhesion of parasites to endothelial cells within the placenta has also been implicated in the increased risk to a pregnant woman and her foetus' (Beeson and Duffy, 2005, Duffy and Fried, 2005). The placenta provides the perfect environment for malaria parasite development as it allows them to avoid elimination by the spleen and to delay detection (Beeson, Amin *et al.*, 2002).

1.1.2.5 Disease Interactions:

The severity of malaria can be affected by co-infection with other diseases. A prime example of this is malaria parasite and HIV co-infection. HIV/AIDS can be contracted in many regions of the world that are endemic for malaria and as this disease spreads, the incidence of malaria and HIV co-infection also increase (as reviewed in Skinner-Adams, McCarthy *et al.*, 2008, ter Kuile, Parise *et al.*, 2004, Nathoo, Serghides *et al.*, 2003, Hochman and Kim, 2012). HIV/AIDS is the leading cause of death in sub-Saharan Africa, with approximately 1.5 million deaths in 2010 (Lozano, Naghavi *et al.*, 2012). While the consequences of HIV and malaria parasite co-infection are not currently understood, available data demonstrate that it is more detrimental to the host than a single infection (Grimwade, French *et al.*, 2004, Franke, Spiegelman *et al.*, 2010). HIV-positive individuals often develop higher malaria parasite burdens than their HIV-negative counterparts (Whitworth, Morgan *et al.*, 2000, Moore, Ayisi *et al.*, 2000, Francesconi, Fabiani *et al.*, 2001) and HIV-1 infection has been shown to be associated with clinical presentation of cerebral malaria in children (Imani, Musoke *et al.*, 2011). Similarly, malaria has also been shown to increase the HIV viral load in individuals (Kamya, Gasasira *et al.*, 2006, Franke, Spiegelman *et al.*, 2010). Treatment of malaria in co-infected individuals has been shown to reduce the elevated HIV viral loads (reviewed in Modjarrad and Vermund, 2010).

Additional infections, such as hookworm and those of a bacterial origin, including aspiration pneumonia and sepsis, can also cause complications in patients with severe malaria. Hookworm infection has been shown to be associated with malaria infection, with children under 15 years of age being more likely to contract a hookworm infection when already infected with malaria (Humphries, Mosites *et al.*, 2011). Bacterial co-infections can result in malaria symptoms such as fever and chills being overlooked, leading to a delay or absence of malaria treatment (reveiwed in Trampuz, Jereb *et al.*, 2003). Complications caused by infection such as metabolic acidosis and failure of oxygen delivery to vital organs can also lead to severe disease and death (Taylor, Borgstein *et al.*, 1993, Marsh, Forster *et al.*, 1995, English, Waruiru *et al.*, 1996).

1.1.3 Vector Control:

A number of methods are in use to control malaria and prevent infection. These methods are primarily based on vector (mosquito) control strategies (Hay, Guerra *et al.*, 2004, WHO, 2008, WHO, 2013). Due to the mosquito's role in Plasmodium transmission, significant efforts have been made to minimise or eliminate the spread of malaria by reducing infective bites. The two main methods currently used to reduce transmission are insecticide-treated bed nets (ITNs) and indoor residual spraying with insecticides. While these methods have proven effective in reducing malaria morbidity and mortality among children in endemic regions (Lengeler, 2004, Eisele, Larsen *et al.*, 2010), particularly in areas such as Zambia's Luangwa district with near-universal mosquito net coverage (Eisele, Miller *et al.*, 2011), there is growing concern of resistance to these insecticides spreading and rendering the ITNs and indoor spraying ineffective. A recent study by Djogbenou et al. (2010) found evidence of resistance to dichlorodiphenyltrichloroethane (DDT) in all but two of the sites tested and to permethrin in over half of the sites tested (Djogbenou, Pasteur *et al.*, 2010). Additionally, mosquito pyrethroid resistance, the drug recommended for use with ITNs, has been increasingly detected (N'Guessan, Corbel *et al.*, 2007, Jones, Machin *et al.*, 2012, Ranson, N'Guessan *et al.*, 2011).

1.1.4 Vaccine Development:

There is no licensed vaccine available for malaria. It has been suggested that a successful vaccine would block the transmission of parasites by impeding sexual reproduction and transmission (Targett and Greenwood, 2008, Penny, Maire *et al.*, 2008) and would ideally be used in conjunction with chemotherapy that is active against

asexual blood stage parasites. However, many of the antigens that have been investigated as vaccine candidates are expressed by intra-erythrocytic stage parasites. A small sub-population of these has progressed to clinical trials, but have had limited success (Targett and Greenwood, 2008, Bejon, Lusingu *et al.*, 2008, Bejon, Mwacharo *et al.*, 2006, Draper, Goodman *et al.*, 2009).

One of the vaccines that has progressed to clinical trials is the RTS,S vaccine. This vaccine targets the circumsporozoite protein of *P. falciparum* (expressed by pre-erythrocytic stages). In a three month study of 894 children in Tanzania, aged 1 to 4 years, RTS,S had a 53% adjusted efficacy rate (Bejon, Lusingu *et al.*, 2008). A later Phase 3 trial in seven African countries followed a much larger study group of 15,460 children in a 6-12 week and a 5-17 month old group. This showed an efficacy of up to 56% against clinical malaria and up to 47% efficacy against severe malaria (Agnandji, Lell *et al.*, 2011). More recently, an analysis of RTS,S/AS01E vaccine efficacy over four years determined the vaccine to be only 16.8% effective. Though more effective in the first 12 months after vaccination, efficacy fell rapidly in following years and was less effective with increasing malaria exposure (Olotu, Fegan *et al.*, 2013). Examples of two other less effective, vaccines are the $MSP1_{42}$-FVO/Alhydrogel and $MSP1_{42}$-3D7/Alhydrogel (Malkin, Long *et al.*, 2007). Both of these vaccines were designed to target the C-terminus of MSP1 (merozoite surface protein 1; expressed by blood stage *P. falciparum* parasites). While both of these vaccines were safe, they were not immunogenic (Malkin, Long *et al.*, 2007).

1.1.5 Chemotherapy:

1.1.5.1 Prevention:

Prophylactic chemotherapeutics are used to prevent severe disease in children, pregnant women in endemic regions and those who travel to these regions (Peters, Thigpen *et al.*, 2007, WHO, 2013). As many of these drugs target asexual stages of parasite development, their aim is to prevent disease rather than infection. Safe and effective chemoprophylaxis is becoming increasingly difficult to achieve as parasite drug resistance spreads. Examples of drugs commonly used for prophylaxis include sulfadoxine-pyrimethamine (Fansidar), mefloquine, amodiaquine and doxycycline (as reviewed in Schlitzer, 2007, Lefevre, Bindschedler *et al.*, 2000, WHO, 2013).

1.1.5.2 Current Treatment Guidelines for Falciparum Malaria:

Malaria treatment relies on prompt and correct use of chemotherapies. However, as with prophylaxis, the increasing spread of parasite drug resistance to current antimalarial agents is having devastating implications for the millions of people who become infected with *P. falciparum* every year. In an effort to reduce the spread of parasites resistant to currently used antimalarial agents, the WHO recommends treatment with a combination of drugs (WHO, 2013). Artemisinin combination therapies (ACT) are recommended as the first-line treatment option in many malaria endemic countries (WHO, 2013). These combinations include at least one of the artemisinin drugs and an additional antimalarial (such as lumefantrine, amodiaquine or mefloquine) (Davis, Karunajeewa *et al.*, 2005, WHO, 2013). Drug groups that are still used to treat malaria include quinolines, the antifolates and artemisinin derivatives.

1.1.5.2.1 Quinoline Antimalarials:

Quinine and quinidine were isolated from the bark of the cinchona tree in 1820 (reviewed in Schlitzer, 2007). Quinine is still used today with intramuscular injections being recommended as initial treatment in severe malaria cases (WHO, 2013). These arylamino alcohols have strong side effects such as nausea, headaches, tinnitus and blurred vision which have led to low compliance issues (as reviewed by Schlitzer, 2007). More importantly, both of these alkaloids can cause life-threatening cardiac arrhythmias and stimulate secretion of insulin causing hypoglycaemia. This is a particular problem during pregnancy (as reviewed by Taylor and White, 2004).

Chloroquine is a 4-aminoquinoline. It is well tolerated, associated with few serious side-effects (as reviewed in Schlitzer, 2007) and has historically been the drug of choice for malaria treatment and prophylaxis (Bloland, 2001). It is a lipophilic weak base with 2 protonation sites. When chloroquine is uncharged it enters the digestive vacuole (DV), primarily facilitated by passive diffusion (Ferrari and Cutler, 1991, Raynes, 1999). Once inside the DV, it becomes protonated by the acidic environment and cannot pass through the membrane again. It is generally accepted that once it accumulates within the DV, it interferes with the process of haemoglobin digestion by binding to the free heme and preventing its detoxification (Gligorijevic, McAllister *et al.*, 2006, Famin, Krugliak *et al.*, 1999). Haemoglobin digestion is an essential process for parasite survival as it provides parasites with amino acids and space within the erythrocyte (Bray, Mungthin *et al.*, 1998, Goldberg, Slater *et al.*, 1991, Jani, Nagarkatti *et al.*, 2008). Due to

widespread use since its discovery in 1934, *P. falciparum* chloroquine resistance is universal (Bloland, 2001) and as a result it is only used to treat sensitive non-falciparum malaria parasites (Mueller, Zimmerman *et al.*, 2007, WHO, 2013). Resistance in *P. falciparum* is primarily mediated by mutations in the chloroquine resistance transporter (*Pf*CRT) gene. These mutations allow parasites to actively pump chloroquine out of the DV, mitigating its effect (Ibrahim, Steenkeste *et al.*, 2009, Price, Uhlemann *et al.*, 2004). Interestingly, chloroquine has been found to increase the number of gametocytes produced by *P. falciparum* cultures *in vitro* (Peatey, Skinner-Adams *et al.*, 2009) a factor which may have been associated with the world-wide spread of resistant parasites.

While chloroquine is now rarely used in the field, other quinoline antimalarial agents including amodiaquine, mefloquine and piperaquine are used in the field to fight disease. These drugs are thought to have modes of action similar to chloroquine (mediated by the haemoglobin digestion pathway) (Hawley, Bray *et al.*, 1998, Ersmark, Samuelsson *et al.*, 2006, Fitch, 2004). However, data describing their modes of action are limited. It is clear however, that these drugs are effective against chloroquine resistant *P. falciparum* (Bhattacharjee and Karle, 1998, Ringwald, Eboumbou *et al.*, 1999).

Amodiaquine, is a 4-aminoquinoline that is believed to inhibit the process of parasite haemoglobin digestion (as reviewed in Schlitzer, 2007). This drug is more toxic than chloroquine and has only recently started being used more widely. It has however, been used in many developing countries for some time, partially due to its affordability (Davis, Karunajeewa *et al.*, 2005, Gupta, Thapar *et al.*, 2002). Amodiaquine is effective against parasites with a low-level resistance to chloroquine, but loses efficacy against high-level resistant *P. falciparum* (as reviewed by Taylor and White, 2004). Combination studies with amodiaquine and other drugs such as artesunate, have demonstrated that these combinations are more effective than each single agent and has lead to the use of these combinations to combat spreading drug resistance (Adjuik, Agnamey *et al.*, 2002, Song, Socheat *et al.*, 2011).

Like other quinolines, the mode of action of piperaquine is believed to be the inhibition of haemoglobin digestion within the DV (O'Neill, Willock *et al.*, 1997). Piperaquine's bulky chemical structure (Bisquinoline) is thought to inhibit the parasites ability to

pump it out of the DV; as they do with chloroquine (Briolant, Henry *et al.*, 2010). Piperaquine has also been investigated as a promising partner drug in ACTs and is now in use. Studies with these low cost combinations suggest minimal toxicity and high efficacy (Davis, Karunajeewa *et al.*, 2005, Mayxay, Thongpraseuth *et al.*, 2006, Krudsood, Tangpukdee *et al.*, 2007, Song, Socheat *et al.*, 2011).

Mefloquine is an arylamino alcohol and is commonly used for prophylaxis (Hoppe, van Schalkwyk *et al.*, 2004). While it is believed to have a similar action to chloroquine, the exact mode of action of this drug remains unknown (Ersmark, Nervall *et al.*, 2006, Hoppe, van Schalkwyk *et al.*, 2004). In a study by Hoppe et al. (2004) it was found that mefloquine inhibited haemoglobin endocytosis. Other studies suggest that mefloquine may inhibit the release of Ca^{2+} ions within the parasite and thus fusion of haemoglobin carrying vesicles with the DV (Fitch, 2004). An alternative mode of action to chloroquine would explain why mefloquine is still active against some chloroquine resistant parasites. Mefloquine treatment failures and parasites resistant to this drug have however, been found and are continuing to spread (Price, Uhlemann *et al.*, 2004, Bloland, 2001, Quashie, Duah *et al.*, 2013). A contributing factor in the increase of mefloquine resistance could be that it appears to induce gametocytogenesis (Peatey, Skinner-Adams *et al.*, 2009).

While additional 8-aminoquinolines are under investigation as antimalarial agents, the only currently licensed 8-aminoquinoline is primaquine. This drug is a unique and valuable tool against malaria parasite infection as it is the only drug effective against liver, asexual and sexual stages of malaria parasite development (Brueckner, Ohrt *et al.*, 2001, Pukrittayakamee, Vanijanonta *et al.*, 1994). Although the relatively short half-life (4-6 hours) of primaquine means that treatment requires daily administration for an extended period, often resulting in poor compliance, this drug is required for radical cure of relapsing malaria (Brueckner, Ohrt *et al.*, 2001). Tafenoquine, a new 8-aminoquinoline with a much longer half-life than primaquine, is also currently under investigation for use as an antimalarial agent (reviewed by Butterworth, Skinner-Adams *et al.*, 2013).

A significant disadvantage to the use of 8-aminoquinolines in the treatment of malaria is that these drugs cause haemolytic anaemia in patients with glucose-6-phosphate dehydrogenase (G6PD) deficiency. G6PD deficiency is common in malaria endemic

regions (as reviewed in Nkhoma, Poole *et al.*, 2009). Although primaquine has been in use for over 50 years the antimalarial mode of action of this drug remains unknown. Treatment with primaquine, however, is known to cause mitochondrial damage and oxidative stress (Bates, Meshnick *et al.*, 1990, Lanners, 1991). Like all other antimalarials, parasites resistant to primaquine exist (Brueckner, Ohrt *et al.*, 2001) and similar to most antimalarials there is some evidence to suggest that low doses of drug can increase gametocyte numbers *in vitro* (Peatey, Skinner-Adams *et al.*, 2009).

Halofantrine was developed by the Walter Reed Army Institute of Research and introduced into therapy in 1988 (Castel, 2003). This drug is active against chloroquine resistant *P. falciparum* but is highly lipophilic and insoluble in water, leading to problems with bioavailability (2005). It is believed to have a similar mode of action to mefloquine and cross-resistance has been found between these two drugs (reviewed by Hyde, 2005). One of the more severe side effects of halofantrine is a high risk of cardiac arrhythmia. For this reason, this drug is not widely used (reviewed in Schlitzer, 2007).

Lumefantrine, also known as benflumatol, is an arylamino alcohol developed in the 1970s by the Chinese academy of Military Science. It is structurally similar to halofantrine and has similar issues with bioavailability, best taken with a fatty meal (as reviewed by Schlitzer, 2007). Lumefantrine, however, does not have the serious side effects of halofantrine, with no dangerous cardiac effects (Ezzet, van Vugt *et al.*, 2000, Bindschedler, Lefevre *et al.*, 2000). Lumefantrine has a synergistic activity with artemether (Hassan Alin, Bjorkman *et al.*, 1999) and this combination (Coartem) has been used as an effective first-line treatment for uncomplicated malaria since the mid-2000s (Makanga and Krudsood, 2009, Makanga, Bassat *et al.*, 2011).

1.1.5.2.2 Antifolates:

Antifolate antimalarials, such as sulfadoxine-pyrimethamine or Fansidar, have been used in combination to treat non-severe *P. falciparum* infections for many years (Ersmark, Nervall *et al.*, 2006, Marks, von Kalckreuth *et al.*, 2005, Bloland, 2001). These drugs work by inhibiting the tetrahydrofolate biosynthesis pathway, a pathway essential for DNA synthesis, by specifically inhibiting the metabolic processes mediated by dihydrofolate reductase and dyhydopteroate synthase (Bloland, 2001). The WHO currently recommends the use of these drugs in pregnancy to prevent the transmission of malaria from mother to foetus (WHO, 2013). While side effects including a severe

skin disease have been linked to the prophylactic use of Fansidar (Bloland, 2001) this drug combination is considered to be safe for use during pregnancy (as reviewed by Peters, Thigpen *et al.*, 2007, WHO, 2013). Unfortunately the spread of drug resistant parasites (90% of all isolates in South America) and the development of multiple mutations conferring resistance to these drugs, has resulted in high failure rates of antifolate treatments (a 60% treatment failure rate has been reported in some locations) (Nzila, Ochong *et al.*, 2005). A contributing factor in this could be the effects of antifolates on increasing gametocytes (Sowunmi, Fateye *et al.*, 2005).

1.1.5.2.3 Artemisinin and derivatives: (Sesquiterpene lactones and metabolites)

The artemisinin drugs, artemether, artesunate and dihydroartemisinin, are derivatives of the parent artemisinin, a natural product from the Chinese herb *Artemisia annua* first extracted and crystallised in 1972 (reviewed in Hoppe, van Schalkwyk *et al.*, 2004, Buonsenso and Cataldi, 2010). Artemisinin drugs are potent endoperoxide antimalarials and arguably the best antimalarials currently available. Artemisinin combination therapies (ACTs) are the recommended first-line antimalarial treatment in the majority of malaria endemic countries (WHO, 2013) as combining artemisinin with other antimalarials such as lumefantrine and piperaquine provides a potent combination and helps counter the short half-life (45-180 minutes; (Eastman and Fidock, 2009, Rijken, McGready *et al.*, 2011)) and high recrudescence rates ($\geq$25%) that can occur when these drugs are administered alone (Hien and White, 1993, De Vries and Dien, 1996). It is believed that ACTs are the best strategy for reducing the development and spread of drug resistant parasites. Despite the wide spread use of artemisinin drugs, their exact mechanism of action remains unknown. The most accepted theory revolves around the endoperoxide moiety of these drugs which is believed to produce free radicals (Lacaze-Dufaure, Najjar *et al.*, 2010). In addition to the original artemisinin compounds, there are drug development programs producing new synthetic derivatives of these drugs (Vivas, Rattray *et al.*, 2007, Charman, Arbe-Barnes *et al.*, 2011). The action of these drugs is likely to be similar to artemisinin but this still needs to be determined.

While artemisinin drugs are potent and fast-acting antimalarials and the mainstay of treatment today, there have been reports of drug resistant parasites and treatment failures with these drugs (Rogers, Sem *et al.*, 2009, Dondorp, Nosten *et al.*, 2009, Borrmann, Sasi *et al.*, 2011). These reports are of great concern and together with the reports of resistance to all of the currently available drugs, serve to highlight the need

for new antimalarials with new and unique molecular targets or modes of action. While there have been reports of artemisinin drugs being active against gametocytes in the field (Bousema, Okell *et al.*, 2010), artemisinin was recently shown to increase the number of gametocytes produced by *P. falciparum* parasites (Peatey, Skinner-Adams *et al.*, 2009).

1.2 Identification of new Plasmodium drug targets

In 2002 the 22.8 Mb genome sequence of *P. falciparum* became available (Gardner, Hall *et al.*, PlasmoDB 2009) (www.plasmodb.org). This information revolutionized malaria research by providing researchers with important information in relation to drug targets and parasite biology. Annotation of the genome suggested that *P. falciparum* encodes approximately 5,300 proteins on its 14 chromosomes (Gardner, Hall *et al.*, 2002) and that a significant proportion of these proteins have no known counterpart in other organisms. This work also revealed that these parasites have an extremely A-T rich genome, a factor that has made the manipulation of *P. falciparum* genes historically very difficult (Gardner, Hall *et al.*, 2002).

With such a wealth of genomic information available, together with the development of suitable manipulation strategies including the development of malaria Gateway™ Transfection vectors (Tonkin, van Dooren *et al.*, 2004, Skinner-Adams, Lawrie *et al.*, 2003) researchers were able to begin investigating the roles of hypothetical *P. falciparum* proteins. This research improved and continues to impact on malaria biology understanding and drug discovery/development research.

1.2.1 Haemoglobin digestion by *P. falciparum*

Our understanding of parasite haemoglobin digestion is one aspect of parasite biology that has improved with the advent of molecular tools and genomic information. As a result of genome mining and molecular tools, we now understand that this essential parasitic process involves numerous parasite proteases and that while some of these enzymes are essential to parasite survival, many are associated with redundancy and can be removed from parasites *in vitro* (Liu, Gluzman *et al.*, 2005, Liu, Istvan *et al.*, 2006, Bonilla, Bonilla *et al.*, 2007, Sijwali, Koo *et al.*, 2006, Sijwali, Kato *et al.*, 2004).

Haemoglobin digestion is essential to the parasite as it is believed to provide essential amino acids and space while maintaining osmotic stability. During the course of asexual

intra-erythrocytic development, malaria parasites consume or digest approximately 75% of their host's red blood cell haemoglobin (Francis, Sullivan *et al.*, 1997, Loria, Miller *et al.*, 1999). The process begins in endocytic vesicles (Abu Bakar, Klonis *et al.*, 2010) and then predominately takes place in the parasite's acidic digestive vacuole (DV) (Goldberg, Slater *et al.*, 1990). Once within the DV, four aspartic proteases called the DV plasmepsins (PMs), (*Pf*PM I, II, IV and HAP), are involved in the initial cleavage of haemoglobin. Additional digestion into smaller peptides is then mediated by the cysteine proteases (*Pf* falcipains 2, 2' and 3) and a metalloprotease (*Pf* falcilysin) (Rosenthal, 2004, Rosenthal, 2002). These small peptides are subsequently digested by *Pf* cathepsin C into dipeptides (Liu, Istvan *et al.*, 2006, Klemba, Gluzman *et al.*, 2004) and then finally by aminopeptidases into individual amino acids (Liu, Istvan *et al.*, 2006, Ragheb, Bompiani *et al.*, 2009). Analyses of the *P. falciparum* genome suggest that *P. falciparum* expresses 10 different aminopeptidases and at least three of these have been validated as prime targets for drug development (Gardiner, Trenholme *et al.*, 2006, Dalal and Klemba, 2007, McGowan, Porter *et al.*, 2009). Of the remaining enzymes, *Pf falcipain 3* (Sijwali, Koo *et al.*, 2006) and *Pf cathepsin C* are believed to be essential (Klemba, Gluzman *et al.*, 2004), while *Pf falcipains 1, 2* and *2'* can be genetically disrupted (Sijwali and Rosenthal, 2004, Sijwali, Koo *et al.*, 2006, Sijwali, Kato *et al.*, 2004). While providing essential resources, haemoglobin digestion also results in the production of toxic free heme. Parasites deal with the production of this toxin by in-activating it or transforming it into hemozoin (Egan, Combrinck *et al.*, 2002, Slater, Swiggard *et al.*, 1991).

1.2.2 *P. falciparum* Plasmepsins

One of the groups of proteases that are involved in haemoglobin degradation and have generated a lot of interest in regards to suitable drug targets are the *Pf*PMs. While 10 PMs have been identified in the genome of *P. falciparum* (PlasmoDB, Coombs, Goldberg *et al.*, 2001) only five of these proteases have been investigated in any detail. Four of these, the *Pf*DV PMs (PMs I, II, IV and HAP), are involved in parasite haemoglobin digestion and have been attractive drug targets for many years. These *Pf*PMs have been found to localise to the DV through immuno-fluorescence assays (IFA) and phase-contrast microscopy techniques (Banerjee, Liu *et al.*, 2002) and found to degrade haemoglobin at a pH optima of 5-5.5 (Banerjee, Liu *et al.*, 2002, Goldberg, Slater *et al.*, 1990, Rosenthal, 2004). *Pf*PMs II and IV can also cleave spectrin, an erythrocyte membrane protein (Wyatt and Berry, 2002). While these *Pf*DV PMs

initially looked very promising as drug targets, recent studies have shown that they have redundant actions and are not essential to parasite survival *in vitro* (Liu, Gluzman *et al.*, 2005). Additionally, a study by Bonilla et al (2007) demonstrated that parasites can survive, *in vitro*, when all of these *Pf*DV PMs are knocked-out (Bonilla, Bonilla *et al.*, 2007). Interestingly three of the *P. falciparum* DV PMs (PMs I, II and HAP) are not found in other malaria species, another factor which has dampened drug discovery research efforts, as this also suggests they are not required for parasite survival (PlasmoDB). While a recent study by Spaccapelo et al. (2010) using a rodent malaria model has demonstrated that the loss of *Pf*PM IV *in vivo* leads to reduced virulence and as a result may be worth investigating further, *Pf*DV PM drug discovery efforts have been dampened (Spaccapelo, Janse *et al.*, 2010).

Today, research is focused on understanding the role/s of the remaining six *P. falciparum* PMs. Studies examining the non-DV PM, *Pf*PM V, for example have demonstrated that this protein is essential to parasite survival and plays a vital role within the endoplasmic reticulum (ER), preparing proteins for export to the red cell cytosol (Boddey, Hodder *et al.*, 2010, Russo, Babbitt *et al.*, 2010, Haase and de Koning-Ward, 2010). These studies have shown that *Pf*PM V is responsible for processing proteins with the Predicted EXport ELement (PEXEL) sequence (xLxE/Q/D) and thereby facilitates their export outside of the parasite into the infected red cell (Haase and de Koning-Ward, 2010, Boddey, Hodder *et al.*, 2010, Russo, Babbitt *et al.*, 2010). Comparatively, little is known about the five remaining *Pf*PMs, but these enzymes are now attracting research interest given that there is evidence that some may well perform essential, non-redundant functions making them ideal drug targets. Preliminary data generated in our laboratories, for example, suggests that at least two of these enzymes, *Pf*PM IX and X are essential to parasite survival and that the inhibition of these enzymes are at least partially responsible for the antimalarial activity demonstrated by the HIV protease inhibitors (HIV PIs).

Studies by our group and others have demonstrated that selected HIV PIs (including saquinavir and ritonavir) can inhibit the growth of asexual malaria parasites at clinically relevant concentrations in laboratory settings (Skinner-Adams, McCarthy *et al.*, 2004, Redmond, Skinner-Adams *et al.*, 2007). HIV PIs have also been shown to inhibit the growth of malaria parasites in mice (Andrews, Fairlie *et al.*, 2006) and sera taken from HIV patients receiving HIV PIs inhibits parasites *in vitro* (Redmond, Skinner-Adams *et*

al., 2007). Additional studies have now also shown that these drugs can inhibit Plasmodium late-stage gametocytes and pre-erythrocytic liver stages (Peatey, Andrews *et al.*, 2010, Hobbs, Voza *et al.*, 2009). As a result of only modest activity (IC_{50} in low μM) together with cost, these HIV PIs are unlikely to be antimalarials in their own right. While there is now evidence that HIV PIs may have positive effects on malaria incidence in children (Achan, Kakuru *et al.*, 2012), studies involving phase III clinical trials on women in sub-Saharan Africa have not discerned any beneficial effect of the use of lopinavir/ritonavir compared to nevirapine based therapy (Porter, Cole *et al.*, 2012), or protection from malaria infection (Skinner-Adams, Butterworth *et al.*, 2012). Possible reasons for this within the study by Porter et al. (2012) may include lopinavir levels in patient sera not reaching sufficient levels to exert antimalarial activity, as well as a reliance on clinically diagnosed malaria (Porter, Cole *et al.*, 2012). Porter et al. (2012) also suggest that their results may not be accurately generalised to areas of low malaria transmission, as all study sites were in high transmission areas where populations are likely to have higher pre-existing immunity (Porter, Cole *et al.*, 2012). The lack of protection from malaria infection determined by Skinner-Adams et al. (2012) may have been the result of low numbers of malaria diagnosis, resulting in an underpowered cohort within this study (Skinner-Adams, Butterworth *et al.*, 2012). Another possible limitation suggested by the authors is the high rate of cotrimoxazole prophylaxis (80%) within the study population. Such a high level may have reduced the episodes of malaria detected, as well as caused potential off-target effects that may have impacted the activity of the HIV PIs tested (Skinner-Adams, Butterworth *et al.*, 2012). While the results of these studies do provide further evidence that HIV PIs are not ideal as antimalarials in their own right, their activity against multiple stages of parasite development is interesting and likely to be the result of a mechanism of action not currently exploited against malaria parasites. To understand this action in detail our group has been investigating the role of the non-*Pf*DV PMs in mediating the antimalarial activity of the HIV PIs.

HIV PIs are active against HIV as they inhibit HIV aspartyl protease, an essential enzyme for production of mature HIV. We hypothesised that these drugs also inhibit an aspartyl protease in malaria parasites. As the four characterised *Pf*DV PMs are not essential for parasite survival (Liu, Gluzman *et al.*, 2005) they are unlikely to be the targets of these drugs in asexual parasites. Additional data demonstrating that the

sensitivity of transgenic parasites over-expressing each of the *Pf*DV PMs to the HIV PIs does not change (unpublished) also supports this idea.

Investigations into the stage-specific expression of the remaining *Pf*PMs suggested that the most likely targets of the HIV PIs in erythrocytic asexual stages were *Pf*PMs V, *Pf*PM IX or *Pf*PM X. While *Pf*PMs V, IX and X are found in asexual blood stage parasites (PlasmoDB; (McGeorge, 2009)), *Pf*PMs VI VII and VIII are not expressed in these stages. Since these original observations, additional data suggested that HIV PIs target both *Pf*PM IX and X and that both enzymes are essential for parasite survival (McGeorge, 2009, Seidens, 2010). Research conducted in our laboratory, prior to the commencement of this project, demonstrated that transgenic *P. falciparum* lines over-expressing *Pf*PM IX and X are less sensitive to selected HIV PIs (Figure 1.3; (McGeorge, 2009, Seidens, 2010)). Additionally, recombinant *Pf*PM V is only weakly inhibited by HIV PIs and transgenic lines over-expressing *Pf*PM V show no change in sensitivity to these drugs (Boddey, Hodder *et al.*, 2010). Prior research also demonstrated that *PfPM IX* and *PfPM X* knock-out constructs could not be integrated into the *P. falciparum* genome (McGeorge, 2009, Seidens, 2010).

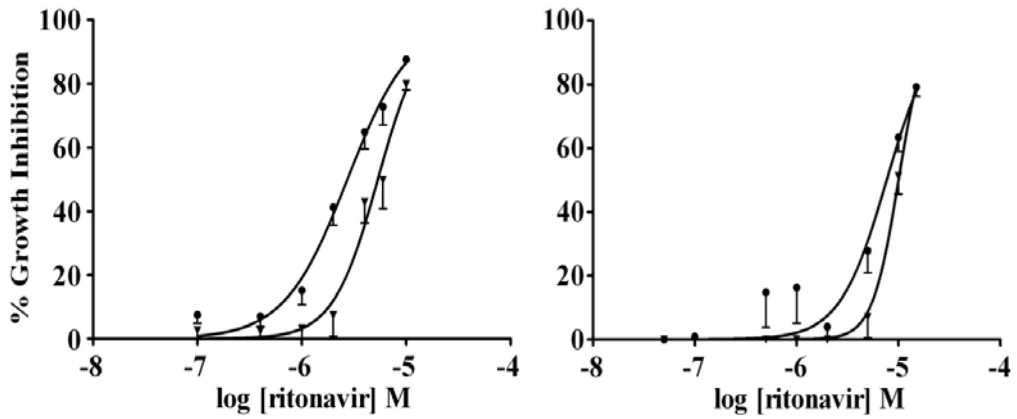


Figure 1.3: ***Pf*PM IX and X transgenic parasites are less sensitive to HIV PIs than control parasites.** Left: Sensitivity of *Pf*PM IX overexpressing transgenic (triangles; 150 nM WR99210) and control D10 (circles) parasites to the HIV PI ritonavir. Mean ± SD of 4 separate triplicate experiments (P=<0.001)(McGeorge, 2009). Right: Sensitivity of *Pf*PM X transgenic parasites (triangles; 50 nM WR99210) and control D10 (circles) parasites to the HIV PI ritonavir. Mean ± SD of 1 triplicate experiment (P=0.008) (Seidens, 2010).

While additional HIV PI targets may exist in malarial parasites, all of the current data suggest that *Pf*PM IX and X perform essential functions in malaria parasites and warrant further investigation as new antimalarial drug targets. To validate *Pf*PM IX and X as targets for drug discovery additional research is now needed to establish the function/s of these enzymes. Specific *Pf*PM IX and *Pf*PM X inhibitors also need to be identified so they can be used as tools to examine enzyme function and provide proof of concept data on the antimalarial potential of inhibitors targeted to these specific PMs.

1.2.3 *Pf*PM IX and X:

*Pf*PM IX, a predicted 627 amino acid protein (~74 kDa; PlasmoDB) is encoded by a 3118 bp gene (PF3D7_1430200) on chromosome 14 of the *P. falciparum* genome. The catalytic aspartic residues of this protease are at positions 31 and 224 in the predicted mature amino acid sequence. *Pf*PM IX contains a PEXEL sequence (Figure 1.4), which marks the protein for processing by *Pf*PM V in the ER and subsequent export into the host red cell (Russo, Babbitt *et al.*, 2010, Boddey, Hodder *et al.*, 2010), a factor that strongly suggests it is transported out of the parasite into the host red cell (Figure 1.4A). The protein has a predicted isoelectric point of 9.63 and as yet, there is no data describing its optimum functioning conditions. Orthologs of this gene can be found in all available Plasmodium genome sequences (PlasmoDB).

*Pf*PM X is encoded by a single exon, 1722 bp gene (PF3D7_0808200) located on chromosome 8. It is predicted to encode a 573 amino acid protein of ~68 kDa (PlasmoDB) with the catalytic aspartic residues at position 266 and 457 in the full length amino acid sequence (Figure 1.4B). Like *PfPM IX*, orthologs of *PfPM X* can be found in all available Plasmodium genome sequences including *P. berghei, P. chabaudi, P. Knowlesi, P. vivax* and *P. yoelii* (PlasmoDB). *Pf*PM X shares 37 % sequence identity with *Pf*PM IX (PlasmoDB) and both are grouped to Group C of the five phylogenetic groups of Apicomplexan aspartic proteases (Shea, Jakle *et al.*, 2007).

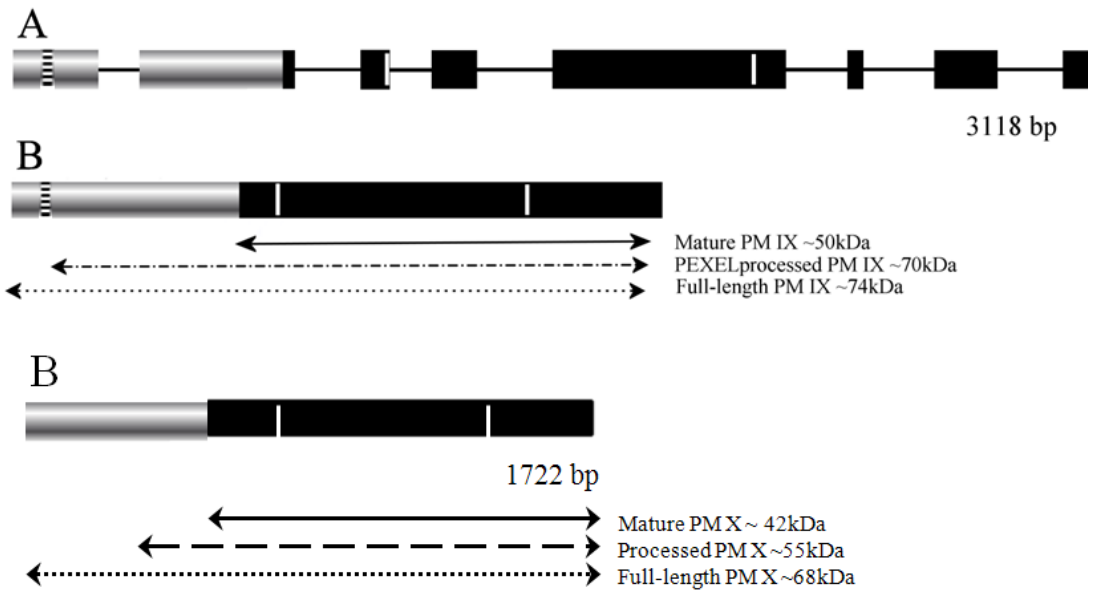


Figure 1.4: **Schematic representations of *Pf*PM IX and X with the respective genes that encode them.** (A) Cartoon of PF3D7_1430200, the gene encoding *Pf*PM IX and a cartoon of *Pf*PM IX. By Tina Skinner-Adams. (B) Cartoon of the single exon gene encoding *Pf*PM X, PF3D7_0808200. The predicted pro-domains are depicted in grey and the *Pf*PM IX PEXEL motif in striped black and white. Solid white regions indicate locations of catalytic aspartic amino acids.

As aspartic proteases, *Pf*PMs cleave proteins via an acid-base mechanism. Briefly, a water molecule becomes 'activated' (converted into a more active nucleophile) via interactions with one of the two active aspartic residues of the protease (Figure 1.5). When the peptide substrate is bound adjacent to this region, the nucleophillic oxygen atom binds to the substrate's amide carbonyl group, causing the formation of a tetrahedral intermediate with both active site ASP residues interacting with the substrate. The nitrogen of the amide bond within the substrate becomes protonated via a proton transfer reaction and the substrates amide bond is cleaved. Inhibitors of aspartic proteases often mimic the tetrahedral intermediate of the cleavage reaction and bind tightly to their targeted enzyme effecting inhibition (reviewed in Ersmark, Samuelsson *et al.*, 2006, Verkhivker, Tiana *et al.*, 2008, Reddy, Ali *et al.*, 2007).

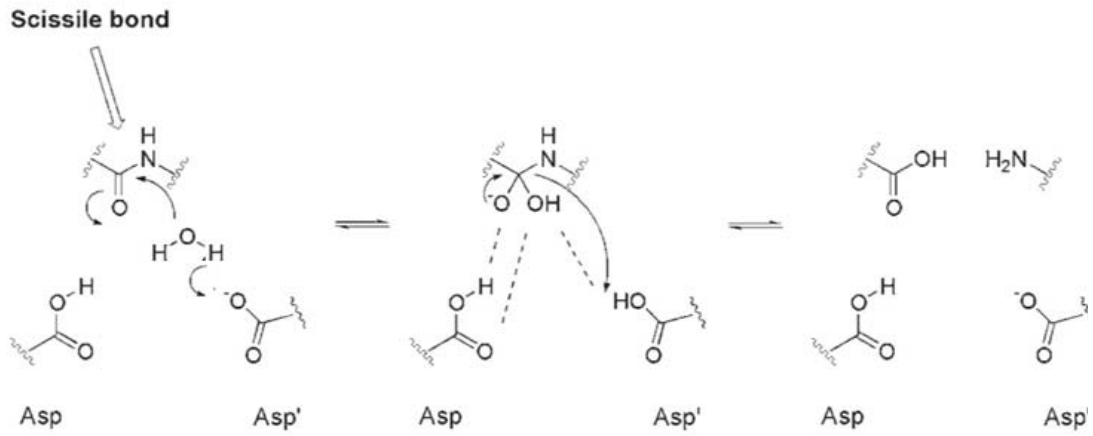


Figure 1.5: **Schematic mechanism of action of aspartic proteases** (adapted from Ersmark et al., 2006). A water molecule facilitates the action of aspartic proteases.

1.3 Methods to Determine Protein Function in *P. falciparum*:

There are a range of approaches available to researchers to investigate protein function in *P. falciparum*. Protein localisation and expression within different life cycle stages can be examined using immunofluorescence. Immunofluorescence assays (IFAs) can also be used to co-localise proteins with known function or location. This process relies on the use of specific antibodies and/or protein tags to determine the location of the protein within different parasite stages. IFAs were used by Banerjee et. al. (2002) to determine the location of *Pf*PMs I-IV in the parasite DV (Banerjee, Liu *et al.*, 2002). Protein specific antibodies can also be used to examine a protein's stage-specific expression through the use of Western blots on parasite protein samples collected at different life-cycle stages. This technique when used in combination with IFAs can also accurately determine the location of proteins within particular parasite organelles or compartments (membrane partitioning; DV location; red cell cytosol) (Goldberg, Slater *et al.*, 1990, Jackson, Spielmann *et al.*, 2007).

An alternative method to determine localisation that is widely used in the malaria field is the generation of transgenic parasites that express tagged proteins. These can then be located either directly (GFP, RFP and YFP) or indirectly using antibodies (Talman, Blagborough *et al.*, 2010). This technique was recently used to locate *Pf*PM V to the endoplasmic reticulum with a tagged *PfPM V* integrated within the parasite genome (Boddey, Hodder *et al.*, 2010). While this technique is limited by the fact that the parasite population needs to be genetically modified it does have the advantage of allowing live cell imaging, meaning the trafficking of a protein can be viewed directly

and in real time. Real time visualization of GFP tagged *Pf*PM II has been useful in determining the trafficking of this protein to the parasitophorous vacuole of the parasite before being transported back to the DV via endocytic vesicles (Klemba, Beatty *et al.*, 2004). Ideally, both IFA with protein specific antibodies and real time visualisation with live cell imaging should be performed to locate and investigate protein trafficking accurately. Tagging proteins for live cell imaging or for fixed IFAs can be difficult when the protein of interest is processed.

Tagging proteins by generating transgenic parasites can also facilitate immunoprecipitation experiments when specific antisera is not available (Boddey et al., 2010). Immunoprecipitation of native protein can provide material for further functional and activity studies. In the event that a recombinant protein is available, these studies also allow native and recombinant protein comparisons to be made. In *P. falciparum* and other biological systems, a substantial number of enzyme studies use either native enzyme or recombinant protein expressed in *E. coli* to investigate enzyme biochemistry (Klemba and Goldberg, 2005, Wyatt and Berry, 2002, Cawley, Olsen *et al.*, 1998, Boddey, Hodder *et al.*, 2010, Friedman, Gordon *et al.*, 1995, Xiao, Tanaka *et al.*, 2007, Stack, Lowther *et al.*, 2007). Through determination of an enzymes optimum activation and functioning conditions, researchers can gain a better understanding of where/what an enzymes role is and ascertain that the recombinant protein that they have produced is representation of the native enzyme. The recombinant *Pf*DV PMs for example, have an optimum pH of 5-6 for substrate cleavage (Banerjee, Liu *et al.*, 2002, Rosenthal, 2004, Goldberg, Slater *et al.*, 1990) which correlates with the optimum pH of 5-5.5 of the DV (Goldberg, Slater *et al.*, 1990).

Various tools are available to investigate the importance of a gene/protein to malaria parasite growth and survival. However, because of their very nature most of these techniques are only useful if the protein or gene of interest is not essential and therefore can be manipulated. Targeted gene disruption studies for example are only successful in the event that integration and disruption is possible (Liu, Gluzman *et al.*, 2005, Bonilla, Bonilla *et al.*, 2007), meaning if a gene is essential to parasite survival, it will not be able to be knocked out. To learn more about the importance of a gene/protein that cannot not be disrupted, additional tools are required. One such tool available to malaria biologists is complementation. In this process, a plasmid containing a copy of the gene of interest, and a new selectable marker, is transfected into cultures containing a knock-

out vector to demonstrate that knock-out is possible if the gene is supplied via an additional mechanism. This has been successful in both *P. falciparum* and *T. gondii* and can be used to examine gene function (Freville, Landrieu *et al.*, 2012, Soldati, Kim *et al.*, 1995, Triglia, Healer *et al.*, 2000). Another tool is the Shld system. This process involves the fusion of a ddFKBP domain to the N or C terminal end of the protein of interest, which promotes degradation of the protein, or causes a knock-down effect (Armstrong and Goldberg, 2007). Addition of Shld 1 stops protein degradation and rescues the parasite. In this way expression of essential genes can be turned on and off, allowing investigation. This method cannot be used if the protein is N and C terminally cleaved as the ddFKBP domain will be removed from the protein. Specific inhibitors of individual proteins have also been used to investigate protein function (Eksi, Czesny *et al.*, 2007). However, such inhibitors are often unavailable or not specific for the protein of interest.

1.4 Inhibitor Discovery

Once a target has been validated, screening and structural studies are required to identify and optimize the activity of potential inhibitors. Structural studies also allow the characterization of a protein. This is particularly important if the target is an enzyme, as these studies allow the binding cavity and the mechanism of inhibitor binding to be identified. This can give important information about substrate specificity and inhibitor designs that can optimize activity while minimizing inhibition of similar host enzymes.

While there is no replacement for the definitive information that can be obtained from X-ray crystallographic analysis of an enzyme, *in silico* modelling and virtual docking has become a powerful tool in drug design with improved efficiency for drug discovery and development (Kirkpatrick, Watson *et al.*, 1999). Studies have demonstrated that *in silico* docking studies can correlate well with actual binding in aspartic proteases (Ripka, Satyshur *et al.*, 2001) and as a result can provide reliable enzyme characterisation data. Molecular-docking-based virtual screening of chemical compounds can also vastly reduce the cost of inhibitor identification studies by reducing the number of potential drugs taken to additional screens. *In silico* screening depends on the availability of a target homology model or X-ray data. Online servers for the prediction of protein structure based on sequence data include SwissModel, 3D Jigsaw and PHYRE. These servers use one of two methods to generate templates for protein

structure prediction. SwissModel and 3D Jigsaw utilise the Basic Local Alignment Search tool (BLAST) and locate solid state templates based on the similarity of the proteins amino acid sequence (Schwede, Kopp *et al.*, 2003, Arnold, Bordoli *et al.*, 2006, Bates, Kelley *et al.*, 2001). Whereas other approaches such as PHYRE utilise secondary and tertiary structural domains as the template source (Kelley and Sternberg, 2009). Templates are then selected based on overall structure similarity (much of the nomenclature for domain clarification is described in the Structural Classification of Proteins (SCOP) database). Protein solid-state structures, unliganded and/or complexed with inhibitors are available to researchers through the Protein Data Bank (PDB – http://www.rcsb.org). All the structures in this repository have been derived from X-ray analysis or via solution state NMR analysis. For example, there are 6335 structures listed that are found using the keyword 'aspartic protease' alone and 29 structures found using the keyword 'plasmepsin' (Berman, Westbrook *et al.*, 2000). There are currently no published structures of *Pf*PM IX and X in the Protein Data Bank.

1.5 The Current Project:

While preliminary evidence suggests that *Pf*PM IX and *Pf*PM X play important roles in malaria parasites, the functions of these enzymes remain unknown. In this project a multifaceted approach was used to gain an understanding of the role of these proteins in asexual intra-erythrocytic *P. falciparum* parasites and to determine if these aspartic proteases are viable drug targets.

The hypotheses of this project were:

1) *Pf*PM IX and X have essential roles in *P. falciparum* and are good targets for anti-malarial drug development.
2) Drugs designed to inhibit these PMs may be potent and selective antimalarials and good tools for additional functional studies.

The aims of this project were:

1) To examine the structure and function of *Pf*PM IX and X.
2) To identify specific inhibitors of *Pf*PM IX and X and use these inhibitors to validate these enzymes as antimalarial drug targets and as tools in functional studies.

Chapter 2: Characterisation of *Plasmodium falciparum* Plasmepsin IX: A potential new target for anti-malaria drug development

2.1 Introduction

All of the drugs that are clinically available for the treatment of malaria have been associated with drug resistant parasites and even the "gold standard" antimalarial agents, the artemisinin combination therapies (ACTs), have now been associated with reduced efficacy and tolerant parasites (Noedl, Socheat *et al.*, 2009, Dondorp, Nosten *et al.*, 2009, Noedl, Se *et al.*, 2008). With an ongoing need to develop new antimalarial chemotherapeutics with unexploited targets within Plasmodium parasites, attention has turned to identifying new essential processes and proteins for drug development strategies. Parasite proteases, including the aspartic proteases or plasmepsins (PMs) have been of particular interest to scientists over the years. However, most of the studies in this area have been performed with the *P. falciparum* digestive vacuole PMs with research demonstrating that these enzymes are redundant and not essential for parasite survival (Liu, Istvan *et al.*, 2006, Bonilla, Bonilla *et al.*, 2007, Liu, Gluzman *et al.*, 2005). While *Pf*PMs I-IV, the digestive vacuole PMs, and *Pf*PM V are the best characterised to date, relatively little is known about the remaining five PMs, *Pf*PMs VI-X (PlasmoDB; www.plasmodb.org, discussed in Chapter 1). Nevertheless recent studies have suggested that these less well characterized PMs may well be good drug targets. Studies examining the antimalarial action of HIV protease inhibitors (HIV-PIs) have shown that these drugs can inhibit the growth of blood stage malaria parasites at clinically relevant concentrations (Redmond, Skinner-Adams *et al.*, 2007, Skinner-Adams, McCarthy *et al.*, 2004, Andrews, Fairlie *et al.*, 2006, Parikh, Gut *et al.*, 2005) and have suggested that they do so by targeting non-DV *Pf*PMs (Bonilla, Bonilla *et al.*, 2007) (discussed in Chapter 1). As *Pf*PM V is relatively insensitive to HIV PIs (Boddey, Hodder *et al.*, 2010) and only *Pf*PMs IX and X are expressed at any significant level during the intra-erythrocytic stages of Plasmodium development, HIV PIs are likely to target *Pf*PM IX and X, making them potential drug targets for the development of new antimalarial agents.

*Pf*PM IX is encoded by the gene PF3D7_1430200 (PlasmoDB; www.plasmodb.org) and little is known about the function and location of this enzyme. Orthologs of *PfPM IX* are present and syntenic in all sequenced Plasmodial species including *P. falciparum*, *P. berghei*, *P. knowlesi*, *P. vivax* and *P. yoelii*. *Pf*PM IX also bares very little resemblance to any of the nine other aspartic proteases in *P. falciparum*, sharing the highest sequence identity with *Pf*PM X at 37% (Papadopoulos and Agarwala, 2007). It is also the only PM that contains a PEXEL sequence, which marks the protein for

processing by *Pf*PM V in the parasite ER and subsequent export into the host red cell (Boddey, Hodder *et al.*, 2010, Russo, Babbitt *et al.*, 2010).

Very few studies into *P. falciparum* PMs have included *Pf*PM IX, however a study by Banerjee et al (2002) determined that *Pf*PM IX has a diffuse pattern of localisation and that it was excluded from the DV (Banerjee, Liu *et al.*, 2002), suggesting an alternative role to the haemoglobin digesting DV PMs. Additionally, Shea et al (2007) reported unpublished data demonstrating a Maurer's Cleft localisation within the host red blood cell (Shea, Jakle *et al.*, 2007). The Maurer's Clefts are structures that are exported into the host red cell and have a role in trafficking exported proteins (Hawthorne, Trenholme *et al.*, 2004, Langreth, Jensen *et al.*, 1978, Trager, Rudzinska *et al.*, 1966). Further evidence of a role distinct from the DV PMs is the peak expression of *Pf*PM IX within mid-schizont parasites (Bozdech, Llinas *et al.*, 2003) and predicted interactions with proteins known to have a role in invasion (Hu, Cabrera *et al.*, 2010). Peak expression within schizonts argues against a role in haemoglobin digestion and has led to speculations that *Pf*PM IX may have a role in merozoite invasion (Bozdech, Llinas *et al.*, 2003, Hu, Cabrera *et al.*, 2010).

To investigate *Pf*PM IX's potential as a new target for antimalarial chemotherapeutic development I sought to characterise this enzyme. The expression and localisation of *Pf*PM IX was investigated and the importance of *Pf*PM IX to parasitic survival examined. Knock-down of *Pf*PM IX was attempted using Peptide Nucleic Acids (PNAs), synthetic nucleic acid analogues, that have been found to be useful in other organisms (Ahn, Lee *et al.*, 2011, Choi, Kim *et al.*, 2011, Macadangdang, Zhang *et al.*, 2011, Kuhn, Sahu *et al.*, 2010, Ahn, Shim *et al.*, 2011) and have only very recently been investigated for specific knock-down ability in *P. falciparum* (Kolevzon, Nasereddin *et al.*, 2014). Previously generated transgenic parasites containing knock-out vectors and C-terminally tagged *Pf*PM IX were also used to investigate this enzyme.

2.2 Materials and Methods

2.2.1 Parasites

P. falciparum clones D10 and 3D7 were cultured essentially as previously described (Trager and Jensen, 1976). In brief, infected O+ human red blood cells (5% haematocrit and 1-10% parasitaemia) were maintained in a low oxygen environment (5% Carbon dioxide and 5% Oxygen in Nitrogen) in complete RPMI medium (Appendix 1) at 37°C.

Parasite growth was monitored microscopically by the preparation and examination of Giemsa stained thin blood smears (100X objective lens under oil) and cultures diluted with fresh uninfected O+ human red blood cells (Australian Red Cross Blood Transfusion Service) when required. Transgenic parasites, pHH1-PMIXcMycB and pHH1-PMIXGFPB, generated before the commencement of this project (McGeorge, 2009) and containing a full length *Pf*PM IX insert with a cMyc or GFP tag respectively (see Appendix 2 for vector maps), were thawed and maintained as described above with the addition of drug selection (Fidock and Wellems, 1997). Briefly, cryopreserved cultures were taken from liquid nitrogen and thawed in a 37°C water bath. The outside of the cryopreservation vial was then cleaned with 70% ethanol and the vial contents transferred to a 10 mL tube. $1/5^{th}$ the pellet volume of 12% NaCl was added drop-wise with gentle shaking. This was left at room temperature for 5 minutes and then 9x the pellet volume of 1.6% NaCl was added drop-wise, mixing gently after each addition. The 10 mL tube was then centrifuged at 1000 rpm for 2-3 minutes and the supernatant was removed and discarded. 9x the pellet volume of 0.9% NaCl was then added drop-wise, resuspending gently and the isolate centrifuged again at 1000 rpm for 2-3 minutes. The supernatant was then removed and the pellet resuspended in warm medium. Red blood cells were added when required. Transgenic parasites containing the 'knock-out' plasmid pCD-PMIXKO (McGeorge, 2009) were maintained in complete medium containing WR99210 alone, or WR99210 and 5'fluoro-uracil to force integration of the plasmid (Maier, Braks *et al.*, 2006).

2.2.2 Expression and Localization of *Pf*PM IX

The stage specific expression and localisation of *Pf*PM IX was investigated using qRT-PCR, immuno-blotting and immuno-fluorescence assays.

2.2.2.1 Real Time PCR

A 100 µl sample of parasite culture (packed cells) at >2% parasitaemia was mixed thoroughly with 1 mL of pre-warmed TRIzol (Invitrogen, USA). After 5 mins at room temperature, samples were either processed immediately or stored at -80°C. Total RNA was extracted from parasite samples and measured by NanoDrop (ThermoScientific, America). Briefly, 1 mL of parasite/TRIzol solution was mixed with 200 µl of chloroform and centrifuged at full speed for 30 mins at 4°C. The aqueous supernatant was taken and RNA precipitated in 500 µl isopropanol at -80°C for 1 hour before being

centrifuged as before. The pellet was then air-dried and resuspended in 20 µl of formamide. RNA was stored at -20°C or -80°C long term.

Reverse transcription was performed with random primers as per the manufacturer's instructions (QuantiTect Reverse Transcription Kit, Qiagen) and cDNA was checked for contaminating genomic DNA by PCR. Briefly, primer sets (Appendix 3) were chosen for genes that contained introns and the PCR product sizes were compared between cDNA and a genomic control, with genomic DNA producing a larger product size. Real time PCR was performed using a Rotor-Gene 6000 real time PCR Cycler (Corbett Research/Qiagen, Australia). Briefly 500ng of cDNA was added to 10 µl SYBR Green PCR Master-mix (Applied Biosystems, Australia) together with the required primer pairs. PM IX specific primers (forward 5'-tcc aga aat gtt atc gac tgg-3' and reverse 5'-ccc ctc caa ata tta atg cag a-3') were designed with the Primer3 online tool (Untergasser, Nijveen *et al.*, 2007). Internal reference genes for stage-specific analysis were *18s rRNA* (PF3D7_0725600) (forward 5'-cgg cga gta cac tat att ctt a-3' and reverse 5'-tta gta gaa cag gga aaa gga t-3'; (Durrand, Berry *et al.*, 2004)) and *Seryl-tRNA synthetase* (PF3D7_0717700) (forward 5'-ata gct acc tca gaa caa cc-3' and reverse 5'-caa gat gag aat cca gcg ta-3'; (Roseler, Prieto *et al.*, 2012)). Each RT-PCR run was performed in quadruplicate and repeated at least twice. Data were analysed with Rotor-Gene 6.0 software (Corbett Research/Qiagen, Australia). *PfPM IX* transcription was calculated relative to ratios of the internal reference genes.

2.2.2.2 Antibody Generation and Purification

Semi-pro *Pf*PM IX (amino acid 143-627; 57kDa; Chapter 3)) was extracted from inclusion bodies by SDS-PAGE and used to immunize 2 rabbits (IMVS Pathology, SA). Rabbits were immunized subcutaneously with an emulsion of semi-pro *Pf*PM IX (200 µg/dose) and adjuvant at 3 weekly intervals (1st dose was emulsified in Freund's Complete Adjuvant, while the 2nd, 3rd and 4th doses were emulsified in Freund's Incomplete Adjuvant). Pre-bleeds and test bleeds were collected at intervals with final bleed-outs occurring 10 days post the 4th immunization. Final bleed-outs were purified/concentrated with a Merck Millipore Montage Antibody Purification Kit as per the manufacturer's instructions.

To remove non-specific binding, α*Pf*PM IX antisera (1:1000 in 5% skim milk/1xPBS) was pre-absorbed against total proteins from RBCs, *E.coli*, and parasites. Briefly,

membranes were soaked in lysed RBC, *E.coli* or parasite pellet for 5 mins at room temperature and then washed with 1x PBS at least 3 times. Membranes were then placed in antisera and incubated with rolling. Membranes were replaced at 15 mins, 1 hour and then after an overnight incubation at 4°C.

2.2.2.3 Immuno-blotting

Packed cell pellets (>5% parasitaemia) were subjected to saponin lysis. Briefly, 40 µl of pellet was mixed thoroughly with 100 µl 0.03% saponin in PBS (Sigma-Aldrich) and incubated on ice for 15 mins. Samples were then centrifuged at full speed (16,000 rpm) for 5 mins at room temperature and the supernatant removed. Pellets were washed 3x with pre-chilled 1x PBS (Appendix 1) and resuspended in 50 µl of 1x PBS. Samples were stored at -20°C. Proteins from saponin-lysed parasite extracts were resolved by reducing 12% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE; 1x running buffer; Appendix 1) and transferred to polyvinylidene fluoride (PVDF) membrane using a semidry electroblotter (Trans-Blot, Bio-Rad) in 1x transfer buffer (Appendix 1). Membranes were probed with either purified rabbit polyclonal anti-sera to recombinant semi-pro-*Pf*PM IX (1:1000; IMVS Pathology, SA), mouse monoclonal antibody to c-Myc (1:2500; Sigma-Aldrich) or mouse monoclonal antibody to GFP (1:500; Roche) overnight at 4°C, followed by a horseradish peroxidase-labelled anti-mouse or anti-rabbit IgG antibody (1:2000 dilution, Chemicon International Inc.). Membranes were stripped and re-probed with an anti-glyceraldehyde-3-phosphate dehydrogenase (GAPDH) rabbit antibody (1:2000 dilution; (Daubenberger, Tisdale *et al.*, 2003)) followed by a horseradish peroxidase-labelled anti-rabbit IgG antibody (1:2000 dilution, Chemicon International Inc.) to assess protein loading.

2.2.2.4 Immuno-fluorescence

Thin blood smears were fixed with 75% acetone/25% ethanol prior to a short Triton X-100 permeabilisation step and exposure to rabbit α*Pf*PM IX (1:150), mouse monoclonal antibody to c-Myc (1:2500 dilution; Sigma-Aldrich), mouse monoclonal antibody to REX1 (1:1000; (Hawthorne, Trenholme *et al.*, 2004)) or mouse monoclonal antibody to GFP (1:500; Roche) in 3% BSA/PBS for 1 hour at room temperature. Smears were washed 5 times with 1x PBS with vacuum aspiration. Bound antibody was visualized with goat anti-mouse IgG-Cy2 (10 µg/ml; Amersham), goat anti-rabbit 488 (1:300; Alexa), goat anti-mouse 555 (1:300; Alexa), goat anti-rabbit IgG-Cy2 (10 µg/ml; Amersham) or goat anti-rabbit IgG-Texas Red (10 µg/ml; Amersham) with Hoechst dye

(0.5µg/ml; Sigma-Aldrich) used for nuclei staining (3% BSA/PBS for 1 hour at room temperature and then washed as above). Parasites were visualised on a Delta Vision Olympus IX71 and images analysed with Corel Paint Shop software.

2.2.3 Examining the role and importance of *Pf*PM IX

Previously generated transgenic parasites containing episomal C-terminally tagged *Pf*PM IX or *Pf*PM IX knock-out vectors were used to investigate the role and importance of *Pf*PM IX. Preliminary experiments investigating the expression and localisation of C-terminally tagged *Pf*PM IX was investigated prior to the commencement of this project (McGeorge, 2009). However, additional follow-up experiments were also performed during the current work. PNAs were also examined as a tool to knock-down *Pf*PM IX transcription/expression.

2.2.3.1 Confirmation of transgenic parasites

Transgenic parasite cultures (>5% parasitaemia) were centrifuged at 1500 rpm for 2-3 minutes. Supernatants were removed and discarded and 200 µl of the remaining pellet was transferred to a clean Eppendorf tube. DNA was extracted using a QIAGEN QIAamp DNA Mini Kit (Qiagen, Australia), as per the manufacturers' instructions. The presence of vector was confirmed by PCR amplification prior to long-term selection. Vector-specific primers located in the heat-shock protein 86 5' (UTR; Appendix 3) and the *P. berghei* dihydrofolate reductase 3' UTR (i.e. flanking the inserted gene; Appendix 3) were used in PCR amplifications to confirm presence of C-terminally tagged *PfPM IX*. Vector and *PfPM IX* specific primers (Appendix 3) were used to determine if the pCD-PMIXKO plasmid was present and had integrated into the parasite genome. PCR cycling conditions were generally 35 cycles of 30 sec at 94°C, 30 sec at 55°C and 65°C for 2.5 minutes, with the annealing temperature varying slightly depending on the melting temperature of the respective primers. PCR products were run on 1% agarose in TAE containing ethidium bromide (Appendix 1) and visualised with a UV transilluminator (BioRad).

2.2.3.2 Over-expression

Transgenic parasites containing the cMyc and GFP C-terminally tagged full-length *Pf*PM IX plasmids, pHH1-PMIXcMycB and pHH1-PMIXGFPB respectively, were maintained on 50-200 nM WR99210. Over-transcription was confirmed with qRT-PCR as 2.2.2.1 using the internal reference gene *β-actin1* (PF3D7_1246200) (forward 5'-aaa

gaa gca agc agg aat cca and reverse 5'-tga tgg tgc aag ggt tgt aa-3'; (Augagneur, Wesolowski *et al.*, 2012)) and transcript levels were calculated using the ΔΔ-CT-method where the fold change = $2^{(-\Delta\Delta\text{-CT})}$ (Livak and Schmittgen, 2001).

*Pf*PM IX expression in transgenic cultures on 50 nM WR99210 and wild-type cultures was examined using Western blot (see section 2.2.2.3) with rabbit polyclonal anti-sera to recombinant semi-pro-*Pf*PM IX (1:1000). Culture samples were all of late-trophozoite to schizont stage of intra-erythrocytic parasite development and anti-rabbit GAPDH antibody was used to assess protein loading ((Daubenberger, Tisdale *et al.*, 2003) section 2.2.2.3) so relative ratios of expression could be determined. Images of multiple Western blot exposures were imported into Image Studio Lite software (Licor) and intensity levels of *Pf*PM IX, relative to GAPDH, were compared.

2.2.3.3 Peptide Nucleic Acid Knock-down of Pf*PM IX*

PfPM IX specific (TAA TTC TAT GTT GAG (targeted 1) and ATT CTG TGC CTC TCC (targeted 2); see Appendix 4 for binding sites) and a control/scrambled PNA with no homology to any *P. falciparum* gene (GAT TCT TAG TTA GTA) were purchased from Biosynthesis, Texas and prepared in incomplete culture media (400 µM stock solutions) before addition to highly synchronised, 0 hour ring stage D10 parasites (0.5% parasitaemia and 5% haematocrit) at a final concentration of 10 µM. All sample cultures were aliquoted from the same diluted culture. At 24 hour intervals, a Giemsa stained thin blood smear was prepared and examined to determine parasitaemia and morphology. Cultures were maintained for approximately 80 hours and were re-treated at 48 hours (media replaced). At the end of each experiment, RNA was obtained and analysed by qRT-PCR (see section 2.2.2.1), with both *β-actin1* and *Seryl-tRNA synthetase* primer pairs being used for qRT-PCR analysis. Smears were assessed by 2 experienced microscopists on two separate occasions and data were pooled. At each time point, the mean parasitemia of treatment groups were compared to untreated controls with Dunnett's multiple comparisons test using GraphPad Prism version 6.00. Differences were considered significant if $p \leq 0.05$. Images of parasite morphology were taken on a Leica DM IRB inverted microscope with NIS Elements software.

2.3 Results:

2.3.1 Localization and expression of *Pf*PM IX

*Pf*PM IX is encoded by a 3118 bp, eight exon gene located on chromosome 14, which conceptually translates into a 627 amino acid protein with a theoretical molecular weight of ~74kDa (PlasmoDB; Figure 2.1A and B). *Pf*PM IX contains an aspartic protease signature domain (active residues at amino acid positions 247 and 387; Figure 2.1B) and is syntenic with other putative aspartic proteases in both human and murine malaria species. Interestingly, the gene that encodes *Pf*PM IX is one of the few genes that encode *Pf*PMs to contain introns (PlasmoDB). *PfPMs I-V* and *PfPM X* are single exon genes, while *PfPMs VI-IX* contain between 7 and 13 introns (PlasmoDB).

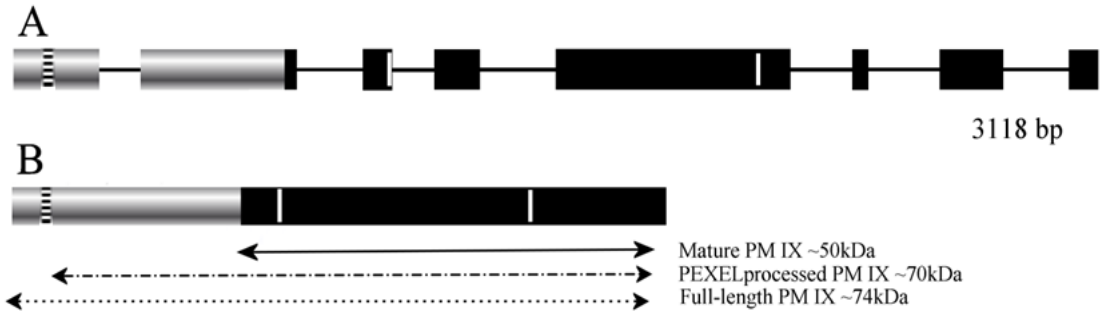


Figure 2.1: **Schematic representation of PF3D7_1430200 and *Pf*PM IX.** A) Cartoon of PF3D7_1430200, the gene encoding *Pf*PM IX and B) cartoon of *Pf*PM IX. The predicted pro-domain of *Pf*PM IX is depicted in grey and the PEXEL motif in checked black and white. Solid white regions indicate locations of catalytic aspartic amino acids. Figure by Tina Skinner-Adams.

qRT-PCR data confirm the global transcription data described in PlasmoDB, demonstrating that transcription of *PfPM IX* occurs at varying degrees throughout the intra-erythrocytic life cycle and is at its maximum in schizont-stage parasites (Figure 2.2). Relative to *18S rRNA* and *Seryl-t transferase*, 68.08% and 77.15% respectively of *PfPM IX* transcription occurs during these parasite stages. This is a marked increase from the combined average 0.64% and 26.75% seen in ring and trophozoite stages (Figure 2.2C).

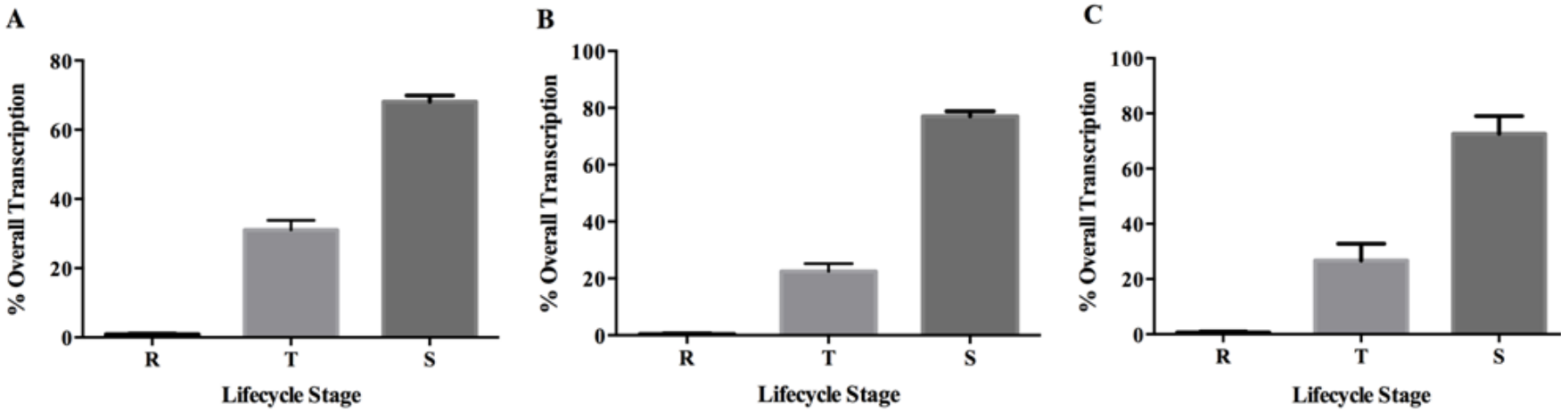


Figure 2.2: **Stage specific transcription of *PfPM IX* as assessed by qRT-PCR analysis of cDNA from synchronised cultures, relative to the internal reference genes.** A) *18s rRNA* and B) *Seryl-t transferase*. C) Combined data. Presented as mean ± SD. N = 5.

Western blot analysis of stage-specific parasite protein samples demonstrated that *Pf*PM IX protein can be detected throughout the asexual intra-erythrocytic stages of parasite development (Figure 2.3). Mature *Pf*PM IX (predicted 50 kDa; Figure 2.1B) does not begin to appear until the trophozoite stage and is primarily found in schizont stage parasites (Figure 2.3). Western blot of transgenic C-terminally tagged *Pf*PM IX, conducted prior to the commencement of this study, failed to detect the transgenic cMyc and GFP tags using specific antibodies (data not shown; (McGeorge, 2009)).

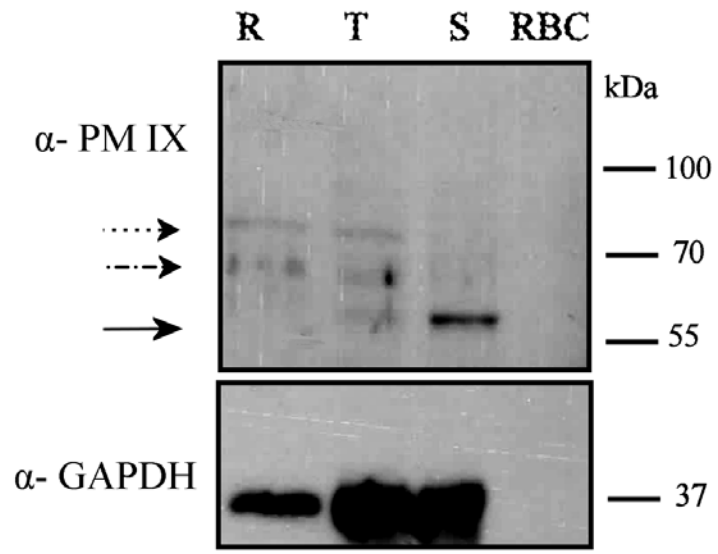


Figure 2.3: ***Pf*PM IX is expressed throughout the intra-erythrocytic cycle.** Western blot of saponin-lysed parasite extracts from synchronized D10 parasites (rings (R); trophozoites (T) and schizonts (S)) and uninfected RBC control preparations probed with *Pf*PM IX anti-sera (upper panel) and antibodies to glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (lower panel).

Conceptually translated *Pf*PM IX contains the PEXEL consensus motif RILSD (amino acid positions 46-50) and IFA data demonstrate that like other PEXEL containing proteins, *Pf*PM IX is exported into the red blood cell cytoplasm (Figure 2.4). While *Pf*PM IX can be found largely in the parasite cytoplasm, in early intra-erythrocytic stages *Pf*PM IX appears to partially co-localise with/next to REX 1 (Figure 2.4B), a known Maurer's Cleft protein (Hawthorne, Trenholme *et al.*, 2004). In late schizont stage parasites, *Pf*PM IX localises within the segmenting merozoites (Figure 2.4A). As with Western blotting, C-terminal tags in the transgenic cultures could not be visualised by IFA (data not shown; (McGeorge, 2009)).

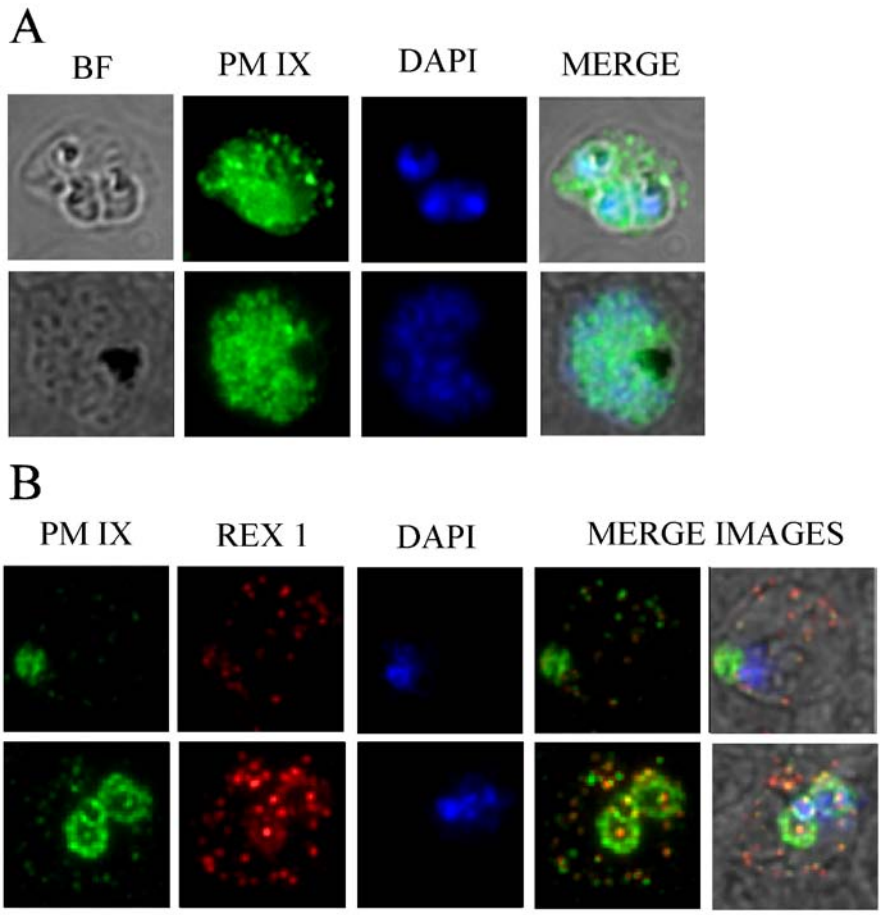


Figure 2.4: ***P. falciparum* PM IX is exported by intra-erythrocytic parasites into the host red cell.** A) The location of *Pf*PM IX in 3D7 parasites examined with poly-clonal antisera generated in rabbits (BF, Bright Field). B) Co-localization of *Pf*PM IX with REX 1. While the majority of *Pf*PM IX localisation is to the parasite cytoplasm, *Pf*PM IX is exported into the red cell where it transiently co-localises with REX 1.

2.3.2 Transfection of parasites with pHH1-PMIXcMycB and pHH1-PMIXGFPB generates transgenic parasites over-expressing *Pf*PM IX

DNA was extracted from transgenic parasite cultures and the presence of the plasmid determined by PCR amplification. Gel analysis of the PCR-generated DNA revealed a band of 1.7kb for pHH1-PMIXcMycB and 2.6kb for pHH1-PMIXGFPB (Figure 2.5), the expected sizes for *PfPM IX* with the C-terminal tag and additional vector sequence (Appendix 2). Transgenic parasite cultures were maintained on 50-200 nM WR99210. Growth rate and intra-erythrocytic development of transgenic parasites were not distinguishable from wild-type parasites.

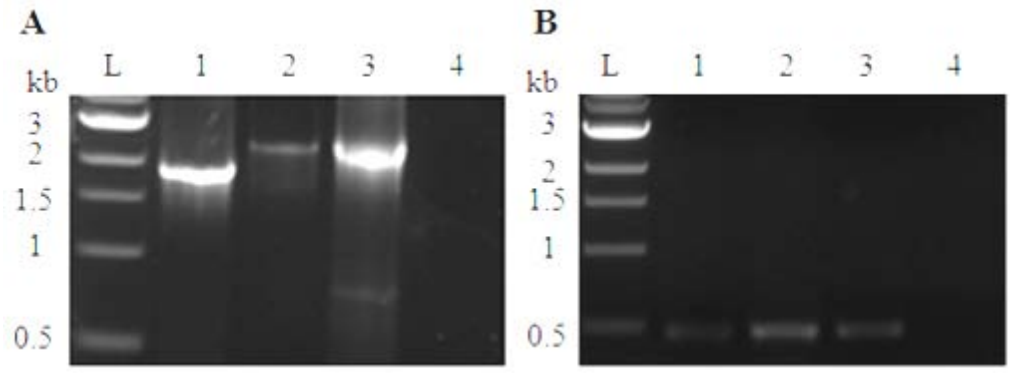


Figure 2.5: **Transfection of D10 parasites with pHH1-PM9cMycB and pHH1-PM9GFPB generates *Pf*PM IX transgenic parasites.** PCR analysis of transgenic parasites demonstrates the presence of vector insert (A) and *hdhfr* (B). Lane 1, DNA isolated from parasites transfected with pHH1-PM9cMycB; Lane 2, DNA isolated from parasites transfected with pHH1-PM9cGFPB; Lane 3, vector control; Lane 4, no template control.

As previously performed assays to assess the sensitivity of pHH1-PM9cMycB transfectants to the HIV protease inhibitors saquinavir and ritonavir had demonstrated that these were significantly less sensitive to these drugs than wild-type parent parasites (IC_{50} 4.0 µM vs 2.8 µM and 5.4 µM vs 2.7 µM for saquinavir and ritonavir respectively; (McGeorge, 2009)) studies to confirm over-expression were pursued in the current project. To confirm increased transcription, qRT-PCR analysis of schizont-stage cDNA samples was performed. Data demonstrated that parasites transfected with pHH1-PMIXcMycB had an increased transcription of *PfPM IX* as compared to wild type controls (Table 2.1). Transgenic parasites over-transcribing *PfPM IX* and exposed to 50 nM WR99210 demonstrated the highest increase in *PfPM IX* transcription (5.5 fold), followed by 100 nM, 150 nM and 200 nM (Table 2.1).

Table 2.1: **Transfection of D10 parasites with pHH1-PM9cMycB generates transgenic parasites that over-transcribe *PfPM IX*.**

Level of WR99210 drug selection	Δ-CT (average of replicates)	ΔΔ-CT	Fold change (Relative to control)
D10 control	1.38 ± 0.55	0	0
50 nM	-1.07 ± 1.28	-2.45	5.5
100 nM	-0.21 ± 0.65	-1.59	3.02
150 nM	0.33 ± 0.29	-1.06	2.07
200 nM	1.51 ± 0.65	0.13	0.91

In preliminary studies, *Pf*PM IX expression levels were examined in transgenic cultures containing plasmids designed to over-express *Pf*PM IX and wild-type controls. Using Image Studio Lite software (Licor), ratios of *Pf*PM IX relative to GAPDH were determined. Transgenic cultures containing pHH1-PM9cMycB were determined to have ~11 times *Pf*PM IX expression compared to D10 control parasites. However, expression levels differed greatly between biological samples.

2.3.3 *Pf*PM IX may be essential to intra-erythrocytic stage *P. falciparum* parasites

P. falciparum clone D10 ring-stage parasites were transfected with pCD-PMIXKO on two separate occasions and selected on 5 nM WR99210 prior to the commencement of this project (McGeorge, 2009). To encourage integration and targeted gene disruption these parasites were retrieved and either cycled on and off 5 nM WR99210 (Fidock and Wellems, 1997) or exposed to 5 nM WR99210 with 1 µM 5'fluoro-uracil (Maier, Braks *et al.*, 2006). However despite >5 attempts of selection integration of the plasmid into the endogenous gene could not be detected by PCR, suggesting that *PfPM IX* may be essential (Figure 2.6).

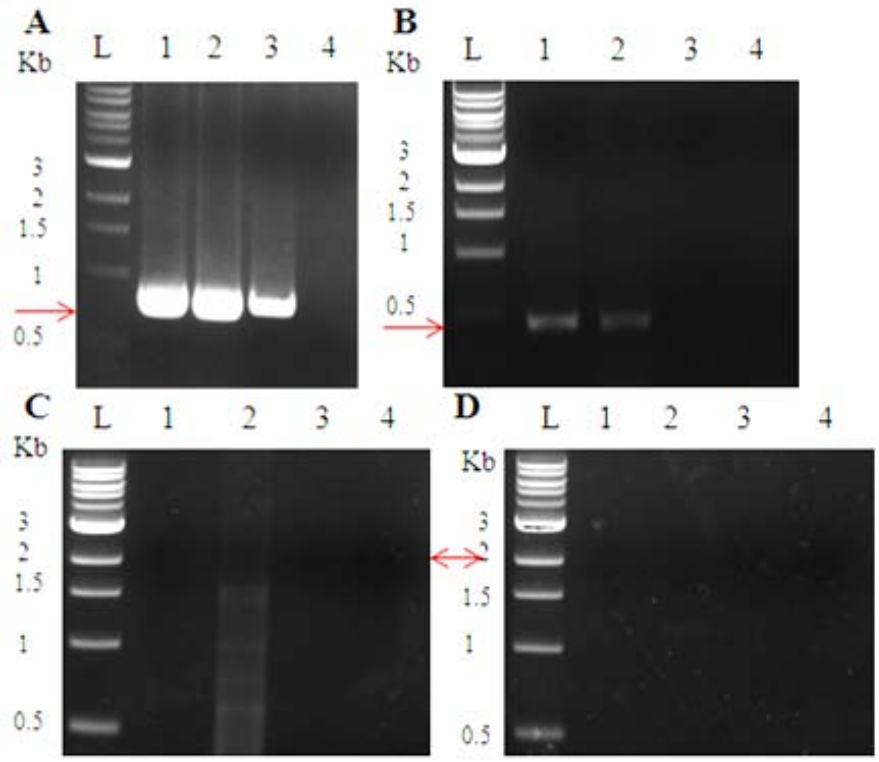


Figure 2.6: **PCR analysis of parasites containing a knock-out vector, pCD-PMIXKO, fails to show genomic integration.** (1) PM IX 'KO' cultures on 4th cycle of WR99210 selection. (2) PM IX 'KO' cultures treated with WR99210 and 5' fluorouracil. (3) D10 control DNA. (4) H_2O negative control. PCR reactions with (A) *Pf*PM V primers to confirm DNA, (B) *hdhfr* primers to confirm vector presence (C) and 5' integration check with *hdhfr* and *PfPM IX* forward primers and (D) 3' integration check with *hdhfr* and *PfPM IX* reverse primers. Arrows indicate expected size.

2.2.4 Peptide Nucleic Acid Knock-down of *PfPM IX*

Parasites treated with *PfPM IX* specific PNAs grew significantly slower than control cultures treated with vehicle and scrambled PNA (Figure 2.7). Parasitaemias at 24 hours suggest starting culture was not diluted to 0.5% parasitaemia as intended. At no stage was there a significant difference in parasitaemia between untreated cultures and PNA controls. Treated cultures also exhibited an altered morphology (Figure 2.8) with notably enlarged DVs. Despite altered morphology and reduced growth, further examination of parasites by qRT-PCR demonstrated that *PfPM IX* transcription was reduced in all PNA treatments including PNA control treatments (Tables 2.2 and 2.3).

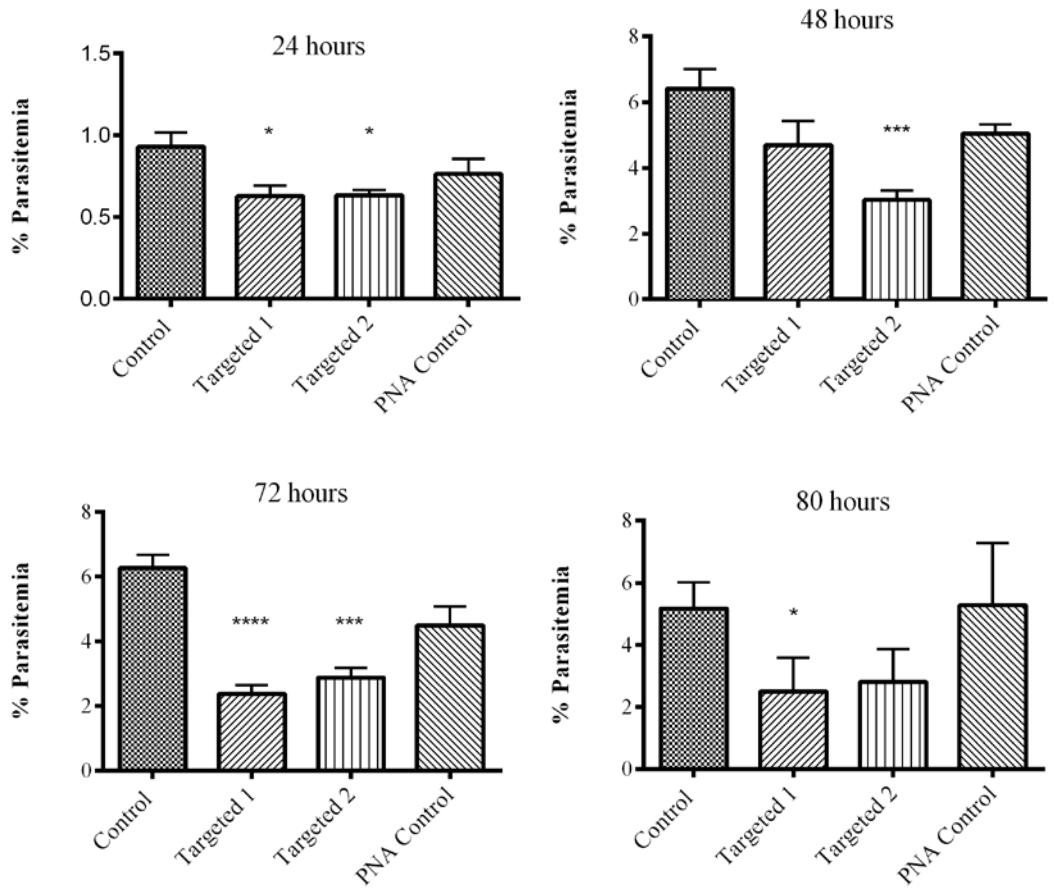


Figure 2.7: **Treatment of *P. falciparum* with *Pf*PM IX targeted PNAs affects *in vitro* parasite growth.** Growth of *P. falciparum* cultures treated with vehicle (control) and scrambled PNA (PNA control) were significantly different to cultures treated with *PfPM IX* specific PNAs (targeted 1 and 2) over 80 hours of exposure. Data are presented as mean parasitaemia ± SD. Significant differences from controls as determined by Dunnett's multiple comparisons test are * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$. N = 3.

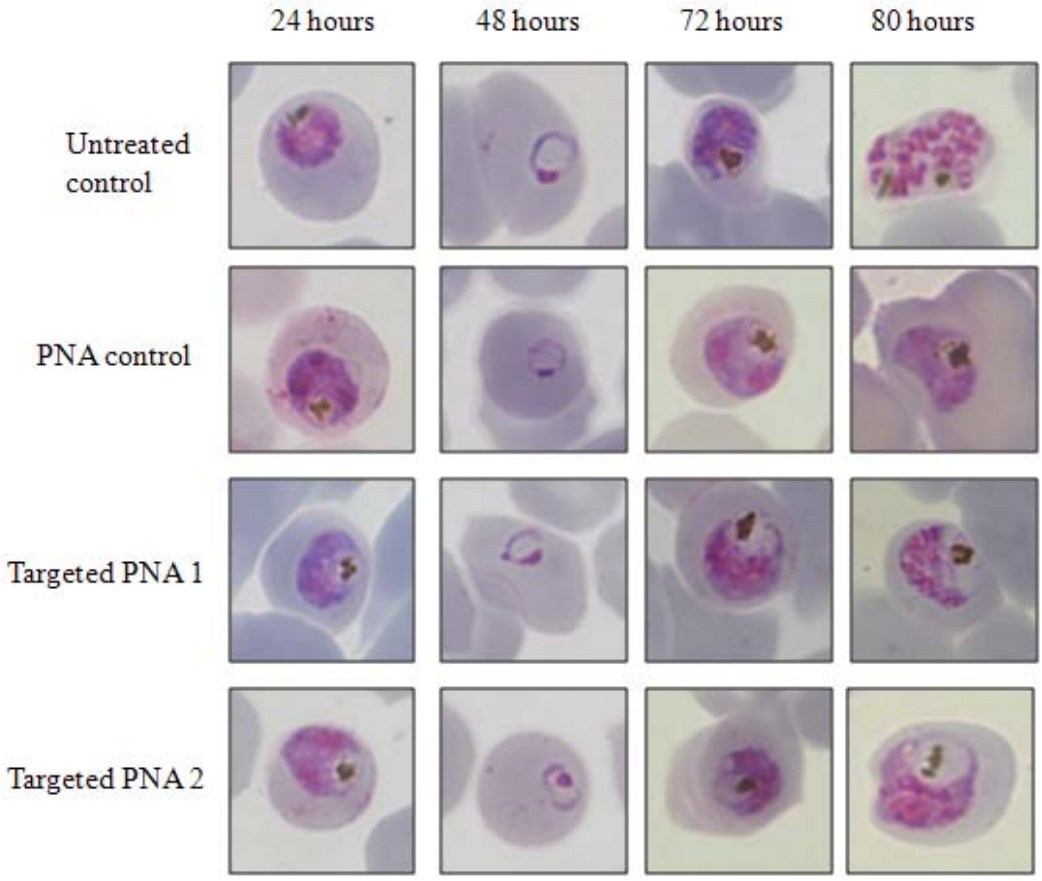


Figure 2.8: ***Pf*PM IX targeted PNA treatment affects parasite morphology.** Giemsa stained blood smears were prepared at time points of 24, 48, 72 and 80 hours (D10 control; PNA targeted treatments (1-2); PNA control and parasite morphology examined.

Table 2.2: **PNA treated cultures demonstrate reduced *PfPM IX* transcription.** qRT-PCR analysis of *PfPM IX* transcription, relative to *β-Actin1*, suggests *PfPM IX* transcription is reduced in all treated samples including the PNA control.

Treatment group	Δ-CT (average of replicates)	ΔΔ-CT	Fold change (relative to Vehicle Control)
Vehicle Control	5.73 ± 2.56	0	0
PNA targeted 1	7.36 ± 2.31	1.62	0.33
PNA targeted 2	7.49 ± 2.36	1.75	0.29
PNA Control	7.37 ± 1.54	1.64	0.32

Table 2.3: **qRT-PCR analysis demonstrates a reduction in *PfPM IX* transcription in PNA treated cultures.** *PfPM IX* transcription, relative to *Seryl-tRNA synthetase*, is reduced in all treated samples.

Treatment group	Δ-CT (average of replicates)	ΔΔ-CT	Fold change (relative to Vehicle Control)
Vehicle Control	0.85 ± 1.16	0	0
PNA targeted 1	0.99 ± 1.56	0.14	0.91
PNA targeted 2	2.36 ± 0.35	1.52	0.35
PNA control	0.95 ± 0.10	0.11	0.92

2.4 Discussion and Conclusion

Drug resistance to all currently available antimalarial agents including artemisinin and its derivatives has been reported (Noedl, Socheat *et al.*, 2009, Dondorp, Nosten *et al.*, 2009, Noedl, Se *et al.*, 2008). Thus there is a need to identify new targets within malaria parasites against which novel antimalarial agents can be developed. Currently available data demonstrate that HIV-PIs are effective against multiple stages of Plasmodium development at clinically relevant concentrations (Redmond, Skinner-Adams *et al.*, 2007, Andrews, Fairlie *et al.*, 2006, Lek-Uthai, Suwanarusk *et al.*, 2008, Peatey, Andrews *et al.*, 2010, Hobbs, Tanaka *et al.*, 2013). While the anti-plasmodial target of these compounds is unknown, it could point to a previously unexploited target for anti-malarial development.

Current data suggest that *Pf*PM IX represents a potential new drug target in malaria parasites, however, very little is known about this protein. This chapter demonstrated that while *Pf*PM IX is present in all intra-erythrocytic asexual stages of *P. falciparum* development, it is primarily transcribed and translated during the late schizont stage of parasite development *in vitro* (Figure 2.3). IFA data indicate that *Pf*PM IX is present in both the parasite cytoplasm and exported into the host red cell (Figure 2.4A), where it appears to partially co-localise with/next to the Maurer's Clefts in early intra-erythrocytic stages (Figure 2.4B). Further co-localisations with the Maurer's Cleft tether protein, MAHRP2 (Pachlatko, Rusch *et al.*, 2010), were attempted, however due to technical issues were not successful. Co-localisations with MAHRP2 may provide further insights into the exact localisation of this protein. Additional co-localisations with HSP 70-x, a protein that localises to J-Dots and which also have a punctate pattern

of localisation (Kulzer, Charnaud *et al.*, 2012, Kulzer, Rug *et al.*, 2010), may also be beneficial. During later stages of the asexual cycle, *Pf*PM IX localises within segmenting merozoites but does not at any stage localise to the DV.

While the function of *Pf*PM IX remains unknown, its transcription and translation within schizont stage parasites in conjunction with its location in segmenting merozoites suggests a role in egress and/or invasion. The location of *Pf*PM IX in the host red cell also supports this hypothesis. Investigation of any interacting proteins through co-immuno-precipitation experiments may provide more insights into *Pf*PM IX's role. Identification of proteins that co-precipitate with *Pf*PM IX and that are known to have a role in egress/invasion would support this hypothesis. While roles other than egress and/or invasion are possible, current data, including an absence of *Pf*PM IX in the DV, suggests that this PM has one or more functions distinct from the already characterised *Pf*PMs.

Immunochemistry data indicate that *Pf*PM IX appears to undergo a number of processing events throughout the asexual life cycle (Figure 2.3). These data suggest that similar to the *Pf*DV PMs (Banerjee, Liu *et al.*, 2002, Dame, Yowell *et al.*, 2003, Kim, Lee *et al.*, 2006), *Pf*PM IX is expressed as a zymogen requiring processing for activity. The loss of any pro-region is likely to be in addition to PEXEL motif cleavage. *Pf*PM IX also appears to undergo processing at its C-terminus with data suggesting the removal of C-terminal tags shortly after protein translation (McGeorge, 2009). The loss of these tags may well explain the large quantity of parasites required by others to isolate *Pf*PM IX-GFP from transgenic parasites (Boddey, Hodder *et al.*, 2010). N and C-terminal processing also limits the techniques available to study this enzyme, with tags being cleaved before they can be used for localisation and expression assays. Until more is known about the amino acid positions of processing, the introduction of fluorescent tags will remain problematic. Interestingly *Toxoplasma gondii* aspartic protease 1 (TgASP1) is processed N and C-terminally (Shea, Jakle *et al.*, 2007) and this parasite is also sensitive to HIV-PIs (Degliesposti, Kasam *et al.*, 2009). Shea et al. (2007) used antibodies specific to different regions of TgASP1 to examine this enzyme's processing (Shea, Jakle *et al.*, 2007). This method would be useful for the investigation of *Pf*PM IX processing and should be conducted in future studies. Additionally, the possibility that generated antisera is cross-reacting with other *Pf*PMs and producing artefacts in Western blot experiments should be investigated. Banerjee et al. (2002) approached this

by testing antisera against purified recombinant forms of *Pf*PMs (Banerjee, Liu *et al.*, 2002). This would confirm *Pf*PM IX antisera sensitivity and the number/size of forms this enzyme is processed to. Unfortunately, this approach is limited by the number of recombinant *Pf*PMs that are available. Another method would be to perform proteomic analysis on material that had been immuno-precipitated from culture using the generated antisera. This would confirm what the antisera was recognising within the parasite and confirm antisera specificity. This method was attempted throughout the course of this project however, due to technical difficulties, could not be completed. To further support the importance of *Pf*PM IX to parasite survival, it has been demonstrated that transgenic parasites over-transcribing *PfPM IX* (Table 2.1) are protected against HIV-PI inhibition (McGeorge, 2009). While additional targets in Plasmodium may exist, these data strongly suggest that *Pf*PM IX inhibition is at least partially responsible for the activity of the HIV-PIs and may be a good target for the development of new anti-malarial chemotherapeutics. Interestingly, these transgenic parasites over-transcribe *PfPM IX* at the highest level at 50nM WR99210 (Table 2.1). The reasons for this remain unclear. It is possible that *Pf*PM IX exhibits some toxicity at higher concentrations and parasites degrade transcripts. Examination of expression levels using densitometry, while supportive of over-expression, was problematic with some variability seen between samples. Similar problems were encountered by Gassmann et al. (2009), who investigated a number of densitometry methods to examine erythropoietin levels in identical samples and found that results differed depending on the method of densitometry used (Gassmann, Grenacher *et al.*, 2009). While further analysis of expression levels in transgenic cultures is warranted, and Western blot may give an approximate idea of expression levels, a more accurate method to quantitate expression is needed. As a broad aspartic protease substrate is used to assess *Pf*PM IX activity, there is currently no way to examine over-expression in parasites that express multiple *Pf*PMs and accurately quantitate levels of increased *Pf*PM IX without more knowledge of the enzyme's specific substrate.

Despite multiple attempts, experiments to genetically disrupt *PfPM IX* were not successful (Figure 2.6). The inability to knock-out *PfPM IX* suggests that it may have vital roles within the parasite and may be essential for parasite survival. While it is impractical to confirm gene essentiality through knock-out techniques, additional transfection of transgenic parasites with complementation vectors (Freville, Landrieu *et al.*, 2012, Triglia, Healer *et al.*, 2000, Soldati, Kim *et al.*, 1995) would be beneficial and

may allow the knock-out plasmid to integrate into the parasite genome. Additional experiments to produce conditional knock-outs would also aid investigations of *PfPM IX* essentiality. During the course of this study, another group used the Shld system (Armstrong and Goldberg, 2007) to produce conditional *Pf*PM IX knock-out cultures. These studies supported the current work, suggesting that *Pf*PM IX is essential for parasite maturation and survival (personal communication Soldati-Favre, 2013).

Experiments using targeted PNAs demonstrated a significant difference in parasitaemia of treated cultures compared to controls (Figure 2.7). PNA treated parasites also demonstrated an altered DV morphology (Figure 2.8; (Magowan, Brown *et al.*, 1997, Glushakova, Mazar *et al.*, 2009)), initially suggesting that they may have knocked-down *PfPM IX* transcription. As Western blotting is not quantitative and there is no enzyme specific assay to quantitate expression of *Pf*PM IX, further analysis was performed using qRT-PCR. This revealed that while *PfPM IX* transcription was reduced in targeted PNA cultures, it was also reduced in PNA control cultures (Tables 2.2 and 2.3). These data suggest that non-specific PNA treatment knocks-down *PfPM IX* expression. However, closer inspection of parasite morphology and parasitaemia (Figure 2.8) suggests that this effect was likely due to the overgrowth and death of control parasites. This resulted in trophozoites rather than schizonts being collected for qRT-PCR analysis. Given *PfPM IX's* marked increase in transcription in schizont stages (Figure 2.2), less mature cultures would have naturally had reduced *PfPM IX* transcription. Interestingly one of the *PfPM IX* targeted treatments decreased *PfPM IX* transcription further than the control PNA treatment (Table 2.3). While still complicated by non-specific PNA effects, this suggests that a *PfPM IX* specific effect may be responsible for the reduced growth seen during these experiments. Examination of *β-actin1* and *Seryl-tRNA synthetase* expression throughout the asexual lifecycle, demonstrates that *β-actin1* is not expressed evenly throughout, with a steady increase in expression to reach a maximum in schizonts (www.plasmoDB.org). While appropriate for comparisons when all cultures are at the same life cycle stage, the difference in maturity between untreated PNA controls and treated samples may have been enough to affect the results. Additionally, while *Seryl-tRNA synthetase* is generally expressed evenly throughout, there is a small dip in expression during late schizonts (www.plasmoDB.org). This may also explain the disparity between relative *PfPM IX* transcription levels in Tables 2.2 and 2.3. Further qRT-PCR analysis with a schizont specific gene, such as subtilisin-like protease 2 (PF3D7_1136900; www.plasmoDB.org), may mitigate these issues and

provide a better insight into the processes occurring during these experiments. Further studies to examine these effects were impacted by the expense of PNA synthesis.

During the course of this study, Kolevzon et al (2014) published work validating PNAs as tools for knocking down gene expression in *P. falciparum* (Kolevzon, Nasereddin *et al.*, 2014). Kolevzon et al. (2014) found that PNAs could significantly knock-down gene expression and suggested that PNAs may cause the down regulation of genes post-transcription (Kolevzon, Nasereddin *et al.*, 2014). Knock-down of gene expression was shown against both a stably expressed transgene containing luciferase and against an essential gene, *Pf*Sec13 (PF3D7_1230700). The study also confirmed our findings that non-specific PNAs do not have a significant effect on parasite parasitaemia. The use of the luciferase transgene allowed for luciferase activity assays which could demonstrate knock-down effects at a protein level (Kolevzon, Nasereddin *et al.*, 2014). This was not possible in the current study given *Pf*PM IX was both N and C-terminally processed, however, further investigation into the use of these molecules to knock-down *PfPM IX* is warranted.

A similar method for knocking-down genes in *P. falciparum* that has been published recently is peptide-morpholino oligomer (PMOs) (Augagneur, Wesolowski *et al.*, 2012). Using a sequence specific to *P. falciparum* gyrase mRNA, Augagneur, et al. (2012) demonstrated that PMOs can enter the parasite and knock-down gyrase mRNA transcription by 60% (Augagneur, Wesolowski *et al.*, 2012). While this study neglected to determine if a non-specific PMO affects transcription with qRT-PCR, it did determine that a PMO specific to an *E.coli* ampicillin resistance gene had no affect on parasite growth (Augagneur, Wesolowski *et al.*, 2012). Given the study's success with gyrase mRNA and the difficulties encountered in this study, PMOs may be a suitable alternative to knock-down *PfPM IX* to confirm it's essentiality to parasite survival. An additional knock-down technique is the use of antisense RNA. Briefly, vectors containing a gene insert of interest in the antisense orientation and a selectable marker can be transfected into *P. falciparum* parasites where they interfere with the normal transcription process (Gardiner, Holt *et al.*, 2000). Importantly, when drug selection is removed, transfected parasites lose vector and gene function returns to wild type levels. This has been used successfully in *P. falciparum* with studies demonstrating the inhibition of *clag9*, a gene involved in cytoadherence (Gardiner, Holt *et al.*, 2000), and the M18 aspartyl aminopeptidase (Teuscher, Lowther *et al.*, 2007) with the subsequent

reversal to wild type once drug selection was removed (Gardiner, Holt *et al.*, 2000). Unfortunately due to time constraints, experiments involving this technique could not be completed during this study. However future investigations of *Pf*PM IX using antisense RNA would provide further insights into the importance of this enzyme to parasite survival.

This chapter has demonstrated that *Pf*PM IX is a unique enzyme with potentially unique functions in the later intra-erythrocytic stages of *P. falciparum*. The inability to knock-out *PfPM IX*, combined with results indicating separate functions to the already well-characterised *Pf*DV PMs suggest that it may be an essential enzyme and that it warrants further investigation as a drug target. With the ever increasing spread of drug resistant parasites, it is imperative that new drug targets be identified across a broad range of parasite stages and new chemotherapeutics that specifically target *Pf*PM IX may prove to be potent and effective antimalarials.

Chapter 3: Expression and Purification of *P. falciparum* Plasmepsin IX

3.1 Introduction:

Aspartic proteases are an important class of enzymes naturally occurring across a variety of species which include plants, parasites, bacteria, mammals and fungi. A large number of these proteins are important factors in disease and are therefore of keen interest to researchers. Examples of these include the human immunodeficiency virus (HIV) protease, an essential HIV enzyme (Cooper, 2002), pepsin, a protease of central importance to peptic ulcer disease (Cooper, 2002) and the plasmepsins of malaria parasites (Coombs, Goldberg *et al.*, 2001). While aspartic proteases have been purified from selected hosts (Karlsen, Hough *et al.*, 1998, Gildberg, Olsen *et al.*, 1990, Boddey, Hodder *et al.*, 2010, Russo, Babbitt *et al.*, 2010) it is often difficult to purify sufficient quantities of these enzymes for kinetic and structural characterisation (Boddey, Hodder *et al.*, 2010, Russo, Babbitt *et al.*, 2010). With this in mind, improved methods to express and purify recombinant proteins are being developed (Kim, Babnigg *et al.*, 2011, Correa and Oppezzo, 2011, as reviewed by Graslund, Nordlund *et al.*, 2008).

Successfully expressing and purifying Plasmodium proteins in *E. coli* can be particularly challenging. A major limitation is that these proteins are frequently insoluble and usually expressed within inclusion bodies (Mehlin, Boni *et al.*, 2006, Flick, Ahuja *et al.*, 2004). It is believed that this is the result of metabolic stress caused by the difference in amino acid usage between Plasmodium and *E. coli* as well as the toxicity of some of these proteins to the bacterial host (Vedadi, Lew *et al.*, 2007). While various conditions such as temperature, enriched media (Graslund, Nordlund *et al.*, 2008, Peti and Page, 2007) and the addition of fusion tags can aid solubility (reviewed in Esposito and Chatterjee, 2006), extensive optimisation is required and is not always successful. Particular characteristics that commonly impact the expression of Plasmodium proteins include a basic pI (>6), the presence of Low Complexity Regions (LCRs; discussed in Chapter 4), signal peptides, export motifs and molecular weights of >56 kDa (Mehlin, Boni *et al.*, 2006, Vedadi, Lew *et al.*, 2007). Two examples of methods being used to combat these issues include codon-optimisation and harmonisation. Codon-optimisation involves altering the codons of the target expression protein to suit the codon preference of *E. coli* (Zhou, Schnake *et al.*, 2004). Whereas codon harmonisation involves altering the genetic sequence such that rare codon motifs are not over-used in the construct (Angov, Hillier *et al.*, 2008). This approach allows the translational machinery to process the protein as they would in the native host, by

pausing in the same positions for example, and the protein's natural refolding processes can be mimicked (Angov, Hillier *et al.*, 2008, Thanaraj and Argos, 1996).

Currently only the digestive vacuole (DV) plasmepsins (PMs) I-IV have been successfully expressed, characterised and their structures elucidated (Bernstein, Cherney *et al.*, 1999, Clemente, Govindasamy *et al.*, 2006, Moon, Tyas *et al.*, 1997, Prade, Jones *et al.*, 2005, Asojo, Afonina *et al.*, 2002, Asojo, Gulnik *et al.*, 2003, Bhaumik, Horimoto *et al.*, 2011, Li, Yowell *et al.*, 2004, Liu, Marzahn *et al.*, 2009, Wyatt and Berry, 2002, Xiao, Tanaka *et al.*, 2007, Xiao, Briere *et al.*, 2010, Hill, Tyas *et al.*, 1994). These have all been expressed as recombinant proteins in *E. coli* in truncated form, typically with N-terminal pro-segments removed. The aim of this work was to express and purify recombinant *Pf*PM IX (r*Pf*PM IX). These studies were performed in conjunction with our collaborators in the Department of Biochemistry and Molecular Biology at the University of Florida. While preliminary studies describing the expression and purification of r*Pf*PM IX have recently been published (Olajuyigbe, 2013) this work was guided by the research described below and were published after the current research was completed without the support of our research group.

3.2 Materials and Methods:

3.2.1 Expression Vectors

3.2.1.1 Semi-pro Pf*PM IX (r*Pf*PM IX):*

Based on previous work with *Pf*DV PMs (Dell'Agli, Parapini *et al.*, 2006, Gutierrez-de-Teran, Nervall *et al.*, 2006), a pET3a expression vector (Novagen, USA) containing a sequence encoding *Pf*PM IX (amino acids 165-627) was constructed. This vector was constructed by collaborators in the United States before the commencement of my project and included a mutation near the C-terminal end of *Pf*PM IX to disrupt an *E.coli* Shine-Delgarno sequence and secondary start codon (Appendix 4).

3.2.1.2 Truncated Pf*PM IX:*

My research suggested that native *Pf*PM IX undergoes post translation truncation at both the N and C termini (Chapter 2). Given these data and the low yields of active protein produced using the semi-pro expression plasmid (above) additional plasmids were constructed. The first of these vectors contained an N and C terminally truncated sequence of *Pf*PM IX (NCtrunc-r*Pf*PM IX). The second contained an N terminally truncated sequence of *Pf*PM IX (Ntrunc-r*Pf*PM IX). These truncated *Pf*PM IX

sequences were codon optimised for expression in *E.coli* (GeneART, Germany). The N terminally truncated *Pf*PM IX sequence for Ntrunc-r*Pf*PM IX was produced by PCR amplification of the codon optimised *Pf*PM IX sequence with the primers PM9Ntrunc and PM9FL (Appendix 3). NCtrunc-r*Pf*PM IX was produced by PCR amplification with primers PM9Ntrunc and PM9Ctrunc (Appendix 3). Each primer contained a BamH1 restriction site which allowed for each insert to be ligated into a pre-digested pET3a expression vector. Each construct was sequenced and transformed into Rosetta 2 (DE3) pLysS following manufacturer's instructions (Novagen, USA). Due to difficulty encountered in constructing the 'Ntrunc-r*Pf*PM IX' vector and time constraints, only 'NCtrunc-r*Pf*PM IX' was used in expression trials.

3.2.2 Expression and purification of semi-pro *Pf*PM IX

3.2.2.1 Expression of insoluble semi-pro Pf*PM IX*

Rosetta 2 (DE3) pLysS containing expression plasmids were grown overnight at 37°C in Luria Broth (LB; Appendix 1) containing ampicillin (50 µg/mL) and chloramphenicol (34 µg/mL). The following day, 2YT medium (Appendix 1) containing ampicillin (50 µg/mL) and chloramphenicol (34 µg/mL) was inoculated with the overnight culture, (10% v/v inoculation) and incubated at 37°C with shaking (250 rpm) until the optical density at 600 nm (OD_{600}) reached 0.8 absorbance units (2YT was used as the blank). A 1 mL aliquot of culture was transferred to an Eppendorf tube, centrifuged and frozen post supernatant removal for SDS-PAGE analysis (T0).

Expression of r*Pf*PM IX was induced with 1 mM IPTG at 37°C for 3 hours with a shaking speed of 250 rpm. Samples for later SDS-PAGE analysis were taken every hour as described above. After 3 hours, the expression culture was transferred to pre-weighed 0.5 L centrifuge bottles and cells harvested by centrifugation at 7,700 x g in a Beckman JA-20 rotor for 10 mins at 4°C. The supernatant was then removed and the cell pellet weight determined. Cell pellets were frozen at -20°C for later purification.

3.2.2.2 Purification of insoluble semi-pro Pf*PM IX*

3.2.2.2.1 Isolation of inclusion bodies

Cell pellets were resuspended in pre-chilled resuspension buffer A (Appendix 1) at 4.2 mL/g. From this point on resuspended cells were kept on ice to prevent proteolysis. A French press (1000 psi; SLM-Aminco) was used to lyse the cells according to the manufacturer's instructions. Cell lysate was layered over a 10 mL 27% sucrose cushion,

up to 18 mL per cushion, in 30 mL Corex tubes. These were then centrifuged at 8,500 x g for 30 mins at 4°C in a swinging-bucket rotor, isolating the inclusion bodies which form the pellet fraction. 100 µL of the supernatant fraction (supernatant 1) was taken and stored for later SDS-PAGE analysis. The remaining supernatant was removed. Pellets were then resuspended in 5 mL of pre-chilled resuspension buffer B (Appendix 1) and once again gently layered over a 10 mL 27% sucrose cushion before being centrifuged as above. 100 µL of the supernatant fraction was taken (supernatant 2) and stored as above and the remaining supernatant removed. Pellets were then resuspended in 15 mLs of pre-chilled resuspension buffer C (Appendix 1) and centrifuged at 8,500 x g for 15 mins at 4°C. 100 µL sample of supernatant (supernatant 3) was taken and the remaining supernatant removed. All pellets were then resuspended in 40 mL of pre-chilled resuspension buffer D (Appendix 1) and centrifuged in pre-weighed 50 mL plastic centrifuge tubes at 8,500 x g for 15 mins at 4°C. A 100 µL sample of the supernatant (supernatant 4) was taken and the rest removed. The resulting inclusion bodies were resuspended in TE buffer pH 8.0 (Appendix 1), at a final concentration of 100 mg/mL and frozen (-20°C) in 1 mL aliquots until required for further purification.

3.2.2.2.2 Solubilisation and refolding

Frozen aliquots of inclusion bodies were thawed on ice before adding drop-wise to an 8 M urea denaturing solution (Appendix 1) to a final protein concentration of 1 mg/mL. The inclusion bodies were allowed to denature with gentle stirring at room temperature for 2 hours after which they were placed in pre-prepared Spectra-Pore MWCO 12000-14000 dialysis tubing (Spectrum, USA) and dialysed against 4 L of refolding buffer, 50 mM sodium phosphate dibasic pH 7.0 (Appendix 1), for 2 hours at room temperature and then 2 hours at 4°C. The dialysis tubing was then placed into 4 L of fresh refolding buffer and dialysed at 4°C overnight. This was followed by 2 further refolding buffer changes and 8 hour dialysis steps. Samples were taken at buffer change to test for activity.

To further investigate buffer conditions on refolding, dialysis experiments were performed with 50 mM sodium phosphate dibasic pH 8.0 buffer and 50 mM bicine pH 8.0 buffer. Post-dialysates were kept at 4°C until further analysis and purification was performed. The effect of concentrating on protein precipitation was also investigated by concentrating dialysate using Centrifugal Concentrators as per manufacturer instructions (Millipore, USA).

An FPLC LCC system (Amersham Pharmacia, USA) was set up according to the manufacturer's instructions. A HiTrap Q Sepharose HP 5 ml cation-exchange column (Amersham Pharmacia, USA) was equilibrated with 5 column volumes of IEX starting buffer (20 mM Tris-Cl, pH 8.0, filtered) at a flow rate of up to 5 mL/min. The post-dialysate was loaded onto the column at a flow rate of 2 mL/min and then eluted with a linear gradient of 0 to 1 M NaCl using a computer-controlled gradient mixer with starting buffer and elution buffer (20 mM Tris-Cl/1M NaCl) over 60 mins. The flow rate was maintained at 2 mL/min and 2 mL elution fractions were collected and kept at 4°C. The absorbance at 280 nm (A_{280}) of each fraction was measured using a 1 cm path-length quartz cuvette. 10-20 µl samples of post-ion exchange fractions were taken for SDS-PAGE analysis. Fractions containing r*Pf*PM IX were pooled and further purified by gel filtration with a Superdex 75 prep-grade gel-filtration resin (Amersham Pharmacia, USA), equilibrated in IEX starting buffer. The proteins were separated using a 1 mL/min flow rate. 1 mL fractions were collected and the A_{280} measured for each with samples taken of fractions of interest for SDS-PAGE analysis and chromogenic activity assays.

3.2.3 Chromogenic activity assays and substrate cleavage

Enzymatic activity of dialysate and purified fractions was assessed using a chromogenic assay and the substrate RS6. The substrate sequence was Lys-Pro-Ile-Glu-Phe*Nph-Arg-Leu (where Nph is a *p*-nitophenylalanine and * represents the cleavage site). This substrate has been shown to be useful for other recombinant PMs (Dell'Agli, Parapini *et al.*, 2006, Gutierrez-de-Teran, Nervall *et al.*, 2006). The rate of cleavage of the substrate can be measured as a decrease in average absorbance over 284 to 324nm. The activity in each fraction or sample of purified r*Pf*PM IX was tested at a range of dilutions using a Cary 50 Bio UV-Visible spectrophotometer (Varian, Mulgrave). Samples were diluted with assay buffer (50 mM sodium acetate, pH 4.5) and water, pre-incubated at 37°C for 30 mins to allow for auto-maturation before the addition of substrate (final concentration of 100 µM). Absorbance was measured immediately on the spectrophotometer and then continuously for a minimum of 200 seconds. Samples which showed activity with a decrease in absorbance over time had their final absorbance measured in a range of 340 to 240 nm. A positive r*Pf*PM II or pepsin control (produced by our collaborators at the University of Florida, prior to the commencement of this study) in addition to a negative control containing substrate only was included in each experiment.

3.2.4 Crystallisation Trials with r*Pf*PM IX

Crystallisation trials were set up using 'Crystal Screen™ – HR2-110', 'Crystal Screen Cryo™ – HR2-122' and 'Crystal Screen 2™ – HR2-112' kits (Hampton Research Corp., CA) as per the manufacturer's instructions. Crystals were grown with the hanging drop vapour diffusion method at room temperature using 2 µL of purified r*Pf*PM IX, 2 µL of precipitant and a 500 µL reservoir volume (Appendix 5).

3.2.5 Investigation of optimum pH for r*Pf*PM IX activation

To investigate the optimum pH for auto-activation of r*Pf*PM IX, 100 µl samples of dialysate in the 50 mM sodium phosphate dibasic pH 7.0 were placed into 0.5 M sodium acetate (pH range of 2-6). These samples were incubated for 30 minutes at 37 °C before SDS-PAGE analysis.

3.2.6 Expression and purification trial of insoluble truncated r*Pf*PM IX (NCtrunc-r*Pf*PM IX)

To evaluate the expression of the NCtrunc-r*Pf*PM IX construct, 1 L expression and purification trials were conducted in 3 expression media (LB, 2YT and Super-Broth (Appendix 1), containing ampicillin (50 µg/mL) and chloramphenicol (34 µg/mL)). Expression of r*Pf*PM IX and NCtrunc-r*Pf*PM IX were also compared in 2YT. Expression trials were carried out as described for r*Pf*PM IX. Inclusion bodies were stored at -20 °C prior to SDS-PAGE analysis.

3.3 Results:

3.3.1 Expression, refolding and purification of semi-pro *Pf*PM IX

A 57 kDa recombinant semi-pro *Pf*PM IX, consisting of amino acids 167-627 was expressed in *E. coli.* On average, 1413 mg of purified inclusion body material was produced per 1 L of expression culture, approximately 30% of the cell mass (Table 3.1). Samples from each expression and purification step were analysed by SDS-PAGE (Figure 3.1).

Table 3.1: **Expression of semi-pro *Pf*PM IX.** Cell pellet and inclusion body yields from 1 L of *E. coli* culture. Dialysate, cation exchange chromatography and gel filtration yields are from 100 mg of inclusion bodies.

Expression/Purification Step	Average yield (mg)
Cell pellet (wet)	4750
Inclusion bodies (wet)	1413
Dialysate	11.6
Cation exchange chromatography	4.62
Gel filtration	1.39

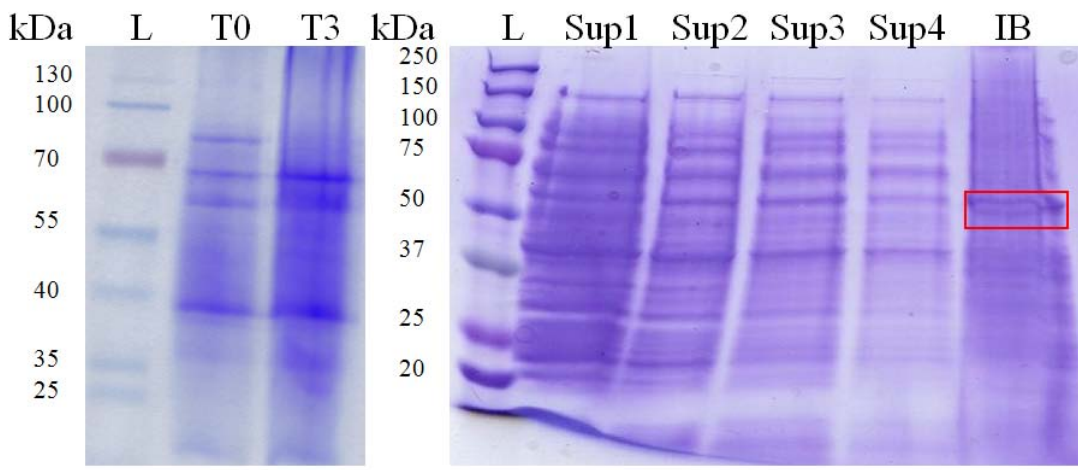


Figure 3.1: **Expression and purification of semi-pro *Pf*PM IX.** Samples were taken at each stage of expression and purification and examined by SDS-PAGE. (TO), culture prior to induction with IPTG. (T3), cultures after 3 hours of expression. (Sup1-4), stages of purification of inclusion bodies (IB). Semi-pro *Pf*PM IX runs at approximately 57 kDa and is highlighted with a red box.

These data demonstrated that concentrating dialysed material resulted in semi-pro *Pf*PM IX precipitation. This was seen both visually and by SDS-PAGE analysis (Figure 3.2) and resulted in all further steps involving un-concentrated material. Following dialysis, refolded material was further purified by cation exchange chromatography before being subjected to gel filtration chromatography to isolate correctly folded semi-pro *Pf*PM IX (fraction 28; 1.38 mg) (Figure 3.3). Approximately 0.46 mg of incorrectly folded protein was present in the 'void volume'. Crystallisation trials using purified semi-pro

*Pf*PM IX were unsuccessful and only produced salt crystals (data not shown). Due to the limited amount of purified semi-pro *Pf*PM IX obtained further crystallisation trials could not be undertaken.

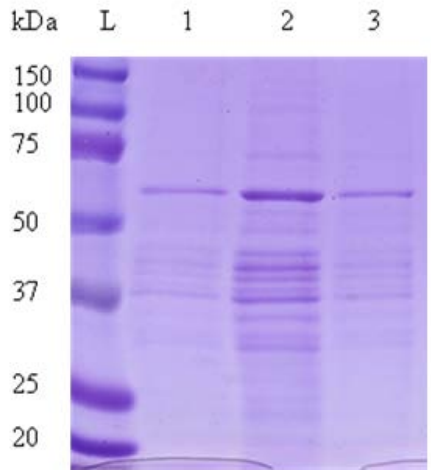


Figure 3.2: **Concentration of dialysis material leads to loss of semi-pro *Pf*PM IX.** Dialysed semi-pro *Pf*PM IX (1) was concentrated to an A_{280} of 0.558 (2) and then 2.167 (3) and subjected to SDS-PAGE analysis.

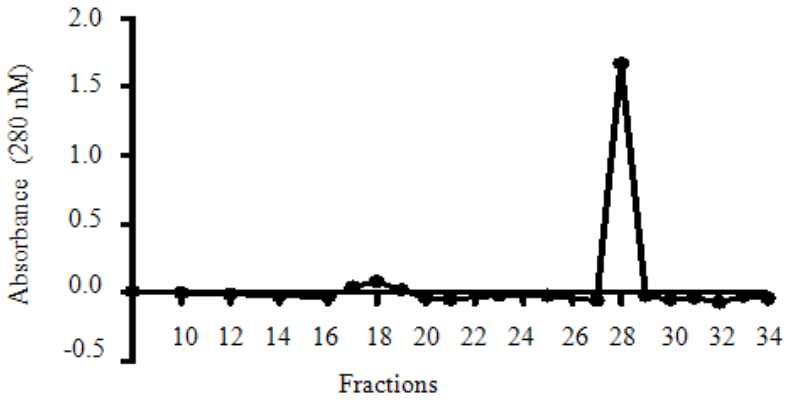


Figure 3.3: **Purification of semi-pro *Pf*PM IX by Gel Filtration and Cation Exchange.** Eluant fractions obtained from cation exchange were further purified by gel filtration. Protein concentration was determined by measuring the A_{280} (absorbance at 280 nm). The 'void volume' was seen in fractions 17-19 and the correctly folded semi-pro *Pf*PM IX was detected in fraction 28.

3.3.2 Active semi-pro *Pf*PM IX cleaves the RS6 substrate

Purified and refolded semi-pro *Pf*PM IX was active in the chomogenic assay at pH 4.5 with a 1.97 factor slope (the slope of absorbance over time). However, there was a delay of approximately 5 minutes before activity was seen (Figure 3.4). The scan of the activity test (Figure 3.5) demonstrated a significant peak shift to 268 when compared to

the negative control (277.1), indicating better cleavage of the substrate than by the r*Pf*PM II positive control (273.0) and validating the activity slope seen in Figure 3.5.

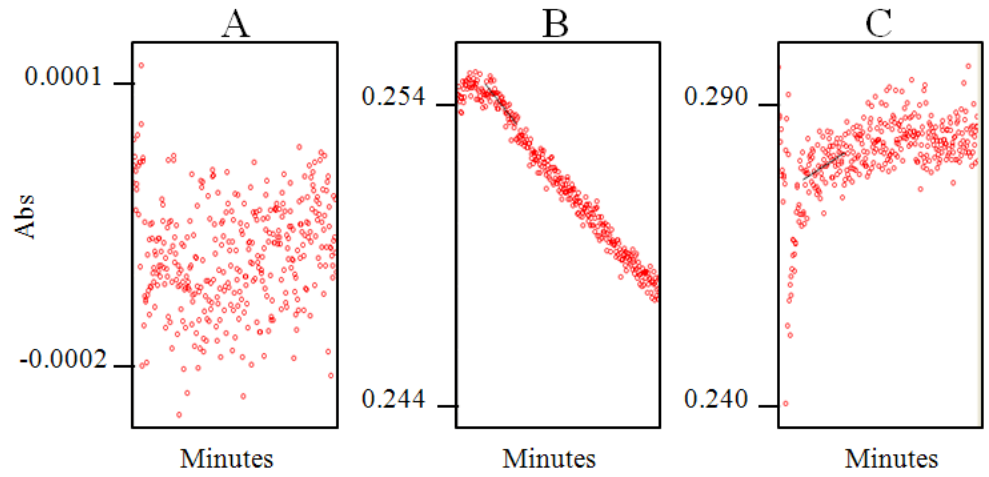


Figure 3.4: **Activity test of Gel Filtration Fraction 28.** The activity of r*Pf*PM IX obtained in fraction 28 was determined in a chromogenic substrate (RS6) activity assay (B) in excess of 40 minutes. Assays were performed alongside an enzyme only, no substrate control (C) and a water blank (A), both of which showed no activity.

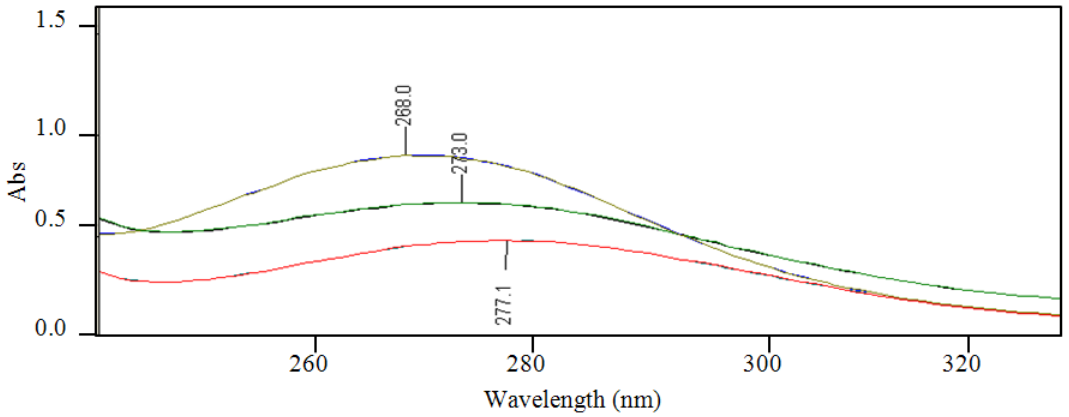


Figure 3.5: **Semi-pro *Pf*PM IX is active.** The cleavage of the chromogenic substrate (RS6) (yellow) by semi-pro *Pf*PM IX obtained in fraction 28 was determined by a scan analysis after activity assay completion and compared to r*Pf*PM II (green) and a negative control (RS6 substrate alone).

3.3.3 Optimal conditions for refolding and purifying semi-pro *Pf*PM IX

Dialysate from all folding conditions contained active enzyme. However, activity assays demonstrated that refolding in 50 mM sodium phosphate dibasic buffer at pH 7.0 produced more active enzyme. While 50 mM sodium phosphate buffer at pH 8.0 resulted in a steeper activity slope, substrate was not cleaved as well achieving a peak

absorbance at 274 nm (Figure 3.7C) in comparison to the pH 7.0 buffer which had a peak absorbance at 273 nm (Figure 3.6C). Enzyme refolded in 50 mM bicine pH 8.0 buffer demonstrated an even better activity slope (Figure 3.8A), however, after an extended period of time this activity decreased (Figure 3.8B). An absorbance scan of this assay also revealed incomplete substrate cleavage (peak of 274 nm) (Figure 3.8C).

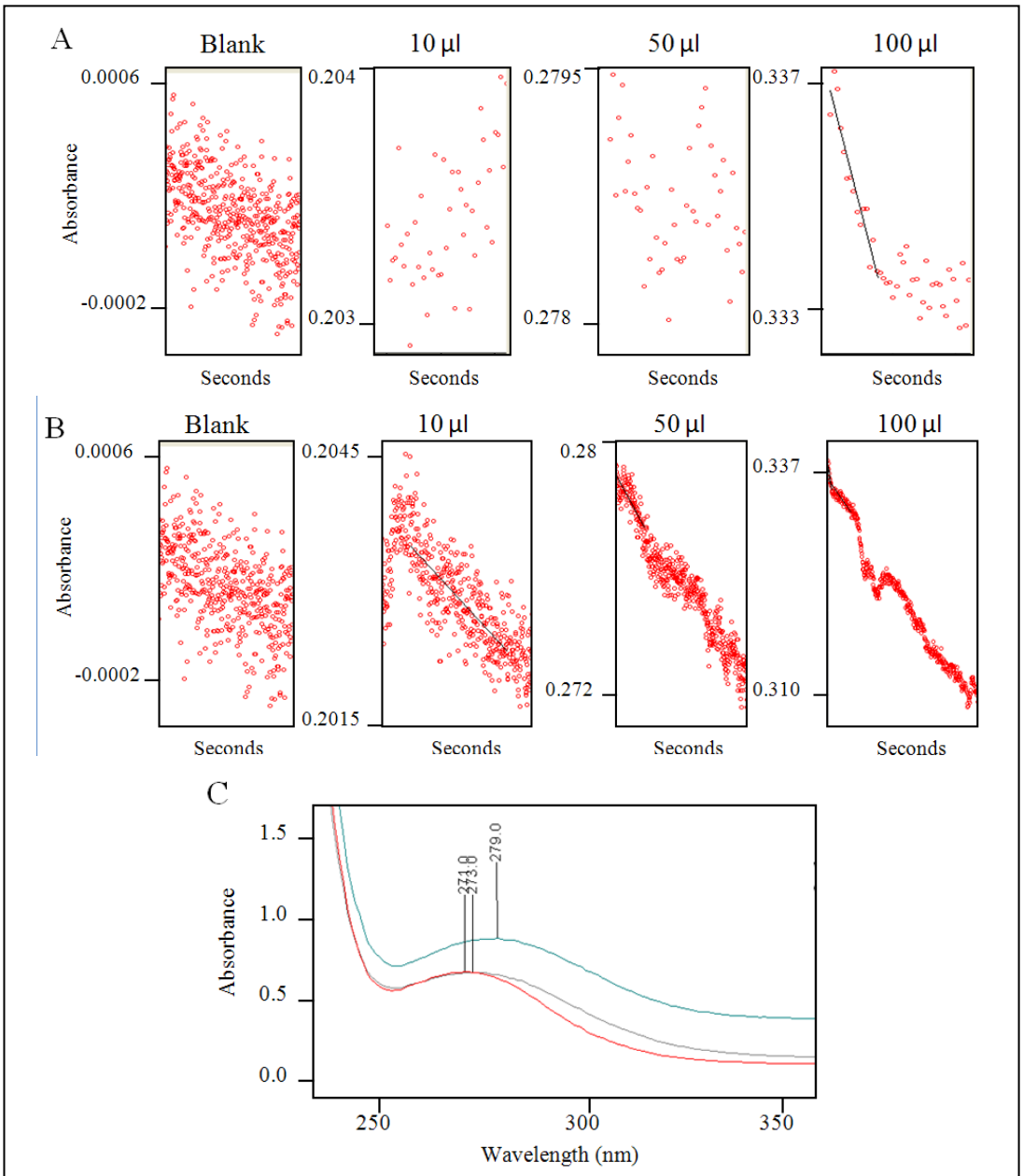


Figure 3.6: **Activity and substrate cleavage by semi-pro *Pf*PM IX refolded in 50 mM sodium phosphate dibasic pH 7.0 buffer.** The activity of semi-pro *Pf*PM IX (10, 50 and 100 µL) against RS6 was assessed for 1000 seconds (A) and in excess of 10, 000 seconds (B) against a water blank control. Data suggest that the enzyme is slow to cleave the substrate. The cleavage of the substrate is seen in the assay scan analysis (bottom) with a peak shift to 273 (red) compared to a negative substrate only control (blue) and positive pepsin control (grey).

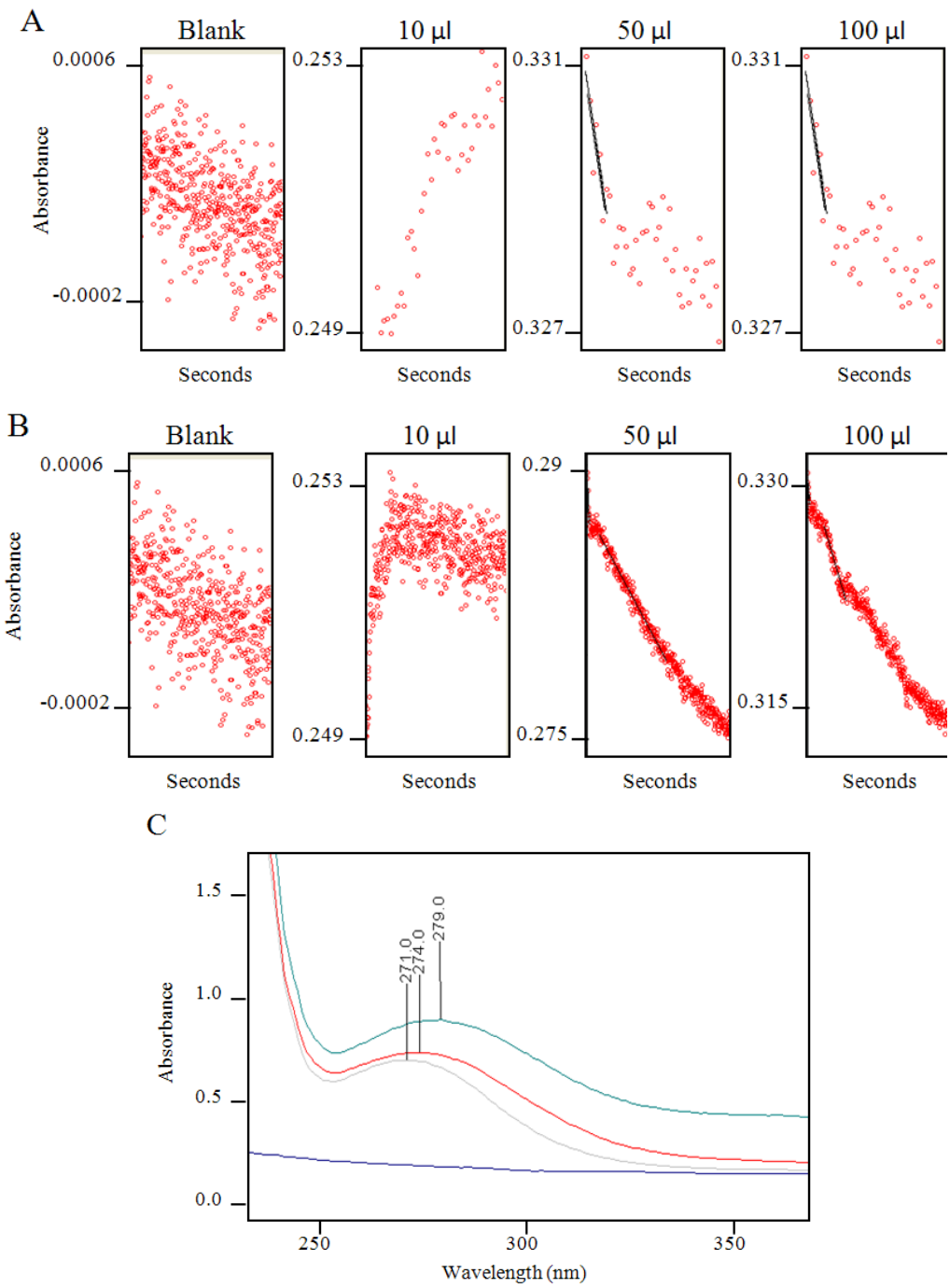


Figure 3.7: **Activity and substrate cleavage of semi-pro *Pf*PM IX refolded in 50 mM sodium phosphate dibasic pH 8.0 buffer.** The activity of semi-pro *Pf*PM IX dialysate at pH 8.0 (10, 50 and 100 µL) against the substrate RS6 was assessed for 1000 seconds (A) and in excess of 10, 000 seconds (B) against a water blank. Data suggest that the enzyme is quicker to cleave substrate than enzyme folded at pH 7.0. However, cleavage is not as effective. The cleavage of the substrate can be seen in the scan analysis of the completed activity assay samples (C) with a peak absorbance at 274 nm (red) compared to a negative substrate only control (blue) and positive pepsin control (grey).

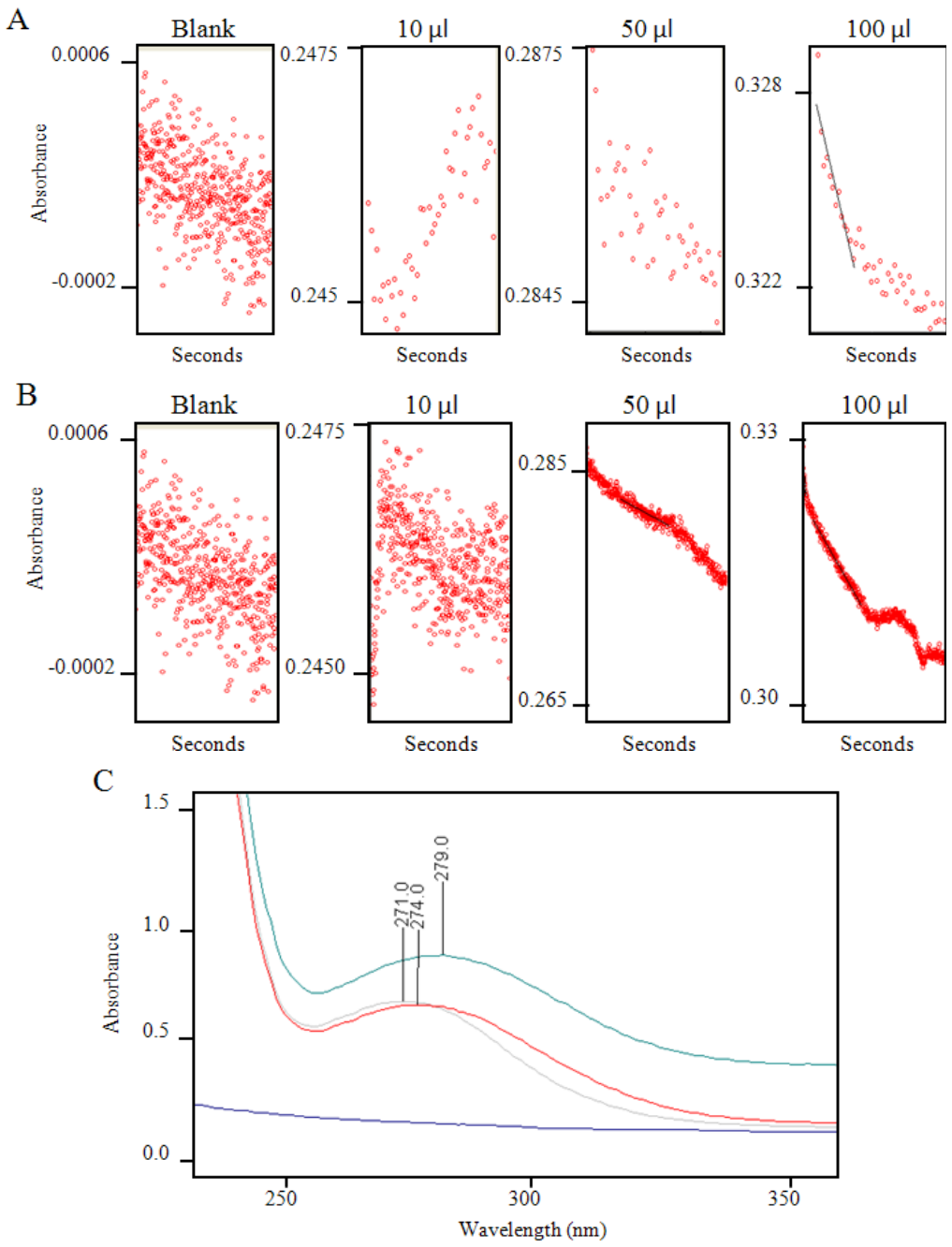
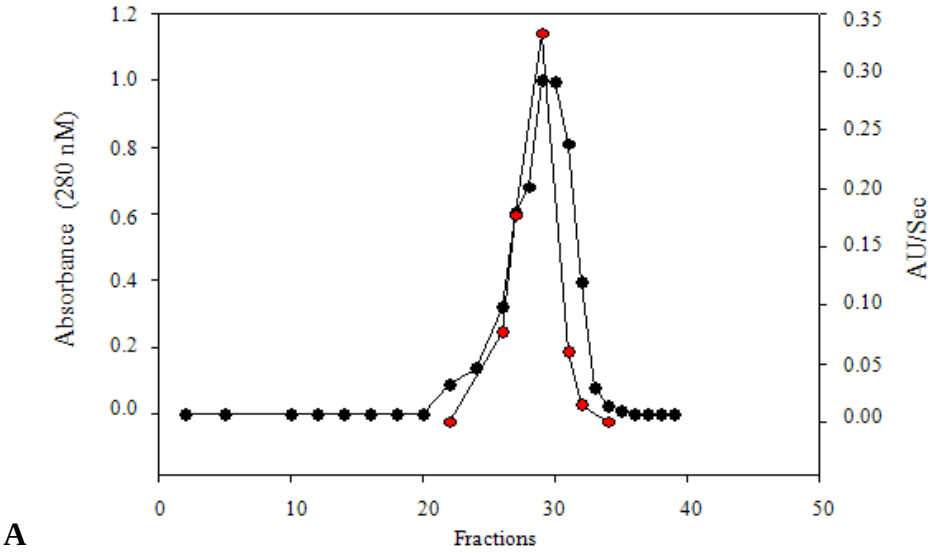
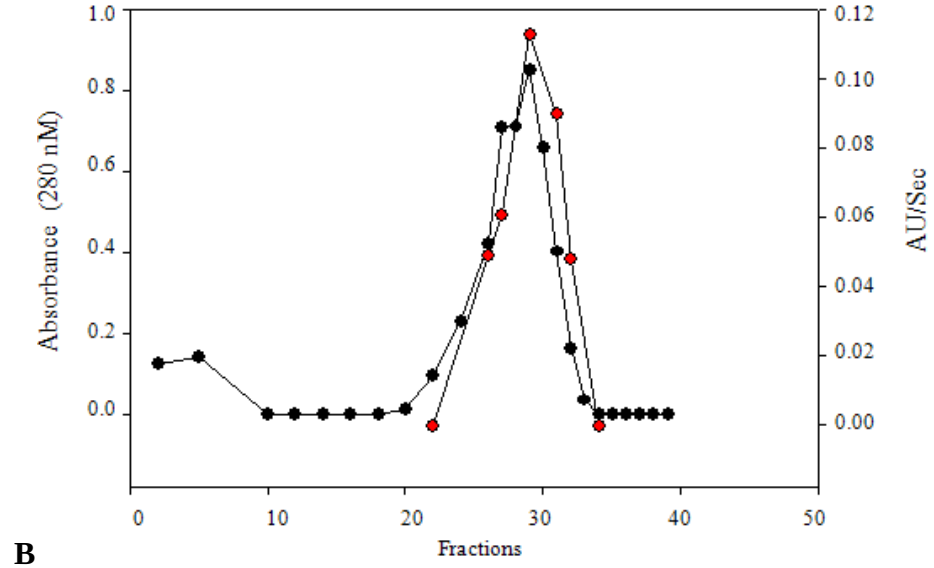


Figure 3.8: **Activity and substrate cleavage of refolded semi-pro *Pf*PM IX dialysed in 50 mM bicine pH 8.0 buffer.** Semi-pro *Pf*PM IX dialysed in a bicine buffer (10, 50 and 100 µL) demonstrates better activity against the substrate RS6 in 1000 seconds (A) compared to >10, 000 seconds (B). The water blank shows no activity. The cleavage of the substrate is seen in the scan analysis of activity assay samples (C) with a peak absorbance at 274 nm (red) compared to a negative substrate only control (blue) and positive pepsin control (grey).

Increased yields of semi-pro *Pf*PM IX were purified from preparations dialysed in 50 mM sodium phosphate dibasic pH 7.0 (Figure 3.9). Interestingly, semi-pro *Pf*PM IX refolded in 50 mM bicine pH 8.0 did not bind to the cation column. All fractions collected had an A_{280} of 0 and the flow through from the column had the same A_{280} as the dialysate that was put onto the column (data not shown).



A



B

Figure 3.9: **Increased yields and activity was seen when semi-pro *Pf*PM IX was refolded at pH 7.0.** Semi-pro *Pf*PM IX dialysed at pH 7.0 and 8.0 in 50 mM sodium phosphate dibasic buffers underwent were purified and the activity against the chromogenic substrate was tested. A) pH 7.0, the A_{280} of tested fractions is shown in black with the majority of semi-pro *Pf*PM IX being present in fractions 29 and 30. Activity of fractions 26-32 against the RS6 substrate is shown in red. B) pH 8.0, the majority of semi-pro *Pf*PM IX is present in fractions (black) 27-30, with the highest activity (red) seen in fraction 29.

3.3.4 Semi-pro *Pf*PM IX is auto-activated at pH 2-3.5

The auto-activation and maturation of semi-pro r*Pf*PM IX to mature r*Pf*PM IX was investigated across a pH range of 2-6 and detected using SDS-PAGE analysis (Figure 3.10). Partial cleavage of the 57 kDa r*Pf*PM IX was seen when pH was between 2 and 3.5 as highlighted in Figure 3.10. The top band in this highlighted region corresponds to mature r*Pf*PM IX ~50 kDa.

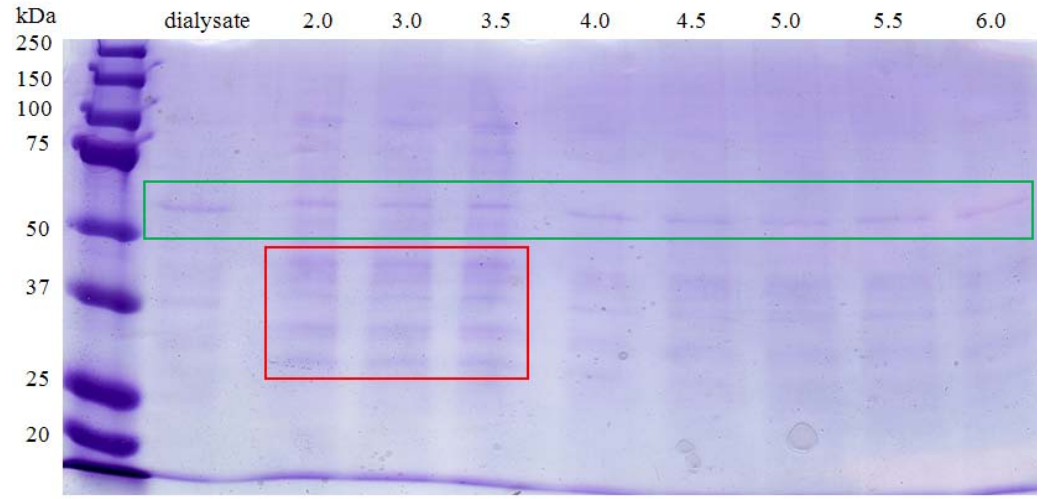


Figure 3.10: **Auto-activation of Semi-pro *Pf*PM IX.** The auto-activation of r*Pf*PM IX was tested across a pH range. Samples of dialysate were incubated at pH 2-6 and analysed by SDS-PAGE. Semi-pro *Pf*PM IX (57 kDa; green box) can be seen in each sample. At pH 2, 3 and 3.5 partial cleavage of the enzyme was seen (highlighted in red).

3.3.5 NCtrunc-r*Pf*PM IX expression and purification trial

To improve protein yields an additional construct and additional growth media were investigated. The construct contained truncations on the N and C termini that brought it in line with published constructs of other *Pf*PMs (Bernstein, Cherney *et al.*, 1999, Clemente, Govindasamy *et al.*, 2006, Moon, Tyas *et al.*, 1997, Prade, Jones *et al.*, 2005, Asojo, Afonina *et al.*, 2002, Asojo, Gulnik *et al.*, 2003, Bhaumik, Horimoto *et al.*, 2011, Li, Yowell *et al.*, 2004, Liu, Marzahn *et al.*, 2009, Wyatt and Berry, 2002, Xiao, Tanaka *et al.*, 2007, Xiao, Briere *et al.*, 2010, Hill, Tyas *et al.*, 1994). The N and C terminal truncated r*Pf*PM IX was expressed in 3 expression media (Figure 3.11) and purified from inclusion bodies. The growth of *E .coli* containing this NCtrunc-r*Pf*PM IX expression construct grown in 2YT was similar to that of *E. coli* containing the semi-pro *Pf*PM IX construct (Table 3.2). While *E. coli* did not grow as quickly in Super Broth, this media produced the largest cell pellet per culture volume (Table 3.2). In

comparison to cell pellet weights, the best weight of inclusion bodies was purified from *E. coli* grown in 2YT. SDS-PAGE analysis of each trial expression and purification (Figure 3.12) revealed that 2YT also produced the most enzyme, however, some loss of NCtrunc-r*Pf*PM IX can be seen in Sup1-4 of the purification process.

Table 3.2: **NCtrunc-r*Pf*PM IX expression and purification compared to Semi-pro *Pf*PM IX expression.**

	NCtrunc-r*Pf*PM IX			Semi-pro *Pf*PM IX
	LB A_{280}	2YT A_{280}	Super Broth A_{280}	2YT A_{280}
Cell Pellet weight (g)	6.69	8.13	12.53	5.9
IB Weight (mg)	120	680	350	1000

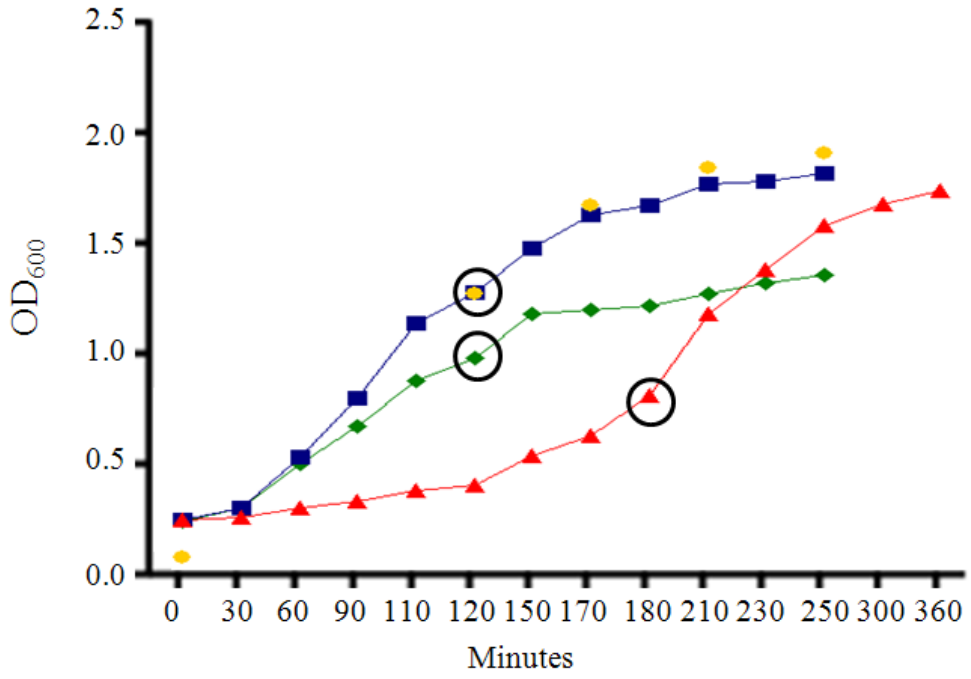


Figure 3.11: **Growth trial of *E. coli* containing the NCtrunc-r*Pf*PM IX expression construct.** The NCtrunc-r*Pf*PM IX construct was expressed in *E. coli* (1 L) in LB (green), 2YT (blue) and Super Broth (red). The OD_{600} of each expression culture was monitored and expression induced with 1 mM IPTG at varying time points (black circle). Data were compared to the growth of *E. coli* (2YT) containing the Semi-pro *Pf*PM IX expression construct (mustard data points).

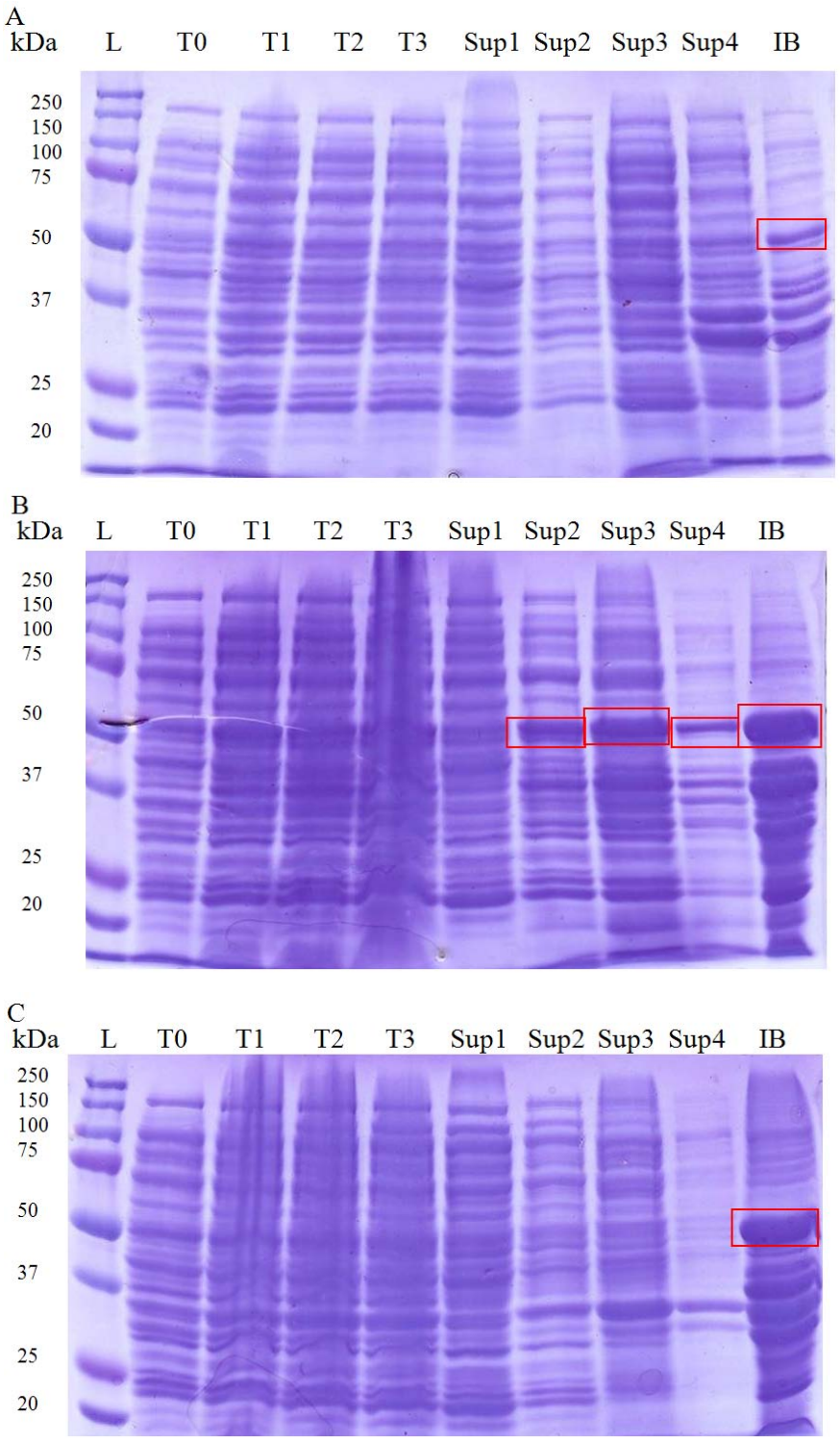


Figure 3.12: **SDS-PAGE analysis of NCtrunc-r*Pf*PM IX expression trials.** Samples were taken at each stage of expression and purification and examined by SDS-PAGE stained with Coomassie Blue. A) LB media. B) 2YT media. C) Super Broth media. NCtrunc-r*Pf*PM IX runs at approximately 50 kDa (red boxes).

3.4 Discussion and Conclusion

Active semi-pro *Pf*PM IX was expressed and purified. However, expression and purification of sufficient material was difficult. Yields were much lower than has been previously achieved with the digestive vacuole PMs (Liu, Marzahn *et al.*, 2009, Li, Yowell *et al.*, 2004, Dell'Agli, Parapini *et al.*, 2006). While inclusion bodies could be purified with yields approximating 30% of the expression culture cell weight, only 1.3% of this (Table 3.1) could be translated into correctly folded and active enzyme. Such low yields prohibited further characterization studies. Factors which may have negatively affected expression include semi-pro *Pf*PM IX pI of 9.34, the number of LCRs present (discussed in Chapter 4) and the presence of an export motif within the semi-pro *Pf*PM IX sequence. These characteristics have been found to impact protein expression and solubility (Mehlin, Boni *et al.*, 2006, Vedadi, Lew *et al.*, 2007, as reviewed in Birkholtz, Blatch *et al.*, 2008). Another factor could be contaminating DNA, with Olajuyigbe (2013) suggesting that the addition of DNAse 1 during inclusion body purification is helpful.

A significant obstacle in purifying semi-pro *Pf*PM IX was the precipitation encountered during concentration steps. This resulted in protein loss (Figure 3.2) and necessitated maintaining the enzyme at low concentrations, an obstacle shared by the Olajuyigbe study (Olajuyigbe, 2013). Working with large volumes resulted in downstream purification difficulties. While absorbance readings and activity assays demonstrated the presence of active semi-pro *Pf*PM IX, the concentration in final samples was too low to be visible via SDS-PAGE analysis (Coomassie or silver staining; data not shown). Alternative methods such as decreasing the elution volumes during cation exchange chromatography and gel filtration may be successful in increasing the final concentration of semi-pro *Pf*PM IX, however, due to time constraints this was not investigated during this study.

Data derived from studies investigating optimal refolding conditions, demonstrated that refolding at pH 7.0, rather than pH 8.0, results in the greatest yields of active protein. While semi-pro *Pf*PM IX refolded at pH 8.0 in both the sodium phosphate dibasic and bicine buffers both had initially faster rates of activity (Figures 3.7 and 3.8 respectively), enzyme refolded at pH 7.0 demonstrated better substrate cleavage (Figure 3.6). Further experiments with dialysates also revealed that larger quantities of purified enzyme, with better activity, are obtained when cation exchange is performed at a pH of

7.0 rather than at pH 8.0 (Figure 3.9). This is not unexpected considering the location of *Pf*PM IX within the red blood cells of infected human hosts (see Chapter 2), which are at a more neutral pH of 7.35-7.45 (Waugh and Grant, 2007). Interestingly, semi-pro *Pf*PM IX refolded in bicine buffer (pH 8.0) for cation exchange failed to bind to the column. Reasons for this remain unclear and given the success of the sodium phosphate dibasic buffer, it was not investigated further.

Experiments to determine the optimum pH for auto-activation of r*Pf*PM IX showed that this may require an acidic environment. While complete auto-activation was not seen, SDS-PAGE analysis revealed the presence of a ~50 kDa protein which corresponds to active mature r*Pf*PM IX (Figure 3.10) Olajuyigbe (2013). It would also appear that complete auto-activation of semi-pro *Pf*PM IX takes longer than 30 minutes. Further experiments to confirm the pH optimum of auto-activation need to be conducted with sufficient quantities of purified and correctly folded semi-pro *Pf*PM IX, and longer periods of time. Research has already shown that in *P. falciparum*, *Pf*PM IX is processed at the N terminal PEXEL motif by *Pf*PM V (Boddey, Hodder *et al.*, 2010) and research in our laboratory (see Chapter 2) also suggests that the native *Pf*PM IX is C terminally processed as well. It may be that *Pf*PM IX is processed by additional enzymes and while demonstrating some activity against the RS6 substrate, does not completely auto-activate with loss of the N-terminal pro-region.

Truncating and codon optimizing the semi-pro *Pf*PM IX expression construct did not result in a marked increase in protein expression. While this result was unexpected others have also shown that codon optimision for expression in *E. coli* does not always improve expression of other proteins (Flick, Ahuja *et al.*, 2004). NCtrunc-r*Pf*PM IX containing *E. coli* had a similar growth pattern in 2YT media to *E. coli* containing the semi-pro *Pf*PM IX construct (Figure 3.11) but inclusion body yields were lower (Table 3.2). SDS-PAGE analysis also revealed that purification of this protein would result in significant loss and would be of little additional benefit to our studies (Figure 3.12). While our NCtrunc-r*Pf*PMIX construct did not increase protein yields in our hands and the quality of produced enzyme has yet to be determined, others have found this construct to produce a higher quality of enzyme (Olajuyigbe, 2013). As Olajuyigbe (2012) included DNAase treatments and a 20 mM Tris-HCI pH 8.0 refolding step before dialysis it seems clear that further experiments with this construct using these conditions is warranted.

Chapter 4: *In silico* modelling of Plasmepsin IX

4.1 Introduction:

4.1.1 *In silico* modelling and docking:

Structures derived from experimental methods such as X-ray crystallography and/or NMR analysis are of great value in aiding the understanding of the form and function of protein targets and are proving to be very useful in the discovery of small molecule-based protein inhibitors. However, these techniques require reasonable amounts of very pure protein, can be limited by the proteins molecular weight or their ability to crystallise and require substantial financial investments in equipment and human infrastructure (as reviewed in Anthonsen, Baptista *et al.*, 1994). A more cost efficient and much less time intensive method of structure determination for many proteins, not currently reported in the Protein Data Base (PDB: http://www.rcsb.org/pdb/home/home.do), is *in silico* modelling. Currently, two different approaches, namely *de novo* prediction and comparative modelling (reviewed in Anthonsen, Baptista *et al.*, 1994) are used to generate *in silico* models of potential protein-based therapeutic targets.

De novo (or *ab initio – Latin:* from first principles) modelling derives *in silico* protein structures, from the 1° protein sequence, without reference to any pre-determined three-dimensional template structure (Leach, 2001, Latek, Ekonomiuk *et al.*, 2007). This modelling technique is performed with the aid of multiple sequence alignments which guide fold selection and secondary and tertiary structure prediction (Ortiz, Kolinski *et al.*, 1999, Skolnick, Kolinski *et al.*, 2001). One of the advantages of *de novo* modelling is that it is not restricted to known protein folds currently reported in the protein database. In cases where the protein of interest has limited similarities to known templates, *de novo* modelling can often be the method of choice as it is more likely to produce models that are closer to the true protein structure (Kolinski, Betancourt *et al.*, 2001, Kolinski, Rotkiewicz *et al.*, 1999) than other approaches. *De novo* modelling, however, does have the disadvantage of being very computationally expensive for proteins with over 100 residues (Kolinski, Betancourt *et al.*, 2001, Kolinski, Rotkiewicz *et al.*, 1999). It is also less accurate than other methods when proteins are large (Kolinski, Betancourt *et al.*, 2001, Kolinski, Rotkiewicz *et al.*, 1999). It is for these reasons that another approach, comparative modelling, is often used in an attempt to generate *in silico* models.

Comparative modelling attempts to predict structure using a homologous template derived from crystal X-ray or solution-state NMR derived data. Whilst comprehensive stand-alone computational packages exist for the non-specialist user, free to use and high quality online web-based systems can also be used to search for suitable templates and to predict protein structure. The first pass predicted structure can then be further refined with free (such as NAMD) or commercially available (such as software provided by Schrodinger Inc.) structural simulation software. Two of the more well-known online servers that predict protein structure based on sequence data include the Swiss-Model system based at the Swiss Institute of Bioinformatics (http://swissmodel.expasy.org/) and the Protein Homology/analogY Recognition Engine V 2.0 - PHYRE2 (http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index). These systems differ in the methods they use to select protein templates. Swiss-Model (Schwede, Kopp *et al.*, 2003, Arnold, Bordoli *et al.*, 2006) uses the amino acid mode of the Basic Local Alignment Search Tool (BLAST: (Altschul, Gish *et al.*, 1990, Mount, 2007)) and locates templates based on sequence similarity, whereas PHYRE2 uses the Structural Classification of Proteins (SCOP) nomenclature to assign protein secondary and tertiary structural domains and selects templates based on structure rather than sequence similarity (Kelley and Sternberg, 2009).

Although advances in software design and the availability of online systems for structural template selection and structure prediction have improved the availability and accuracy of *in silico* protein modelling, making it a powerful tool in biology (Ripka, Satyshur *et al.*, 2001) it is still essential that structure predictions be validated during and after model generation to ensure the quality of the prediction. Programmes such as ProCheck (Laskowski, 1993) can be used during this process. ProCheck (freely available at http://www.ebi.ac.uk/thornton-srv/software/PROCHECK/), while still being actively improved, checks the stereo-chemical quality of protein models and provides an analysis of the overall structure and residue-by-residue geometry.

In silico drug discovery often involves using a protein model in inhibitor docking studies (Ripka, Satyshur *et al.*, 2001, Degliesposti, Kasam *et al.*, 2009, as reviewed in Warren, Do *et al.*, 2012). Despite relatively high computational demands, high-throughput virtual molecular docking screens can vastly reduce costs by facilitating the screening of large online chemical libraries and the identification of credible molecular targets before embarking on a comprehensively more expensive laboratory-based

synthetic program (Ripka, Satyshur *et al.*, 2001, Kasam, Salzemann *et al.*, 2009). A number of commercial companies such as Sigma-Aldrich (www.sigmaaldrich.com) and Ambinter (www.ambinter.com) have virtual small molecule structure chemical compound libraries freely available for download, enabling researchers to perform virtual screens with relative ease before obtaining compounds for confirmation assays. The Medicines for Malaria Venture (MMV) has also recently made lead compounds with known antimalarial activity from an extensive screening campaign of around 4 million compounds available for researchers. Most of these libraries are supplied in 2D format and can be easily used in *in silico* screening (Gagaring, Borboa *et al.*, 2013, Guiguemde, Shelat *et al.*, 2010, Gamo, Sanz *et al.*, 2010). Furthermore, 400 of these compounds, both drug-like and probe-like, have been made physically available to researchers (Spangenberg, Burrows *et al.*, 2013).

4.1.2 *Pf*PM IX and Orthologs:

PM IX is present in all sequenced Plasmodium species including *P. falciparum, P. berghei, P. knowlesi, P. vivax* and *P. yoelii.* It is one of 10 aspartic proteases in *P. falciparum,* and it bares very little resemblance to the remaining nine aspartic proteases in this organism (Table 4.1). The highest homology is seen when *Pf*PM IX is compared to *Pf*PM VIII (29%) and *Pf*PM X (37%), with *Pf*PM X being listed as a possible paralog by PlasmoDB (www.plasmodb.org). A phylogenetic tree of the PMs, generated with the online server Phylogenic.fr can be found in Appendix 8. Despite obvious differences, *Pf*PM IX does share key areas of conservation with all nine remaining *P. falciparum* PMs. Regions of highest sequence homology occur predominantly around the key catalytic aspartic residues (Figure 4.1; shown in red). *Pf*PMs I-VII have predicted transmembrane domains (blue) located at their N-termini with *Pf*PM V's transmembrane domain being located within this proteins first Low Complexity Region (LCR) (Figure 4.1). *Pf*PMs V-X have a signal peptide at their N-termini which is absent in *Pf*PMs I-IV. Additionally, *P. falciparum* PM IX is the only PM that contains a PEXEL motif in its sequence (xLxE/Q/D; amino acid positions 45-50), which marks the protein for export into the host red cell (Boddey, Hodder *et al.*, 2010, Russo, Babbitt *et al.*, 2010). When compared to *Pf*PMs I-VIII, *P. falciparum* PMs V, IX and X also have comparatively large low complexity insert regions. The number of LCRs (purple) within *Pf*PM IX (4 LCRs) is greater than the majority of LCR regions in other malarial PMs. Interestingly *Pf*PM IX has a LCR proximal to its second catalytic ASP residue

(one residue removed), while the other malarial PMs either have no LCRs or LCRs that are at least 12 (*Pf*PM VII) or 13 (*Pf*PM V) residues removed from their active site.

Table 4.1: **Comparison of *P. falciparum* PM coding sequences (full length transcripts).** *Pf*PM sequences were paired and aligned using Clustal Omega (Sievers, Wilm *et al.*, 2011) and ClustalX 2.1 software (Larkin, Blackshields *et al.*, 2007) and the level of identity determined. (*PMs for which structural data has been obtained).

*Pf*PM	Sequence identity to *Pf*PM IX
I	24%
II*	23%
III (HAP)*	23%
IV*	24%
V	21%
VI	25%
VII	22%
VIII	29%
X	37%

LCRs contain non-random, limited alphabet amino acids (Wootton, 1994, Wootton J., 1993, reviewed in Zilversmit, Volkman *et al.*, 2010) and are usually located in the solvent exposed hydrophilic loops of proteins (Aravind, Iyer *et al.*, 2003). They are extremely common in *P. falciparum*, with approximately 87% of the proteins in this species having one or more of these regions. (DePristo, Zilversmit *et al.*, 2006). LCRs are also extremely divergent and evolve rapidly (as discussed by Zilversmit, Volkman *et al.*, 2010), lending to the ongoing antigenic variation of *P. falciparum.* All PM IX orthologs contain LCRs (Figure 4.2). While the length of these LCRs varies between species they are, for the most part, located in similar regions of the protein. Of particular concern from a homology modelling perspective is that all PM IX orthologs in human infecting species contain either one large LCR of 31 or 28 residues (*P. falciparum* and *P. vivax*) or 2 smaller LCRs (19 and 10 residues) *(P. knowelsi*) immediately before the second catalytic ASP residue (Figure 4.2; red), which are often difficult to simulate (reviewed in Zilversmit, Volkman *et al.*, 2010).

Although all PM IX orthologs contain LCRs, there are an increased number of these regions in the proteins encoded by species that infect humans. In addition, while there is a high level of conservation between all species that have been compared (Figure 4.2), only those species that infect mice (*P. berghei*, *P. yoelli* and *P. chabaudi*) contain a region thought to be a transmembrane domain at the N terminal region of the protein (underlined and in bold). The transmembrane domains in *Pf*PMs I-V are thought to mediate localisation to the DV (*Pf*PMs I-IV; (Francis, Gluzman *et al.*, 1994)) and ER (*Pf*PM V; (Russo, Babbitt *et al.*, 2010)). While there is no experimental evidence that it is a true transmembrane domain, and the localisation of PM IX in mouse species remains unknown, it may suggest that PM IX has a different function in different species.

Although there are 6301 aspartic protease structures derived from experimental data (X-ray and/or NMR analysis) or from *in silico* simulations in the Protein Data Bank (PDB) (Berman, Westbrook *et al.*, 2000), the database contains no published structures of *Pf*PM IX. The aim of this computational study was to attempt to simulate a model of the *Pf*PM IX structure and so determine the limits of what is achievable using the current modelling approaches, given the presence of the large LCRs in the *Pf*PM IX sequence. It was also hoped that a *Pf*PM IX homology model may aid our understanding of this protein and fast-track inhibitor identification studies.

```
PM I     1   ------MALSIKE-DFSSA----------FAKNESAVNSSTFNNNMKTWKIQKRFQILYVFFFLLITGALF----YYL--  57
PM II    1   ------MDITVREHDFKHG----------FIKSNSTFDGLNIDNSKNKKKIQKGFQILYVLLFCSVMCGLF----YYV--  58
PM HAP   1   ------MNLTIKEEDFTNT----------FMKNEESFNTFRVTKVKRW-NAKRLFKILFVTVFIVLAGGFS----YYI--  57
PM IV    1   ------MALTVKEEEFSNT----------LIKNASAFDRLKLGNLKNL-KIQKKLQFLYLILFVLITGVFF----FFL--  57
PM V     1   ------------------------------------------------MNNYFLRKENFFILFCFVFVSIF---------  23
PM VI    1   ------------------------------------------------------MPYFHIFLYILIFC--V----LVH--  18
PM VII   1   ------------------------------------------------MNKNIIQIYLFVFILLLKQHI----VILKN  26
PM VIII  1   ------------------------------------------------------MN--------ILFCLFV----ITN--  12
PM IX    1   MFFINFKKIKKKQFPIYLTQHRIITVFLIFIYFINLKDCFHINNSRILSDVDKHRGLYYNIPKCNVCHK---CSICTHEN  77
PM X     1   -----MKRISPLNTLFYLS-------LFFSYTFKGLKCTRIYKIGTKA----------LP-CSECHDVFDCTGCLFEE  55

PM I     58  --IDNVLFPK--NKKINEIMNTSKHVI----------------IGFSIENSHDRIMKTVKQHRLKNYIKESLKFFKTGLT  117
PM II    59  --YENVWLQR--DNEMNEILKNSEHLT----------------IGFKVENAHDRILKTIKTHKLKNYIKESVNFLNSGLT  118
PM HAP   58  --FENFVFQK--NRKINHIIKTSKYST----------------VGFNIENSYDRLMKTIKEHKLKNYIKESVKLFNKGLT  117
PM IV    58  --IGNFYSHR----KLYQVIKNTKHTT----------------IGFKIDRPHDKVLSSVLKNKLSTYVKESFKFFKSGYA  115
PM V     24  -FVSNVTIIKC-NNVENKIDNVGKKIE------------------------NVGKKIGDMENKNDNVEN---------  66
PM VI    19  --ICPIHTLN--IFKNDENEKGSLNIP----------------LGKENNLFFNEI-------KLENRFKNNIKGYIQNVQ  71
PM VII   27  EEFTNPYSIRK---------KDIKAIVNV----------------------NNKLKSINIHKLDN-----------IN  62
PM VIII  13  --LYNIIAVK--AFK--ENLRVSKYYA----------------GGK------HKL-------NLENKYI----GISTIVL  53
PM IX    78  GEAQNVIPMVAIPSKRKHIQDINKERE-----ENKYPLHIFE--EKDIYNNKDNVVKKEDIYKLRKKKKQK----KNCL-  145
PM X     56  KESSHVIPLKLNKKNPNDHKKLQKHHESLKLGDVKYYVNRGEGISGSLGTSSGNTLDDMDLINEEINKKRT----NAQLD  131

PM I     118 QKPHL-----------------------------------------------------------GNAGD-SV------TLND  133
PM II    119 KTNYL-----------------------------------------------------------GSSND-NI------ELVD  134
PM HAP   118 KKSYL-----------------------------------------------------------GSEFD-NV------ELKD  133
PM IV    116 QKGYL-----------------------------------------------------------GSEND-SI------ELDD  131
PM V     67  KNDNV-----------------------------------------------------------GNKND-NVKN----ASSD  84
PM VI    72  KFHYL-----------------------------------------------------------MEKNKPNVLSYIQEDLLN  94
PM VII   63  KKDLL-----------------------------------------------------------GSYNENYI--LIKLKKQD  83
PM VIII  54  KGGYI-------------------------------------------------------------------------------  58
PM IX    146 --NFLEKDTMFLSPSHDKETFHINHMNKIKDEKYKQEYEEEKEIY------DNTNTSQEKNETNNEQNL-NINLINNDK  215
PM X     132 EKNFLDFTTYNKNKAQDIS----DHLSDIQKHVYEQDAQKGNKNFTNNENNSDNENNSDNENNSDNENNLDNENNLDNEN  207

PM I     134 VAN---V-------------------------------MYYGEAQIGDNKQK-FAFIFDTGSANLWVPSAQCNTIGCKT  177
PM II    135 FQN---I-------------------------------MFYGDAEVGDNQQP-FTFILDTGSANLWVPSVKCTTAGCLT  178
PM HAP   134 LAN---V-------------------------------LSFGEAKLGDNGQK-FNFLFHTASSNVWVPSIKCTSESCES  177
PM IV    132 VAN---L-------------------------------MFYGEGQIGTNKQP-FMFIFDTGSANLWVPSVNCDSIGCST  175
PM V     85  LYKYKLY-----------GDID-------------EYAYYFLDIDIGKPSQR-ISLILDTGSSSLSFPCNGCKDCGIHM  138
PM VI    95  FHN---S-------------------------------QFIADIGVGNPPQV-FKVVFDTGSSNLAIPSTKCIKGGCAS  138
PM VII   84  IFSKK-L-------------------------------STYYGEVQIGEQSENNMNVLFDTGSSQVWILNDTCKNSLCNN  131
PM VIII  59  --N---R-------------------------------QFIGEINIGNPPQT-FKVLFDTGSTNLWIPSKNCFTRACYN  100
PM IX    216 VT--------------------------LPLQQLEDSQYVGYIQIGTPPQT-IRPIFDTGSTNIWIVSTKCKDETCLK  266
PM X     208 NSDNSSIEKNFIALENKNATVEQTKENIFLVPLKHLRDSQFVGELLVGTPPQT-VYPIFDTGSTNVWVVTTACEEESCKK  286


PM I     178 -KNLYDSNKSKTY------EKDGTK--V---------EMNYVSGT-VSGFFSKDIVTIAN-LSFPYKF-IEVTDTN----  232
PM II    179 -KHLYDSSKSRTY------EKDGTK--V---------EMNYVSGT-VSGFFSKDLVTVGN-LSLPYKF-IEVIDTN----  233
```

```
PM HAP  178  -KNHYDSSKSKTY------EKDDTP--V---------KLTSKAGT-ISGIFSKDLVTIGK-LSVPYKF-IEMTEIV----  232
PM IV   176  -KHLYDASASKSY------EKDGTK--V---------EISYGSGT-VRGYFSKDVISLGD-LSLPYKF-IEVTDAD----  230
PM V    139  -EKPYNLNYSKTSSI-LYCNKSNCPYGLKCVGNKCEYLQSYCEGSQIYGFYFSDIVTLPS-YNNKNKISFEKLMGC----  211
PM VI   139  -HKKFNPNKSRTFTKNLKNNQESVY--T---------YIQYGTGT-SILEQSYDDVYLKG-LKIKHQC-IGLAIEE----  199
PM VII  132  IHSKYKRTKSFVY----KYDKKGLPSVI---------EIFYLSGK-IVAFEGYDTIYLGKKLKIPHTN-ISFATKV----  192
PM VIII 101  -KRKYDHKISKNY----KLVKKKDP--V---------EILFGTGE-IHIAYVTDDIHLGD-IKVKNQE-FGLASYI----  157
PM IX   267  -VHRYNHKLSSSF----KYYEPHTN--L---------DIMFGTGI-IQGVIGVETFKIGP-FEIKNQS-FGLVKREKASD  327
PM X    287  -VRRYDPNKSKTF----RRSFIEKN--L---------HIVFGSGS-ISGSVGTDTFMLGK-HLVRNQT-FGLVESE-SNN  346

PM I    233  ----GFEP-AYTLGQFDGIVGLGWKD----LSIGSVDPVVVELKNQNKIE----QAVFTFYLPFDDKHKGYLTIGGIEDR  299
PM II   234  ----GFEP-TYTASTFDGILGLGWKD----LSIGSVDPIVVELKNQNKIE----NALFTFYLPVHDKHTGFLTIGGIEER  300
PM HAP  233  ----GFEP-FYSESDVDGVFGLGWKD----LSIGSIDPYIVELKTQNKIE----QAVYSIYLPPENKNKGYLTIGGIEER  299
PM IV   231  ----DLEP-IYSGSEFDGILGLGWKD----LSIGSIDPVVVELKKQNKID----NALFTFYLPVHDKHVGYLTIGGIESD  297
PM V    212  HMH---EESLFLHQQATGVLGFSLTKPN------GVPTFVDLLFKHTPSL----KPIYSICVS---EHGGELIIGGYEPD  275
PM VI   200  ----SLHP-FSDL-PFDGIVGLGFSDPDFRSQNKYASPLIETIKKQNLLK----RNIFSFYVPKKLEKSGAITFGKANKK  269
PM VII  193  -----DIP-ILEEFKWDGIIGLGFQNGD--SIKRGIKPFLDILKDDKILTNKNYKNQFGYYLS---DKEGYITLGGIDNR  261
PM VIII 158  ----SDDP-FSDM-QFDGLFGLGISD------DKKKKQLIYDSIPKNILE----KNMFAIYYPKNVDDDGAITFGGYDKK  221
PM IX   328  NKS---NV-FERI-NFEGIVGLAFPE----MLSTGKSTLYENLMSSYKLQ----HNEFSIYIGKDSK-YSALIFGGVDKN  393
PM X    347  NKNGGDNI-FDYI-SFEGIVGLGFPG----MLSAGNIPFFDNLLKQNPNV----DPQFSFYISPYDG-KSTLIIGGISKS  415

PM I    300  -FY--E-----------------------------------------------GQLTYEKLN-HDLYWQVDL-DLH---FG  325
PM II   301  -FY--E-----------------------------------------------GPLTYEKLN-HDLYWQITL-DAH---VG  326
PM HAP  300  -FF--D-----------------------------------------------GPLNYEKLN-HDLMWQVDL-DVH---FG  325
PM IV   298  -FY--E-----------------------------------------------GPLTYEKLN-HDLYWQIDL-DIH---FG  323
PM V    276  -YFLSNQKEKQKMDKSDNNSSNKGNVSIKLKNNDKNDDEENNSKDVIVSNNVEDIVWQAIT-RKYYYYIKIYGLD---LY  350
PM VI   270  -YTVEG-----------------------------------------------KSIEWFPVI-SLYYWEINLLDIQ---LS  298
PM VII  262  LKNTPD-----------------------------------------------EEIIWTPVSTEMGYWTIQIMGIRKEYVN  295
PM VIII 222  -FIREN-----------------------------------------------SSIEWFDVT-SSKYWAIQMKGLK---IN  250
PM IX   394  -FF--E-----------------------------------------------GDIYMFPVV-KEYYWEIHFDGLY---ID  420
PM X    416  -FY--E-----------------------------------------------GDIYMLPVL-KESYWEVKLDELY---IG  442

PM I    326  NLT-------------------------------------------------------V--------EKATAIVDSGTSS  342
PM II   327  NIM-------------------------------------------------------L--------EKANCIVDSGTSA  343
PM HAP  326  NVS-------------------------------------------------------S--------KKANVILDSATSV  342
PM IV   324  KYV-------------------------------------------------------M--------QKANAVVDSGTST  340
PM V    351  GTNIMD-----------------------------------------------------------KKELDMLVDSGSTF  370
PM VI   299  HKN-----------------------------------------------------LFLCESKK---CRAAIDTGSSL  320
PM VII  296  NHF----------------EEN---------------------------KEEEEVIVKYEAFHDGGKNSIIDTGTYL  329
PM VIII 251  DVF-----------------------------------------------------LDVCSKNHEGFCQAVIDTGTSS  275
PM IX   421  HQKFCCGVNSIVYDLKKKDQENNKLFFTRKYFRKNKFKTHLRKYLLKKIKHQKKQKHSNHKKKKLNKKKNYLIFDSGTSF  500
PM X    443  KERICC------------DEES-----------------------------------------------YVIFDTGTSY  462


PM I    343  ITAPTEFLNKFFEGLDVVKIPFLPLYITTCNNPKLPTLEF------------------------RSATNV--------  388
PM II   344  ITVPTDFLNKMLQNLDVIKVPFLPFYVTLCNNSKLPTFEF------------------------TSENGK--------  389
PM HAP  343  ITVPTEFFNQFVESASVFKVPFLSLYVTTCGNTKLPTLEY------------------------RSPNKV--------  388
PM IV   341  ITAPTSFLNKFFRDMNVIKVPFLPLYVTTCDNDDLPTLEF------------------------HSRNNK--------  386
```

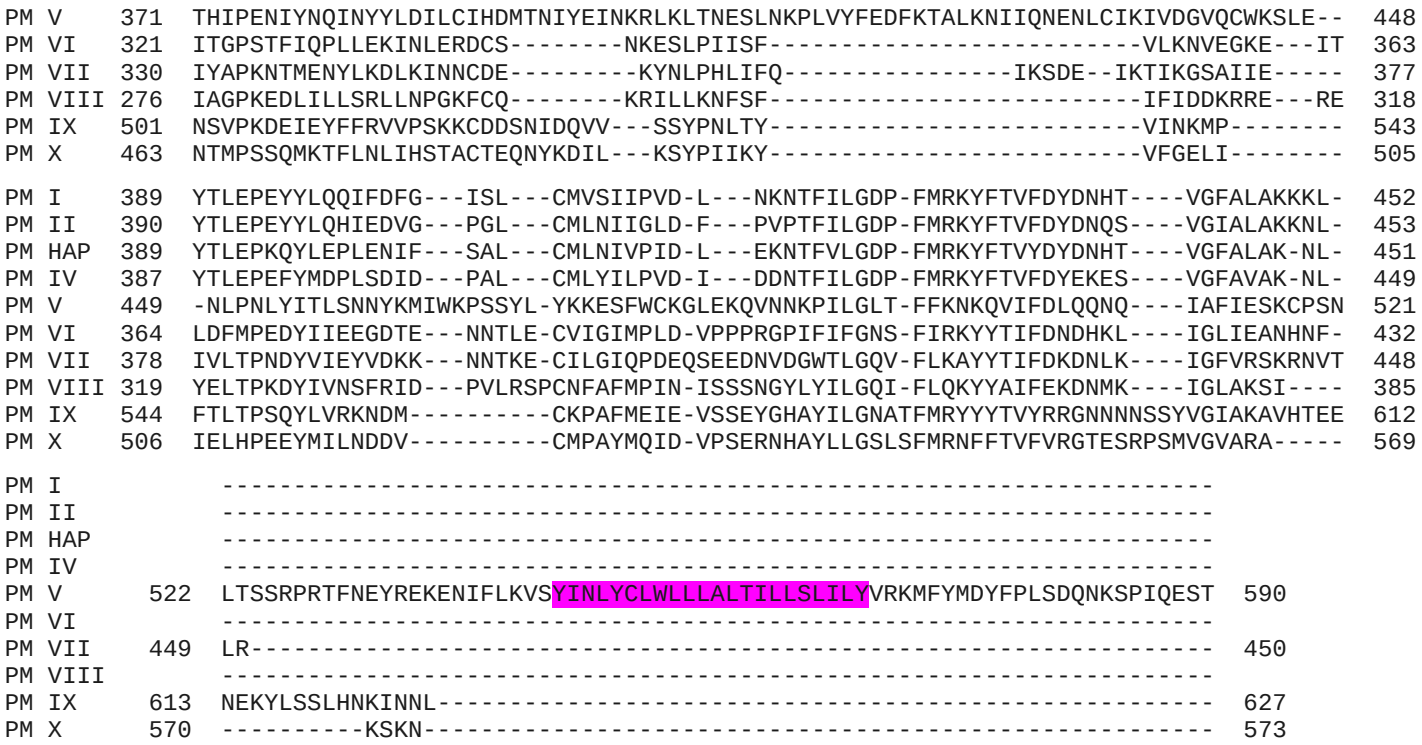


Figure 4.1: **Alignment of the 10 *P. falciparum* PMs.**

An alignment of the 10 *P. falciparum* PMs was generated with the Cobalt Multiple Protein Alignment Tool (Papadopoulos and Agarwala, 2007). Key areas of conservation are seen around the catalytic aspartic residues (red). *Pf*PMs I-V contain transmembrane domains (blue) at their N terminal ends which are not seen in the other *Pf*PMs, however they lack the signal peptides (bold and underlined) that exist in the remaining *Pf*PMs (PlasmoDB, 2009). *Pf*PM V's transmembrane domain is within a Low Complexity Region (LCR) (purple). *Pf*PM IX contains a PEXEL motif (green) at its N-terminal end which is not seen in the nine other *Pf*PMs.

```
P. falciparum 1    MFFINFKKIKKKQFPIYLTQHRIITVFLIFIYFINLKDCFHIN-NSRILSDVDKHRGLY-YNIPKCNVCHKCSICTHENG  78
P. vivax          ------------------------------------------------------------------------------
P. knowlesi   1    MPPHRFQKLGNRLLSTLLIHPKASLLFVVHLFLFRQGTCLRPS-GSNTFAT-----DLT-WEEGRKLTCQHVTRVTRMMP  73
P. berghei    1    MFFLTLKKLRKKCFFVFLTHPTITALFFIYIFNFVKSYHVNFNQNPNNLPSLKNNQEYYKQKIQPCNSCVNCFVCIHENG  80
P. chabaudi   1    MFFLNFKKLRKNYFLALLTHPTITVLFFIYIFNFVTSDYAHLNKKSNNLPSLKNDQEYNKQNIQPCNSCVNCSVCIHENA  80
P. yoellii    1    MFFLTLKKLRKKYFLLFLTHPTITTLFFIYIFNLVKSDYPNFNKNKYNLPSLKKNQKYFKQKIQPCNSCINCSICIHENG  80

P. falciparum 79   EAQNVIPMVAIPSKRKHIQD----INKEREENKYPLHIFEEK------DIYNNKDNVVKKEDIYKLRKKKKQK-K-----  142
P. vivax      1    -------MMAIPSKRKYLQYKIGKLNSELLQN-LPTKKLKKKKKKGSYSFFEGEDEDEGEEEQENEQDEQQQK-EEEGDP  71
P. knowlesi   74   VPCNIIPMVAIPSKRKYLQEKIEKIKSELHQN-LPEQKWKKKKKKESYSFFEGEDDDKGEVD-----------EEEGDS  140
P. berghei    81   EPQNILPLVAIPSKRHYFYEQ-DMSKNSNLNG------FPVK----------NKMDNST----NYFQKELNKK-KKN---  135
P. chabaudi   81   EPQDILPLVAVPSRRKYFYEQ-DRSKDDDLND------FPVEDKI-------NENDERTEYDNDYSQNELSKK-KKK---  142
P. yoellii    81   EPQNILPLVAIPSKRHYFYEQ-DMSKNSNLNG------SPVK----------NKREDSIKFDRNYSQKELNKKNKKN---  140

P. falciparum 143  -------NCLNFLEKDTMFLSPSHDKETFHIN-----HMNKIKDEKYKQEYEEEKEIYDNTNTSQEKNETNNEQNLN-IN  209
P. vivax      72   PSDATMDNHISHHNKGTHYGGEDHHPDEFPNCVNSDCHMSN--DASAEPDYLGQ---FMNGSGEKARAQTQRSNWSN-WN  145
P. knowlesi   141  TSHATMDNQIFHHNKGTHYEGEDKHPDEFQKCATSDCHMNK--DASGIPDYLRH---FMDGSEEKAQ--------TS-WS  206
P. berghei    136  ------YNFIENHT-------------AISNID------NDITDEY--KESESD---LDEENIVKDNFNNLRSSYENIYN  185
P. chabaudi   143  -----MYNFIENHN-------------EMPDMD------NDVTDE----EYEEN---MEEENMEEENTTNLESNYENTFN  191
P. yoellii    141  ------YNFIENHI-------------AMSNIN------NDITDGD--RETEDG---LNQENIAKDNFNNLISSYENIYN  190

P. falciparum 210  LIN-----------NDKVTLPLQQLEDSQYVGYIQIGTPPQTIRPIFDTGSTNIWIVSTKCKDETCLKVHRYNHKLSSSF  278
P. vivax      146  SWSGAFKKKEVSS-TDQVTLPLQQLQDSQYVGYIQIGNPPQTIRPIFDTGSTNIWVVSTKCKDDTCLKVHRYNYKLSRSF  224
P. knowlesi   207  SWSSAFKKKEVSSSTDQVTLPLQQLQDSQYVGYIQIGNPPQTIRPIFDTGSTNIWVVSTKCKDDTCLKVHRYDYKLSKSF  286
P. berghei    186  Q------KKEHSI-DSKVILPLQQLQDSQYVGSIQIGNPPQTIRPIFDTGSTNIWVVSTKCKDRTCLKVHRYNHKLSNTF  258
P. chabaudi   192  S------EKEDST-DSKVILPLQQLKDSQYVGFIQIGNPPQTIRPIFDTGSTNIWIVSTKCKDKTCLKVHRYNHKLSDTF  264
P. yoellii    191  Q------KTEHSI-DNKVILPLQQLQDSQYVGSIQIGNPPQTIKPIFDTGSTNIWVVSTKCRDKTCLKVHRYNHKLSNTF  263

P. falciparum 279  KYYEPHTNLDIMFGTGIIQGVIGVETFKIGPFEIKNQSFGLVKREKASDNKSNVFERINFEGIVGLAFPEMLSTGKSTLY  358
P. vivax      225  RYYKPHTNLDIMFGTGIIQGVIGVETFRIGPFKVFNQPFGLVKREKRSEAKSNVFERINFEGIVGLAFPAMLSTGKTTIY  304
P. knowlesi   287  RYYKPRTNLDIMFGTGIIQGVIGVENFRIGPFKLFNQPFGLVKREKRSEAKSNVFERINFEGIVGLAFPAMLSTGKTTIY  366
P. berghei    259  KYYTPRTNLDIMFGTGIIQGTIGIDTFKIGPFKIENQSFGLVKREKGTNKKSNVFKRINFEGIIGLAFSTMLSTGGNPIY  338
P. chabaudi   265  KYYTPRTNLDIMFGTGIIQGTIGIETFKIGPFKIENQSFGLVKREKGTDDKSNVFERINFEGIVGLAFPAMLSTGGNPIY  344
P. yoellii    264  KYYSPRTNLDIMFGTGIIQGTIGIDTFKIGPFKIENQSFGLVKREKGSNKKSNVFERINFEGIIGLAFPTMLSTGGNPIY  343


P. falciparum 359  ENLMSSYKLQHNEFSIYIGKDSKYSALIFGGVDKNFFEGDIYMFPVVKEYYWEIHFDGLYIDHQKFCCGVNSIVYDLK--  436
P. vivax      305  ENLMNTYKFSHNEFSIYIGKDNKHSALIFGGVERRFFEGDIYMFPVVREYYWEIQFDGLYIDHQKFCCDSSSIVYDMRKK  384
P. knowlesi   367  ENLMDTYKLSHNEFSIYIGKDNKHSALIFGGVDRRFFEGDIYMFPVVKEYYWEIEFDGLYIDHQKFCCDSSSIVYDMRKK  446
P. berghei    339  ENLMSNYNFPHNEFSIYIGMDNKY-----------------------------------------------SALIFGGVEKK  373
P. chabaudi   345  ENLMSSYNFPHNEFSIYIGMDNKY-----------------------------------------------SALIFGGVEKK  379
P. yoellii    344  ENLMASYNFPHNEFSIYIGMDNKY-----------------------------------------------SALIFGGVEKK  378
```

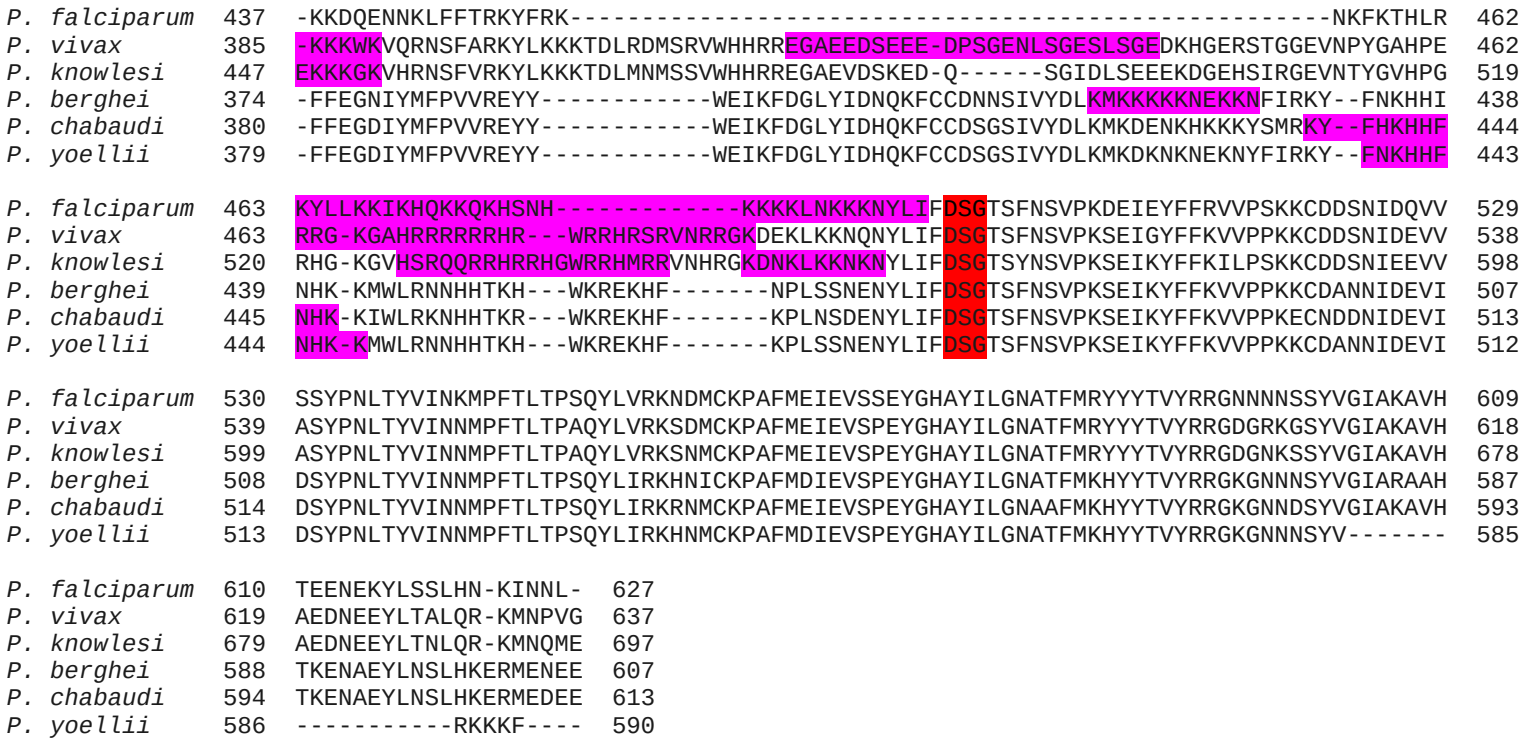


Figure 4.2: **Comparison of PM IX orthologs from six Plasmodium species.**

Alignments of PM IX orthologs were prepared using the Cobalt Multiple Protein Alignment Tool (Papadopoulos and Agarwala, 2007). There is a high degree of sequence conservation between orthologs, proximal to the catalytic sites (red). However, significant differences including the presence of predicted transmembrane domains (bold and underlined) in rodent malaria species and the PEXEL motif (green) in *P. falciparum* PM IX exist. In addition, while LCRs (purple) can be seen in all orthologs they are in greater numbers in species specifically infecting humans.

4.2 Materials and Methods:

4.2.1 *Pf*PM IX homology model generation:

4.2.1.1 Pf*PM IX model template and docking model selection:*

The protein coding sequence of *Pf*PM IX (PF3D7_1430200) was obtained from the online database www.plasmodb.org and entered, unmodified, into online structure prediction applications, namely: SwissModel (Arnold, Bordoli *et al.*, 2006), CPH model (Nielsen, Lundegaard *et al.*, 2010), 3D Jigsaw (Bates, Kelley *et al.*, 2001) and Protein Homology/analogY Recognition Engine (PHYRE) (Kelley and Sternberg, 2009). Each approach provided a template structure based on the specific alignment method utilised by that particular application.

4.2.1.2 Preparation and refinement of the docking model:

The most appropriate model generated by the public domain prediction systems, ranked on sequence identity and Expect (E) value was prepared for further study using the 'Protein Preparation Wizard' (incorporated into the Schrodïnger Inc. modelling suite of software; (Schrodinger, 2011). The E value describes the number of 'hits' or alignment matches that can be expected by chance when searching a database - a lower E value is indicative of a better template with query match. During this process, the model was assessed for major structural errors including incomplete side chains and loops, flipped residues and missing hydrogen atoms. The automated wizard corrected any observable errors whereupon an initial energy minimisation was performed. This minimisation step 'relaxes' the structure of the protein, removing high energy hot-spots from the model (such as VDW clashes, unfavourable bond and torsion angles and non-ideal bond lengths) and locates a local energy by gradually changing atom coordinates to conformations possessing decreasing total energies (Leach, 2001). Following on from the automated analyses a further minimisation of one thousand iterations was carried out in explicit solvent using a OPLS2005 force-field using the Desmond package (Maestro-Desmond Interoperability Tools, 2010, Bowers, Chow *et al.*, 2006, Shivakumar, Williams *et al.*, 2010). When the minimisation runs were complete, the model underwent a 2 nanosecond molecular dynamics simulation under NTP conditions (fixed temperature and pressure with variable volume) in explicit solvent containing 0.15 M NaCl again using Desmond. This initial simulation was carried out to relax the model to more accurately resemble the protein environment under physiological conditions. A downstream docking model for further refinement was taken at random

from frame 304 of 400 during the latter stages of the simulation experiment once the system had reached thermal equilibrium.

To generate a model for docking studies, further simulation was performed on the system derived from above using NAMD software (Phillips, Braun *et al.*, 2005). An initial NAMD explicit solvent minimisation (0.15 M NaCl) was carried out at 298 K, over 10000 steps with pressure equilibrated to 1 atm (Appendix 6). After this minimisation, the *Pf*PM IX model was equilibrated for pressure and temperature inside a water box for 3000 steps with an initial additional minimisation of 200 steps (Appendix 6). Molecular dynamics simulations were carried out for 20 ps at 310 K in a water box (Appendix 6) and then a full production run was obtained over 20 ns under the conditions described above (Appendix 6). The final *Pf*PM IX 'docking model' was taken at random (time step 1201) in the latter stages of the molecular dynamic simulation.

4.2.2 Validating the model:

4.2.2.1 Examining the model:

The *Pf*PM IX 'docking model' was examined using Virtual Molecular Docking (VMD) (Humphrey, Dalke *et al.*, 1996) and compared to published solid state data for *Pf*PM II and the chosen template structure. The models were aligned using STAMP (Russell and Barton, 1992), a structural alignment tool in VMD.

4.3 Results:

4.3.1 Choice of a template and selecting a *Pf*PM IX 'working' model

Online applications (SwissModel, CPH Model, 3D Jigsaw and PHYRE) were used to generate preliminary *Pf*PM IX models. The templates used for each of these models were then examined and the most appropriate template and model, based on sequence identity and E value was selected for further refinement. Approaches that use sequence alignment methods (Swissmodel, 3D Jigsaw and CPH model) produced templates found to have low sequence identity to the query (<31%; Table 4.2). The model generated by the Swiss server was based on 1PFZ.pdb – a proPM II protein from *P. falciparum* with an amino acid identity of only 19% to *Pf*PM IX. Models derived from the other online applications had similar levels of identity and were also based on aspartic protease structures (CPH model used 1QDM – a prohytepsin, 3D Jigsaw used 1HTR – a human pro-gastrocyn, Table 4.2). With this is mind, and as tertiary protein structure is normally

more conserved than sequence (Doolittle, 1992), the solid state template chosen for alignment and model production was the one produced by the threading application PHYRE. This program, which selects templates based on secondary structure (Kelley and Sternberg, 2009), produced a model based on the template (PDB code 1AM5 – an aspartic protease from a species of Atlantic cod (*Gadus morhua;* Figure 4.3) with a 29% sequence identity and an Expect (E) value of $1.4e^{-29}$ (Table 4.2).

Table 4.2: **Online server template choices for *Pf*PM IX structure prediction.**

Server	Template	Sequence Identity (%)	E-value
SwissModel	1PFZ (proPM II)	19	
CPH Model	1QDM (prohytepsin)	17.97	$1e^{-23}$
3D Jigsaw	1HTR (human pro-gastrocyn)	30.79	-199
PHYRE	1AM5 (Atlantic cod aspartic protease)	29	$1.4e^{-29}$

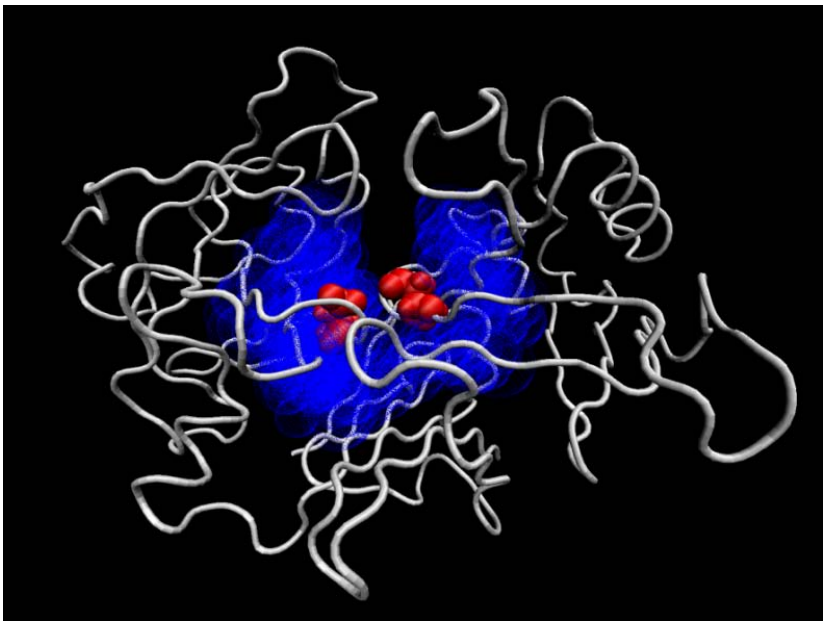


Figure 4.3: **Homology model of *Pf*PM IX.** The C-α backbone of the model is shown in tube (grey). The hydrophobic residues (blue dotted) within 5 Å of the catalytic residues (VDW; red) give an indication as to the shape and size of the active site.

4.3.2 Refinement of the model required for docking studies.

Minimizations and subsequent dynamic simulations were performed using Schrödinger Inc. and NAMD modelling software to reduce the conformational energy of the system and predicted equilibrium conformations in explicit solvent. Analysis of trajectory data

demonstrated no major fluctuations at equilibrium. Analysis of a Ramachandran diagram depicting the torsional angles of the generated *Pf*PM IX model demonstrates no major structural abnormalities (Figure 4.4). The small proportion of residues outside the optimal regions is similar to those seen in well characterised crystal structures of other PMs.

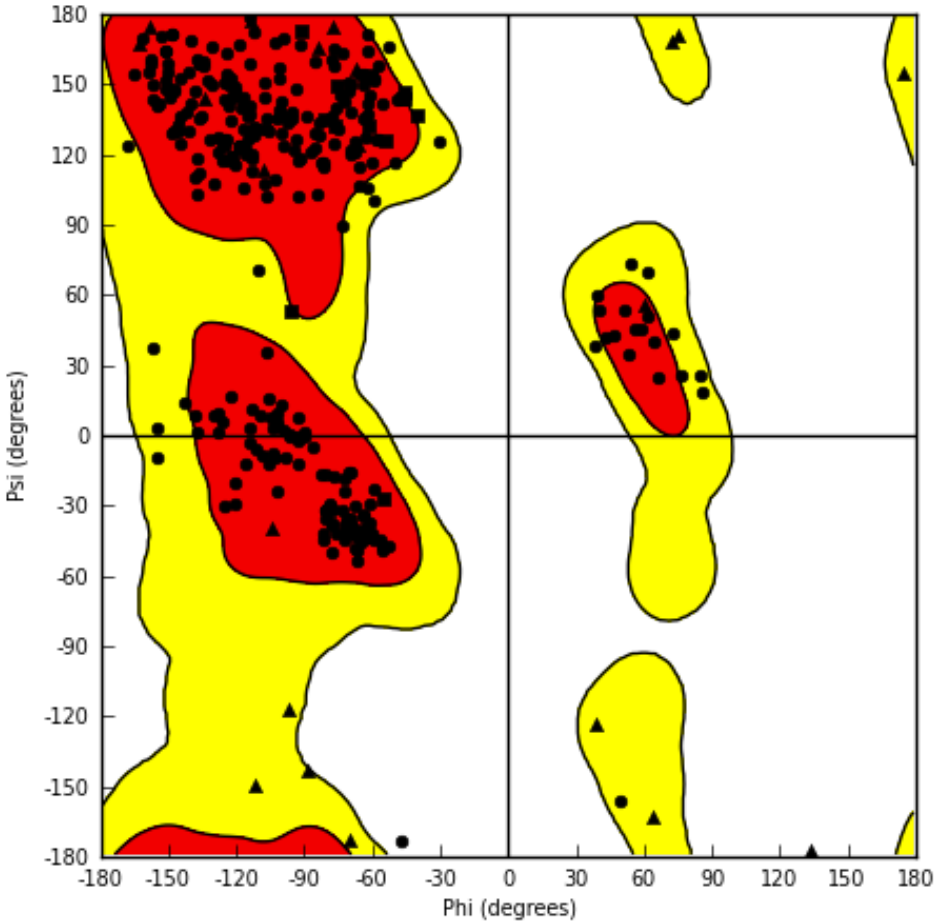


Figure 4.4: **Ramachandran diagram of the Φ-Ψ dihedral angles of the *Pf*PM IX homology model.** Individual residues (black) are placed according to their torsional angles with the majority being placed in either the most favoured (red) or additionally allowed (yellow) regions. This plot was generated with Schrödinger software.

4.3.3 Examination of the model:

The minimised *Pf*PM IX model (Figure 4.5C) was compared to structures of other aspartic proteases obtained from the PDB (Figure 4.5A/B; Figure 4.6). Analysis of aspartic protease solid state structures (PMII: PDB code 1LF3 and model template: PDB code 1AM5) possessing some homology (23-29%) to *Pf*PM IX was carried out using VMD. These comparisons revealed similarities in tertiary structure and the

relationship of residues between structures, proximal to the active site. All of the structures chosen for comparative purposes demonstrated the characteristic bilobal structure of the aspartic protease family, with a catalytic aspartic residue resident within each lobe (as reviewed by Dunn, 2002). The catalytic aspartic residues also share a similar location within the catalytic binding domain to comparative proteases, with both aspartic residues being directed towards the binding cleft (Figure 4.5). This suggested that the positioning of the catalytic residues is highly conserved within the different *Pf*PMs, an observation which is to be expected with crucial catalytic residues (Banerjee, Liu *et al.*, 2002). Whilst there are differences in the overall tertiary structures of the molecules, the hydrophobic residues proximal to each binding cleft are also well conserved (Figures 4.5 and 4.6).

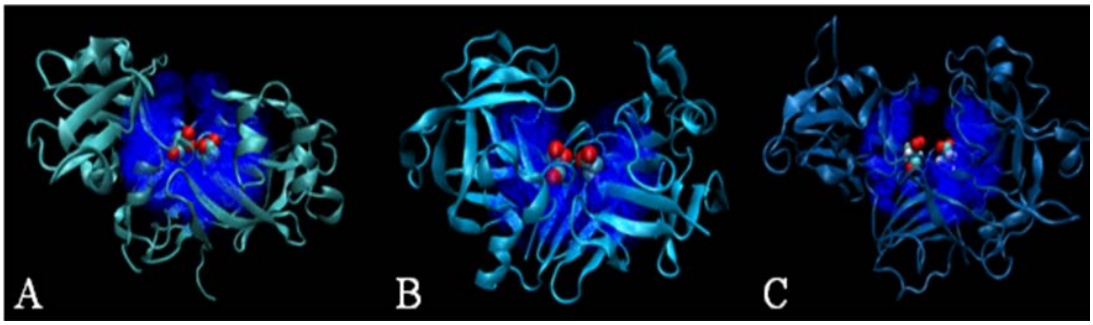


Figure 4.5: ***Pf*PM IX in silico homology model shares similar active site architecture with aspartic protease structures.** All C-α backbones are shown in new cartoon representation. The hydrophobic residues (blue dotted) within 10 Å of the catalytic residues (VDW) of each enzyme give an indication as to the location and size of the respective active sites. A) A liganded *Pf*PM II derived from X-ray data (PDB code: 1LF3) (Asojo, Gulnik *et al.*, 2003). Catalytic aspartic residues are 34 and 214. B) An aspartic protease from *Gadus morhua* (PDB code: 1AM5) which was used as a template structure (Karlsen, Hough *et al.*, 1998). Catalytic aspartic residues are 32 and 215. C) The generated *Pf*PM IX *in silico* homology model. Catalytic aspartic residues are 31 and 224.

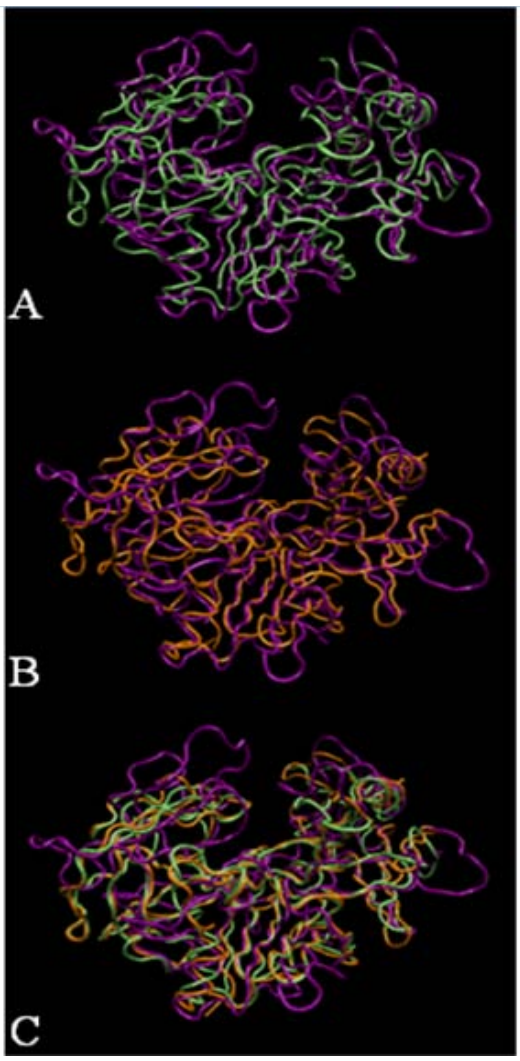


Figure 4.6: **Structural alignment of the *Pf*PM IX in silico homology model with 1AM5 and *Pf*PM II.** A) The *Pf*PM IX model (purple) is superimposed with its Atlantic cod aspartic protease template (lime) (PDB Code: 1AM5). B) The *Pf*PM IX model (purple) is aligned with a crystal structure of *Pf*PM II (orange; PDB Code: 1LF3). C) Merge of A and B.

While the structural similarities between the *Pf*PM IX homology model and other aspartic proteases was encouraging suggesting the generation of a realistic model, further examination of the model revealed a serious flaw in the assembly of the protein. During the process of generating the model, PHYRE had omitted an essential region of the *Pf*PM IX structure (Figure 4.7). Subsequent analysis of both the model and *Pf*PM IX sequence suggested that this deleted region included the second catalytic aspartic residue located within a region of low complexity. A BLAST search of the deleted region (Table 4.3) showed that this insert region is highly specific to PM IX, with only

one additional hit being identified (a hypothetical protein predicted to exist in *P. chabaudi*).

```
PM IX:KFKTHLRKYLLKKIKHQKKQKHSNHKKKKLNKKKNYLIFDSGTSFNSVPKDEIEYFFRVVPSKKCDDSNI
Model:KFKTHL----------------------------------------------------PSKKCDDSNI
```

Figure 4.7: **Structure prediction removes a vital region of *Pf*PM IX.** The deleted region (residues 462-515, red) containing the second *Pf*PM IX catalytic site (highlighted in green) deleted from the PHYRE prediction.

Table 4.3: **Top BLAST results for deleted region of *Pf*PM IX.**

Name/Description	**Accession**	**Query Coverage (%)**	**Identity (%)**	**E value**
Plasmepsin IX (*P. falciparum* 3D7)	XP_001348455.1	100	100	$2e^{-26}$
Hypothetical Protein (*P. chabaudi*)	XP_731757.1	100	61	$8e^{-12}$
Aspartyl Protease (*P. chabaudi* PM IX ortholog)	XP_741066.1	100	61	$2e^{-10}$
Aspartyl Protease (*P. berghei* PM IX ortholog)	XP_680314.1	85	65	$8e^{-10}$
Pepsinogen A (*P. yoelii* 17XNL PM IX ortholog)	XP_72866.1	85	65	$1e^{-09}$
Aspartyl Protease (*P. vivax* SaI-1 PM IX ortholog)	XP_001616672.1	64	77	$4e^{-09}$
Aspartyl Protease (*P. knowlesi* PM IX ortholog)	XP_002260721.1	46	76	$2e^{-05}$

4.4 Discussion and Conclusions:

Although a *Pf*PM IX model was generated using well documented automated model-building approaches and possessed a high degree of structural similarity to several well

characterised aspartic proteases, critical analysis of this model revealed serious omissions in the protein primary sequence. As a result of this flaw the model was not used in further *in silico* studies. The deletion of an amino acid sequence containing the second catalytic aspartic residue, a region key to the function of *Pf*PMs (Banerjee, Liu *et al.*, 2002), would ensure model inaccuracies. The deletion remained initially undetected due to the PHYRE algorithms positioning of a downstream aspartic residue in the active site. The deletion of a low complexity region consisting of 54 amino acid residues, not seen in any other species, may have also given rise to the high degree of structural similarity observed between the active site of the *Pf*PM IX model and other aspartic proteases (Figures 4.5 and 4.6). As this region is unique to PM IX presumably it was removed due to its lack of homology to any other known aspartic protease structures within the structural database. As discussed in Zilversmit et al. (2010), low complexity regions, which are highly flexible and generally exposed to external solvent, are usually located within loop regions and are often difficult to simulate. Low complexity regions are common in *P. falciparum* proteins and are believed to be due to the rapid evolution of this parasite (Zilversmit, Volkman *et al.*, 2010).

During the course of these studies, another group generated *in silico* models of *Pf*PMs I, V-X (Guruprasad, Tanneeru *et al.*, 2011). This study used the FUGUE fold prediction method (Shi, Blundell *et al.*, 2001) and a single template structure for all predictions (1QS8: *Pf*PM IV bound with pepstatin) (Guruprasad, Tanneeru *et al.*, 2011). The study used a highly truncated sequence for *Pf*PM IX, starting just 33 amino acids before the first catalytic ASP residue. It is well known that the digestive vacuole *Pf*PMs have a pro-segment that is lost during processing and begins approximately 30 amino acids before their first catalytic site (Banerjee, Liu *et al.*, 2002) and all successfully expressed recombinant *Pf*DV PMs share a similar truncation (see Chapter 3). In addition to this, research in our laboratory has shown that *Pf*PM IX undergoes multiple processing events throughout the *P. falciparum* life cycle, resulting in a mature protein size of approximately 50 kDa (see Chapter 2). Truncating the N termini would have also aided template matching, as the *Pf*PM sequences become less similar away from the active site (Figure 4.1).

Guruprasad et al. (2011) did not encounter automated deletions when generating a *Pf*PM IX homology model, which may have been due to the purposeful deletion of the *Pf*PM IX insert which had no similarity to the template sequence between residues 205-

206 of their model (Guruprasad, Tanneeru *et al.*, 2011). Comparisons between the sequences of the two generated models show that much of the region deleted by Guruprasad et al. (2011) coincides with the LCR which was deleted during the generation of the model in this study. Despite the author's assurances that this region is far enough away from the second catalytic site to significantly affect function (Guruprasad, Tanneeru *et al.*, 2011), investigation of the studies generated model's sequence shows that the deleted region ends just three residues short of the second catalytic ASP and includes regions of *Pf*PM IX, not identified as an LCR (PlasmoDB). While deletion of problematic sections does aid model generation, a deletion in such proximity to the catalytic site may well affect the accuracy of the model and any downstream applications such as inhibitor identification. The authors also neglected to perform free energy minimisations or molecular dynamic simulations on their structure. These steps relax the model to more accurately resemble the protein environment under physiological conditions (Leach, 2001). In the absence of any known structures similar to this region in *Pf*PM IX, in our view, *in silico* modelling of this type is unlikely to be able to produce an accurate and reliable homology model that can be confidently used for docking and screening purposes.

An alternative approach to structure prediction, *de novo* modelling, was investigated in an attempt to solve this dilemma. *De novo* modelling was performed by Dr Bela Stantic at Griffith University, using the same sequence data. However, this avenue of structure prediction did not result in a viable structure that conformed to known aspartic protease characteristics. The size of *Pf*PM IX and its complex folding behaviour proved too complex for current software applications and despite a number of attempts was not successful. *De novo* modelling has been found to be useful for small proteins (less than 150 amino acid residues) (Skolnick, Kolinski *et al.*, 2001, Latek, Ekonomiuk *et al.*, 2007, Ortiz, Kolinski *et al.*, 1999) and can be more accurate than fold recognition methods (such as PHYRE) in some cases (Skolnick, Kolinski *et al.*, 2001). Unfortunately, as the protein size increases so do *de novo* structural prediction difficulties.

With an increasing number of structures becoming available in the PDB it is possible that a template able to facilitate an accurate model of *Pf*PM IX will become available in the future. Until this is the case, it appears that the only way forward in terms of

understanding the structure of *Pf*PM IX is to generate recombinant enzyme and determine this proteins crystal structure.

Chapter 5: Characterisation of *Plasmodium falciparum* Plasmepsin X

5.1 Introduction:

The HIV-PIs are active at clinically relevant concentrations across a broad range of *P. falciparum* life cycle stages (Andrews, Fairlie *et al.*, 2006, Skinner-Adams, McCarthy *et al.*, 2004, Redmond, Skinner-Adams *et al.*, 2007). While the anti-plasmodial target of the HIV PIs is currently unknown, their activity could point to a previously unexploited target for antimalarial chemotherapeutic development. Data from studies investigating the activity of the HIV PIs ((Andrews, Fairlie *et al.*, 2006, Skinner-Adams, McCarthy *et al.*, 2004, Redmond, Skinner-Adams *et al.*, 2007, Bonilla, Bonilla *et al.*, 2007); discussed in chapters 1 and 2) suggest that *Pf*PM X may be a target of the HIV-PIs and a new antimalarial target.

*Pf*PM X is encoded by the single exon gene PF3D7_0808200 (www.plasmodb.org) and similar to *Pf*PM IX, very little is known about the location and expression of this aspartic protease. PM X orthologs are present and syntenic in all sequenced Plasmodium species including *P. falciparum, P. vivax, P. berghei, P. knowlesi* and *P. yoelii.* As with *Pf*PM IX, *Pf*PM X bears little resemblance to other *Pf*PMs, sharing the highest sequence identity with *Pf*PM IX (Papadopoulos and Agarwala, 2007) which is listed as a possible paralog by PlasmoDB.

Studies that have included *Pf*PM X are scarce, however Banerjee et al. (2002) demonstrated with an antibody against a *Pf*PM X peptide, that *Pf*PM X had a diffuse pattern of localisation and is excluded from the DV (Banerjee, Liu *et al.*, 2002). In contrast, Moura et al. (2009) determined the localisation of a C-terminally tagged *Pf*PM X, under control of the *hrp3* promoter which is maximally active in early ring stages, to be predominantly in the DV and secondarily to the parasitophorous vacuolar space (Moura, Dame *et al.*, 2009). This study also determined that transgenic parasites over-expressing *Pf*PM X were protected from selected aspartic protease inhibitors designed to target the *Pf*DV PMs I and II (Moura, Dame *et al.*, 2009). In another study, Hu et al. (2010) demonstrated that an episomal copy of the first 1 kb of *PfPM X*, which had been C-terminally GFP tagged and was under control of the schizont-stage *ama-1* gene (PF3D7_1133400) promoter, had an apical localisation within merozoites (Hu, Cabrera *et al.*, 2010). While these localization data are conflicting, current transcription data have demonstrated that in intra-erythrocytic stages, *PfPM X* is primarily transcribed in schizont stage parasites (Bozdech, Llinas *et al.*, 2003). Interestingly, *PfPM X* is also transcribed in late-stage gametocytes, ookinetes and sporozoites (www.plasmoDB.org)

and has been found to interact with proteins that localise to the red blood cell (LaCount, Vignali *et al.*, 2005). Peak expression in schizont stage asexual parasites, together with expression in sporozoites (Florens, Washburn *et al.*, 2002) and interaction with proteins within the red cell cytoplasm (LaCount, Vignali *et al.*, 2005) and other genes known to play a role in merozoite invasion (Hu, Cabrera *et al.*, 2010), argues against a role in haemoglobin digestion within the DV and has led to speculations that *Pf*PM X has a vital role in egress and/or invasion (Bozdech, Llinas *et al.*, 2003, Hu, Cabrera *et al.*, 2010). Should *Pf*PM X prove to have a vital role within *P. falciparum*, it is likely to be a potential drug target.

To examine the potential of *Pf*PM X as a new target for antimalarial chemotherapeutic development and further explore the role of *Pf*PM X in asexual intra-erythrocytic stage *P. falciparum* parasites this chapter sought to clarify the location of *Pf*PM X in asexual erythrocytic parasites and investigate the importance of *Pf*PM X to parasite survival.

5.2 Methods:

5.2.1 Parasites

P. falciparum clones D10 and 3D7 were cultured as previously described (Trager and Jensen, 1976). Transgenic parasites containing the vector pHH1-PMXGFPB were generated before the commencement of this project (Seidens, 2010, Ljunggren, 2006) and contained a full length *PfPM X* insert with a C-terminal GFP tag (see Appendix 2 for plasmid maps). Cultures were thawed and maintained as previously described (see section 2.2.1). Transgenic parasites containing the 'knock-out' plasmid pCD-PMXKO ((Seidens, 2010, Ljunggren, 2006); see Appendix 2 for plasmid map) were maintained in complete medium containing WR99210 alone, or WR99210 and 5'fluoro-uracil to encourage plasmid integration (Maier, Braks *et al.*, 2006).

5.2.2 Expression and Localization of *Pf*PM X

*Pf*PM X stage specific expression and localisation was investigated using qRT-PCR, immuno-fluorescence and immuno-blotting.

5.2.2.1 Real time PCR (qRT-PCR)

Total RNA was extracted from parasites and reverse transcription performed as described (section 2.2.2.1). qRT-PCR was performed as section 2.2.2.1 with *PfPM X* specific primers (forward 5'- ttg tat gcc tgc cta cat gc - 3' and reverse 5'- cta ctt tct gtg

cct cgt aca a - 3') which were designed with the Primer3 online tool (Untergasser, Nijveen *et al.*, 2007). Internal reference genes used for stage-specific analysis were *18s rRNA* (Augagneur, Wesolowski *et al.*, 2012) and *Seryl-t transferase* (Roseler, Prieto *et al.*, 2012) (see 2.2.2.1). Each qRT-PCR run was performed in quadruplicate and repeated at least twice. Data were analysed with Rotor-Gene 6.0 software (Corbett Research/Qiagen, Australia). *PfPM X* transcription was calculated relative to ratios of internal reference genes as described (section 2.2.2.1).

5.2.2.2 Antibody Generation

Semi-pro *Pf*PM X inclusion bodies, expressed and purified as Chapter 3, (amino acids 113-573; supplied by our collaborators at the University of Florida) were used to immunise four mice (C57BL/6). Briefly, mice were immunised with an emulsion of *Pf*PM X inclusion bodies (1.25 mg/dose) and adjuvant at 3 weekly intervals (1^{st} dose was emulsified in Freund's Complete Adjuvant, while the 2^{nd}, 3^{rd} and 4^{th} doses were emulsified in Freund's incomplete Adjuvant). 50 µl of the initial injection was given subcutaneously at the base of the tail and the remaining 50 µl given subcutaneously in the peritoneum. The remaining immunizations were given inter-peritoneally. Pre-bleeds and test bleeds were collected at intervals with final bleed-outs with cardiac puncture occurring 10 days post the 4^{th} immunisation.

5.2.2.3 Immuno-blotting

Proteins from saponin-lysed parasite pellet extracts (see section 2.2.2.3) were resolved by reducing 12% SDS-PAGE and transferred to polyvinylidene fluoride (PVDF) membrane using a semidry electroblotter (Trans-Blot, Bio-Rad; see 2.2.2.3). Membranes were probed with mouse anti-sera to *Pf*PM X (1:200) overnight at 4°C in 5% skim milk/PBS. Membranes were washed with 1xPBST (Appendix 1), followed by incubation with a horseradish peroxidase-labelled anti-mouse IgG antibody (1:2000 dilution, Chemicon International Inc.) for 1 hour at room temperature. Membranes were stripped and re-probed with an anti-glyceraldehyde-3-phosphate dehydrogenase (GAPDH) rabbit antibody (1:2000 dilution; (Daubenberger, Tisdale *et al.*, 2003)) followed by a horseradish peroxidase-labelled anti-rabbit IgG antibody (1:2000 dilution, Chemicon International Inc.) to assess protein loading.

5.2.2.4 Immuno-fluorescence

Thin blood smears were fixed on Conclavin A coated slides (0.5 mg/mL) with 4% formalin/0.005% gluteraldehyde and exposed to mouse anti-*Pf*PMX (1:100) or mouse monoclonal antibody to GFP (1:500; Roche) in 3% BSA/PBS for 1 hour at room temperature. Smears were washed 5 times with 1xPBS with vacuum aspiration. Bound antibody was visualized with goat anti-mouse IgG-Cy2 (10 µg/ml; Amersham), goat anti-rabbit IgG-Cy2 (10 µg/ml; Amersham) or goat anti-rabbit IgG-Texas Red (10 µg/ml; Amersham) with Hoechst dye (0.5 µg/ml; Sigma-Aldrich) used for nuclei staining (3% BSA/PBS for 1 hour at room temperature and then washed as above). Parasites were visualised on a Delta Vision Olympus IX71 and images analysed with Corel Paint Shop Pro software.

5.2.3 Examining the role and importance of *Pf*PM X

Previously generated transgenic parasites containing an episomal plasmid encoding C-terminally tagged *Pf*PM X or a *PfPM X* knock-out plasmid were used to investigate the role and importance of *Pf*PM X to asexual stage *P. falciparum* parasites (Seidens, 2010). A previously generated anti-sense vector (Ljunggren, 2006) was also transfected into parasites so that knock-down of *PfPM X* could be examined.

5.2.3.1 Confirmation of transgenic parasites

Parasite DNA was extracted from cultures >5% parasitaemia (see section 2.2.3.1) and the presence of plasmid confirmed by PCR amplification. To confirm the presence of the C-terminally tagged *Pf*PM X plasmid, a *PfPM X*-specific primer (Appendix 3) and a *P. berghei* dihydrofolate reductase 3' UTR (i.e. flanking the inserted gene; Appendix 3) primer were used in PCR amplifications. To confirm the presence the pCD-PMXKO plasmid PCR amplification using *PfPM X* specific and vector-specific *hdhfr* primers (Appendix 3) was performed. PCR cycling conditions were 35 cycles of 30 sec at 94°C, 30 sec at 55°C and 65°C for 2.5 minutes, with the annealing temperature varying slightly depending on the melting temperature of the respective primers. PCR products were examined using agarose (1%) electrophoresis (in TAE containing ethidium bromide; Appendix 1). PCR products were visualised with a UV transilluminator (BioRad).

5.2.3.2 Over-expression of Pf*PM X*

Transgenic parasites containing the full-length C-terminally tagged *Pf*PM X plasmid, pHH1-PMXGFPB, were maintained on 50-200 nM WR99210 to increase plasmid copy number and expression of the *Pf* PM X transgene. Increased *PfPM X* transcription was investigated with qRT PCR (see section 2.2.2.1) using the internal reference genes *18s rRNA* (Durrand, Berry *et al.*, 2004) and *β-actin* (forward 5'-aaa gaa gca agc agg aat cca and reverse 5'-tga tgg tgc aag ggt tgt aa-3') (Augagneur, Wesolowski *et al.*, 2012). Transcript levels were calculated using the ΔΔ-CT-method where the fold change $= 2^{(-\Delta\Delta\text{-CT})}$ (Livak and Schmittgen, 2001).

In preliminary studies, *Pf*PM X expression in transgenic and wild-type cultures was examined using Western blot (see section 2.2.2.3) with mouse antisera to *Pf*PM X (1:200). A rabbit GAPDH antibody was used to assess protein loading ((Daubenberger, Tisdale *et al.*, 2003) section 2.2.2.3). Expression levels of *Pf*PM X relative to GAPDH were compared using Image Studio Lite software (Licor) as section 2.2.3.2.

5.2.3.3 Antisense knock-down of Pf*PM X*

An antisense plasmid containing a full-length *PfPM X* insert in anti-sense orientation with a c-Myc tag, pHHC-HX_AS_cMycB, was generated prior to the commencement of this project ((Ljunggren, 2006) see Appendix 2 for vector map). This was transfected into *P. falciparum* clone 3D7. Briefly, 500 µl packed red blood cells were mixed with 5 ml pre-warmed cytomix (Appendix 1) before centrifugation at 1500 rpm for 3 mins. The supernatant was removed and the treated red blood cells mixed with 30 µl concentrated transfection vector. This was electroporated at 310 v, with 950 µF capacitance and ∞ resistance in a Genepulser Xcell (Biorad). Trophozoites at 10% parasitaemia were mixed with electroporated red blood cells and 10 ml complete media and incubated at 37°C. After 6 h media was changed and after 24 h, cultures were maintained on 5 nM WR99210 drug selection. Parasites resistant to WR99210 were obtained 18-22 days later. To confirm the presence of anti-sense plasmid, PCR was performed on DNA extracted from transfected cultures with vector specific *hdhfr* primers (Appendix 3). PCR amplification with *PfPM X* specific forward and reverse primers (Appendix 3) was used as a DNA control. Transgenic parasites were maintained on 5-50 nM WR99210 drug selection to increase the expression of transfected plasmid and knock-down of *PfPM X* transcription was confirmed with qRT-PCR as 5.2.3.2.

5.2.4 Characterisation of *Pf*PM X

To investigate *Pf*PM X function and activity, an expression plasmid was produced and introduced into *E. coli*. This work was performed with Dr Liam St Pierre, a post-doc within the Malaria Biology Laboratory at QIMR Berghofer. Transgenic C-terminally tagged *Pf*PM X was also extracted from cultures and investigated in activity assays.

5.2.4.1 Recombinant enzyme

5.2.4.1.1 Expression of soluble rPfPM X

A pGEX-6P-1 expression plasmid (GE Healthcare) containing full-length *PfPM X* codon-optimised for expression in *E. coli* (GeneART) was constructed (see Appendix 2 for pGEX-6P-1 vector map) and transformed into *E. coli* ER2566 by Dr Liam St Pierre following manufacturer's instructions (New England Biolabs).

E. coli containing the expression plasmid were grown overnight at 37°C in LB media (Appendix 1) containing ampicillin (50 μg/mL). The following day, pre-warmed 2YT media (Appendix 1) was inoculated with the overnight culture (1:50) and incubated at 37°C with shaking (250 rpm) until the optical density at 600 nm (OD_{600}) reached 0.6 absorbance units (2YT was used as a blank). A 1 mL aliquot was transferred to an Eppendorf tube, centrifuged and the pellet frozen for SDS-PAGE analysis (T0).

Protein expression was induced with 0.5 mM IPTG at 37°C for 2 hours with shaking (250 rpm) before the culture was transferred to 16°C for expression overnight, also with shaking. The next day a sample was taken for SDS-PAGE analysis (TE) and cultures were harvested at 4,500 x g for 10 mins at 4°C. Supernatant was removed and cell pellets were either purified immediately or stored at -20°C. From this point, resuspended cells were kept on ice to prevent proteolysis. Recombinant *Pf*PM X (r*Pf*PM X) was purified by either GST or Ni-NTA purification.

5.2.4.1.2 GST purification

Pellets were resuspended in STE buffer (50 mL; Appendix 1) containing 1 mg/mL egg white lysozyme (Sigma-Aldrich) and 4.5 μl DNAse (20 mg/ml; New England Biolabs) and incubated on ice for 1 hr. Following this, 10% sarcosyl (Sigma-Aldrich) in STE buffer (Appendix 1) was added to a final concentration of 0.7% and the expression solution sonicated at 4°C (Branson Sonifier 250). Samples were then pelleted at 25,000 x g for 25 mins at 4°C. A sample of supernatant and pellet were kept for SDS-PAGE

analysis. Triton X-100 (Sigma-Aldrich) was added to the soluble supernatant fraction to a final concentration of 2% and then mixed with pre-washed GST beads (GE Healthcare) on a rotating wheel at 4°C overnight.

The following day the GST bead suspension was centrifuged at 600 x g for 7 mins at 4°C and the supernatant removed and kept (unbound). The GST beads were then washed with STE buffer with three increasing salt concentration (0-1M) for 45 mins each, followed by two washes (15 min each) with PreScission cleavage buffer (Appendix 1). PreScission protease was then added (10 µL; GE Healthcare; final volume of 2 mL). This was incubated at 4°C for 48 hours on a rotating wheel before beads were centrifuged at 600 x g for 10 mins and the supernatant removed (Eluate 1). Beads were washed 2 more times and the supernatants kept from each wash (Eluates 2 and 3). Eluates were kept at -20°C.

5.2.4.1.3 Ni-NTA purification

Expression pellets were resuspended in 15 mL Ni-NTA Binding buffer (Appendix 1) with egg white lysozyme (1 mg/mL; Sigma-Aldrich) and DNAse (4.5 µl of 20 mg/ml; New England Biolabs). This was incubated on ice for no less than 30 mins before the addition of 10% sarcosyl in STE (0.7% final concentration; Appendix 1) and sonication. The sample was then centrifuged as described above and samples of pellet and supernatant kept for SDS-PAGE analysis. Triton X-100 was added to the supernatant fraction at a final concentration of 1% and the suspension mixed with pre-washed Nickel beads (Qiagen) for overnight incubation on a rotating wheel at 4°C.

The following day the sample was fractionated by centrifugation at 600 x g for 5 mins at 4°C. The supernatant was kept for SDS-PAGE analysis (unbound) and the beads washed in wash buffers 2-4 (Appendix 1) for 45 mins each. Protein was eluted from the beads with Imidizole elution buffer (Appendix 1). Eluates were stored at -20°C.

5.2.4.1.4 Enzyme assays and substrate cleavage.

Enzymatic activity of r*Pf*PM X was assessed using a fluorogenic substrate M-2445 obtained from Bachem, Switzerland. This substrate (sequence Mca-Gly-Lys-Pro-Ile-Leu-Phe-*Phe-Arg-Leu-Lys(Dnp)-D-Arg-NH_2; *represents cleavage site, Mca is the fluorescent moiety and Dnp is the quenching group) is a general aspartic protease substrate which was originally designed for human Cathepsin D and E (Yasuda,

Kageyama *et al.*, 1999). The rate of cleavage of the substrate was measured by the increase in fluorescence (excitation 320 nm, emission 405 nm). The activity of each sample was tested at a range of dilutions in a 96 well plate (Sarstedt) using a FluoStar Optima (BMG Labtech). Samples were diluted with assay buffer (100 mM sodium acetate, pH 4.5 or 100 mM Tris pH 6.8) and water and the fluorescence measured immediately after the addition of fluorogenic substrate (final concentration of 25 μM; final assay volume 200 μL). The gain set was 1059 and fluorescence was measured every 30 seconds for 50 mins at 37°C. Every experiment included a positive control (r*Pf*PM II; produced by our collaborators at the University of Florida) in addition to a negative substrate only control.

5.2.4.2 Pull down of GFP tagged Pf*PM X*

Transgenic parasites containing *Pf*PM X C-terminally tagged with GFP were maintained on 50 nM WR99210. When cultures were >4% parasitaemia cell pellets were collected and frozen at -20°C. *Pf*PM X-GFP was extracted from culture pellets using a GFP-Trap_A kit (Chromotek). Briefly, parasite pellets were thawed and resuspended in ice-cold kit lysis buffer for 30 mins on ice with frequent pipetting before centrifugation at 20,000 x g for 5-10 mins at 4°C. Supernatants were transferred to pre-cooled tubes before being diluted in 500-1000 μL of kit dilution buffer. A 50 μL sample was taken for SDS-PAGE analysis. GFP-Trap beads were equilibrated in dilution buffer and added to the cell lysate. After 2 hrs on a roller at 4°C, the sample was centrifuged at 2,000 x g for 2 mins at 4°C and the supernatant removed. This was kept for SDS-PAGE analysis (unbound). The beads were washed at least 3 times with dilution buffer before being resuspended again in dilution buffer (30 μL beads/50 μL buffer). A 20 μL sample of the bead/buffer suspension was used in enzyme assays with the fluorogenic substrate M-2445. Western blot analysis of samples was performed as previously described (section 5.2.2.3) with mouse anti-*Pf*PMX (1:200).

Enzyme assays were performed as previously described (section 5.2.4.1.4) in a final volume of 200 μL, but with a gain of 1000. Cleavage of the substrate M-2445 was assessed in sodium acetate buffer (100 mM; pH 4.5-7.5).

5.3 Results:

5.3.1 *Pf*PM X expression and localisation

*Pf*PM X is encoded by a 1722 bp, single exon gene (PF3D7_0808200; PlasmoDB) located on chromosome 8. It is predicted to encode a 573 amino acid protein of ~68 kDa (PlasmoDB) with the catalytic aspartic residues at position 266 and 457 in the full length amino acid sequence (Figure 5.1). Orthologs of the gene encoding PM X can be found in all available Plasmodium genome sequences including *P. berghei, P. chabaudi, P. knowlesi, P. vivax* and *P. yoelii* (PlasmoDB).

Confirming global transcription data available in PlasmoDB, qRT-PCR data demonstrate that the vast majority of intra-erythrocytic *PfPM X* transcription occurs in schizont stage parasites (Figure 5.2) with 95% and 92% occurring during these stages, relative to *Seryl-tRNA synthetase* and *18s rRNA* respectively (Figure 5.2 A/B). Less than 5% of the transcription that occurs during intra-erythrocytic asexual development occurs in ring and trophozoite stages (Figure 5.2).

Stage-specific Western blot analysis demonstrated that *Pf*PM X is not expressed in ring stage parasites, but is expressed by trophozoite and schizont stage parasites (Figure 5.3). Data suggest that *Pf*PM X exists in three forms in these life cycle stages, with a ~70 kDa, ~55 kDa and ~42 kDa form being identified in Western blots (Figure 5.3).

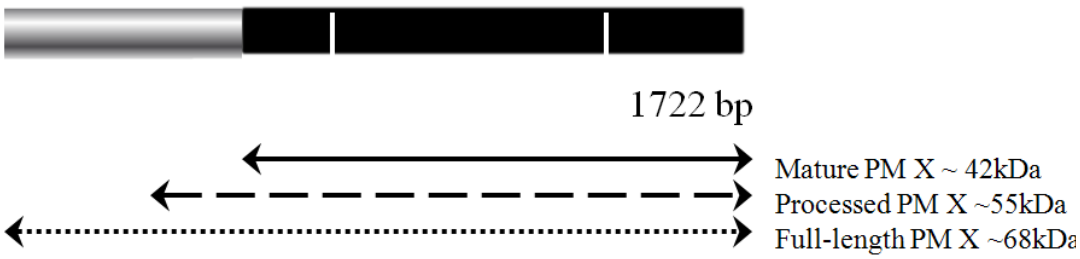


Figure 5.1: **Schematic representation of PF3D7_0808200 and *Pf*PM X.** The predicted pro-domain of *Pf*PM X is depicted in grey. Solid white regions indicate locations of catalytic aspartic amino acids.

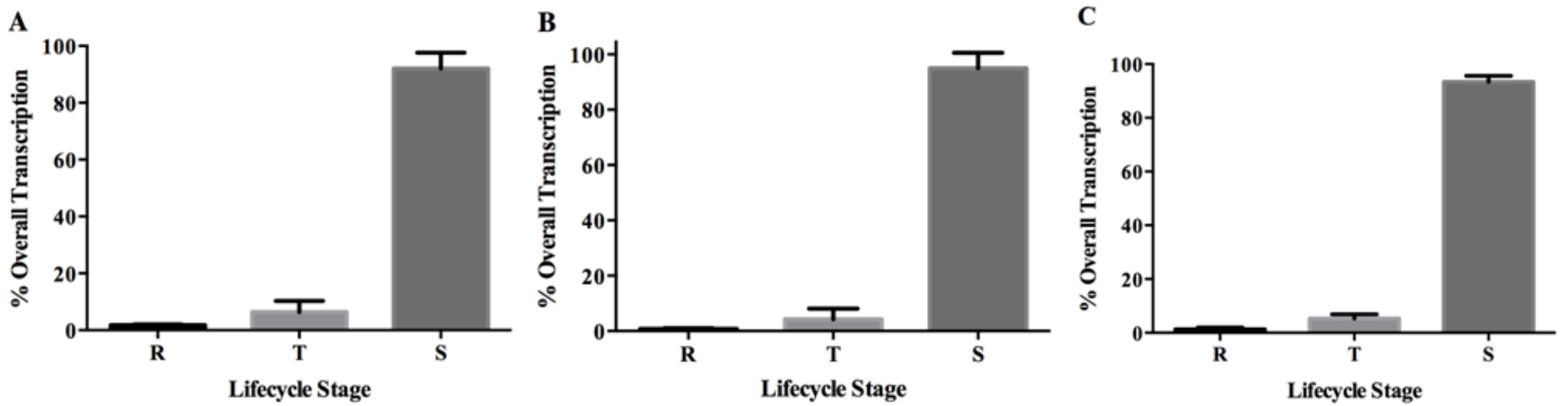


Figure 5.2: ***PfPM X* is primarily transcribed in schizont stage asexual intra-erythrocytic *P. falciparum* parasites.** The stage specific transcription of *PfPM X* was assessed using qRT-PCR. Data are presented relative to internal reference genes A) 1*8s rRNA* and B) *Seryl-t transferase*. Combined data are present in C). All data presented as mean + SD. N = 5.

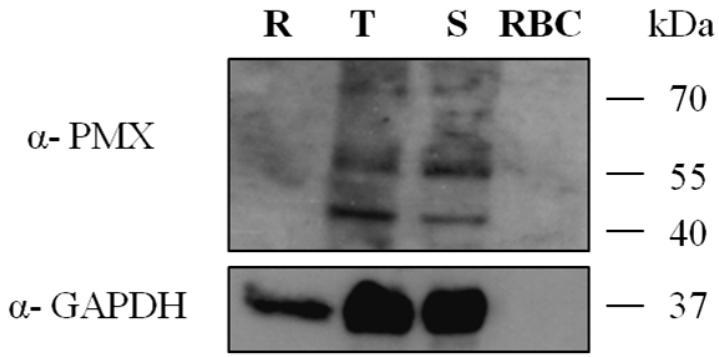


Figure 5.3: **Asexual intra-erythrocytic stage specific expression of *Pf*PM X by *P. falciparum*.** Western blot of saponin-lysed parasite pellet extracts from synchronised D10 parasites (rings (R); trophozoites (T) and schizonts (S)) and uninfected RBC as a control. These were probed with mouse *Pf*PM X anti-sera (top panel) and antibodies to GAPDH (lower panel).

IFA data of both native (Figure 5.4A) and C-terminally tagged *Pf*PM X (Figure 5.4B) demonstrated that *Pf*PM X localises to the parasite cytoplasm in early trophozoite intra-erythrocytic life cycle stages.

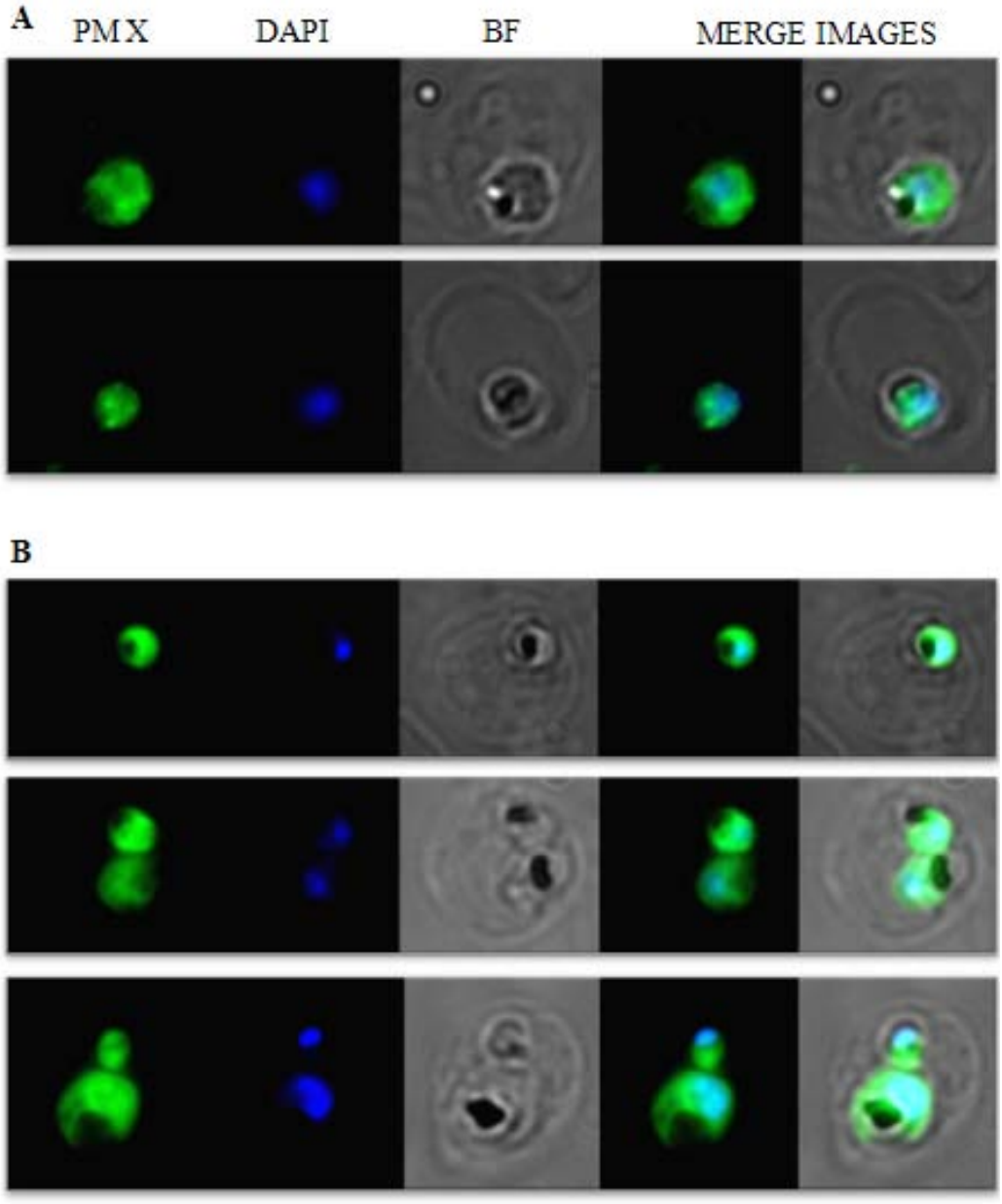


Figure 5.4: ***Pf*PM X is localised to the parasite cytoplasm.** A) The localisation of native *Pf*PM X in D10 parasites examined with mouse *Pf*PM X antisera (BF, Bright Field). B) Transgenically tagged *Pf*PM X detected with mouse GFP antisera.

5.3.2 *Pf*PM X is over-transcribed in transgenic parasites

DNA was extracted from thawed transgenic parasite cultures and PCR amplification was used to confirm the presence of plasmid. Gel analysis of PCR-generated DNA revealed a band of ~2.5 kb for pHH1-PMXGFP (Figure 5.5A), the expected size for

PfPM X with the C-terminal GFP tag and additional vector sequence. The vector specific *hdhfr* gene at the expected size of ~500 bp was also detected (Figure 5.5B). Transgenic parasites were maintained on 50 nM WR99210 drug selection. There was no distinguishable difference in growth and intra-erythrocytic development between transgenic and wild-type parasite cultures (data not shown).

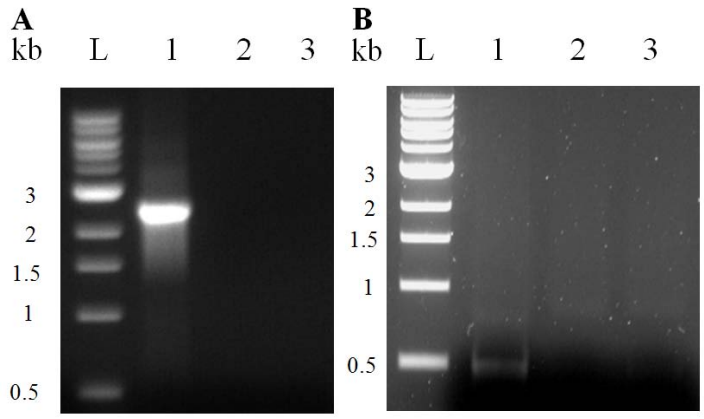


Figure 5.5: **Transgenic cultures contain pHH1-PMXGFP.** PCR analysis of transgenic parasites demonstrates the presence of vector containing the *PfPM X* insert. (A) DNA generated by PCR with *PfPM X* and *PbDT* primers and (B) DNA generated using *hdhfr* primers (B). 1) pHH1-PMXGFP culture DNA, 2) D10 wild-type control DNA, 3) no template control.

Western blot of asynchronous cultures containing C-terminally tagged *Pf*PM X with antisera to GFP, conducted prior to the commencement of this study, confirmed the presence of the transgenic GFP tagged *Pf*PM X (Seidens, 2010).

Transgenic cultures were investigated by others (Seidens, 2010) for altered sensitivity to rtionavir. Data demonstrated that transfected parasites were less sensitive to ritonavir than wild-type parasites (IC_{50} 6.6 ± 0.1 μM vs 4.8 μM; (P=0.008) (Seidens, 2010)). To confirm increased transcription of *PfPM X*, qRT-PCR analysis of schizont-stage cDNA was performed. Data demonstrated that transgenic parasites on 5-50 nM drug selection had increased levels of *PfPM X* transcription compared to wild-type controls (Table 5.1). The highest increase was seen in parasites on 25 nM and 50 nM drug selection (4.184 and 2.717 fold change), relative to *β-actin* and *18s rRNA* respectively (Table 5.1 A/B).

Table 5.1: **Parasites transfected with pHH1-PMXGFP have increased levels of *PfPM X* transcription.** qRT-PCR analysis of *PfPM X* transcription, relative to *β-actin* (A) and *18s rRNA* (B).

(A)

Level of WR99210 drug selection	Δ-CT (average of replicates)	ΔΔ-CT	Fold change (Relative to control)
D10 control	5.78 ± 0.39	0	0
5 nM	3.76 ± 1.1	-2.03	4.08
25 nM	3.72 ± 0.52	-2.06	4.18
50 nM	4.18 ± 0.30	-1.60	3.05

(B)

Level of WR99210 drug selection	Δ-CT (average of replicates)	ΔΔ-CT	Fold change (Relative to control)
D10 control	12.97 ± 0.14	0	0
5 nM	11.83 ± 2.14	-1.13	2.19
25 nM	12.52 ± 1.23	-0.44	1.36
50 nM	11.53 ± 0.56	-1.44	2.72

In preliminary studies, *Pf*PM X expression levels in transgenic cultures containing the plasmid pHH1-PMXGFP were examined and compared to wild-type controls. Using Image Studio Lite software (Licor), ratios of *Pf*PM X relative to GAPDH were determined. Transgenic cultures appeared to have decreased levels of *Pf*PM X expression, compared to the D10 comparison.

5.3.3 Is *Pf*PM X essential to *P. falciparum* growth in red blood cells?

5.3.3.1 Genetic disruption of PfPM X

P. falciparum clone D10 parasites were transfected with pCD-PMXDKO and selected on 5 nM WR99210 prior to the commencement of this project (Seidens, 2010). To encourage integration of this plasmid and targeted *PfPM X* gene disruption, these parasites were retrieved and either cycled on and off 5 nM WR99210 (Fidock and Wellems, 1997) or exposed to 5 nM WR99210 with 1 μM 5' fluoro-uracil (Maier, Braks *et al.*, 2006). PCR analysis of DNA extracted from cultures demonstrated presence of endogenous *PfPM X* and vector (Figure 5.6A and 5.6B). Further analysis suggested

there may have been 3' integration of the knock-out plasmid at some point; however this was not detectable in cultures after further drug cycling (Figure 5.6D).

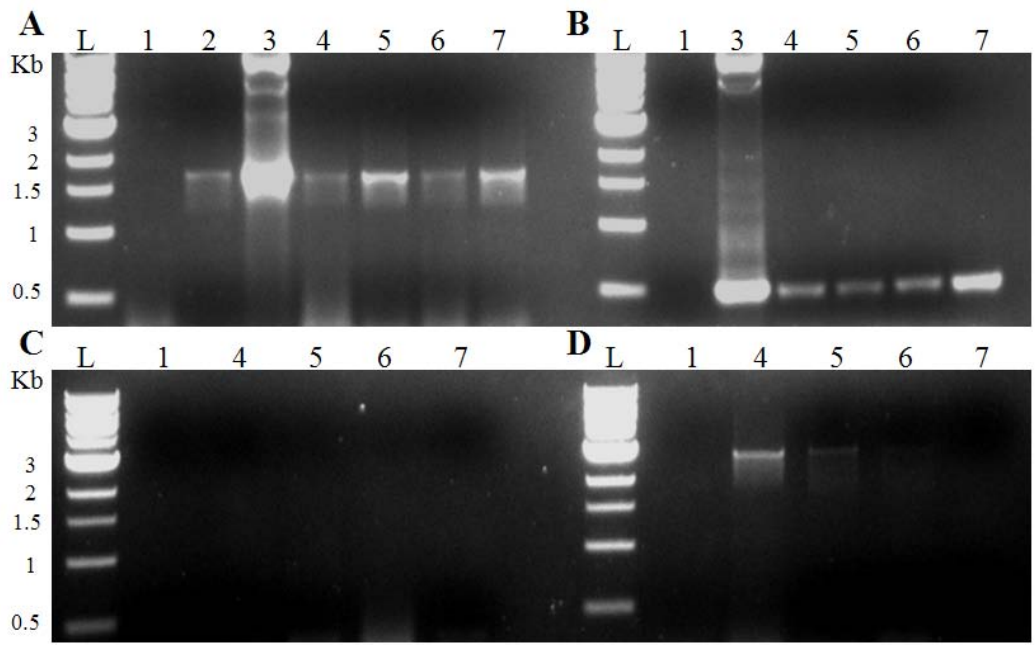


Figure 5.6: **PCR analysis of transgenic parasites containing pCD-PMXDKO.** 1) H_2O negative control. 2) D10 wild-type control. 3) Vector control (pHH1-PMXGFP). 4) PM X 'KO' cultures on WR99210 selection. 5) PM X 'KO' cultures on WR99210 and 5' fluoro-uracil. 6) PM X 'KO' cultures on 2nd cycle of selection. 7) PM X 'KO' cultures on 4th cycle of double drug selection. A) Endogenous *PfPM X* primers to confirm DNA B) *hdhfr* primers to confirm vector presence, C) 5' integration check with *PfPM X* forward and *hdhfr* forward primers and D) 3' integration check with *PfPM X* and *hdhfr* reverse primers.

5.3.3.2 Knock-down of PfPM X transcription

P. falciparum clone 3D7 was transfected with an antisense vector, pHHC-HX-_AS_cMycB, generated prior to the commencement of this project and selected on 5 nM WR99210. Transgenic cultures were maintained on 5 nM and 50 nM WR99210 to increase any knock-down effect. To confirm the presence of vector, DNA was extracted and PCR amplification performed with *PfPM X* full-length primers and *hdhfr* primers. Gel analysis of the PCR-generated DNA revealed the presence of *PfPM X*, confirming DNA, as well as the vector specific *hdhfr* (Figure 5.9). While the intra-erythrocytic development and morphology of transgenic cultures resembled those of wild-type parasites, transgenic parasites appeared to have a slower growth rate (data not shown).

To investigate *PfPM X* knock-down, qRT-PCR was performed on schizont-stage cDNA. Data demonstrated that transgenic parasites had decreased levels of *PfPM X* transcription relative to *β-actin* and *18s rRNA*, with those on 50 nM drug selection having less *PfPM X* than controls (Tables 5.2 A/B).

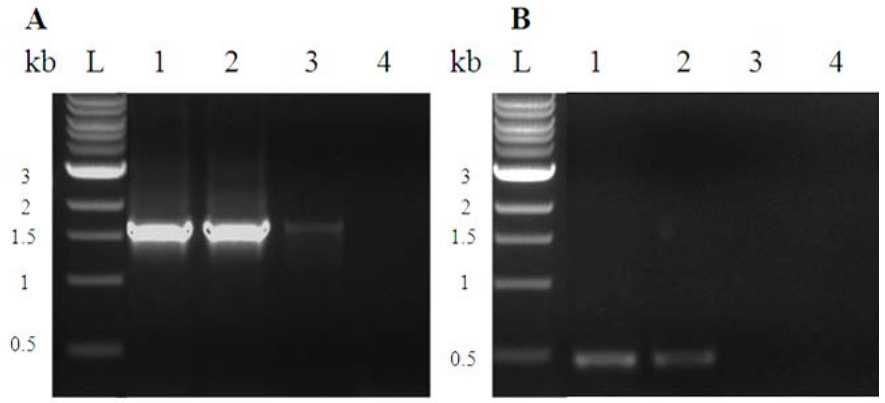


Figure 5.7: **Parasites transfected with pHHC-HX_AS_cMycB contain vector.** To confirm the presence of DNA, samples were assessed for *PfPM X* (A) and to confirm the presence of vector in transgenic parasites, samples were assessed for *hdhfr* (B). Transgenic culture on 5 nM WR99210 drug selection (1) and 50 nM WR99210 (2) 3D7 untransfected control (3) and (4) a no template control.

5.3.4 *Pf*PM X characterisation

5.3.5.1 Recombinant enzyme

A ~100 kDa recombinant full-length *Pf*PM X, with GST and HIS tags, was expressed in *E. coli*. Subsequent purification steps and cleavage of GST and HIS tags resulted in a ~68 kDa protein. Samples from expression and purification steps were analysed by SDS-PAGE (Figure 5.8). Enzymatic activity of r*Pf*PM X was assessed using the fluorogenic substrate M-2445 at pH 4.5 and 6.8. No activity was detected (Figure 5.9).

Table 5.2: **Transgenic parasites containing the *Pf*PM X antisense plasmid demonstrate reduced *PfPM X* transcription.** qRT-PCR analysis of *PfPM X* transcription, relative to *β-Actin* (A) and *18s rRNA* (B), suggests *PfPM X* transcription is reduced in transfected cultures. Each was performed three times.

(A)

Level of WR99210 drug selection	Δ-CT (average)	ΔΔ-CT	Fold change (Relative to control)
3D7 Control	6.15 ± 0.82	0	0
5 nM	6.40 ± 0.	0.25	0.84
50 nM	6.60 ±	0.45	0.73

(B)

Level of WR99210 drug selection	Δ-CT (average)	ΔΔ-CT	Fold change (Relative to control)
3D7 Control	8.58 ± 0.80	0	0
5 nM	10.60 ± 1.35	2.02	0.25
50 nM	11.25 ± 0.74	2.67	0.16

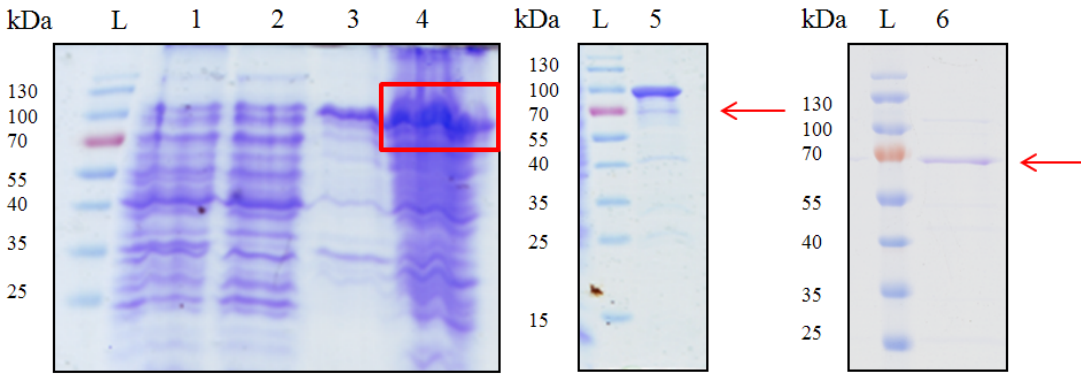


Figure 5.8: **Expression and purification of r*Pf*PM X.** Samples were taken throughout the expression and purification process and analysed by SDS-PAGE. (1) Un-induced culture sample, insoluble fraction. (2) Soluble fraction from un-induced culture sample. r*Pf*PM X is present in both the insoluble (3) and soluble (4) samples of induced cultures and is highlighted by a red box in the soluble fraction. r*Pf*PM X containing both GST and HIS tags post-purification (5). r*Pf*PM X post-cleavage (6) as performed by Dr Liam St Pierre. r*Pf*PM X is indicated by red arrows.

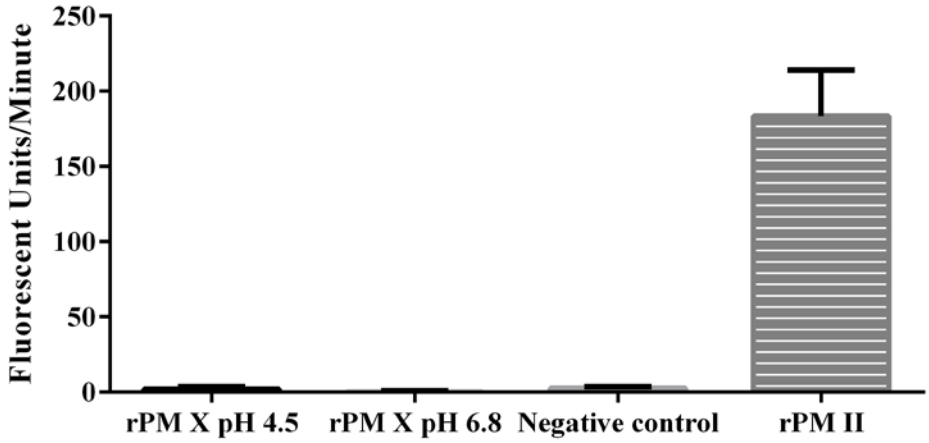


Figure 5.9: **r*Pf*PM X does not cleave the fluorogenic substrate M-2455.** The enzymatic activity of r*Pf*PM X was assessed at pH 4.5 and 6.8 and compared to a negative and positive (r*Pf*PM II) controls. Assay data presented as mean + SD of at least 2 experiments.

5.3.5.2 Pull down transgenic enzyme

To further investigate *Pf*PM X and determine if enzyme activity could be detected, C-terminally tagged enzyme was extracted from cultures using a GFP-Trap_A kit (Chromotek). Extracted enzyme was analysed by SDS-PAGE (Figure 5.10) and *Pf*PM X was detected by Western blot at ~70, 55 and 40 kDa in purified samples (Figure 5.11). *Pf*PM X was detected in unbound samples at ~40, 55 and 60 kDa (Figure 5.11). The activity of pulled down enzyme was tested against the fluorogenic substrate M-2445 across a pH range (4.5-7.5) and greatest activity was detected at pH 4.5 (Figure 5.12). Enzyme stored overnight at 4°C had significantly reduced activity and could not be used (data not shown).

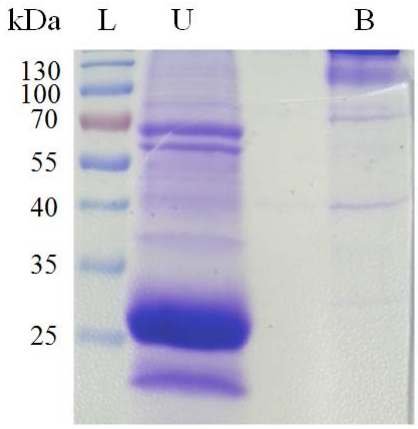


Figure 5.10: **SDS-PAGE analysis of GFP tagged *Pf*PM X extracted from transgenic cultures.** Pull down material that did not bind to the GFP beads (Unbound; U) and purified beads (B) analysed by SDS-PAGE.

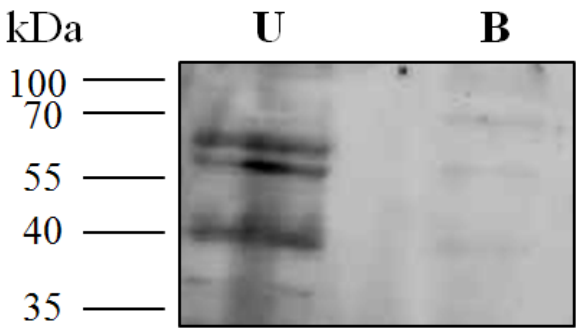


Figure 5.11: **Western blot analysis of pull down material.** Western blot of both the unbound (U) and purified sample (B) from a *Pf*PM X-GFP extraction. Membranes were probed with mouse anti-sera to *Pf*PM X.

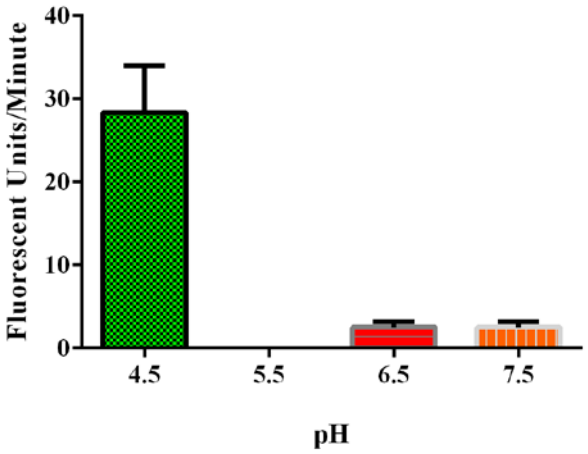


Figure 5.12: **GFP tagged *Pf*PM X is active.** Enzyme activity was measured by cleavage of the fluorescent substrate from pH 4.5-7.5. Data is represented as mean + SD of at least two experiments.

5.4 Discussion and Conclusion:

There is an urgent need for new and unique antimalarial drug targets in order to combat spreading drug resistance. Current data suggest that *Pf*PM X may represent a new drug target for the development of new antimalarial agents; however information on this aspartic protease is scarce. In this study I have demonstrated that *Pf*PM X does represent a potential new drug target. While it is expressed in both trophozoite and

schizont stages (Figure 5.3), the vast majority of *PfPM X* transcription occurs during the late schizont stage of intra-erythrocytic *P. falciparum* development (Figure 5.2), confirming what has been found previously (PlasmoDB, (Bozdech, Llinas *et al.*, 2003, Hu, Cabrera *et al.*, 2010)). Immunochemistry data suggests that similar to the *Pf*DV PMs (Dame, Yowell *et al.*, 2003, Banerjee, Liu *et al.*, 2002, Kim, Lee *et al.*, 2006) and *Pf*PM IX (Chapter 2) *Pf*PM X is expressed as a zymogen (~70 kDa) and undergoes processing to form a ~55 kDa protein, followed by the mature enzyme of ~42 kDa (Figure 5.3). This mature size is only slightly larger than the *Pf*DV PMs (~37 kDa (Banerjee, Liu *et al.*, 2002)) and suggests that despite the much larger N-terminal pro-region, *Pf*PM X is processed in a similar position ~30 amino acids before the first catalytic aspartic residue. The additional form of *Pf*PM X, ~55 kDa, suggests that multiple processing events are likely to occur. Antibodies specific to different regions of *Pf*PM X would allow the investigation of the processing of *Pf*PM X in more detail, as was done with TgASP1 which undergoes multiple processing events at both the N and C-terminal end (Shea, Jakle *et al.*, 2007). As discussed in section 2.4, while this antisera is believed to be *Pf*PM X specific, cross-reactivity with other *Pf*PMs is possible. Cross-reactivity may have led to artefacts in the current Western blot experiments. To further investigate this possibility cross-reactivity to other *Pf*PMs should be assessed, essentially as previously assessed by Banerjee et al. (2002) for plasmepsins I-IV. Proteomic analysis of antisera binding with immuno-precipitation experiments should also be performed, as previously discussed in section 2.4.

IFA data indicate that *Pf*PM X localises to the parasite cytoplasm of ring and trophozoite stage parasites (Figure 5.4) and is not exported into the host red cell. While confirmatory co-localisations with *Pf*DV PMs or *Pf*PM V need to be conducted, *Pf*PM X does not appear to localise to a specific organelle within the parasite cytoplasm. This localisation is in contrast to that found by a previous study with similar life cycle stages that used a C-terminal red fluorescent protein (RFP) tag that suggested a parasite vacuolar space and DV localisation (Moura, Dame *et al.*, 2009). While the use of fluorescent tags is useful for investigating protein localisation, it is possible that the addition of a tag of such size may have altered the protein's native trafficking and localisation. Confirmation of localisation with protein specific antibodies, as performed in this chapter, is important to ensure the transgenic tag has not inadvertently affected the target protein's localisation. While the reasons behind the difference seen in localisation of GFP and RFP tagged *Pf*PM X remains unknown, the comparable

localisations seen in this chapter between IFAs using antibodies to *Pf*PM X and the C-terminal GFP tag is encouraging and suggests the GFP tag has not altered protein trafficking. Additionally, the diffuse *Pf*PM X localisation seen in this work was similar to the observations made by Banerjee et al (2002) who suggested that *Pf*PM X was excluded from the DV (Banerjee, Liu *et al.*, 2002). Further investigation of *Pf*PM X localisation in schizont-stage parasites is required and was attempted during the current study. However due to a number of technical issues and time constraints, this could not be completed. Future studies examining the localisation of *Pf*PM X in schizont-stage parasites may demonstrate similar results to Hu et al. (2010) who showed an apical localisation in merozoites (Hu, Cabrera *et al.*, 2010).

Experiments to genetically disrupt *PfPM X* and investigate the importance of this enzyme to parasite survival, initially suggested that the knock-out plasmid may have integrated into the endogenous gene. PCR analysis suggested that in early drug selection cycles a single homologous recombinant event may have occurred (Figure 5.6). However, as this was not seen in cultures that underwent further cycling, it may have been a lethal event. Given the impracticality of confirming gene essentiality through knock-out techniques, with parasites unable to survive when an essential gene is disrupted, experiments to produce conditional knock-outs were warranted. While antisense knock-down was instigated (discussed below), additional conditional knock-down systems such as the Shld technique (Armstrong and Goldberg, 2007) should be implemented and would permit additional investigation into the role and function or *Pf*PM X in parasites. In this system, proteins of interest are fused with a domain that promotes protein degradation in the absence of the compound Shld1 (Armstrong and Goldberg, 2007). This allows investigation of the effect of protein knock-down on cultures, with the ability to rescue the wild-type phenotype. In this way, the importance of *Pf*PM X to parasite survival, and therefore suitability of this enzyme as a potential drug target, could be further investigated without the complications of disrupting a potentially essential gene.

To further investigate the role of *Pf*PM X in a conditional fashion, antisense technology was used to reduce *PfPM X* transcription. Transfected cultures transcribed 15% less *PfPM X* than untransfected controls (Table 5.2). While investigation of *Pf*PM X protein levels needs to be performed, no significant morphological changes were observed in in these cultures. However, transgenic cultures appeared to grow slower than wild-type

D10 parasites. A reduced growth rate, once quantified, would further support the hypothesis that *Pf*PM X is important to parasite survival. Quantification of growth together with determining the effect of knock-down on egress/invasion will provide additional information of *Pf*PM X's role in Plasmodium parasites. Additional knock-down experiments with techniques including PNAs (Kolevzon, Nasereddin *et al.*, 2014), PMOs (Augagneur, Wesolowski *et al.*, 2012) and/or Shld (Armstrong and Goldberg, 2007) (discussed in 2.4) would further elucidate the effect of knocking-down *PfPM X* and provide valuable clues as to the function of *Pf*PM X in malaria parasites.

Data demonstrating that transgenic parasites containing a plasmid designed to increase *Pf*PM X expression, are significantly less sensitive to the HIV PI ritonavir (Seidens, 2010), supports the hypothesis that *Pf*PM X is a potential new drug target. While additional targets exist, such as *Pf*PM IX (Chapter 2), these data suggested that *Pf*PM X inhibition is at least partially responsible for the activity of ritonavir in Plasmodium. Investigation of *PfPM X* transcription in these cultures demonstrated that cultures under plasmid selection pressure (50 nM WR99210) had a ~3 and 2.7 fold increased transcription of *PfPM X*, relative to internal gene controls (Table 5.1). Although densitometry suggested that *Pf*PM X expression was reduced in transgenic cultures these data were derived from only one biological replicate and densitometry of ECL membranes have been shown to be unreliable (Chapter 2; (Gassmann, Grenacher *et al.*, 2009)). Further investigation with an alternative method is required to confirm these results, however, there is currently no method to accurately quantitate increased or decreased levels of *Pf*PM X in parasites as current substrates are not specific for this enzyme.

While the function of *Pf*PM X within asexual intra-erythrocytic parasites remains unknown at this time, the transcription and translation profile of *Pf*PM X suggests a role in egress and/or invasion. The identification of *Pf*PM X in sporozoites (Florens, Washburn *et al.*, 2002) and its interaction with proteins that localise to the red cell (LaCount, Vignali *et al.*, 2005) and have a known role in invasion (Hu, Cabrera *et al.*, 2010), support this hypothesis. Experiments to confirm the presence of *Pf*PM X in *P. falciparum* merozoite stages would provide additional support to this hypothesis, however could not be completed in this study due to time constraints. As with *Pf*PM IX, further investigations into interacting partners with co-immunoprecipitation experiments may provide more of an insight into the function of *Pf*PM X. Any additional interacting

partners with a known role in egress/invasion would further support this view. The parasite cytoplasm localisation of *Pf*PM X does suggest that this enzyme has function/s distinct from the well-characterised *Pf*DV PMs and *Pf*PM IX and as a result may present a previously unexploited and important parasite function for the development of antimalarial agents.

Although the current project was successful in generating r*Pf*PM X, it was not active (Figure 5.9). Inactivity may have been associated with purification difficulties and handling. While GST-r*Pf*PM X-HIS could be readily purified from expression cultures, removal of the GST tag proved difficult and appeared to impact auto-catalysis. Exposing the tagged enzyme to low pH in attempts to induce auto-catalysis and removal of the pro-region, with the GST tag, proved unsuccessful (data not shown). Unfortunately, as a result of poor yields, further investigations into refolding of the produced protein into an active form could not be performed (Chapter 3). Construction of an expression plasmid without the GST tag was initiated, however, due to technical difficulties and time constraints, was not completed. Additional strategies that may have been useful include the use of expression plasmids successfully used to express recombinant forms of the *Pf*DV PMs (Wyatt and Berry, 2002, Moon, Tyas *et al.*, 1997, Xiao, Tanaka *et al.*, 2007, Beyer, Goldfarb *et al.*, 2004). These plasmids do not contain the GST tag which was problematic in this study. Truncating the *Pf*PM X insert on the N-termini to approximately 30 amino acids before the first catalytic aspartic residue, and subsequent removal of the inhibitory pro-region, may also result in active enzyme.

Enzymatically active C-terminally tagged *Pf*PM X was successfully extracted from transgenic *P. falciparum* cultures (Figure 5.12). SDS-PAGE and Western blot analysis demonstrated that while *Pf*PM X-GFP was extracted from cultures, additional enzyme remained in unbound samples (Figures 5.10 and 5.11). To increase yields of extracted enzyme, larger volumes of GFP beads could be used to bind more *Pf*PM X-GFP enzyme. Alternatively, the binding incubation time of 2 hours could be increased to ensure saturation of the GFP beads. Extracted enzyme demonstrated activity was optimal at pH 4.5, with only small levels of activity seen at higher pH (Figure 5.12). Interestingly, this is lower than the pH of active *Pf*PM V, pH 5-7 (Boddey, Hodder *et al.*, 2010, Russo, Babbitt *et al.*, 2010), and more in line with the optimal pH of r*Pf*PMs I and II (Xiao, Tanaka *et al.*, 2007, Banerjee, Liu *et al.*, 2002). Future studies with this enzyme should investigate a more comprehensive pH range. A set-back in using pulled

down *Pf*PM X enzyme was the instability of the protein, meaning it needed to be used immediately after extraction. Samples stored overnight at 4°C demonstrated significantly reduced activity (data not shown). While GFP-tagged *Pf*PM X was suitable for small scale activity assays, such as screening a small number of *in silico* docking hits (Chapter 6), large quantities of parasite culture and pull down reagents would be required for large scale testing of inhibitors and further characterization studies. An active recombinant enzyme would be more cost effective and efficient for screening purposes.

This study has demonstrated that *Pf*PM X is a unique enzyme that could be a potential new drug target for antimalarial drug development. The results suggesting a separate/additional function to the well-characterised *Pf*DV PMs and that it is a target of the HIV PIs warrant further investigation. Should further experimentation demonstrate *Pf*PM X is essential to parasite survival, much-needed drugs targeting this enzyme may prove to be potent and selective antimalarial agents.

Chapter 6: Generation of a *Pf*PM X *in silico* model and the identification of new inhibitors

6.1 Introduction:

6.1.1 *In silico* screening and inhibitor identification:

Lead inhibitor identification via high-throughput *in silico* docking screens can be more efficient and cost effective than laboratory-based *in vitro* methods (Ripka, Satyshur *et al.*, 2001, Degliesposti, Kasam *et al.*, 2009). Despite high computational demands, this process allows millions of compounds to be screened rapidly, dramatically reducing the time, equipment and finances required for laboratory based high-throughput screening (Bleicher, Bohm *et al.*, 2003). *In silico* drug discovery often involves using a protein model generated by X-ray crystallography, NMR or *in silico* protein modelling (Degliesposti, Kasam *et al.*, 2009, Ripka, Satyshur *et al.*, 2001, Bouillon, Giganti *et al.*, 2013) and has been used successfully to identify potent and selective Plasmodium inhibitors including those targeted to *P. falciparum* PM II (Degliesposti, Kasam *et al.*, 2009) and *P. vivax* subtilisin-like serine protease (SUB1) (Bouillon, Giganti *et al.*, 2013).

With continuing advances in software design, there are now a number of software options available for lead inhibitor identification. While not as computationally efficient as computational world-wide grid infrastructures such as the World-wide *In Silico* Docking On Malaria II (WISDOM II), which allow the screening of millions of compounds against multiple drug targets in significantly less computational time (Kasam, Salzemann *et al.*, 2009), software packages such as those provided by Schrödinger Inc. provide a platform for users to go from protein structure prediction to lead inhibitor identification. With the increasing availability of small molecule chemical compound libraries freely available to download and with this software, researchers can perform virtual screens for specific inhibitors before obtaining compounds for confirmation in biological assays.

In addition to performing *in silico* screens for inhibitors using a target protein structure, potential inhibitors can be identified using pharmacophore modelling and screening. This approach involves identifying the spatial arrangement and physicochemical attributes of active ligands that facilitate target binding and using this information to predict the binding of additional compounds (Leach, 2001). This can be done without knowledge of the protein structure, instead examining common features of ligands that actively bind to the protein (Leach, 2001). In essence, once the pharmacophore model has been generated, chemical compound libraries can be screened for ligands with the

same spatial arrangement of chemical features as known inhibitors. While this method does not take into account all of the features of potential ligands, it does allow for more rapid screening of chemical libraries and has been successful in identifying inhibitors (Leach, 2001, Dixon, Smondyrev *et al.*, 2006b, Dixon, Smondyrev *et al.*, 2006a).

6.1.2 PM X and its orthologs:

Like PM IX, PM X can be found in the genomes of all Plasmodium species including *P. falciparum, P. knowlesi, P. vivax* and *P. yoelii.* The sequence of *P. falciparum* PM X also bears little resemblance to other known *Pf*PMs (Table 6.1) sharing 24-26% sequence identity with the well characterised digestive vacuole PMs (*Pf*PM I-IV). *Pf*PM X demonstrates the highest homology with *Pf*PM IX (37%) with PlasmoDB listing *Pf*PM IX as a possible paralog. The similarities and differences between the *Pf* PMs are discussed *vide supra* (Chapter 4 (4.1.2)) and the phylogenetic groupings can be seen in Appendix 8.

Table 6.1: **Comparison of *Pf*PM X protein coding sequence to that of the remaining *Pf*PMs (full length transcripts).** Plasmespin sequences were paired and aligned using Clustal Omega (Sievers, Wilm *et al.*, 2011, Kasam, Salzemann *et al.*, 2009) and ClustalX 2.1 software (Larkin, Blackshields *et al.*, 2007, Degliesposti, Kasam *et al.*, 2009, Kasam, Salzemann *et al.*, 2009) and level of identity determined.

*Pf*PM	Sequence identity to *Pf*PM X
I	25%
II	24%
III (HAP)	24%
IV	26%
V	21%
VI	26%
VII	25%
VIII	27%
IX	37%

An alignment of six PM X orthologs (Figure 6.1) demonstrates a high level of conservation between species, particularly at the N terminal signal peptide (underlined and in bold, Figure 6.1) present in all orthologs and the regions surrounding the catalytic site (red). However, a marked difference between the orthologs can be seen within the LCRs (purple). While LCRs are in similar locations, their number varies between species. In addition *P. knowlesi* has a 4th LCR in PM X that stands alone at the

C-terminal end. Interestingly, there are more LCRs in human pathogens compared to those species that infect mice, which was also seen with *Pf*PM IX and its orthologs (Chapter 4). The decreased number and size of LCRs in *Pf*PM X compared to *Pf*PM IX, as well as the absence of any in close proximity to the catalytic site, is encouraging and suggests finding a suitable solid state template for structure prediction may be successful.

This chapter describes the generation of a *Pf*PM X *in silico* homology model, which in the absence of an X-ray crystal structure, can potentially be used to better understand the tertiary structure of this potential drug target. The structure generated in this study was used to identify potential inhibitors of *Pf*PM X through virtual screening and docking studies with compound libraries and to investigate the binding of HIV PIs to *Pf*PM X. Lead compounds were evaluated *in vitro* and in enzyme activity assays for antimalarial and anti-*Pf*PM X activity.

```
P. falciparum 1    MKRISPLNTLFYLSLFFSYTFKGLKCTRIYKIGTKALPCSECHDVFDCTGCLFEEKES-SHVIPLKLNKKNP--NDHKKL  77
P. vivax      1    MKHMGGFRTLFCGALFLLQFWGEATCHKVYKVGSEPIPCSQCKDVRECSACLFEEQDS-PHAIHLKLNKKKP--SDHSNL  77
P. knowlesi   1    MKDTKVLRTLLCGALFLLHLCQDARCHKVYKVGPEPIPCSQCKDLRECSACLFEEEES-PYAIHLKLNKNKP--NDHSNL  77
P. bergei     1    MKSIKILPVFYLVT-FFLHNYNEIKCNKVYISGDNFMPCSECKDINKCGGCLFDNNEFLPSAIELKLTKKNRDNNNHDNL  79
P. chabaudi   1    MKSVKIVPIFYLVA-FFFHNYNEVTCGKVYTTGDNFMPCSECKDINKCGGCLFEHNESLPSAIELKLTKKN-DNSIHDNL  78
P. yoellii    1    MKSIKILPVFYLVA-FFLHNYNEITCNKLYITGDNFMPCSECKDINKCGGCLFDNNEALPSAIELKLTQKNRDNNNHDNL  79

P. falciparum 78   QKHHESLKLGDVKYYVNRGEGISGSLGTSSGNTLDDMDLINEEINKK--RTNAQLDEKNFLDFTTYNKNKAQDISDH-LS  154
P. vivax      78   KKHHDSLKLGGVKYYVKRGEGISGSLGNPSGHTLDDIDSINEEIRIR--REESAGVGRNGAEVTAYGGS---GPSDR-FA  151
P. knowlesi   78   KKHHDSLKLGGVKYYVKRGEGISGSLGNPLGNTMDDIDSINEEIQNR--RKESAGGGRNFIEMSNYKKD---SLSDY-FS  151
P. bergei     80   KIHHDSLKLGNVKYYVKRGEGVSGSFGNASGNDINSMAKIHNEIKNRTENKEENKSSLSFIDHINNNLGE--GKGGYNFS  157
P. chabaudi   79   KIHHDSLKLGNVKYYVNRGEGVSGSLGHASGNDIRNMAEIHNEIKNRTEKKEKNKSSLSFIDR---NNPE-EGKGGNNLS  154
P. yoellii    80   KIHHDSLKLGNVKYYVKRGEGVSGSFGNVSGNDINSMAEIHNEIKNRKEKKEENKSSLSFIDHSNNNNSE-EGKGDYNFS  158

P. falciparum 155  DIQKHVYEQDAQ--KGNKNFTNNENNSDNEN-NSDNENNSDNENNLDNENNLDNENNSDNSSIEKNFIALENKNATVEQT  231
P. vivax      152  GVQRHAH---AGGVSGDMASGEQRGSDEIAK-GE------------------------------SFIDVKNAHAVVEQT  196
P. knowlesi   152  GVQKHAH---SE--VGRVNMGDEKGRRENEG-GEGGEGHASEHAD---------------RVERNFIDLKNTNAVVEQT  209
P. bergei     158  RIQKH----EQDG--DKINAQEEFEKIKSQVVDKSVVSFSDRVLD-------ENEGQTQ-----------PSKGVTVEET  213
P. chabaudi   155  KIQKH----EQDE--DKVNTQEKFEKVKSQT-DKDASSSSGSAID-------DNANKPG-----------PSKGVTIEET  209
P. yoellii    159  KIQKH----EQDG--DKINTQEEFEKIQSQAVNKSGVSFSDRVLD-------ENGNQAQ-----------PSKGVTIEET  214

P. falciparum 232  KENIFLVPLKHLRDSQFVGELLVGTPPQTVYPIFDTGSTNVWVVTTACEEESCKKVRRYDPNKSKTFRRSFIEKNLHIVF  311
P. vivax      197  EENVFLIPLKHLRDSQFVGTLLVGVPPQEIHPIFDTGSTNLWVVTTDCEEESCKKVKRYNPYKSKTFRRSFIGKNLHIVF  276
P. knowlesi   210  EENVFLIPLKHLRDSQFVGKLLVGVPPQEIHPIFDTGSTNLWVVTTDCEEKSCKKVQRYNPYKSKTFRRSFIGKNLHIVF  289
P. bergei     214  SDNVFLVPLQHLRDSQFVGKLLVGTPPQEIHPIFDTGSTNLWVVTTECKEDSCKKVHQYNPNKSKTFRRSFIKQNLHIVF  293
P. chabaudi   210  SDNVFLVPLQHLRDSQFVGKLLVGTPPQEIHPIFDTGSTNLWVVTTECKEESCRKVHQYNPNKSKTFRRSFVNQNLHIVF  289
P. yoellii    215  SDNVFLVPLQHLRDSQFVGKLLVGTPPQEIHPIFDTGSTNLWVVTTECKEDSCKKVHQYNPNKSKTFRRSFIKQNLHIVF  294

P. falciparum 312  GSGSISGSVGTDTFMLGKHLVRNQTFGLVESESNNNKNGGDNIFDYISFEGIVGLGFPGMLSAGNIPFFDNLLKQNPNVD  391
P. vivax      277  GSGSISGSIGKETFVLGDHTVRNQTFGLVESESNDSLNG-DNIFDYIDFEGIVGLGFPEMLSAGKVSFFDNLLSQNKNLS  355
P. knowlesi   290  GSGSISGSIGKETFVLGNHTVRNQTFGLVESESNDSLNG-DNIFDYIDFEGIVGLGFPEMLSAGKVSFFDNLLKQNKNLS  368
P. bergei     294  GSGSITGTLGKDNFILGNHIIRNQTFGLVKSETSDNLNNADNVFEYINFEGIVGLGFPGMLTAGNIPFFDNLLKQYKNMT  373
P. chabaudi   290  GSGSISGTLGKDNFILGNHTIRNQIFGLVKSESSDNLNTSDNVFEYINFEGIVGLGFPGMLTAGNIPFFDNLLKQHENIT  369
P. yoellii    295  GSGAITGTLGKDNFILGNHIIRNQTFGLVKSETSDNLNNSDNVFEYINFEGIVGLGFPGMLTAGNIPFFDNLLKQYKNMT  374


P. falciparum 392  PQFSFYISPYDGKSTLIIGGISKSFYEGDIYMLPVLKESYWEVKLDELYIGKERICCDEESYVIFDTGTSYNTMPSSQMK  471
P. vivax      356  PQFSFYISPEDNTSTFLVGGVSKSFYEGSIYMLPVVKEYYWEVELDGIYVGEKKICCEEKSYAIFDTGTSYNTMPSAQMK  435
P. knowlesi   369  PQFSFYISPDDNTSTFIIGGLSKSFYQGSIYMLPVIKEYYWEVELDGIYVGEKKICCEEKSYAIFDTGTSYNTMPSAQIK  448
```

```
P. bergei      374  PQFSFYISPNDETSTFIVGGISKSYYEGDIYMLPVVKEYYWEVKLDAIYIGDEKICCEEESYAIFDSGTSYNTMPSTQIS  453
P. chabaudi    370  PQFSFYISPNDESSTFIVGGINKSYYEGDIFMLPVIKEYYWEVKLDAIYIGDEKICCDEESYAIFDSGTSYNTIPSIQTD  449
P. yoellii     375  PQFSFYISPNDETSTFIVGGISKSYYEGDIYMLPVIKEYYWEVKLDAIYIGDEKICCEEESYAIFDSGTSYNTMPSTQIN  454

P. falciparum 472  TFLNLIHSTACTEQNYKDILKSYPIIKYVFGELIIELHPEEYMILNDDVCMPAYMQIDVPSERNHAYLLGSLSFMRNFFT  551
P. vivax       436  GFFDVVPSAPCTEENYQEVLKNYPVIKYLFGDLVIELLPEEYMILNEESCIPAYMQIDVPSEKNHAYLLGSIAFMRHYYT  515
P. knowlesi    449  NFFDVVPSVACTEENYQDVLKNYPIIKYVFGDLIIELMPEEYMILNEDNCIPAYMQIDVPSEKNHAYLLGSLAFMRHYYT  528
P. bergei      454  NFFKIVSSKPCNEENYNNILKEYPTIKYVFGKLVIELLPNEYMIVNDDLCVPAYMQIDVPSENNNAYLLGTIAFMRHYFT  533
P. chabaudi    450  NFFKLIPSKPCNEENYNDILKDYPTIKYVFGKLVIELLPSEYMIVNDDLCAPAYMQIDVPSENNNAYLLGTIAFMRHYFT  529
P. yoellii     455  NFFKIVSSKPCNEENYNDILKEYPSIKYVFGKLVIELLPNEYMIVNDDLCVPAYMQIDVPSENNNAYLLGTIAFMRHYFT  534

P. falciparum 552  VFVRGTESRPSMVGVARAKSKN--------  573
P. vivax       516  VFVRGAGGQPSMVGVAKARAAAEAAQKVAE  545
P. knowlesi    529  VFVRGVNGKPSMVGVAKAKSAASAVN----  554
P. bergei      534  VFVRGQEGNPSMVGVAKAKQV---------  554
P. chabaudi    530  IFVRGQEGRPSMVGVAKAKRN---------  550
P. yoellii     535  IFVRGQEGNPSMVGVAKAKRV---------  555
```

Figure 6.1: **Comparison of PM X orthologs from six Plasmodium species.**

Alignments of six PM X orthologs as prepared by the Cobalt Multiple Alignment Tool (Papadopoulos and Agarwala, 2007). A high degree of conservation can be seen between the orthologs, particularly at the N terminal signal peptide (bold and underlined) and proximal to the catalytic sites (red). In comparison, while LCRs (purple) are present in all species they are more numerous in species that infect humans.

6.2 Methods:

6.2.1 *Pf*PM X homology model generation:

6.2.1.1 Pf*PM X template and model selection:*

The full length protein coding sequence of *Pf*PM X (PF3D7_0808200) was obtained from the online database www.plasmodb.org and entered into the online structure prediction application on the servers, SwissModel (Arnold, Bordoli *et al.*, 2006), CPH model (Nielsen, Lundegaard *et al.*, 2010), 3D Jigsaw (Bates, Kelley *et al.*, 2001) and Protein Homology/analogY Recognition Engine (PHYRE) (Kelley and Sternberg, 2009). Each application selected a template structure for structure prediction based on the specific alignment method of that application (discussed in Chapter 4).

6.2.1.2 Preparation and Refinement of the working model:

The models generated by the prediction servers were ranked according to their Expect (E) value and sequence identity and the most appropriate *in silico* model was prepared for further investigation with the 'Protein Preparation Wizard' (incorporated into Schrödinger Inc. modelling software (Schrodinger, 2011)). A description of E values and the 'Protein Preparation Wizard' can be found in section 4.2.1.2. When no errors were found, the model underwent explicit solvent modelling in the sub-program Desmond (Maestro-Desmond Interoperability Tools, 2010, Bowers, Chow *et al.*, 2006) in an orthorhombic solvent box with solvent containing 0.15M NaCl to mimic physiological conditions. The docking model then underwent a minimisation of 5000 iterations using an OPLS2005 force field in Desmond.

Once the minimisation was complete, a molecular dynamics simulation was performed using Desmond. This was performed under NTP conditions (fixed temperature and pressure with variable volume) in explicit solvent containing 0.15M NaCl. There was 50 ps of molecular dynamics relaxation time before the initiation of the production molecular dynamics simulation. A random frame selected from the 1.2 ns production simulation at thermal equilibrium was used for later docking and simulation experiments (frame 544 of 833). This docking model was then further minimised for 2000 iterations in Desmond to further relax the model to resemble its natural state.

6.2.2 Examining the model:

The *Pf*PM X docking model was examined using Virtual Molecular Docking (VMD) (Humphrey, Dalke *et al.*, 1996). To examine the similarities and differences between the

generated model and known aspartic protease structures, the docking model was compared to a published model of *Pf*PM II (PDB code: 1LF3 (Asojo, Gulnik *et al.*, 2003)) and the chosen template structure. The models were aligned by STAMP (Russell and Barton, 1992) and their structural similarities compared, in particular the active sites.

6.2.3 Preparation of the *Pf*PM X model for docking:

In order to perform docking studies and screen virtual chemical compound libraries for hits, a receptor grid was generated using Glide in the Schrödinger modelling suite (Friesner, Banks *et al.*, 2004, Halgren, Murphy *et al.*, 2004, Friesner, Murphy *et al.*, 2006). The centre of the generated 3D grid (size 12^3 Å) was specified as between the two catalytic aspartic residues (ASP 266 and 457; Schrödinger numbering) and the remaining settings as default.

6.2.4 Preparation of Ligand libraries:

Before ligand libraries could be screened for lead inhibitors they were first prepared in 3D format. The library of HIV Protease Inhibitors (HIV PIs) utilised in this study was prepared by Dr Brown (Griffith University) prior to the commencement of this project. Briefly, the 3D coordinates of each HIV PI was obtained from the Cambridge Structural Database (Groom, Allen *et al.*, 2013) and any structural errors were corrected within Schrödinger software. The ligand structures were minimised to a convergence threshold of 0.05 and then prepared for docking as described below.

Chemical compound libraries from GlaxoSmithKline (www.ebi.ac.uk/chemblntd; (Gamo, Sanz *et al.*, 2010)), Sigma-Aldrich (http://www.sigmaaldrich.com/chemistry/chemistry-services/selected-structure.html), the National Cancer Institute (NCI) (http://cactus.nci.nih.gov/ncidb2.2/) and Ambinter (http://ambinter.com/catalog/SDF) were obtained in 2D format and prepared for 3D docking with the LigPrep application within the Schrödinger modelling suite (Schrodinger, 2013). LigPrep converts 2D structures to energy minimised 3D molecular structures which can then be used for 3D *in silico* docking (Schrodinger, 2013). Due to the large size of the Ambinter archive (>5,000,000 compounds), the library required separation into smaller libraries of 50,000 compounds before ligands could successfully be prepared. Ligands were prepared using an OPLS_2005 force field, with no desalting at pH 7 ± 2.0 using Epik to assign ligand protonation states (Shelley, Cholleti *et al.*,

2007, Greenwood, Calkins *et al.*, 2010). Specified chiralities were retained and tautomers generated with all other settings as default.

6.2.5 Inhibitor docking:

6.2.5.1 HIV PI library

The Glide application in Schrödinger (Friesner, Banks *et al.*, 2004, Friesner, Murphy *et al.*, 2006, Halgren, Murphy *et al.*, 2004) was used to perform an induced fit (Sherman, Day *et al.*, 2006) standard precision (SP) dock of the prepared HIV PI ligands into the *Pf*PM X model. The scaling factor was reduced to 0.8 and all other settings were left as default.

6.2.5.2 GlaxoSmithKline

Prepared GlaxoSmithKline ligands were docked within the *Pf*PM X model as 6.2.5.1.

6.2.5.3 Sigma-Aldrich

The Sigma-Aldrich compound library was screened for top hits using the Virtual Screening Workflow within the Schrödinger modelling suite. The Lipinksi's Rule of 5 filter was removed to ensure all docking compounds were identified and result thresholds were changed to the top 20% for both HTVS and SP docking and 25% of XP docks with a maximum of two poses allowed for each ligand. All other settings were default.

6.2.5.4 NCI

The NCI compound library was screening using the Virtual Screening Workflow as 6.2.5.3.

6.2.5.5 Ambinter

Ambinter libraries were screened using the Virtual Screening Workflow as 6.2.5.3.

6.2.6 Sensitivity of *P. falciparum* to docking hits *in vitro*

Top docking hits that satisfied Lipinski's Rule of 5 and were financially viable to purchase were examined for their activity against *P. falciparum.* The *in vitro* sensitivities of *P. falciparum* lines D10 and 3D7 to these compounds was determined using [^{3}H]-hypoxanthine incorporation as previously described (Skinner, Manning *et al.*, 1996). Inhibitors that demonstrated activity were also examined against transgenic

cultures that over-transcribed *PfPM X* on 50 and 150 nM drug selection (Chapter 5). Serial dilutions of each inhibitor were prepared in culture media and added with [^{3}H]-hypoxanthine (0.5µCi/well; PerkinElmer) to asynchronous cultures at a 1% parasitaemia and 1% haematocrit. The amount of [^{3}H]-hypoxanthine incorporated into parasites was measured after an incubation of 48 hours. Each assay was performed in triplicate on at least three separate occasions, data were pooled and growth curves analysed (non-linear regression curve fit and IC_{50} determination). Data are presented as mean $\pm$ SD. IC_{50} values were compared using the extra sum of squares F-test (GraphPad Prism).

6.2.7 Enzyme Activity Assays

The ability of top docking "hits" to inhibit *Pf*PM X was examined in enzyme activity assays with transgenic C-terminally tagged *Pf*PM X (Chapter 5). Enzyme activity assays were performed as 5.2.4.2, with the addition of 10 or 20 µM inhibitor. Activity assays were performed with at least two biological replicates. Vehicle, enzyme only and substrate only controls were included in each assay. The general aspartic protease inhibitor Pepstatin A (Sigma-Aldrich) was used as a positive inhibitor control and recombinant *Pf*PM II (supplied by our collaborators at the University of Florida) was used as a positive control. Inhibition of *Pf*PM X by "hits" was compared to untreated controls and analysed by Dunnett's Multiple Comparison test (GraphPad Prism). Differences were considered significant if $P \leq 0.05$.

6.2.8 Pharmacophore hypothesis generation and screening

A pharmacophore hypothesis of *Pf*PM X inhibition was generated using Phase (Dixon, Smondyrev *et al.*, 2006a, Dixon, Smondyrev *et al.*, 2006b) and the 'Develop Common Pharmacophore Hypothesis' workflow resident within the Schrödinger modelling suite. The HIV PI ligands were used to generate the pharmacophore hypothesis and in the absence of substantial enzyme activity data, due in part to the low yields of extracted C-terminally tagged *Pf*PM X, 'activity' values were substituted with docking scores obtained as reported above. Activity thresholds were set above a value of 6 to include all HIV PIs with antimalarial IC_{50} values in the low µM range, with values below 5.9 being considered inactive. All other settings remained as default and once ligand 'sites' were created and 16 active groups determined, pharmacophore settings were changed so that generated hypotheses had to match at least 16 active groups with a minimum 3 and maximum 7 matching 'ligand interaction sites'. Active hypotheses were scored and the

top-ranking hypothesis was used to perform a pharmacophore screen of the Sigma-Aldrich compound library in the module 'Advanced Pharmacophore Screening' using Phase.

The Sigma-Aldrich library was first generated to a format compatible with Phase as described in the Schrödinger user manual (Schrodinger, 2009). The maximum number of conformers for each ligand was set to 10 and the number of conformations per rotatable bond was limited to 8. All other settings were default. In the advanced pharmacophore screen existing conformers were kept and the number of conformers per rotatable bond was changed to 8. The maximum number of conformers per structure was changed to 50 and all other settings were default.

6.3 Results

6.3.1 Choice of a template and selecting a *Pf*PM X 'docking model'

Online applications (SwissModel, CPH Model, Phyre and 3D Jigsaw) were used to generate preliminary structures of *Pf*PM X. As with *Pf*PM IX, the template chosen was produced by the threading application PHYRE (Kelley and Sternberg, 2009), which selects templates for structure prediction based on secondary structural alignments. This was also the only application that generated a realistic structure, *i.e.* it predicted a structure without the deletion of the LCR insert regions in the final model. The predicted structure for *Pf*PM X was based on a template structure derived from the aspartic protease human Cathepsin E (PDB code 1TZS; Figure 6.2) (Ostermann, Gerhartz *et al.*, 2004) with a 30% sequence identity and an E value of $2.45e^{-25}$.

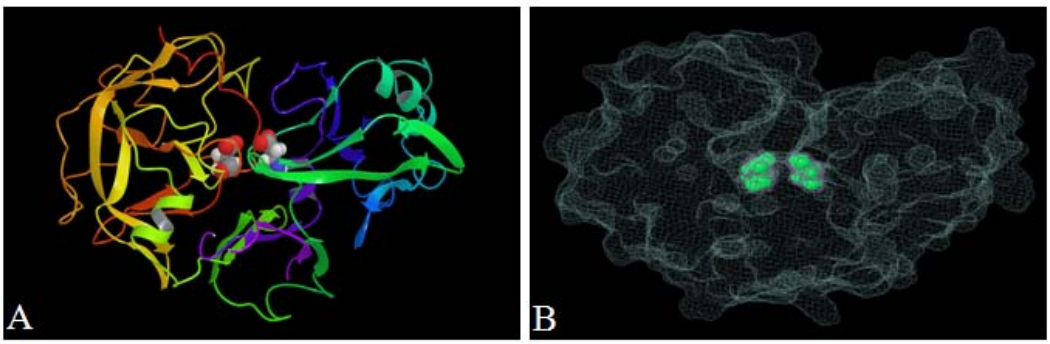


Figure 6.2: **Homology model of *Pf*PM X.** A) The C-α backbone of the model is shown in ribbon (rainbow). The catalytic aspartic residues (266 and 457; VDW) are positioned within the active site. B) The surfaces of the *Pf*PM X model (grey) give an indication of its shape and size, in particular the active site. The catalytic aspartic residues (green) can be seen within the active site.

6.3.2 Refinement of the *Pf*PM X 'docking model'

Minimisations and Molecular Dynamic simulations were performed (Schrödinger Inc.) in explicit solvent to reduce the conformational energy of the system and to predict equilibrium conformations. A Ramachandran plot, describing the torsional angles of the model, suggested no major structural abnormalities with most of the non-glycine resides located in favourable secondary structural motifs (Figure 6.3).

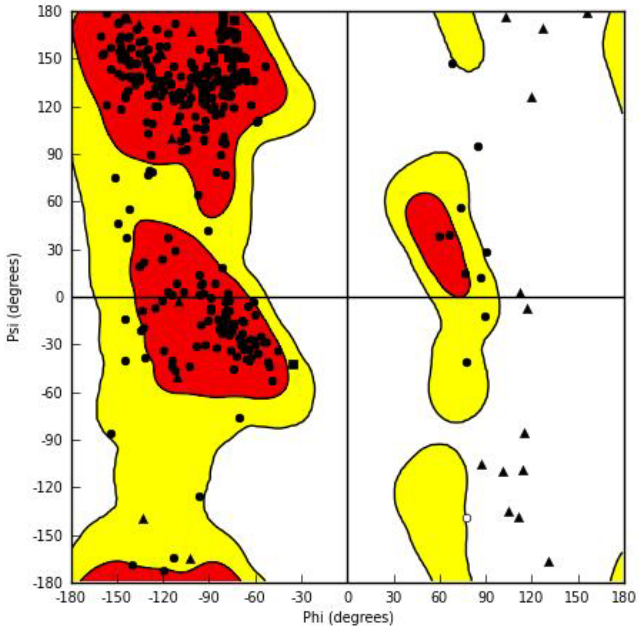


Figure 6.3: **Ramachandran diagram of the Φ-Ψ dihedral angles of the *Pf*PM X homology model.** Individual residues (black) are placed according to their torsional angles with the majority being placed in either the most favoured (red) or additionally allowed (yellow) regions. Triangles represent glycine residues. This plot was generated with VMD software.

6.3.3 Examination of the model

The generated *Pf*PM X model (Figure 6.4A; purple) was aligned with solid state structures of other aspartic proteases (Model template human Cathepsin E: PDB code 1TZS (green) and *Pf*PM II: PDB code ILF3 (orange)) obtained from the PDB (Figure 6.4) using VMD software. All structures demonstrated the characteristic bilobal structure of All structures demonstrated the characteristic bilobal structure of the aspartic protease family (Dunn, 2002). While there are a number of differences between the structures, primarily loop regions at the top and bottom of the model, a high degree of similarity can be seen in the proteases tertiary structures proximal to the active site (Figure 6.4).

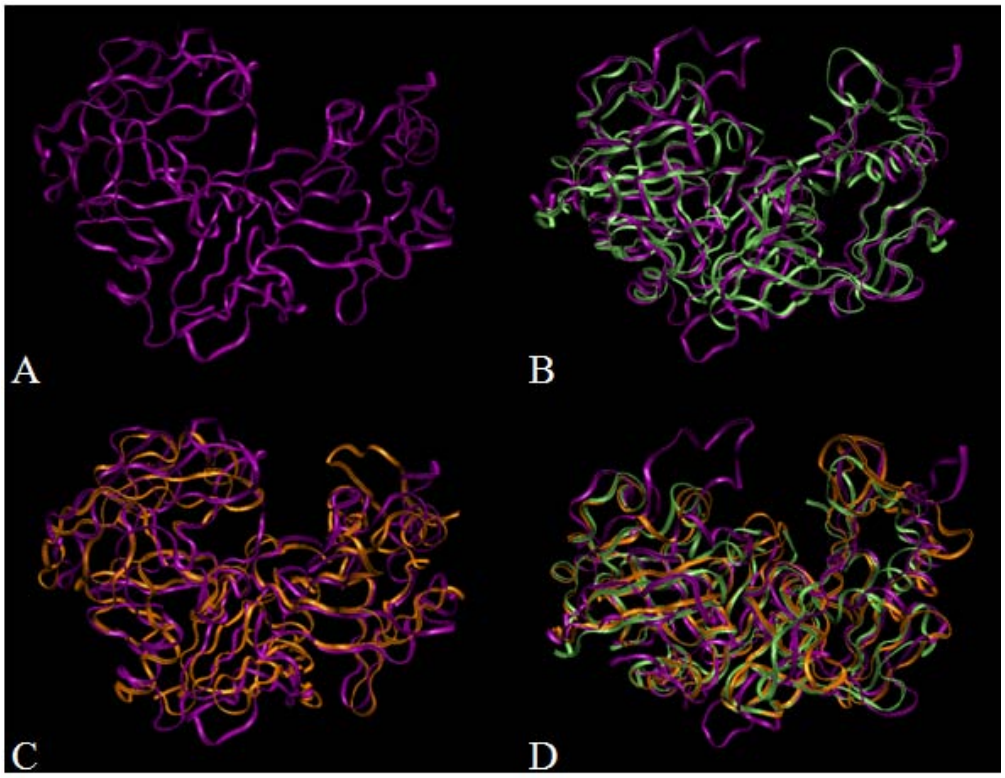


Figure 6.4: **Structural alignment of the *Pf*PM X homology model with 1TZS and *Pf*PM II.** A) The *Pf*PM X model (purple). B) The *Pf*PM X model (purple) superimposed with its human cathepsin E template (green; PDB code: 1TZS). C) The *Pf*PM X model (purple) is aligned with the crystal structure of *Pf*PM II (orange; PDB code: 1LF3). D) Merge of B and C.

6.3.4 Inhibitor Docking

A HIV PI library was docked within the *Pf*PM X model to assess binding of inhibitors. The inhibitors are ranked by Glide docking score and shown in Table 6.2 with their respective IC_{50} data (Skinner-Adams, McCarthy *et al.*, 2004, Peatey, Andrews *et al.*, 2010) or EC_{50} data (Parikh, Gut *et al.*, 2005, Parikh, Liu *et al.*, 2006) available within the literature and logP, as calculated with ACD/ChemSketch software. Chemical structures are also supplied (Figure 6.6). The Glide docking score is an approximation of the ligand binding free energy (Friesner, Banks *et al.*, 2004, Friesner, Murphy *et al.*, 2006) with top binding inhibitors having a lower numerical value. The general aspartic protease inhibitor Pepstatin A was included as a control. The top predicted binding inhibitor was indinavir (Table 6.2; Figure 6.6), with a docking score of -9.48. Analysis of the docking geometries revealed that the compound was able to adopt several conformations within the binding site (Figure 6.5), each having a different Glide docking score.

Table 6.2: **HIV PIs dock within the *Pf*PM X homology model.** Inhibitors are ranked by Glide docking score and are displayed with their IC_{50}/EC_{50} data.

HIV PI	'Glide' Docking Score	Structure number	*In vitro* sensitivity of *P. falciparum*	LogP
Indinavir	-9.48	1	1-3.8 µM ^	2.88
Tipranavir	-8.09	2	10-23 µM ^^	7.21
Nelfinavir	-7.89	3	21.8 ± 2.4 µM **	6.98
Lopinavir	-7.89	4	0.9-2.1 µM *	7.44
Saquinavir	-7.26	5	0.6-7.2 µM *	4.44
Ritonavir	-7.19	6	0.6-1.8 µM ^	5.28
Pepstatin A	-6.34	7	7.5 ± 0.4 µM **	2.16
Darunavir	-5.90	8	70 µM ^^	3.94
Amprenavir	-5.89	9	71.24 ± 16.1 µM **	4.20
Atazanavir	-5.79	10	35 ± 7.4 µM **	5.20

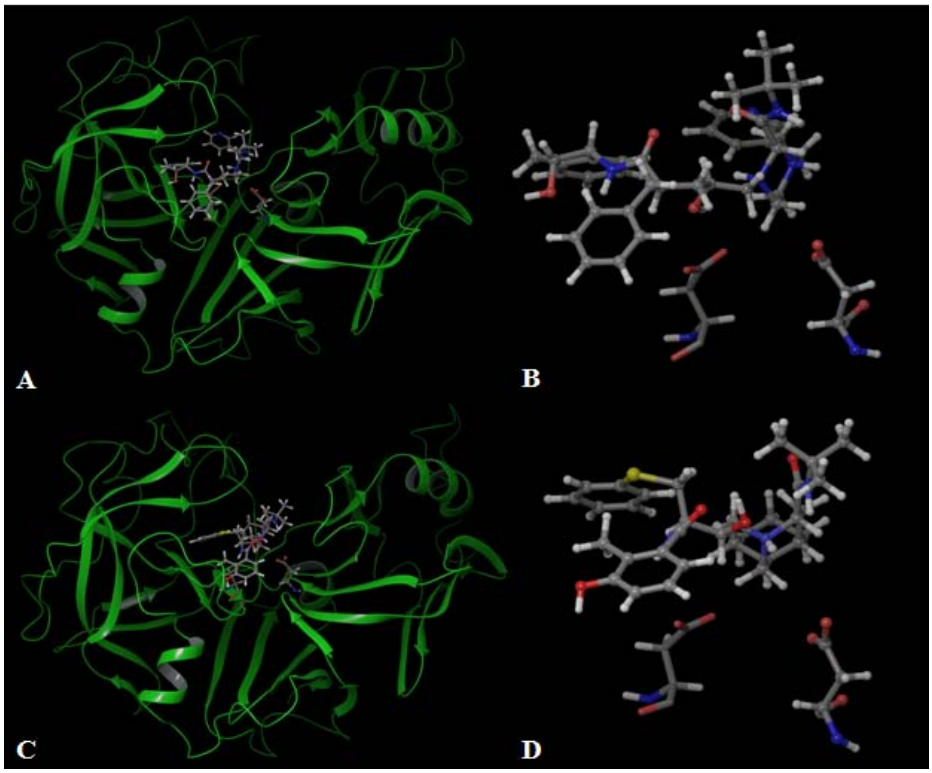
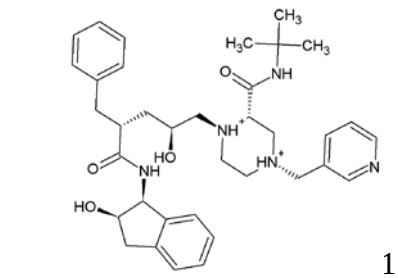
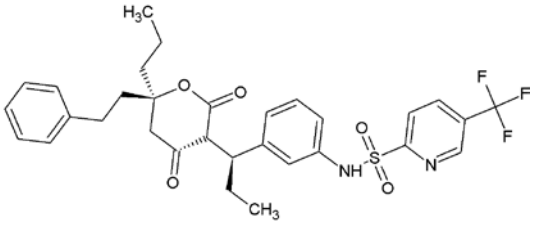


Figure 6.5: **HIV PIs dock within the *Pf*PM X homology model binding cleft.** indinavir (ball and stick) binds within the model binding cleft (A) and interacts with the catalytic aspartic residues (ball and stick), (B). nelfinavir (ball and stick) docks within the model binding cleft (C) and interacts with the catalytic aspartic residues (ball and stick), (D). *Pf*PM X model backbone is displayed in ribbons (green).

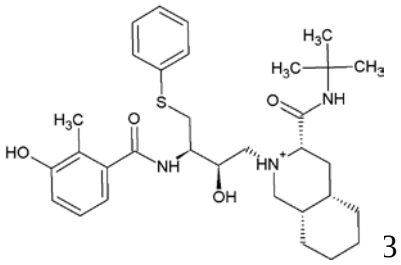
*IC_{50} (Parikh, Gut *et al.*, 2005), **IC_{50} (Parikh, Liu *et al.*, 2006), ^EC_{50} (Skinner-Adams, McCarthy *et al.*, 2004), ^^ EC_{50} (Peatey, Andrews *et al.*, 2010).



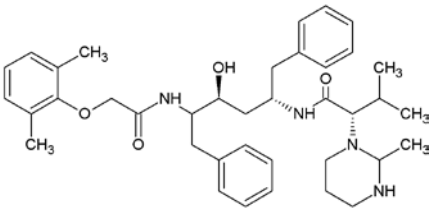
1



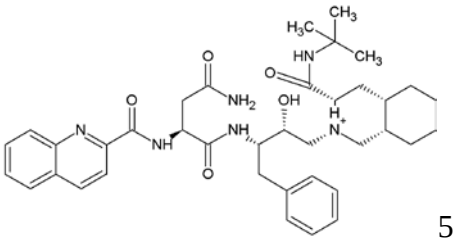
2



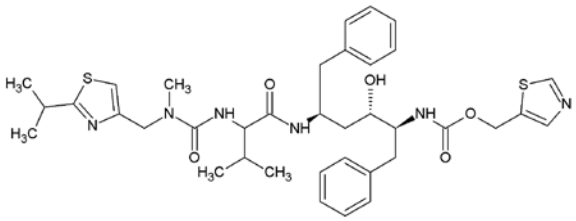
3



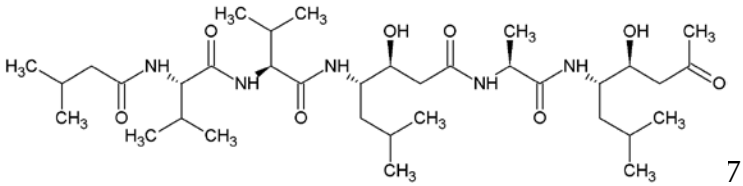
4



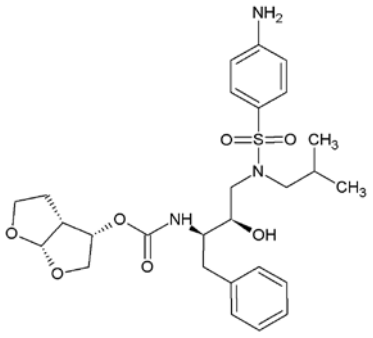
5



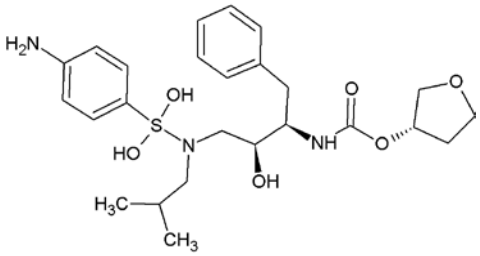
6



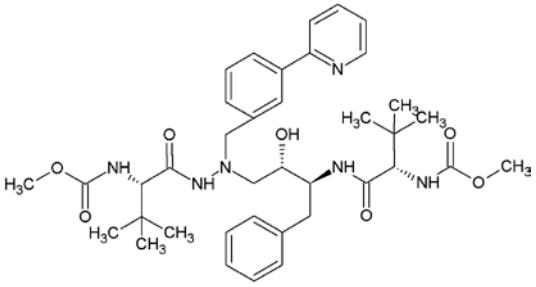
7



8



9



10

Figure 6.6: **Structures of the HIV PIs.** Refer to Table 6.2 for corresponding compound information.

6.3.5 Identification of lead inhibitors

6.3.5.1 GlaxoSmithKline library screen

The GlaxoSmithKline library (6781 compounds) was docked with the *Pf*PM X model to identify potential inhibitors; the top 14 being examined further. As above, the inhibitors were ranked by Glide docking score with the respective published IC_{50} data (Table 6.3; (Gamo, Sanz *et al.*, 2010)). The logP, the calculated value for the lipophilicity of the molecule (Ghose, 1998), is also presented (Table 6.3). According to Lipinksi's rule of five, the logP of a good drug like compound should be within a range of -0.4 to 5.6 (Ghose, Viswanadhan *et al.*, 1999). Structures for these compounds can be found in Appendix 7. The top docking hit had a Glide docking score of -11.13 and IC_{50} of 0.82 μM (Table 6.3). Interestingly, the majority of top docking "hits" had a calculated logP above 5, suggesting that they may not be highly soluble in water and are outside the Lipinski's rule of 5 for suitability as orally active drugs (Ghose, Viswanadhan *et al.*, 1999).

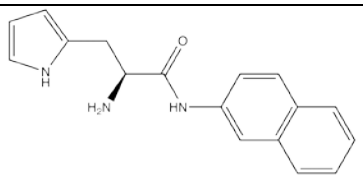
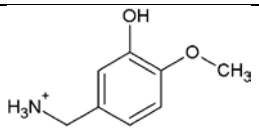
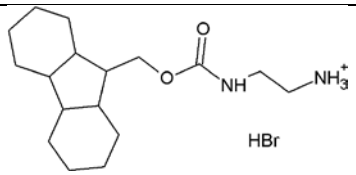
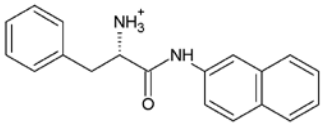
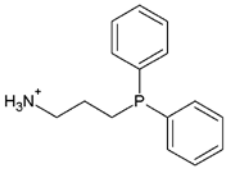
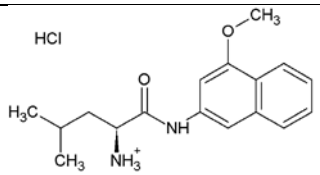
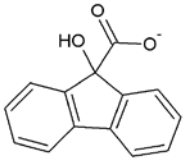
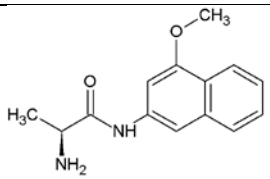
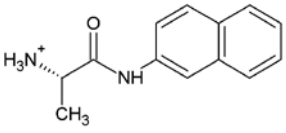
Table 6.3: **Top docking "hits" from the GlaxoSmithKline library.** Inhibitors are ranked by Glide docking score and are displayed with their respective ChEMBL identifier, LogP and IC_{50} data. *(Gamo, Sanz *et al.*, 2010)

Compound	**Schrödinger 'Glide' Docking Score**	***P. falciparum in vitro* IC_{50} (μM)***	**LogP**
TCMDC-131544	-11.13	0.82	7.35
TCMDC-139334	-10.65	0.4	8.28
TCMDC-140588	-10.58	0.35	7.87
TCMDC-135362	-10.38	0.78	6.29
TCMDC-141769	-10.19	0.76	6.20
TCMDC-141484	-10.18	0.8	7.32
TCMDC-137201	-10.16	1.73	8.11
TCMDC-140556	-10.06	0.95	6.11
TCMDC-139797	-9.96	0.78	7.28
TCMDC-141567	-9.96	0.78	9.64
TCMDC-136458	-9.95	0.81	6.73
TCMDC-139016	-9.94	0.94	9.94
TCMDC-141514	-9.92	0.94	4.74
TCMDC-141101	-9.89	0.2	6.27

6.3.5.2 Sigma-Aldrich library screen

The Sigma-Aldrich library was screened for top docking hits with the virtual screening workflow in the Schrödinger modelling suite. The top 20% of hits from a high-throughput screen were then screened with an SP dock and the top 20% from those went on to extra precision (XP) docking. Hits were ranked according to their Glide docking score and the top nine hits are shown in Table 6.4 along with the corresponding structure and logP calculated with ACD/ChemSketch. The logP for four Sigma-Aldrich hits were not able to be calculated by ACD/ChemSketch and were subsequently calculated with InstantJChem by Associate Professor Rohan Davis, Griffith University. The top binding compound was L-histidine β- napthylamide, with a Glide docking score of -8.23 and a logP of 0.55. All remaining molecules had a logP that fell within the range 0 - 5 (Table 6.4).

Table 6.4: **Top docking hits from a screen of the Sigma-Aldrich library.** Top ranking binders are presented with the respective structure, Glide docking score and LogP. *LogP calculated by Associate Professor Rohan Davis using InstantJChem.

Compound	Structure	Glide Docking Score	LogP
L-Histidine β-Napthylamide		-8.23	0.55
3-Hydroxy-4-methoxybenzylamine hydrochloride		-8.18	0.21*
N-Fmoc-ethylenediamine hydrobromide		-7.99	2.20*
L-Phenylalanine-2-naphthylamide		-7.87	2.93
3-(Diphenylphosphino)-1-propylamine		-7.86	3.97
L-Leucine 4-methoxy-β-naphthylamide hydrochloride		-7.82	2.94*
9-Hydroxy-9-fluorenecarboxylic acid		-7.79	1.28
L-Alanine 4-methoxy-β-naphthylamide		-7.66	1.69*
L-Alanine β-Napthylamide		-7.65	1.07

6.3.5.3 NCI library screen

The NCI library, consisting of >260,000 compounds was screened for binding to the *Pf*PM X model. The top scoring compound was an NSC 19478 with a Glide docking score of –16.21 (Table 6.5; Figure 6.7).

Table 6.5: ***In silico* screening of the NCI library reveals top hits.** 14 of the top docking "hits" along with the corresponding Glide docking scores and logP values. Not all compounds have a CID. Structures are presented in Figure 6.7.

Compound and CID	**Structure Number**	**Glide Docking Score**	**LogP**
NSC 19478	1	-16.21	-4.22
Glucosulfone CID 10440240	2	-15.91	-8.14
NSC 30359	3	-15.71	-3.38
Liriodendrin CID 6453461	4	-13.85	-8.60
Myricetrin CID 5281673	5	-12.82	1.98
NSC 20273 CID 54605098	6	-12.55	-5.26
NSC 19780 CID 9561109	7	-12.03	-3.13
Deamino dpn CID 4481	8	-11.92	-6.53
Neohesperidin CID 45358138	9	-11.85	2.44
3-pyridinealdehyde adenine dinucleotide CID 409335	10	-11.75	-7.99
NSC 35122 CID 234724	11	-11.77	-2.34
NSC 21559 CID 54608322	12	-11.74	1.70
NAD coenzyme 1 CID 5892	13	-11.68	-5.72
Asiaticoside CID 5351525	14	-11.37	2.82

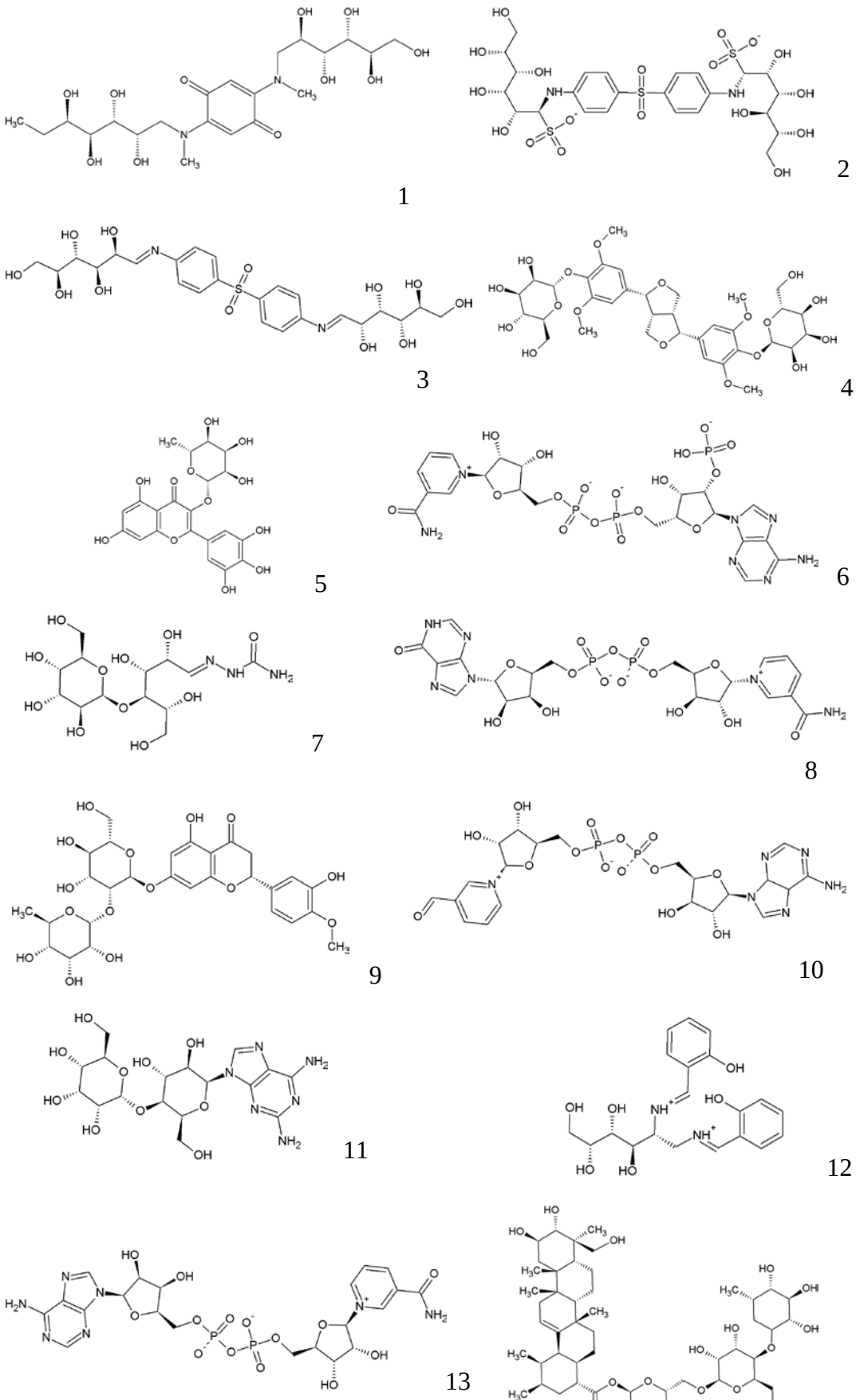


Figure 6.7: **Structures of top docking NCI hits.** Numbers correspond to compounds in Table 6.5.

6.3.5.4 Ambinter library screen

A library of 700,000 Ambinter compounds was screened in a Virtual Screening Workflow for binding within the *Pf*PM X homology model. The top binding hit identified in this library had a Glide docking score of -11.25 and LogP value of 5.2 (Table 6.6). The top 11 compounds from this screen are shown in Table 6.6 with their corresponding LogP scores (www.ambinter.com; (Ambinter, 2014)). The chemical structures of these hits are presented in Appendix 7.

Table 6.6: **Top docking "hits" from an *in silico* screen of the Ambinter library.** The top 11 inhibitors are presented with their Glide docking score and LogP. *(Ambinter, 2014).

Compound code	Glide docking score	LogP*
Amb17039013	-11.25	5.2
Amb17456959	-10.72	3.47
Amb18502744	-10.66	-0.74
Amb17151146	-10.64	2.67
Amb17320578	-10.38	4.26
Amb16943777	-10.25	3.3
Amb17308586	-10.09	3.58
Amb17228674	-10.07	3.56
Amb16940839	-10.05	5.21
Amb16943683	-10.04	3.44
Amb17092319	-10.03	4.15

6.3.6 Inhibition studies with lead docking "hits"

6.3.6.1 The effect of top docking "hits" on parasite growth in vitro

Lead docking "hits" that were commercially available and financially viable were investigated for their ability to inhibit *P. falciparum* growth *in vitro*. These included three compounds from the Sigma-Aldrich library and two from the NCI library. Of these, *N*-Fmoc-ethylenediamine hydrobromide (NFmoc) and L-histidine β-napthylamide (L-His) achieved IC_{50} values of 14.2 ± 2.4 μM and 19.6 ± 2.4 μM respectively (Figure 6.8, Table 6.7), with L-Alanine β-Napthylamide, Neohesperidin and Myricetrin failing to inhibit parasite growth more than 23% at the maximum concentration, 100 μM (Table 6.7).

Table 6.7: ***In vitro* sensitivities of *P. falciparum* to selected docking "hits".** Data are presented as the mean ± SD of three triplicate experiments.

Docking hit	*P. falciparum* growth inhibition
N-Fmoc-ethylenediamine hydrobromide	IC_{50} 14.2 ± 2.4 μM
L-histidine β-napthylamide	IC_{50} 19.6 ± 2.4 μM
L-Alanine β-Napthylamide	4.96% at 100 μM
Neohesperidin	17.17% at 100 μM
Myricetrin	22.28% at 100 μM

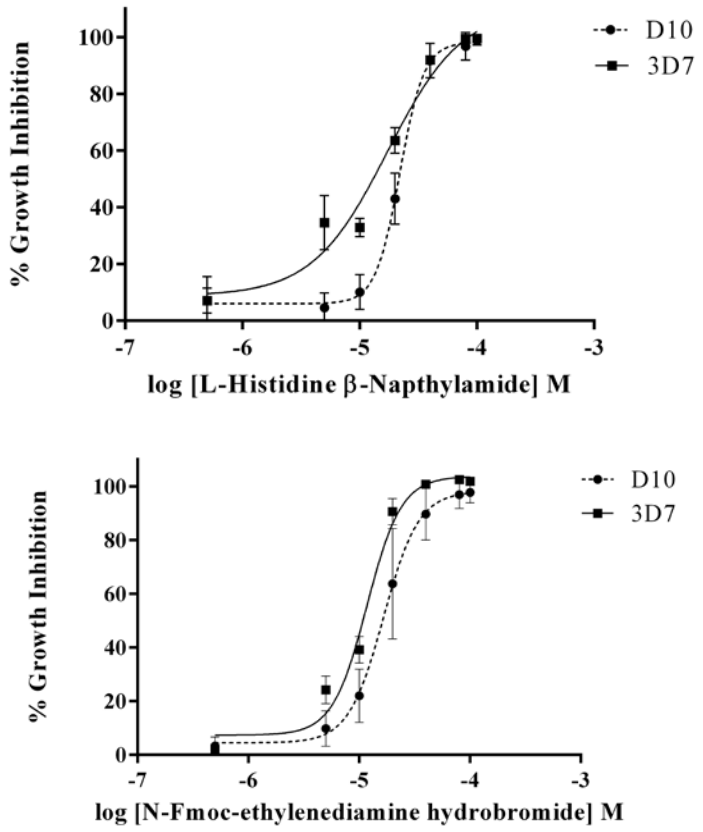


Figure 6.8: ***In vitro* sensitivity of D10 and 3D7 to L-His and NFmoc.** The i*n vitro* sensitivity of D10 (circles, dashed line) and 3D7 (square, solid line) to L-His and NFmoc was determined by ^{3}H-hypoxanthine incorporation. Data are presented as mean ± SD of 3 experiments in triplicate.

The sensitivity of transgenic cultures over-transcribing *Pf*PM X to NFmoc and L-His was examined. Transgenic parasites which had been maintained on 50 nM of WR99210 did not have significantly different sensitivity to L-His or N-Fmoc, compared to D10 parasites (Figure 6.9). These transgenic parasite populations exhibited IC_{50} of 21.4 µM compared to the wild type parasites which had an IC_{50} value of 20.8 µM for L-His (P=0.57) and an IC_{50} of 16.8 µM for N-Fmoc compared to the D10 IC_{50} of 20.1 µM (P=0.22) (Figure 6.9).

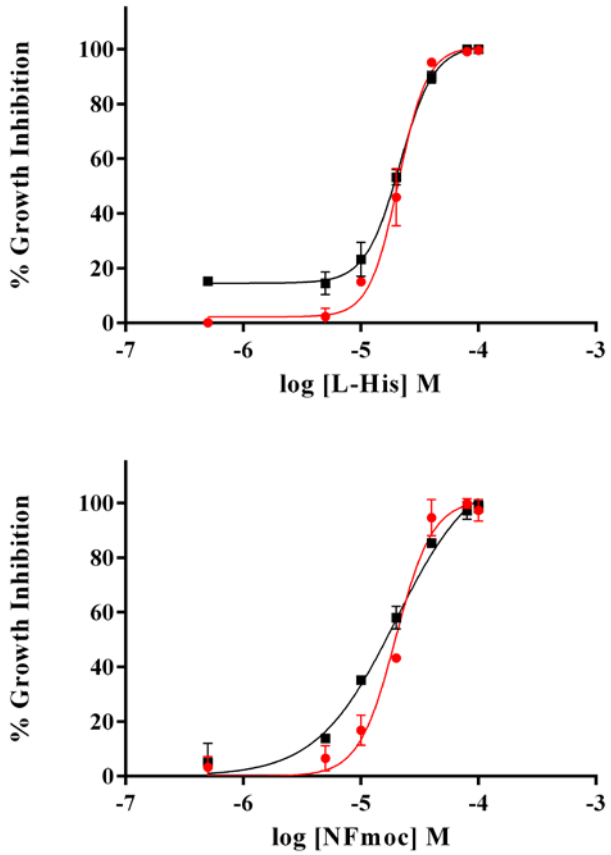


Figure 6.9: **The *in vitro* sensitivities of transgenic cultures containing a C-terminally tagged *Pf*PM X to NFmoc and L-His.** The sensitivity of transgenic cultures over transcribing *Pf*PM X (50nM of WR99210) (squares, black) to L-His and NFmoc, compared with D10 wild type (red circles, red line) parasites. Data are presented as mean ± SD of four triplicate experiments.

6.3.6.2 The effect of lead docking hits on enzyme activity

Enzyme activity assays with HIV PIs and top docking hits were performed to examine inhibition of *Pf*PM X transgenic enzyme (Chapter 5). In contrast to HIV PI docking scores, lopinavir, saquinavir and ritonavir significantly inhibited *Pf*PM X enzyme activity at 10 µM ($P \leq 0.0001$) while tipranavir, nelfinavir and duranavir did n ot (Figure

6.10). Pepstatin A, a general aspartic protease inhibitor included as a control, also significantly inhibited *Pf*PM X enzyme activity (P≤0.0001; Figure 6.10).

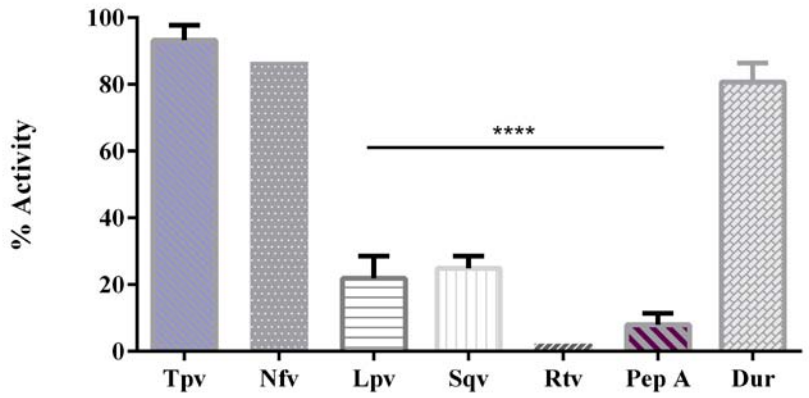


Figure 6.10: **Selected HIV PIs inhibit *Pf*PM X-GFP at 10 µM.** HIV PIs lopinavir (Lpv), saquinavir and ritonavir, together with Pepstatin A (Pep A), significantly inhibit *Pf*PM X enzyme activity using the substrate M-2455 whereas tipranavir (Tpv), nelfinavir (Nfv) and duranavir (Dur) have no activity against this enzyme. Data are presented as mean ± SD of at least 2 experiments. P≤0.0001.

Top docking compounds failed to significantly inhibit *Pf*PM X transgenic enzyme at 10 (Figure 6.11A) or 20 µM (Figure 6.11B). Whereas Pepstatin A significantly inhibited cleavage of the fluorogenic substrate M-2455 at 20 µM (P≤0.01; Figure 6.11B). Vehicle only controls had no effect on *Pf*PM X enzyme activity (Figure 6.11). Positive *Pf*PM II assay controls demonstrated activity in all assays (data not shown).

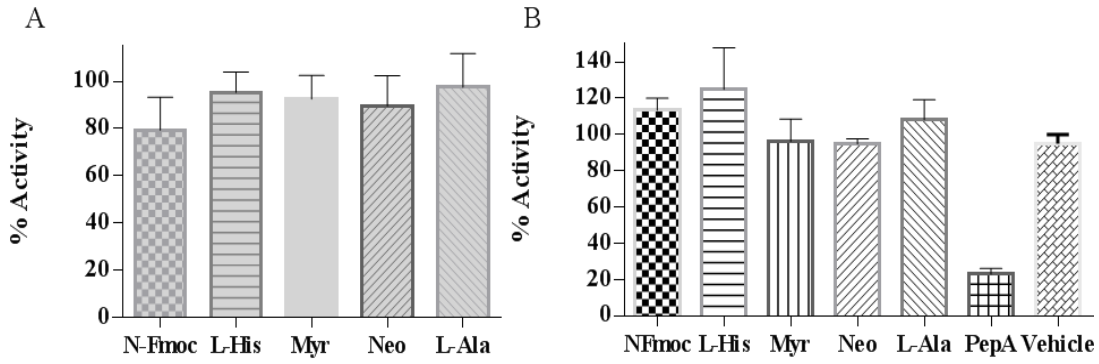


Figure 6.11: **Lead docking compounds do not inhibit *Pf*PM X-GFP.** Lead compounds did not significantly inhibit *Pf*PM X at 10 (A) or 20 µM (B). Pepstatin A, included as a positive control at 20 µM, significantly inhibited enzyme activity (P≤0.01). Vehicle controls did not significantly inhibit enzyme (B). Data are presented as mean ± SD of at least 3 experiments.

6.3.7 Pharmacophore model generation and library screening

HIV PI docking scores were used to generate a pharmacophore model of *Pf*PM X binding in the Schrödinger modelling suite. The top scoring pharmacophore hypothesis consisted of 4 key features. These were 2 hydrogen-bond acceptors (pink), 1 hydrogen-bond donor (blue) and 1 aromatic ring (brown; Figure 6.12). This model was then used to screen for inhibitors that had the same key features within the Sigma-Aldrich library using the Phase package within Schrödinger. Results were ranked by goodness of fit by Phase and the top 15 are shown in Table 6.8 with corresponding LogP. Structures for these compounds are in Figure 6.13. Due to time constraints hits from this screen could not be investigated further.

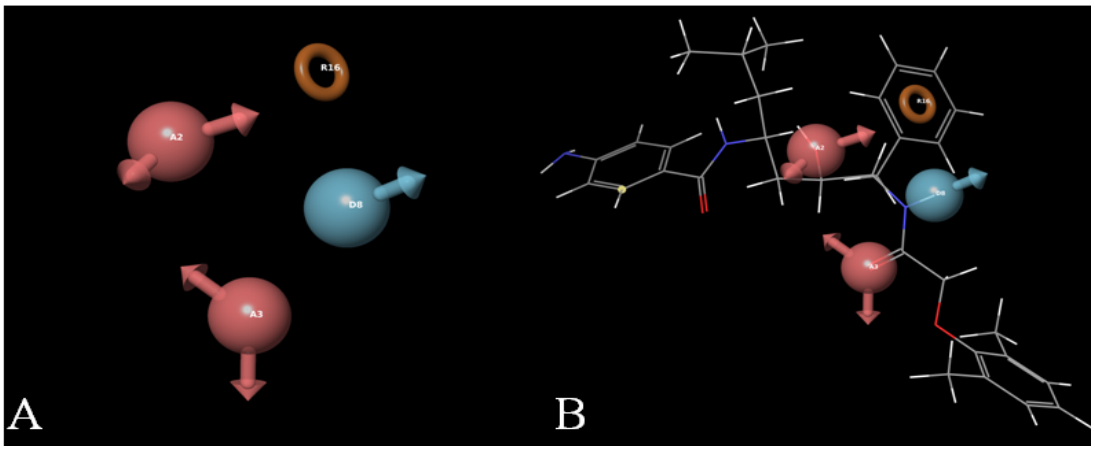


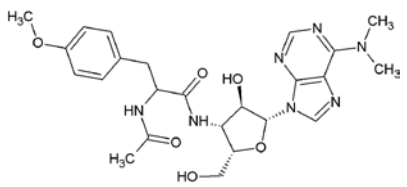
Figure 6.12: **Pharmacophore model for *Pf*PM X inhibition.** The generated pharmacophore alone (A) and shown on an inhibitor (B) consists of 4 key features in particular spatial arrangement. The key features are 2 hydrogen-bond acceptors (pink), a hydrogen-bond donor (blue) and an aromatic ring (brown) displayed on an example HIV PI.

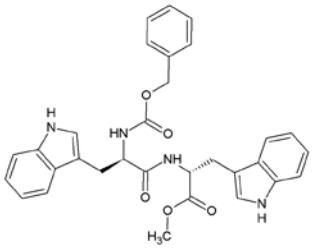
Table 6.8: ***In silico* pharmacophore screen of the Sigma-Aldrich library.** Results are ranked by goodness of fit by Phase. Where no common name exists for a compound, database ID is used.

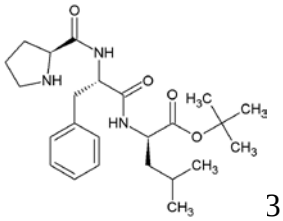
Compound and CID	**Structure Number**	**LogP**
NSC400640 CID 343981	1	0.30
NSC627875 CID 363200	2	4.50
NSC360207 CID 434889	3	2.60
NSC163805 CID 294760	4	4.00
NSC408367 CID 348867	5	2.40
n|A-acetyl-n-(2-methoxyphenyl) phenylalaninamide CID: 247780	6	2.40
NSC711715 CID 400197	7	1.80
NSC717470 CID 402997	8	3.40
Emetine CID 419025	9	7.50
NSC111148 CID 269688	10	6.40
NSC94213 CID 261593	11	2.50
ethyl n-acetyl-|A-oxophenylalaninate CID 227287	12	1.40
enkephalin-Leu, des-Tyr(1)- CID 2994	13	-1.90
NSC62317 CID 247505	14	2.60
NSC333733 CID 333315	15	2.10

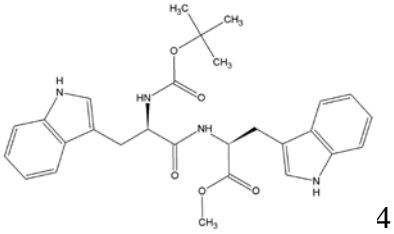
6.4 Discussion and Conclusion:

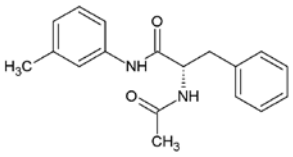
Using well documented automated model-building approaches this study generated a tangible *in silico Pf*PM X homology model. Observations that give credence to the model include; the model possessed a high degree of structural similarity to other well characterised aspartic proteases (Figure 6.4) and analysis of this model revealed no structural abnormalities (Figure 6.3); docking experiments with suspected inhibitors revealed that these compounds docked within the binding cleft of the *Pf*PM homology model. In addition, this computational study avoided the issues inherent to LCRs, as the LCRs within the *Pf*PM X protein sequence are of smaller size, rendering them more amenable to accurate modelling.

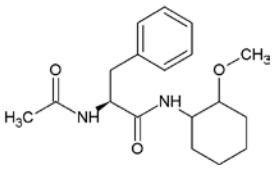

1

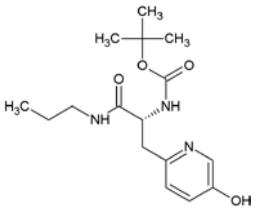

2

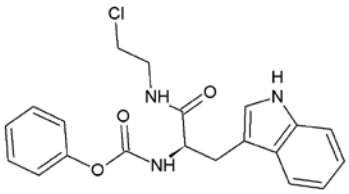

3

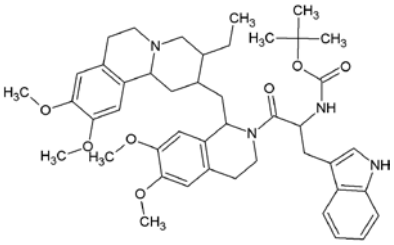

4

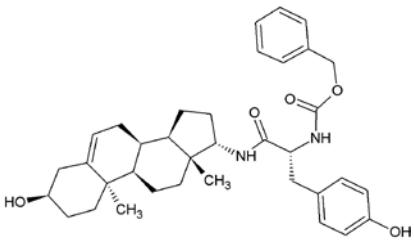

5

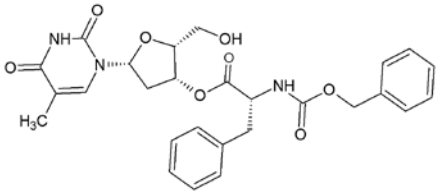

6

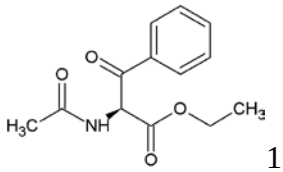

7

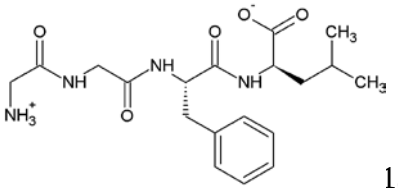

8

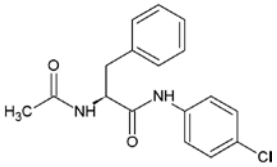

9


10


11


12


13


14

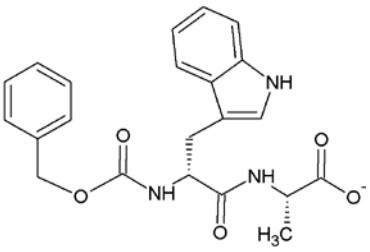

15

Figure 6.13: **Structures of the top ranked pharmacophore search results.** Numbers correspond to compounds in Table 6.8.

As discussed previously, during the course of this study another group published work on the generation of *in silico* models of *Pf*PMs I, V-X (Guruprasad, Tanneeru *et al.*, 2011). Guruprasad et al. (2011) used a truncated *Pf*PM X sequence, starting just 33 amino acids before the first catalytic aspartic residue (Guruprasad, Tanneeru *et al.*, 2011) for their model prediction. Given only one of the online structure prediction applications used in this study was able to generate a complete *Pf*PM X preliminary structure, this N-terminal truncation, as is done for all successfully expressed recombinant DV PMs (see Chapter 3), may have aided template matching. Additionally, research in our laboratory has shown that *Pf*PM X is N-terminally processed (Chapter 5), resulting in a mature protein of ~ 42 kDa in size. Despite the potentially beneficial aspects of their approach, Guruprasad et al. (2011) neglected to perform free energy minimisations or molecular dynamic simulations on their structure. Additionally, there appeared to be no validation performed. Without access to their *Pf*PM X model, direct comparisons between the two generated *in silico* models could not be performed. However, while using a truncated sequence may have been beneficial, without further minimisation and refinement the structure may not accurately resemble the protein environment under physiological conditions.

To aid in validating the *Pf*PM X model and to investigate *Pf*PM X as a target of the HIV PIs, *in silico* docking of a HIV PI library was performed. Failure of these inhibitors to dock within the binding cleft would suggest an inaccurate model as current data strongly suggests *Pf*PM X is a target of the HIV PIs (discussed Chapter 1 and 5). While Glide docking scores cannot be directly compared between systems, the favourable numerical values obtained in our study (Table 6.2) were encouraging and resembled those achieved when a HIV Protease model is used. This is the protease that the HIV PIs are designed to inhibit (as reviewed by De Clercq, 2009). Additionally, all inhibitors docked within the binding cleft of *Pf*PM X, with the top binders interacting directly with the catalytic aspartic residues (Figure 6.5). These data provide further support that *Pf*PM X is targeted by the HIV PIs and suggests the generated model may be suitable for *in silico* screening to identify potential inhibitors of this enzyme.

In silico screens of four chemical compound libraries revealed top binding hits possessing a wide variety of molecular structures (Tables 6.3-6.6; Figures 6.7; Appendix 8). While several molecules possess structures similar to HIV PIs and general aspartic protease inhibitors, a significant number of docking hits had increasingly

divergent and complex structures. This suggests that the binding cleft of the *Pf*PM X homology model may be fairly non-specific. All top binders from the GlaxoSmithKline library had IC_{50}s in the low μM range when tested against *P. falciparum* (Gamo, Sanz *et al.*, 2010). Interestingly, the second ranked docking hit from a screen of the NCI library was glucosulfone (Table 6.6, Figure 6.7), a component in the drug Promin that has been investigated for anti-malarial, anti-tuberculosis and anti-leprosy efficacy (as reviewed by Barr, 2011). Promin is metabolized into the well-known antifolate dapsone, however like primaquine, this metabolite has adverse indications for those with glucose-6-phosphate dehydrogenase deficiency (GlaxoSmithKline, 2008, Barr, 2011). Unfortunately the target of Promin within these organisms remains unknown. Despite a large number of docking hits being identified, many of the lead compounds from these screens were either not commercially available or not financially viable for the budget of this project. Despite this, five compounds from the Sigma-Aldrich and NCI libraries (three and two respectively), that had relatively good docking scores were obtained and investigated for their anti-malarial efficacy *in vitro*.

In vitro growth inhibition assays against *P. falciparum* demonstrated that NFmoc and L-His had anti-malarial activity, achieving IC_{50} values of 14.2 ± 2.4 μM and 19.6 ± 2.4 μM respectively (Table 6.7; Figure 6.8). While not in the nM range, these data were encouraging. The remaining compounds (L-alanine β-napthylamide, neohesperidin and myricetin) failed to inhibit parasite growth by more than 23%, even at drug concentrations of 100 μM (Table 6.7). However, it is important to remember that the docking scores of these compounds are only a measure of the ligand binding within the model and don't take into account other very important factors when using whole cell assays such as poor cellular transport, *in vitro* chemical modification, or cellular excretion.

In order to further investigate if the parasite target of NFmoc and L-His is *Pf*PM X, the sensitivity of transgenic cultures containing plasmids designed to over-express *Pf*PM X were examined. While parasites on 50 nM drug selection were determined to be less sensitive to L-His than D10 wild-type parasites (Figure 6.9), this difference was not statistically significant. Analysis of NFmoc activity also failed to demonstrate a significant difference in sensitivity of transgenic parasites as compared to wild-type controls (Figure 6.9). These transgenic parasites have been shown to be significantly

protected against selected HIV PI treatment (Seidens, 2010) and these results suggest *Pf*PM X may not be the *P. falciparum* target of NFmoc and L-His.

To examine the ability of docking hits to inhibit *Pf*PM X, HIV PIs and selected inhibitors were investigated in enzyme activity assays. In contrast to Glide docking score rankings, lopinavir, saquinavir and ritonavir significantly inhibited enzyme activity at 10 μM while higher ranked HIV PIs, nelfinavir and tipranavir, demonstrated no inhibition (Figure 6.10). This anti-*Pf*PM X activity is similar to the activity of the HIV PIs against *P. falciparum in vitro*, with lopinavir, saquinavir and ritonavir all exhibiting higher levels of growth inhibition as compared to nelfinavir and tipranavir (Table 6.2). In addition to this, top binding hits that were tested *in vitro* had no anti-*Pf*PM X activity at 10 or 20 μM (Figure 6.11). While this does suggest the *Pf*PM X model may not be optimal and needs further refinement, only a small proportion of top docking hits were investigated experimentally. A wider investigation of top docking hits may identify inhibitors with increased anti-plasmodial activity. Additionally, further refinement of the generated *Pf*PM X homology model including incorporating the N-terminal truncation performed by Guruprasad et al (2011) may prove beneficial.

An alternative approach to *in silico* screening with a target protein model derived from experimental or *in silico* prediction methods is pharmacophore modelling and screening (Leach, 2001). Given the above results, this was briefly investigated as an alternative in identifying inhibitors of *Pf*PM X. Without active recombinant *Pf*PM X (Chapter 5) and only small amounts of transgenic *Pf*PM X protein from GFP pull downs (Chapter 5), the generated pharmacophore was based on common features of the HIV PIs and docking data with the *Pf*PM X homology model. While not ideal, all HIV PIs used in the generation of the pharmacophore had known anti-malarial activity (Table 6.2). A screen of the Sigma-Aldrich library with the generated pharmacophore identified a number of inhibitors with the same common features (Table 6.8; Figure 6.13). Due to time constraints these could not be investigated further with *in vitro* or enzyme activity assays, of note, however, is the significantly reduced computational time this process involved compared to *in silico* screens with the *Pf*PM X homology model. A pharmacophore model derived from ligands with *Pf*PM X enzyme inhibition would significantly expedite the search for potent and selective inhibitors of *Pf*PM X.

While the *Pf*PM X homology model generated in this study may not be optimal, it was successfully used to identify two inhibitors with anti-malarial activity *in vitro* and importantly, determine that selected HIV PIs do potently inhibit *Pf*PM X. This inhibition of *Pf*PM X provides further support to the hypothesis that *Pf*PM X is a good drug target for future development of anti-malarial chemotherapeutics. *In silico* docking has been shown to be effective at identifying potent and selective inhibitors (Bouillon, Giganti *et al.*, 2013, Degliesposti, Kasam *et al.*, 2009) and with further investigation, lead inhibitors with anti-*Pf*PM X activity may be identified.

Chapter 7: Discussion and Conclusion

All of our current antimalarials are being gradually rendered ineffective by spreading drug resistance. Of particular concern is the emerging resistance to artemisinin combination therapies (ACT), our most potent antimalarials (Dondorp, Nosten *et al.*, 2009, Noedl, Socheat *et al.*, 2009, Noedl, Se *et al.*, 2008). Therefore there is an urgent need for the identification of new and unique targets within the malaria parasite against which novel chemotherapeutics can be developed. In this thesis, the *P. falciparum* aspartic proteases PMs IX and X were investigated as potential new targets for drug development. Studies to identify potential inhibitors of these enzymes and examine their antimalarial efficacy were also conducted.

Data generated in studies examining the role of *Pf*PM IX revealed that this enzyme is present throughout the intra-erythrocytic asexual stages of *P. falciparum* development and is likely to have one or more vital function/s distinct from other characterised *Pf*PMs. While *Pf*PM IX is expressed in all asexual stages it is primarily transcribed and translated during the late schizont stage of development *in vitro* (Figures 2.2 and 2.3). Immunochemistry data indicate that *Pf*PM IX undergoes a number of processing events throughout the asexual life cycle, with the smallest and most mature form being present in trophozoite stages, but being predominantly found in schizont stages (Figure 2.3). These data suggest that *Pf*PM IX has a primary role in schizonts. Consistent with the presence of a PEXEL motif, IFA data demonstrate that *Pf*PM IX is exported into the host red cell. Localisation data suggest that *Pf*PM IX is deposited in the host cell in a punctate fashion. Co-localization with the Maurer's Cleft protein REX 1 also suggests transient co-localisation to these parasite derived structures (Figure 2.4). *Pf*PM IX is the only PM that has a PEXEL motif and is thus far the only reported *Pf*PM to be exported in to the host red cell. Further investigations into the transport and localisation of *Pf*PM IX within the host red cell cytosol using MAHRP2 (a Maurer's Cleft tether protein (Pachlatko, Rusch *et al.*, 2010)), and HSP 70-x (which localises to J-dots and is involved in chaperones with HSP40 (Kulzer, Charnaud *et al.*, 2012)), were attempted, but were unsuccessful (data not shown). Assays with new aliquots of antibody (supplied by Dr Hans-Peter Beck (Swiss Tropical and Public Health Institute) and Associate Professor Alex Maier (The Australian National University) respectively) may prove successful and provide a better insight into the location and role of *Pf*PM IX within the red cell. *Pf*PM IX was also found within parasites, with these data suggesting location in the parasite cytoplasm and within merozoites in segmenting schizonts (Figure 2.4). Interestingly, and though co-localisations with *Pf*DV PMs should be conducted to

confirm, at no stage was *Pf*PM IX found in the DV of parasites, again suggesting that this enzyme has a role distinct from the well-characterised *Pf*DV PMs I-IV.

*Pf*PM IX knock-out and knock-down experiments suggested that this enzyme has an important role in parasite survival. Knock-out experiments to genetically disrupt *Pf*PM IX were not successful, despite multiple attempts. Though not confirmatory, this suggests that *Pf*PM IX may have vital roles within the malaria parasite and supports the hypothesis that this enzyme is a good drug target. Experiments to knock-down *Pf*PM IX using gene targeted peptide nucleic acids (PNAs), although preliminary, also suggested that *Pf*PM IX is important to parasites, with reduced transcription resulting in a significant decrease in parasitaemia (Figure 2.7). PNA treated cultures also displayed an altered morphology with a distinct enlargement of the DV (Figure 2.8). Whilst non-specific PNA effects may have complicated these data, including control cultures starting to become overgrown in later stages of the assay, the reduced growth of cultures treated with *Pf*PM IX specific PNAs suggested this observation may have been *Pf*PM IX specific. Further studies which include determining the amount of active enzyme present will be required to confirm this. While there is currently no enzyme specific assay available to quantitate the activity of *Pf*PM IX, and Western blot data is rarely accurately quantifiable (Gassmann, Grenacher *et al.*, 2009), PNAs have recently been validated as a tool to investigate protein function in malaria parasites (Kolevzon, Nasereddin *et al.*, 2014).

Although the function of *Pf*PM IX remains unknown, it is clear that *Pf*PM IX has one or more function/s distinct from the already well-characterised *Pf*DV PMs and could be an unexploited target for drug development. The bulk transcription and translation within schizont stage parasites, in combination with the localisation within segmenting merozoites, strongly suggests a role in egress and/or invasion. The export of *Pf*PM IX into the host red cell also supports this idea. Further co-localisation studies are essential to determine *Pf*PM IX's role within the host red cell. *P. falciparum* extensively modifies the host red cell through a number of exported structures. The punctate patterning and transient co-localisation with the Maurer's Clefts may indicate that *Pf*PM IX is associated with Maurer's Clefts or their tethers. Alternatively, *Pf*PM IX may simply be redirected to other exported structures via the Maurer's Clefts. It is also possible that *Pf*PM IX is directed to the J-dots within the host red cell. J-dots have been shown to exhibit a similar punctate localisation and are believed to play a role in protein

trafficking (Kulzer, Charnaud *et al.*, 2012, Kulzer, Rug *et al.*, 2010). As discussed in Chapters 1 and 2, assays with transgenic parasites containing plasmids designed to over-express *Pf*PM IX were protected from HIV PI treatment. Combined with evidence that *Pf*PM IX has a potentially unexploited and essential function within the malaria parasite these data suggest that *Pf*PM IX is an excellent target for future drug development.

To further characterise *Pf*PM IX, a recombinant *Pf*PM IX was expressed in *E. coli*. This enzyme was N-terminally truncated and while active recombinant enzyme could be obtained, it was at yields of <2% of inclusion body preparation from expression culture (Table 3.1). Such low yields prohibited further characterization studies. However, studies with less purified material found that refolding at pH 7.0 produces greater yields of active enzyme than refolding at pH 8.0 (Figure 3.9). Additional truncations of the *Pf*PM IX sequence, in conjunction with codon optimisation for expression in *E. coli*, did not demonstrate any marked increase in expression yields (Table 3.2). During the course of these studies, preliminary studies describing the expression and purification of *Pf*PM IX were published (Olajuyigbe, 2013). This work was guided by the research completed in this thesis and achieved greater yields of active enzyme. Incorporation of Olajuyigbe's (2013) modifications to refolding conditions and the inclusion of DNAse during the purification process should be pursued in future studies.

In the absence of crystal structure data and in an effort to identify specific inhibitors of *Pf*PM IX, attempts were made to generate an *in silico* homology model. Although a *Pf*PM IX model was generated using well documented automated model-building approaches and possessed a high degree of structural similarity to several well characterised aspartic proteases, critical analysis of this model revealed serious omissions in the protein primary sequence (Figure 4.7). As a result of this flaw the model was not used in further *in silico* studies. The deletion of a low complexity region, composed of a peptide sequence not seen in any other species, may have also given rise to the high degree of structural similarity observed between the active sites of the *Pf*PM IX model and other aspartic proteases structures. As this region is unique to *Pf*PM IX, it is thought that it was removed due to its lack of homology to any other known aspartic protease structures. In an attempt to solve this dilemma, *de novo* modelling of *Pf*PM IX was investigated by Dr Bela Stantic at Griffith University. Unfortunately the size of *Pf*PM IX and its complex folding meant this was also unsuccessful. These observations made it clear that *in silico* modelling is unlikely to be able to produce an accurate and

reliable homology model of *Pf*PM IX for further investigation. Without a reliable template structure, the only way forward in terms of understanding the structure of *Pf*PM IX and screening for inhibitors *in silico* is to increase yields of recombinant enzyme and determine the protein solid-state crystal or solution-state structure.

Investigations into the location, expression and function of *Pf*PM X supported the hypothesis that this enzyme is a good drug target for chemotherapeutic development. Similar to *Pf*PM IX, *Pf*PM X appears to perform a unique function within malaria parasites. *Pf*PM X is present in trophozoite and schizont stages of intra-erythrocytic *P. falciparum* development with no detectable expression in ring stage parasites (Figure 5.3). Similar to *PfPM IX*, the vast majority of transcription takes place in the later schizont stages (Figure 5.2). Immunochemistry data suggests that similar to *Pf*PM IX (Chapter 2) and the *Pf*DV PMs (Banerjee, Liu *et al.*, 2002, Kim, Lee *et al.*, 2006, Dame, Yowell *et al.*, 2003), *Pf*PM X is expressed as a zymogen and undergoes processing of the pro-region on the N-terminus to form a mature *Pf*PM X enzyme of ~42 kDa (Figure 5.3). Similar to *Pf*PM IX, data suggest *Pf*PM X undergoes multiple processing events, all of which appear to occur in all stages in which it is expressed. IFA data indicate that *Pf*PM X localises to the parasite cytoplasm (Figure 5.4). In contrast to what has been found previously (Moura, Dame *et al.*, 2009), *Pf*PM X does not appear to localise exclusively to any specific organelles, instead having a more diffuse localisation in trophozoite stages. Due to time constraints this could not be investigated further, however, co-localisation studies with *Pf*DV PMs would confirm our findings. Additionally, co-localisations with *Pf*PM V would determine if *Pf*PM X is trafficked to the endoplasmic reticulum at any stage as IFA data demonstrated regions of increased signal intensity within the parasite cytoplasm (Figure 5.4).

The importance of *Pf*PM X to parasite survival requires further investigation. While PCR analysis suggests that 3' integration of the knock-out plasmid may have occurred, it was not seen in cultures that underwent further drug cycling (Figure 5.6). There was also no evidence of 5' integration at any stage. This suggests that the initially detected 3' integration event may have been lethal or of such a disadvantage that parasites did not survive in the mixed culture. Examination of the gene product that would have resulted from a 3' integration event suggests this would have resulted in disruption of the gene immediately after the first catalytic site. Combined with additional vector sequence, it is unlikely this could have produced a functional protein. Experiments to

knock-down *Pf*PM X in the form of generating and transfecting an anti-sense plasmid were conducted. These studies were successful in reducing *PfPM X* transcription (Table 5.2; 15%) and while no significant morphological changes were observed in these parasites, growth appeared to be slower than the wild-type parasites, suggesting that *Pf*PM X may potentially be important to parasite growth. These data further support the hypothesis that *Pf*PM X may have a role in parasite egress and/or host red cell invasion. Additional knock-down experiments, including increasing the level of drug selection on transgenic parasites, could further elucidate the effect of knocking-down *PfPM X*.

To further characterise *Pf*PM X attempts were made to express a recombinant enzyme in bacteria. The expression construct contained a full-length *Pf*PM X insert, codon-optimised for expression in *E. coli*. Unfortunately these experiments failed to produce active recombinant enzyme. As discussed in Chapter 5, reasons for this may have been associated with the difficulties encountered in purifying r*Pf*PM X as increased yields would have allowed further investigations into refolding. This inactivity may also have been the result of the full-length construct given auto-catalysis may not occur with this enzyme. Studies to examine the effect of N-terminally truncating r*Pf*PM X were instigated but could not be completed.

While the current studies did not did not deduce the function of *Pf*PM X within intra-erythrocytic stages, data suggest that this enzyme has distinct role/s from the *Pf*DV PMs. *Pf*PM X was not localised to parasite DVs in our studies. Additionally, the bulk transcription of *Pf*PM X in schizont stage parasites and the reduced growth seen with knock-down experiments suggests an important role outside of haemoglobin digestion which occurs preferentially in trophozoites (Aikawa, Hepler *et al.*, 1966, Abu Bakar, Klonis *et al.*, 2010). As speculated by others, our data support a role in parasite egress and or invasion. However, further investigations examining the effect of knock-down on merozoite invasion is required to provide further insights into this possibility. As with *Pf*PM IX, assays with transgenic parasites containing a plasmid designed to over-express *Pf*PM X were protected from treatment by ritonavir (discussed in Chapters 1 and 5). In conjunction with evidence that *Pf*PM X has a potentially important function, these data indicate that *Pf*PM X may be a promising potential drug target for future antimalarial agent development.

To identify lead inhibitors of *Pf*PM X and in the absence of any known crystal structure, an *in silico* homology model of *Pf*PM X was generated. This model shared a high degree of structural similarity to other well-characterised aspartic proteases (Figure 6.4) and demonstrated no major structural abnormalities (Figure 6.3). An *in silico* dock of a HIV PI library demonstrated that all inhibitors bound within the binding cleft of the model, with the top binding inhibitors interacting directly with the catalytic aspartic residues. *In silico* screens with four chemical compound libraries identified a series of compounds predicted to bind to the active-site of *Pf*PM X. While several of these compounds had structures similar to HIV PIs and other known broadly active aspartic protease inhibitors, many also exhibited divergent and complex structures. Although docking data must be assessed with some caution, these results suggested that the binding cleft of the *Pf*PM X may be able to accommodate a wide variety of inhibitors. Encouragingly, all of the compounds with high docking scores from the GlaxoSmithKline library have reported *P. falciparum* EC_{50}s in the low micro-molar range (Table 6.3; (Gamo, Sanz *et al.*, 2010)).

Many of the best docking hits could not be obtained for further studies (discussed in Chapter 6). However five compounds from the Sigma-Aldrich and NCI libraries (three and two respectively) that possessed good docking scores, were obtained commercially and investigated for *P. falciparum* inhibition *in vitro*. Two of these compounds achieved IC_{50}s <20 μM (Table 6.7; Figure 6.8). The poor antimalarial activity of these compounds may have been associated with biological factors such as transport problems (i.e. the ability of the compounds to cross membranes) not assessed in the docking studies. However, it is also possible that the *Pf*PM X model may not be accurate. Interestingly, although the antimalarial activity of the two active compounds was poor, assays with transgenic parasites containing an over-expression plasmid revealed no significant difference in drug sensitivity compared to the wild-type comparison. This suggests that while L-histidine β-napthylamide and *N*-Fmoc-ethylenediamine hydrobromide exhibit antimalarial activity *in vitro*, *Pf*PM X may not have been their primary target.

To investigate the ability of docking hits to inhibit *Pf*PM X, enzymatically active C-terminally tagged *Pf*PM X was extracted from transgenic *P. falciparum* cultures (Figure 5.12). While not viable for large scale inhibitor assays and problems with protein stability meant extracted enzyme needed to be used almost immediately, this approach

did allow some scope for the investigation of enzyme inhibition by selected lead docking hits in the absence of active recombinant enzyme. In contrast to Glide docking score rankings, 10 μM solutions of lopinavir, saquinavir and ritonavir significantly inhibited enzyme activity while higher ranked HIV PIs, nelfinavir and tipranavir, demonstrated no inhibition (Figure 6.10). This anti-*Pf*PM X activity is similar to the activity of the HIV PIs against *P. falciparum in vitro*, with lopinavir, saquinavir and ritonavir all exhibiting higher levels of growth inhibition than nelfinavir and tipranavir (Table 6.2). Additionally, the lead docking hits tested *in vitro* exhibited no anti-*Pf*PM X-GFP activity at 10 or 20 μM. While these data suggest the generated *Pf*PM X homology model may not be optimal, a very small portion of top docking hits were investigated experimentally. Importantly however, activity data with the HIV PIs demonstrate that lopinavir, saquinavir and ritonavir do inhibit *Pf*PM X and this taken together with data demonstrating that parasites over-expressing *Pf*PM X are less sensitive to selected HIV PIs, suggest that this enzyme may be the/an antimalarial target of HIV PIs and support the idea that *Pf*PM X is important to asexual parasite development. A wider investigation of docking hits may identify additional compounds with increased anti-plasmodial and anti-*Pf*PM X activity.

Given the above results, pharmacophore modelling was briefly investigated as an alternative for identifying potent *Pf*PM X inhibitors. A pharmacophore model was generated based on the common features of the HIV PIs and docking data with the *Pf*PM X homology model. While not ideal, all HIV PIs used in the generation of the pharmacophore had known anti-malarial activity *in vitro* (Table 6.2). The generated pharmacophore identified a number of inhibitors from the Sigma-Aldrich library that shared common features (Table 6.8; Figure 6.13) and while these could not be investigated further, due to time constraints, it did demonstrate the significantly reduced computational time involved in this screening process when compared to the docking *in silico* screens. A practical pharmacophore model derived from ligands with *Pf*PM X enzyme inhibition data may significantly expedite the search for potent and selective inhibitors of *Pf*PM X.

A number of difficulties were encountered during this project. Characterisation of *Pf*PM IX was extremely challenging and hampered by the N and C-terminal processing that occurs throughout the intra-erythrocytic parasite life cycle. Without fluorescent tags, localisation studies were dependent on *Pf*PM IX antisera. This antisera was far from

ideal and required extensive processing to remove non-specific binding and extensive optimisation for Western blot experiments. Additionally, while *Pf*PM IX and *Pf*PM X share low sequence identity with other *Pf*PMs, each respective antisera may have cross-reacted, leading to artefacts in Western blot experiments. With a limited number of recombinant *Pf*PMs available for cross-reactivity screening, reactivity to other *Pf*PMs should be assessed in future studies by proteomic analysis of material immuno-precipitated with the generated antisera.

Unfortunately, experiments to identify potent inhibitors of *Pf*PM IX were unsuccessful. Attempts to generate a *Pf*PM IX homology model for inhibitor library screening were complicated by the presence of hard to model low complexity regions within the protein primary sequence. These regions were not amenable to structure prediction and resulted in the applications being used producing inaccurate structures with a key section of the protein's sequence missing. Attempts to generate a 3D structure of *Pf*PM IX through X-ray crystallography were also unsuccessful. Whilst some recombinant *Pf*PM IX enzyme was generated and used in a small crystallisation trial, it failed to produce any protein crystals suitable for solid-state X-ray analysis. Unfortunately yields of active recombinant *Pf*PM IX were insufficient to allow large scale crystallisation trials with an increased number of conditions.

Although a homology model of *Pf*PM X was successfully generated and a number of top docking inhibitors identified, the vast majority of these were not available for further study. Of the compounds that were commercially available, many of them were far in excess of what was financially feasible for this study. The production of active recombinant *Pf*PM X for validation of the *in silico* model with enzyme inhibition studies and the production of an X-ray crystal structure also encountered several difficulties. While soluble full-length recombinant enzyme could be expressed and purified, there were ongoing problems with enzyme activity. Unfortunately due to time constraints, this could not be investigated thoroughly.

Although the research presented here has increased the knowledge of *Pf*PMs IX and X, there is much that can be done in future studies. While data strongly suggests that *Pf*PM IX is essential to parasite survival, and another group has recently generated conditional knock-outs supporting this hypothesis (Soldati-Favre, 2013), further investigation is needed to determine *Pf*PM X's importance to parasite survival. Given that it is

impractical to confirm gene essentiality through knock-out techniques, with parasites unable to survive disruption of essential genes, a technique such as the Shld system (Armstrong and Goldberg, 2007) may be more successful. Further studies with knock-down techniques, such as PNAs and other antisense technologies as previously discussed may also aid in the determination of *Pf*PM IX and X's function/s. Though altered morphology was evident after *Pf*PM IX specific PNA treatment, additional studies are required to control for non-specific PNA effects. Further examination of transgenic parasites containing a *Pf*PM X antisense vector should be conducted to determine if knock-down of *PfPM X* has an effect on merozoite egress and invasion.

Further investigations into *Pf*PM IX trafficking through fluorescent tags may provide more insights into this enzymes localisation in the host red cell. A series of constructs containing different length inserts of *Pf*PM IX with a C-terminal GFP tag, or a full length *Pf*PM IX insert with an internal GFP tag, would allow further assessment of *Pf*PM IX trafficking and processing throughout the intra-erythrocytic life cycle.

Investigation of *Pf*PM IX and X localisation and expression in other Plasmodium species may also shed more light on their respective functions. As discussed in Chapter 4, *Pf*PM IX orthologs in species infecting mice contain a transmembrane domain that is not seen in human infecting species. While these have not been experimentally confirmed, this suggests that *Pf*PM IX is a dynamic protein with potentially multiple functions. Investigation of the role of PM IX in species more amenable to genetic manipulation, such as *P. berghei*, may shed more light on the potential roles of *Pf*PM IX. Additionally, confirmation of the expression and localisation of these enzymes in other *P. falciparum* life cycle stages, such as gametocytes, should be conducted. Enzymes with roles in a broad range of life cycle stages are promising as drug targets, as they could facilitate the development of drugs that have activity against disease, transmission and relapse.

As discussed in Chapter 3, inclusion of the conditions used in the preliminary study of expression and purification of recombinant *Pf*PM IX by Olajuyigbe (2013) may aid in increasing yields of active recombinant *Pf*PM IX (Olajuyigbe, 2013). Future studies to express recombinant *Pf*PM X should utilise an expression vector that does not contain a GST tag. Additionally, N-terminally truncating *Pf*PM X to ~30 residues before the first

catalytic aspartic residue may produce active enzyme, particularly if auto-catalysis does not occur.

Continued *in silico* screens with the *Pf*PM X homology model may identify additional compounds for further study. Additionally, should an X-ray crystal structure of *Pf*PM IX become available, *in silico* screens for inhibitors of *Pf*PM IX should be conducted and top-docking hits tested further for antimalarial efficacy *in vitro*.

With the ever increasing spread of drug resistant parasites, it is imperative that new drug targets be identified across a broad range of parasite stages. This thesis has sought to characterise *Pf*PM IX and X and increase the knowledge available on these potentially important *P. falciparum* enzymes. While this research was not successful in identifying the function of these enzymes and has not identified inhibitors with *in vitro* IC_{50} values in the low nM range, it has confirmed *Pf*PM IX and X's potential as new targets for antimalarial chemotherapeutic development with data supporting the hypothesis that these enzymes play an important role in parasite growth and development. The tools and resources developed during this project will also continue to be useful for further studies investigating the role of these and other PMs in malaria parasites. Inhibitors designed to target *Pf*PMs IX and X may become potent and selective antimalarial agents that are not only effective against multi-drug resistant parasite, but also effective in preventing disease and transmission.

Appendices

Appendix 1: Solutions

2YT

Triptone	80 g
Yeast Extract	50 g
NaCl	45 g

Made up to 5 L in Millipore H_2O and autoclaved. Media stored at 4 °C.

10% APS (Ammonium persulfate)

Ammonium persulfate	10 g
Millipore H_2O	100 ml

This was aliquoted and stored at -20 °C.

5% Blocking solution:

Skim-milk powder	5 g
PBS pH 7.2-7.4	100 ml

Complete culture medium

RPMI-Hepes 1640 incomplete medium (see below)	85 ml
Normal Human Serum (heat inactivated)	10 ml
5% Sodium Bicarbonate Solution	4.2 ml
Gentamycin (5 mg/ml)	200 µl

Media was wrapped in foil and stored at 4 °C.

Coomassie Blue Stain:

10% Acetic Acid

0.25% Coomassie Brilliant Blue R250

50% Methanol

The stain was dissolved, and then filtered with Whatman number 1 filter paper

Cytomix:

120 mM KCL

0.15 mM $CaCl_2$

10mM K_2HPO_4/KH_2PO_4 (at pH 7.6)

25 mM Hepes (at pH 7.6)

2 mM EGTA

5 mM MgCl2

DNA loading buffer (6X)

Bromophenol blue	0.05 g
Xylene cyanol	0.05 g
Glycerol	6 ml
Millipore H_2O	14 ml

Aliquot and store at -20 °C.

Ethidium Bromide (EtBr) 1 mg/ml

EtBr	100 mg
Millipore H2O	100 ml

Freezing Solution:

3% Sorbitol

0.65% NaCl

28% Glycerol

Make up in H_20 and filter sterilize. Store at 4 °C.

5M Imidazole

Imidazole	34.04 g

Made up to 100 ml with Millipore H_2O

LB

Tryptone	10 g
Yeast Extract	5 g
NaCl	10 g

Make up to 1 litre with Millipore H_2O and adjust to pH 7.0. Autoclave and store at 4 °C.

10X Ni-NTA binding buffer

200 mM Tris pH 8.0

3 M NaCl

100 mM imidazole

Make up to desired volume in H_2O and store at room temperature.

5x Ni-NTA elution buffer (500 mM)

10x Ni-NTA binding buffer	50 ml
5M Imidazole	50 ml

Ni-NTA wash buffer 2

1x Ni-NTA elution buffer	4 ml
1x Ni-NTA binding buffer	96 ml

Ni-NTA wash buffer 3

1x Ni-NTA elution buffer	12.5 ml
1x Ni-NTA binding buffer	87.5 ml

Ni-NTA wash buffer 4

1x Ni-NTA elution buffer	21 ml
1x Ni-NTA binding buffer	79 ml

10X Phosphate buffer saline (PBS)

NaCl	80 g
KCl	2.0 g
KH2PO4	2.4 g
Na2HPO4	14.4 g

All components were dissolved in 800 ml Millipore H_2O, the pH was adjusted to 7.4 and the volume adjusted to 1 l with additional Millipore H_2O.

Working solution (1X PBS) was made up with Millipore H_2O.

PBST

1x PBS	999 ml
Tween-20	1 ml

PreScission cleavage buffer

50 mM Tris pH 7.0

150 mM NaCl

1 mM EDTA

1 mM DTT

Make up to desired volume in H_2O and store at room temperature.

Resuspension Buffer A

0.01M Tris pH 8.0

0.02M $MgCl_2$

0.005M CaCl

Make up to desired volume in H_2O and store at room temperature.

Resuspension Buffer B

0.01M Tris pH 8.0

0.001M EDTA

0.002M BME

0.1M NaCl

Make up to desired volume in H_2O and store at room temperature.

Resuspension Buffer C

0.05M Tris pH 8.0

0.005M EDTA

0.005M BME

0.5% Triton X-100

Make up to desired volume in H_2O and store at room temperature.

Resuspension Buffer D

0.05M Tris pH 8.0

0.005M EDTA

0.005M BME

Make up to desired volume in H_2O and store at room temperature.

RPMI-Hepes

RPMI 1640 powder	20.8 g
Millipore H_2O	1.6 l
When dissolved, add Hepes	11.88 g

Adjusted pH to 6.75 at room temperature, adjusted volume to 2 litres, filter sterilised (0.22 μm) and stored it at 4°C in 1 litre aliquots wrapped in aluminium foil.

10% sarcosyl in STE

Sarcosyl 10 g

Make up to 100 ml in STE buffer (see below).

10% Sodium dodecyl sulphate (SDS)

SDS 100 g

Millipore H2O 1 litre

The solution was heated to 68°C, the pH adjusted to 7.2 with HCl and stored at 4 °C.

SDS Loading Buffer (10%):

1M Tris-Cl (pH 6.8) 10 ml

Dithiothreitol 3.08 g

Sodium dodecyl sulphate 2 g

Bromophenol Blue 0.2g

Make up to 10 ml in Glycerol and store at 4 °C.

50 mM sodium phosphate dibasic buffer

Sodium phosphate dibasic 28.4 g

Make up to 4 litre with H_2O and adjust to required pH.

12% SDS polyacrylamide separating gel:

H_2O 4.335 ml

30% acrylamide 5.2 ml

10% SDS 130 µl

1.5 M Tris-HCl pH 8.8 3.25 ml

10% APS 65 µl

TEMED 6.5 µl

SDS polyacrylamide stacking gel:

H_2O 3 mL

30% acrylamide 660µL

0.5 M Tris-HCl pH 6.8 1.26 ml

10% APS 25 µL

TEMED 5µL

SDS-PAGE destaining solution

80% H2O

10% acetic acid

10% methanol

SDS Running Buffer (10X):

10% SDS	10 ml
Glycine	14.4 g
Tris-base	3.03 g
Methanol	200 mL
H2O	800 mL

5% D-Sorbitol:

D-Sorbitol	5 g
Millipore H_2O	100 ml

Stir at room temperature to dissolve

Filter sterilize and refrigerate

STE buffer

100 mM Tris pH 8.0

1500 mM NaCl

10 mM EDTA

Make up to desired volume and store at room temperature.

Super-Broth Media

Tryptone	32 g
Yeast extract	20 g
Sodium chloride	5 g

Make up to 1 litre in Millipore H_2O and bring to pH 7.0 with NaOH. Autoclave and store at 4 °C.

1XTAE buffer:

10 M Tris

1 mM EDTA

Bring to pH 8

1 L H_2O

TE:

10 mM Tris

1 mM EDTA

Make up to 1 litre in H2O and autoclave. Store at room temperature.

1x Transfer Buffer:

Tris	3.03 g
Glycine	14.4 g
Methanol	200 ml

Make up to 1 litre in H_2O and store at room temperature.

8 M urea denaturing solution

Urea	128.12 g

Dissolve in 100 ml H_2O with heat

Amberlite	1 scoop, stir for 30 mins in fume hood

0.05 M Caps

0.005 M EDTA

B-mercaptoethanol	3.55 ml

Make up to 250 ml with H_2O

Appendix 2: Vector Maps

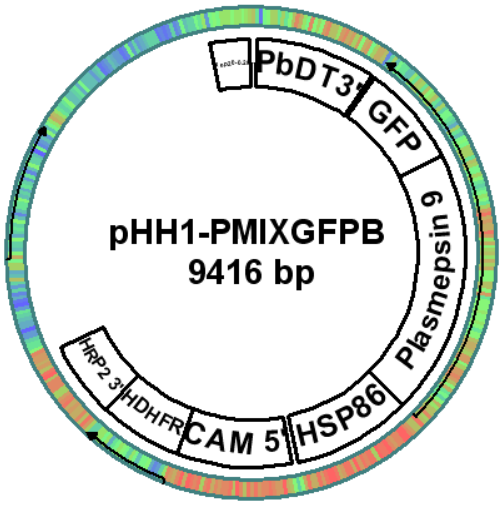


Figure A2.1: **pHH1-PMIXGFPB containing GFP tagged *PfPM IX* (McGeorge 2009).**

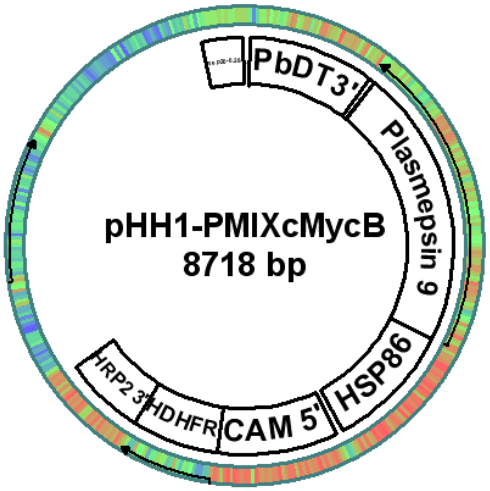


Figure A2.2: **pHH1-PMIXcMycB containing cMyc tagged *PfPM IX* (McGeorge 2009).**

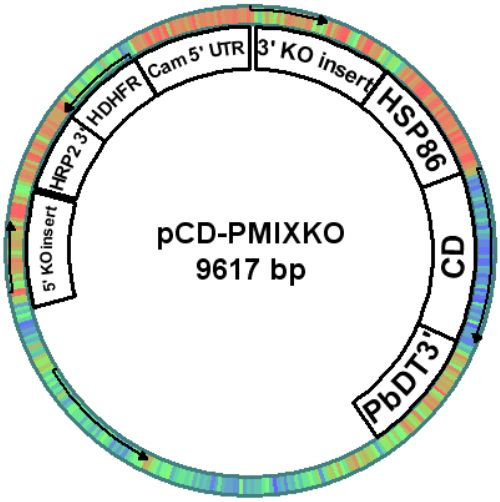


Figure A2.3: **Double cross-over knock-out transfection vector for targeted gene disruption of *PfPM IX* (McGeorge 2009).**

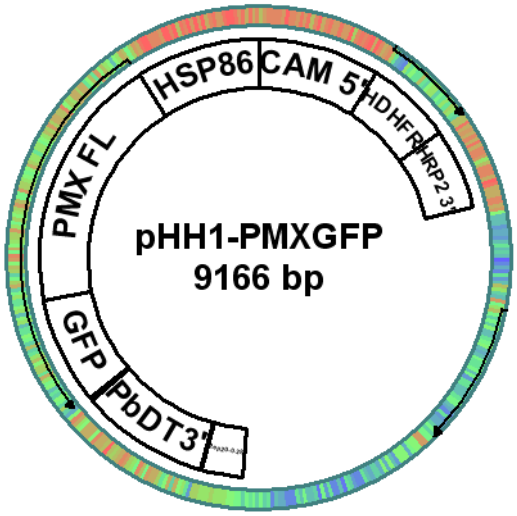


Figure A2.4: **pHH1-PMXGFP containing GFP tagged *PfPM X* (Seidens 2010).**

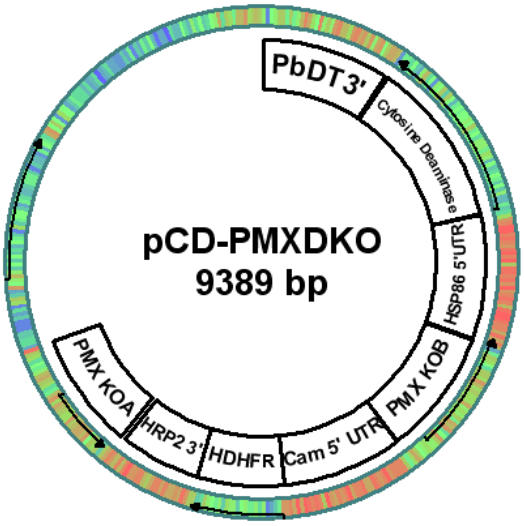


Figure A2.5: **Double cross-over knock-out transfection vector for the targeted gene disruption of *PfPM X* (Seidens 2010).**

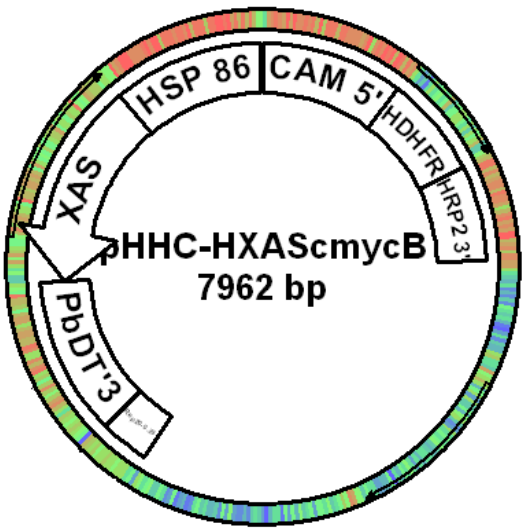


Figure A2.6: **Antisense transfection vector for the knock-down of *PfPM X* (Ljunggren 2006).**

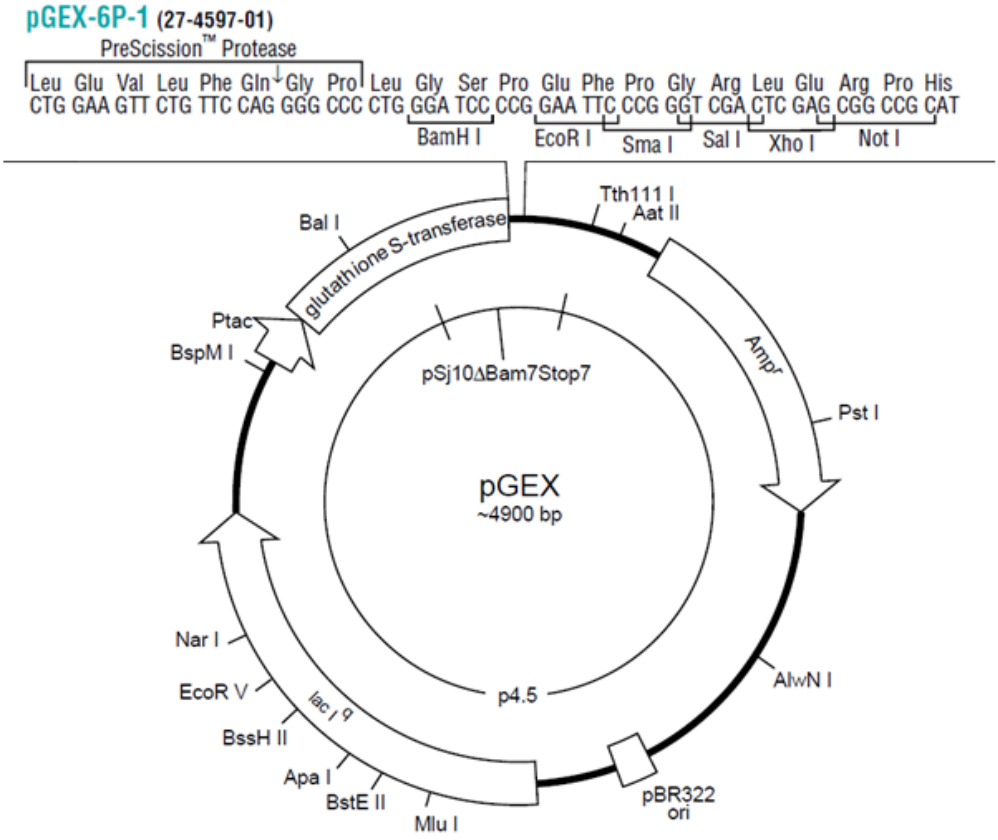


Figure A2.7: **pGEX-6P-1 expression vector (GE Healthcare).**

Appendix 3: Primers

Table A3.1: **Primers used and the respective sequences.**

Primer	Sequence
HSP86x	5’ – TATAAATATTTACATAAATATT – 3’
β-actin1 Forward (RT-PCR)	5’-AAA GAA GCA AGC AGG AAT CCA – 3’
β-actin1 Reverse (RT-PCR)	5’-TGA TGG TGC AAG GGT TGT AA-3
Seryl-tRNA synthetase Forward (RT-PCR)	5’- ATA GCT ACC TCA GAA CAA CC - 3’
Seryl-tRNA synthetase Reverse (RT-PCR)	5’- CAA GAT GAG AAT CCA GCG TA - 3
PfPM X Forward (RT-PCR)	5’- TTG TAT GCC TGC CTA CAT GC - 3’
PfPM X Reverse (RT-PCR)	5’- CTA CTT TCT GTG CCT CGT ACA A - 3’
PM IX Forward (RT-PCR)	5’-TCC AGA AAT GTT ATC GAC TGG-3’
PM IX Reverse (RT-PCR)	5’-CCC CTC CAA ATA TTA ATG CAG A-3’
18s rRNA Forward (RT-PCR)	5’-CGG CGA GTA CAC TAT ATT CTT A-3’
18s rRNA Reverse (RT-PCR)	5’-TTA GTA GAA CAG GGA AAA GGA T-3’
PM9Ctrunc	5’ – GGA TCC TTA TTA TTC GGT ATG AAC TGC TTT GG – 3’
PM9FL	5’ – GGA TCC TTA TTA CAG ATT ATT AAT TTT GTT ATG – 3’
PM9Ntrunc	5’ – GGA TCC AAC CAT ATG AAT AAA ATC AAA GAC G – 3’
PM9 BF	5’ – AGA TCT ATG TTT TTT ATA AAT TTT AAG – 3’
PM9 PR	5’ – CTG CAG TAA ATT ATT TAT TTT ATT ATG – 3’
PM X Forward	5’ – AGA TCT ATG AAA CGC ATT AGC CCT CTA – 3’
PM X Reverse	5’ – CTG CAG GTT TTT ACT TTT TGC TCT TGC – 3’
Pbdt Reverse	5’ – GCA CAC AAC ATA CAC ATT TTT ACA G - 3’
PMV Forward	5’ – AGA TCT ATG AAT AAT TAT TTT TTA AGG – 3’
PMV Reverse	5’ – CTG CAG TGT TGA TTC CTG TAT GGG AGA – 3’
hdhfr Forward	5’ – ACT GCA TCG TCG CTG TGT C – 3’
hdhfr Reverse	5’ – TGC CTT TCT CCT CCT GGA C – 3’

Appendix 4: *Plasmodium falciparum* Plasmepsin IX and X sequences

A4.1 *Pf*PM IX

A4.1.1 *PfPM IX* genomic sequence

ATGTTTTTTATAAATTTTAAGAAAATAAAAAAGAAACAATTTCCGATATATT
TAACTCAACATAGAATTATAACAGTGTTTCTTATTTTTATTTATTTTATAAAC
TTAAAAGATTGTTTCCATATAAATAATTCGCGTATATTAAGTGATGTAGATA
AACATAGAGGATTGTATTATAATATACCAAAATGTAACGTATGTCATAAGT
GTTCGATATGTACCCACGAAAATGGAGAGGCACAGgtagttcaaaaataatatgtattattaat
atgtacctatgtatgaactgaaaatataggtctataatacacaccattgtttatatatatatatatatatatttttttttttttttttttttt
ttgatagAATGTCATTCCTATGGTTGCTATACCGAGCAAGAGAAAACATATTCA
AGATATAAATAAGGAAAGGGAGGAAAACAAATATCCTTTACATATATTTGA
AGAAAAAGACATTTATAATAATAAAGATAATGTTGTTAAAAAAGAAGATAT
ATATAAATTAAGAAAAAAAAAAAAACAGAAAAAAAATTGTTTAAATTTTCT
TGAAAAGGATACAATGTTTTTGTCACCTTCACATGATAAAGAAACGTTTCAT
ATAAATCATATGAATAAAATAAAAGATGAAAAATATAAACAAGAATATGA
AGAAGAGAAAGAAATATATGATAATACAAATACATCACAAGAAAAAAATG
AAACAAACAATGAACAGAATCTAAATATTAATTTAATAAACAATGATAAAG
TGACATTACCCTTACAACAATTGGAAGATgtaatatataaaataataaattaaaatataaatataaata
aatataaatataaatatatataaatataaatatatataaatataaataaatataaatatatataaatataaatgaataatgttaagttatta
gtacgtatgaatactttgtatttttttttttttttccacctccttttcacagAGTCAATATGTTGGCTACATTCA
AATAGGAACTCCTCCACAAACCATAAGACCCATATTTGATACAGGAAGCAC
gtaacattttaacataaataaataaataaatatatatatatatatatatatatatatatatatatatatatatatatttatttatatgtgtgatttat
ttcattttatttttttttttttaaagAAATATATGGATTGTAAGCACAAAGTGCAAAGATGAAAC
ATGCTTAAAAGTACATAGGTACAACCATAAACTGTCTAGCAGTTTTAAATAT
TACGAACCTCATACAAACCTAGATATTATGgtaaccacacatagatacatacatgcatacatatat
atatatatatatatatatatatatatatgtgtacatgaaaaaattcctagttcatcatttaattaatattttgttctccctttttattttttattttttt
ttattttatgttatttatttattttatttatttattttttttattttttttttttattttttttttttattttttattttttcccagTTTGGCACAG
GAATAATACAAGGGGTAATAGGAGTGGAGACATTTAAGATTGGACCGTTTG
AAATAAAAAATCAATCTTTTGGTTTGGTAAAAAGAGAAAAAGCTAGCGATA
ATAAATCTAATGTTTTTGAAAGAATAAATTTTGAAGGTATTGTTGGATTAGC
TTTTCCAGAAATGTTATCGACTGGTAAGAGTACGCTTTATGAGAATTTGATG
TCTTCATATAAATTACAACATAATGAATTTTCCATATATATTGGTAAGGATA
GTAAATATTCTGCATTAATATTTGGAGGGGTAGATAAAAATTTTTTTGAAGG
AGATATATATATGTTTCCTGTTGTTAAGGAATATTATTGGGAAATCCATTTT

GATGGTTTATATATTGATCATCAGAAATTTTGTTGTGGTGTTAATTCTATTGT
ATATGATTTAAAAAAAAAAGATCAAGAAAATAATAAATTATTTTTTACGAG
AAAATATTTTAGAAAAAATAAATTCAAAACTCATTTAAGAAAATATCTTCTT
AAAAAAATAAAACATCAAAAAAAACAAAAACATTCTAATCATAAAAAAAA
AAAATTAAACAAGAAAAAAAATTACTTAATTTTTGATTCTGGAACATCTTTT
AATAGTGTCCCAAAGGATGAAATCGAATATTTCTTTCGTGTCGTTCCTTCAA
AGgtaaaatataaaacatggcacattatatgtaatatatataaataaatatacaactaccatgcatatatatatgtatatatgtatat
atgtatatatgtatatatgtatatatatatatatatatatatatatatatatatatatatttttttttttttttttttttttttttagAAATGT
GATGATAGTAACATTGATCAGGTGGTTTCGAGTTACCCCAACTTGACTTACG
TCATAgtattaaaaaataaaaaaataaaactacaaacacatctacattatatatatatatatatatatatatatatatatatatatat
atatatatatatatatacatatacatatatacatatatacatatacatatatacatatatacatatacatatatacatatatacatatacat
atgtatccatatttatggcccatttttttgtagAACAAAATGCCGTTCACGTTGACCCCCTCACAAT
ACTTGGTTCGTAAGAATGATATGTGTAAACCAGCTTTTATGGAAATAGAAGT
TTCATCTGAATATGGACATGCGTACATACTAGGGAATGCAACGTTTATGAG
ATATTATTACACTGTTTATAGACGAGGAAATAATAACAATAGTTCATATgtaag
atattataaaaataaagaaaaacaaaagaacaaaaataataatatacacacaaaaatttataatgcgaaacatttatatgttatatat
atatatatatatatatatatatatatatatatatatatttatttatttatttatttatttatatttatttatttatatttattttttagGTGGGTA
TTGCTAAAGCTGTTCATACAGAGGAAAACGAAAAATATCTCAGTTCCTTAC
ATAATAAAATAAATAATTTATAA

Sequence Length: 3118 bp

Exons: Uppercase   Introns: Lowercase

Targeted PNA 1 binding site: green

Targeted PNA 2 binding site: yellow

A4.1.2 *Pf*PM IX protein sequence

MFFINFKKIKKKQFPIYLTQHRIITVFLIFIYFINLKDCFHINNSRILSDVDKHRGLY
YNIPKCNVCHKCSICTHENGEAQNVIPMVAIPSKRKHIQDINKEREENKYPLHIF
EEKDIYNNKDNVVKKEDIYKLRKKKKQKKNCLNFLEKDTMFLSPSHDKETFHI
NHMNKIKDEKYKQEYEEEKEIYDNTNTSQEKNETNNEQNLNINLINNDKVTLPL
QQLEDSQYVGYIQIGTPPQTIRPIFDTGSTNIWIVSTKCKDETCLKVHRYNHKLSS
SFKYYEPHTNLDIMFGTGIIQGVIGVETFKIGPFEIKNQSFGLVKREKASDNKSNV
FERINFEGIVGLAFPEMLSTGKSTLYENLMSSYKLQHNEFSIYIGKDSKYSALIFG
GVDKNFFEGDIYMFPVVKEYYWEIHFDGLYIDHQKFCCGVNSIVYDLKKKDQE
NNKLFFTRKYFRKNKFKTHLRKYLLKKIKHQKKQKHSNHKKKKLNKKKNYLI
FDSGTSFNSVPKDEIEYFFRVVPSKKCDDSNIDQVVSSYPNLTYVINKMPFTLTP

SQYLVRKNDMCKPAFMEIEVSSEYGHAYILGNATFMRYYYTVYRRGNNNNSS
YVGIAKAVHTEENEKYLSSLHNKINNL

Sequence Length: 627 aa

Catalytic sites: green Hypothesised pro-segment: underlined

A4.1.3 r*Pf*PM IX protein sequence

NCLNFLEKDTMFLSPSHDKETFHINHMNKIKDEKYKQEYEEEKEIYDNTNTSQE
KNETNNEQNLNINLINNDKVTLPLQQLEDSQYVGYIQIGTPPQTIRPIFDTGSTNI
WIVSTKCKDETCLKVHRYNHKLSSSFKYYEPHTNLDIMFGTGIIQGVIGVETFKI
GPFEIKNQSFGLVKREKASDNKSNVFERINFEGIVGLAFPEMLSTGKSTLYENLM
SSYKLQHNEFSIYIGKDSKYSALIFGGVDKNFFEGDIYLFPVVKEYYWEIHFDGL
YIDHQKFCCGVNSIVYDLKKKDQENNKLFFTRKYFRKNKFKTHLRKYLLKKIK
HQKKQKHSNHKKKKLNKKKNYLIFDSGTSFNSVPKDEIEYFFRVVPSKKCDDS
NIDQVVSSYPNLTYVINKMPFTLTPSQYLVRKNDMCKPAFMEIEVSSEYGHAYI
LGNATFMRYYYTVYRRGNNNNSSYVGIAKAVHTEENEKYLSSLHNKINNL

Catalytic sites: green Met to Leu mutation: yellow

Hypothesised pro-segment: underlined

A4.2 *Pf*PM X

A4.2.1 *PfPM X* genomic sequence

ATGAAACGCATTAGCCCTCTAAACACCCTTTTTTATTTAAGTTTATTTTTTTC
ATACACATTTAAAGGGTTAAAATGCACCAGAATATATAAAATCGGAACGAA
AGCGTTACCGTGTTCTGAGTGTCACGATGTATTTGATTGTACCGGGTGTTTA
TTCGAGGAAAAAGAATCTTCTCATGTGATACCTTTAAAATTAAACAAGAAG
AATCCAAACGATCATAAAAAATTACAAAAACATCATGAGTCTCTAAAATTA
GGGGACGTAAAATATTATGTAAACAGAGGAGAAGGAATTTCCGGAAGTTTA
GGAACGTCGTCTGGTAATACATTAGATGATATGGATTTAATAAATGAAGAA
ATAAATAAAAAACGAACAAATGCACAATTAGATGAAAAAAACTTTTTAGAT
TTTACTACGTATAATAAAAATAAAGCACAAGATATATCTGACCATTTATCTG
ATATTCAGAAACATGTGTATGAACAAGATGCTCAAAAGGGAAATAAAAATT
TTACAAATAATGAAAACAATTCAGATAATGAAAACAATTCAGATAATGAAA
ACAATTCAGATAATGAAAACAATTTAGATAATGAAAACAATTTAGATAATG
AAAACAATTCAGATAATAGTTCGATTGAAAAAAATTTTATTGCTTTAGAAA
ATAAAAACGCTACAGTAGAGCAAACAAAAGAAAATATTTTTCTAGTACCCT
TAAAACACTTAAGAGATAGCCAATTTGTAGGAGAATTATTAGTTGGTACCC

CCCCTCAAACTGTATATCCAATATTTGATACAGGAAGTACAAATGTATGGGT
AGTGACAACAGCTTGTGAAGAAGAGTCTTGTAAAAAAGTCCGAAGATATGA
TCCTAATAAATCAAAAACATTTAGGAGATCATTCATAGAAAAAAATTTACA
CATAGTATTTGGATCCGGTTCTATATCAGGTTCTGTAGGTACAGATACATTT
ATGTTAGGAAAGCACCTAGTAAGAAATCAGACTTTTGGATTAGTAGAGAGT
GAATCAAATAATAATAAAAATGGTGGGGATAATATATTTGATTATATATCCT
TTGAAGGTATAGTAGGCTTAGGATTTCCAGGAATGTTATCAGCTGGGAATAT
ACCATTTTTTGATAATTTATTAAAACAAAATCCAAATGTAGATCCTCAATTT
TCTTTTTATATATCTCCTTATGATGGGAAGTCAACTTTAATAATTGGTGGGAT
TAGTAAATCGTTTTATGAGGGCGATATATACATGTTGCCAGTATTAAAAGAA
TCCTACTGGGAAGTAAAATTAGATGAATTATATATTGGTAAAGAAAGAATA
TGTTGTGATGAAGAAAGTTACGTTATATTTGACACAGGTACATCTTATAATA
CAATGCCTAGTAGTCAGATGAAAACATTTTTGAATCTAATACATTCAACTGC
ATGTACTGAACAAAACTATAAAGATATATTAAAATCATATCCTATAATTAA
ATATGTTTTTGGTGAGCTAATCATTGAATTACATCCAGAAGAATATATGATT
TTAAACGATGATGTATGTATGCCTGCCTACATGCAAATCGATGTACCATCCG
AAAGAAACCATGCATATTTATTAGGCAGTTTATCTTTTATGAGGAATTTTTT
TACGGTATTTGTACGAGGCACAGAAAGTAGACCTTCTATGGTTGGAGTAGC
AAGAGCAAAAAGTAAAAACTAA

Sequence length: 1722 bp

A4.2.2 *Pf*PM X and r*Pf*PM X protein sequence

MKRISPLNTLFYLSLFFSYTFKGLKCTRIYKIGTKALPCSECHDVFDCTGCLFEEK
ESSHVIPLKLNKKNPNDHKKLQKHHESLKLGDVKYYVNRGEGISGSLGTSSGN
TLDDMDLINEEINKKRTNAQLDEKNFLDFTTYNKNKAQDISDHLSDIQKHVYE
QDAQKGNKNFTNNENNSDNENNSDNENNSDNENNLDNENNLDNENNSDNSSI
EKNFIALENKNATVEQTKENIFLVPLKHLRDSQFVGELLVGTPPQTVYPIFDTGS
TNVWVVTTACEEESCKKVRRYDPNKSKTFRRSFIEKNLHIVFGSGSISGSVGTDT
FMLGKHLVRNQTFGLVESESNNNKNGGDNIFDYISFEGIVGLGFPGMLSAGNIP
FFDNLLKQNPNVDPQFSFYISPYDGKSTLIIGGISKSFYEGDIYMLPVLKESYWEV
KLDELYIGKERICCDEESYVIFDTGTSYNTMPSSQMKTFLNLIHSTACTEQNYKD
ILKSYPIIKYVFGELIIELHPEEYMILNDDVCMPAYMQIDVPSERNHAYLLGSLSF
MRNFFTVFVRGTESRPSMVGVARAKSKN

Sequence Length: 573 amino acids

Catalytic sites: green Hypothesised pro-segment: underlined

Appendix 5: Crystal tray buffer conditions used in r*Pf*PM IX crystal screens.

Table A5.1: **Buffer conditions used in r*Pf*PM IX crystal screens.**

Screen name:	**Buffer**
Crystal Screen™ -HR2-110	0.4 M Ammonium phosphate monobasic
	0.1 M TRIS hydrochloride pH 8.5, 2.0 M Ammonium sulphate
	0.2 M Sodium citrate tribasic dihydrate, 0.1M HEPES sodium pH 7.5, 30% v/v (+/-)-2-Methyl-2,4-pentanediol
	0.1 M Sodium cacodylate trihydrate pH 6.5, 1.4 M Sodium acetate trihydrate
	0.2 M Sodium citrate tribasic dihydrate, 0.1 M Sodium cacodylate trihydrate pH 6.5, 30% v/v 2-Propanol
	0.2 M Ammonium acetate, 0.1 M Sodium acetate trihydrate pH 4.6, 30% w/v Polyethylene glycol 4,000
	0.1 M Sodium citrate tribasic dihydrate pH 5.6, 1.0 M Ammonium phosphate monobasic
	0.2 M Magnesium chloride hexahydrate, 0.1 M HEPES sodium pH 7.5, 30% v/v 2-Propanol
	0.2 M Sodium citrate tribasic dihydrate, 0.1M TRIS hydrochloride pH 8.5, 30% v/v Polyethylene glycol 400
	0.1 M HEPES sodium pH 7.5, 1.5 M Lithium sulphate monohydrate
	0.2 M Ammonium acetate, 0.1 M TRIS hydrochloride pH 8.5, 30% v/v 2-Propanol
	0.2 M Ammonium sulphate, 0.1 M Sodium acetate trihydrate pH 4.6, 24% w/v Polyethylene glycol 4,000
	0.2 M Calcium chloride dihydrate, 0.1 M Sodium acetate trihydrate pH 4.6, 20% v/v 2-Propanol
	0.1 M Imidazole pH 6.5, 1.0 M Sodium acetate trihydrate

	0.1 M HEPES sodium pH 7.5, 0.8 M Potassium sodium tartrate tetrahydrate
	0.2 M Ammonium sulphate, 30% w/v Polyethylene glycol 4,000
	0.1 M Sodium acetate trihydrate pH 4.6, 2.0 M Sodium formate
	0.1 M HEPES sodium pH 7.5, 0.8 M Sodium phosphate monobasic monohydrate, 0.8 M Potassium phosphate monobasic
	0.1 M TRIS hydrochloride pH 8.5, 8% w/v Polyethylene glycol 8,000
	0.1 M HEPES sodium pH 7.5, 2% v/v Polyethylene glycol 400, 2.0 M Ammonium sulphate
	0.1 M HEPES sodium pH 7.5, 10% v/v 2-Propanol, 20% w/v Polyethylene glycol 4,000
	0.05 M Potassium phosphate monobasic, 20% w/v Polyethylene glycol 8,000
	0.2 M Calcium acetate hydrate, 0.1 M Sodium cacodylate trihydrate pH 6.5, 18% w/v Polyethylene glycol 8,000
	0.1 M Sodium acetate trihydrate pH 4.6, 2.0 M Ammonium sulphate
	0.1 M TRIS hydrochloride pH 8.5, 2.0 M Ammonium phosphate monobasic
Crystal Screen™ -HR2-112	2.0 M Ammonium sulphate, 5% v/v 2-Propanol
	1.0 M Imidazole pH 7.0
	1.5 M Sodium chloride, 10% v/v Ethanol
	0.1 M Sodium acetate trihydrate pH 4.6, 2.0 M Sodium chloride
	0.2 M Sodium chloride, 0.1 M Sodium acetate trihydrate pH 4.6, 30% v/v (+/-)-2-Methyl-2,4-pentanediol
	0.1 M Cadmium chloride hydrate, 0.1 M Sodium acetate trihydrate pH 4.6, 30% v/v Polyethylene

	glycol 400
	0.2 M Ammonium sulphate, 0.1 M Sodium acetate trihydrate pH 4.6, 30% w/v Polyethylene glycol monomethyl ether 2,000
	0.5 M Ammonium sulphate, 0.1 M Sodium citrate tribasic dihydrate pH 5.6, 1.0 M Lithium sulphate monohydrate
	0.01 M Iron (III) chloride hexahydrate, 0.1 M Sodium citrate tribasic dihydrate pH 5.6, 10% v/v Jeffamine M-600®
	0.1 M MES monohydrate pH 6.5, 2.0 M Sodium chloride
	0.05 M Cesium chloride, 0.1 M MES monohydrate pH 6.5, 30% v/v Jeffamine M-600®
	0.01 M Cobalt (II) chloride hexahydrate, 0.1 M MES monohydrate pH 6.5, 1.8 M Ammonium sulphate
	0.2 M Ammonium sulphate, 0.1 M MES monohydrate pH 6.5, 30% w/v Polyethylene glycol monomethyl ether 5,000
	0.5 M Ammonium sulphate, 0.1 M HEPES pH 7.5, 30% v/v (+/-)-2-Methyl-2,4-pentanediol
	0.1 M Sodium chloride, 0.1 M HEPES pH 7.5, 1.6 M Ammonium sulphate
	0.1 M HEPES pH 7.5, 2.0 M Ammonium formate
	0.05 M Cadmium sulphate hydrate, 0.1 M HEPES pH 7.5, 1.0 M Sodium acetate trihydrate
	0.1 M HEPES pH 7.5, 70% v/v (+/-)-2-Methyl-2,4-pentanediol
	0.1 M HEPES pH 7.5, 20% w/v Polyethylene glycol 10,000
	0.01 M Nickel (II) chloride hexahydrate, 0.1 M Tris pH 8.5, 1.0 M Lithium sulphate monohydrate
	1.5 M Ammonium sulphate, 0.1 M Tris pH 8.5, 12% v/v Glycerol

	0.1 M Tris pH 8.5, 20% v/v Ethanol
	0.01 M Nickel (II) chloride hexahydrate, 0.1 M Tris pH 8.5, 20% w/v Polyethylene glycol monomethyl ether 2,000
	0.1 M BICINE pH 9.0, 2.0 M Magnesium chloride hexahydrate
Crystal Screen Cryo™ - HR2-122	0.26 M Ammonium phosphate monobasic, 35% v/v Glycerol
	0.17 M Ammonium acetate, 0.085 M Sodium acetate trihydrate pH 4.6, 25.5% w/v Polyethylene glycol 4,000, 15% v/v Glycerol
	0.17 M Ammonium sulphate, 0.085 M Sodium cacodylate trihydrate pH 6.5, 25.5% w/v Polyethylene glycol 8,000, 15% v/v Glycerol
	0.16 M Ammonium sulphate, 0.08 M Sodium acetate trihydrate pH 4.6, 20% w/v Polyethylene glycol 4,000, 20% v/v Glycerol
	0.14 M Calcium chloride dihydrate, 0.07 M Sodium acetate trihydrate pH 4.6, 14% v/v 2-Propanol, 30% v/v Glycerol
	0.07 M Imidazole pH 6.5, 0.7 M Sodium acetate trihydrate, 30% v/v Glycerol
	0.085 M HEPES sodium pH 7.5, 1.7% v/v Polyethylene glycol 400, 1.7 M Ammonium sulphate, 15% v/v Glycerol
	0.04 M Potassium phosphate monobasic, 16% w/v Polyethylene glycol 8,000, 20% v/v Glycerol
	0.1 M Magnesium formate dihydrate, 50% v/v Glycerol
	0.8 M Lithium sulphate monohydrate, 1.6% w/v Polyethylene glycol 8,000, 20% v/v Glycerol

Appendix 6: NAMD Scripts for *in silico* modelling of *Pf*PM IX

A6.1: Step 1, initial minimisation of *Pf*PM IX in a water box

```
#############################################################
### JOB DESCRIPTION                                   ########
#############################################################
#
## Minimisation and Equilibration of
## PM9 in a Water Box
##
## to run: namd filename.conf > filename.log &
## to run multi CPU    charmrun namd2 ++local +p<procs> <configfile>
# to run GPU enabled  charmrun ++local +p4 ./namd2 +idlepoll <configfile>
# charmrun +p4 /usr/local/bin/namd2 PM9_H2O_Minimise_PBC_Step_1.namd >
PM9_H2O_Minimise_PBC_Step_1.namdrun.log &
#
#################################################################
### ADJUSTABLE PARAMETERS                                    ##
#################################################################

### input Options
coordinates             Desmond_MD_frame304_NO_H_renum_autopsf.pdb
structure               Desmond_MD_frame304_NO_H_renum_autopsf.psf
parameters              Applications/VMD1.8.7.app/Contents/vmd/plugins/
noarch/tcl/readcharmmpar1.1/par_all27_prot_lipid_na.inp
paratypecharmm          on

#Output Options

outputname              PM9_Phyre_model_H2O_Min.output
dcdfile                 PM9_Phyre_model_H2O_Min.dcd
xstFile                 PM9_Phyre_model_H2O_Min.xst
dcdfreq                 50
xstFreq                 50
restartfreq             500     ;# 500steps = every 1ps
```

```
####  Temperature Setup
set temperature         298
temperature             $temperature;  # initial temperature
firsttimestep           0

### IMD Setup

IMDon                   on ;#
IMDport                 3000    ;# port number (enter it in VMD)
IMDfreq                 1       ;# send every 1 frame
IMDwait                 no      ;# wait for VMD to connect before running?

#####################################################################
### SIMULATION PARAMETERS                                         ##
#####################################################################

## Force-Field Parameters
exclude                 scaled1-4
1-4scaling              1.0
cutoff                  12.0
switching               on
switchdist              10.0
pairlistdist            13.5

## Integrator Parameters
timestep                2.0  ;# 2fs/step
rigidBonds              all  ;# needed for 2fs steps
nonbondedFreq           1
fullElectFrequency      2
stepspercycle           10

## Constant Temperature Control
langevin                off    ;# do langevin dynamics
langevinDamping         5     ;# damping coefficient (gamma) of 5/ps
```

```
langevinTemp            $temperature
langevinHydrogen        off    ;# don't couple langevin bath to hydrogens

## Periodic Boundary Conditions
##
## you get the info to make the following from:
## set sel [atomselect top all]
## cell basis vectors:
## set m [measure minmax $sel]
## foreach {j1 j2} $m {}
## foreach {x2 y2 z2} $j2 {}
## foreach {x1 y1 z1} $j1 {}
## expr $x2 - $x1
## expr $y2 - $y1
## expr $z2 - $z1
## cellOrigin:
## measure center $sel
##

cellBasisVector1        71.7    0.0   0.0
cellBasisVector2        0.0   92.3   0.0
cellBasisVector3        0.0    0.0  89.4
cellOrigin              -15.2  -14.1  -10.46
wrapAll                 on

## PME (for full-system periodic electrostatics)
PME                     yes
PMEGridSizeX            64
PMEGridSizeY            64
PMEGridSizeZ            64

## Constant Pressure Control (variable volume)
useGroupPressure        yes ;# needed for rigidBonds
useFlexibleCell         no
useConstantArea         no
```

```
langevinPiston          off
langevinPistonTarget    1.01325 ;#  in bar -> 1 atm
langevinPistonPeriod    100.
langevinPistonDecay     50.
langevinPistonTemp      $temperature

binaryoutput            no      ;# give me the pdb instead of the .coor

#####################################################################
### EXTRA PARAMETERS                                                ##
#####################################################################
#
#
#####################################################################
### EXECUTION SCRIPT                                                 ##
#####################################################################

## Minimisation
minimise                10000
```

A6.2: Step 2, equilibrate temperature and pressure

```
#####################################################################
### JOB DESCRIPTION                                                 ##
#####################################################################

### Equilibrate Tempearture & Pressure  ********

# restart of a XXXXXXXXX
#
# to run SIngle CPU:  namd filename.conf > filename.log &
# to run multi CPU    charmrun namd2 ++local +p<procs> <configfile>
```

```
# charmrun +p4 /usr/local/bin/namd2 RM_Equil_temp_pressure_Step_2.namd > RM_Equil_temp_pressure_Step_2.namd.log &

#  Equilibrate with incrementally increasing temp steps of 20 K.
#############################################################
## ADJUSTABLE PARAMETERS                                   ##
#############################################################
## CHANGE ME FOR YOUR FILENAMES

structure          Desmond_MD_frame304_NO_H_renum_autopsf.psf
coordinates        PM9_Phyre_model_H2O_Min.output.coor
outputname         RM_PM9_equil_temp_pressure_step2.out

set temp_step      20
firsttimestep      0

### IMD Setup

IMDon              on ;#
IMDport            3000    ;# port number (enter it in VMD)
IMDfreq            5       ;# send every 1 frame
IMDwait            no      ;# wait for VMD to connect before running?

#############################################################
## SIMULATION PARAMETERS                                   ##
#############################################################

# Input
paraTypeCharmm     on
parameters         /Applications/VMD1.8.7.app/Contents/vmd/plugins/noarch/tcl/readcharmmpar1.1/par_all27_prot_lipid_na.inp

# NOTE: Do not set the initial velocity temperature if you
# have also specified a .vel restart file!
```

```
temperature             $temp_step

#keep temp if want to assign new velocitites

# Periodic Boundary conditions
# NOTE: Do not set the periodic cell basis if you have also
# specified an .xsc restart file!
## Periodic Boundary Conditions
##
## you get the info to make the following from:
## set sel [atomselect top all]
## cell basis vectors:
## set m [measure minmax $sel]
## foreach {j1 j2} $m {}
## foreach {x2 y2 z2} $j2 {}
## foreach {x1 y1 z1} $j1 {}
## expr $x2 - $x1
## expr $y2 - $y1
## expr $z2 - $z1
## cellOrigin:
## measure center $sel
################################### OUTOUT
#########################################

#>Main< (bin) 27 % set sel [atomselect top all]
#atomselect194
#>Main< (bin) 28 % set m [measure minmax $sel]
#{-51.700538635253906          -60.96098709106445          -55.79301834106445}
{21.141117095947266 32.80785369873047 34.98427200317383}
#>Main< (bin) 29 % foreach {j1 j2} $m {}
#>Main< (bin) 30 % foreach {x2 y2 z2} $j2 {}
#>Main< (bin) 31 % foreach {x1 y1 z1} $j1 {}
#>Main< (bin) 32 % expr $x2 - $x1
#72.84165573120117
#>Main< (bin) 33 % expr $y2 - $y1
```

#93.76884078979492

#>Main< (bin) 34 % expr $z2 - $z1

#90.77729034423828

measure center $sel -15.220047950744629 -14.028674125671387 -10.426637649536133

################################## END OF OUTOUT

##

#Note: Add 0.1 A And Round Up To Nearest 0.1 A To Each Cell Dimension Obtaind # From Above To Allow For Changing Box Sizes. note XYZ

if (Noedl, Se *et al.*) {

cellBasisVector1 73.1 0.0 0.0

cellBasisVector2 0.0 93.9 0.0

cellBasisVector3 0.0 0.0 90.9

cellOrigin -15.4 -14.2 -10.6

}

wrapWater on

wrapAll on

Force-Field Parameters

exclude scaled1-4

1-4scaling 1.0

cutoff 12.0

switching on

switchdist 10.0

pairlistdist 13.5

Integrator Parameters

timestep 2.0 ;# 2fs/step

rigidBonds all ;# needed for 2fs steps

nonbondedFreq 1

fullElectFrequency 2

stepspercycle 10

```
#PME (for full-system periodic electrostatics)
if (Noedl, Se et al.) {
PME                     yes
PMEGridSizeX        64
PMEGridSizeY        64
PMEGridSizeZ        64
}
# temp rescaling actually brings the temp down.
# perform incremental heating
if (Noedl, Se et al.) {
reassignTemp            $temp_step ;# starting temp
reassignIncr            $temp_step ;# temp increment
reassignHold            300 ;# final temp
reassignFreq            100 ;# num timesteps to temp incr.
#       step     temp
# -------- -------
#       0               20
#       200             40
#       400             60
#       600             80
#       800             100
#       1000            120
#       1200            140
# 1000 steps, add 100 deg.  So do 3000 steps and get 300 $ init_temp
}

# Use BerendsenPressure to equilibrate
if (Noedl, Se et al.) {
useGroupPressure      yes ;# needed for 2fs steps
useFlexibleCell         no  ;# no for water box, yes for membrane
useConstantArea         no  ;# no for water box, yes for membrane
BerendsenPressure                       on
BerendsenPressureTarget                 1.01325
BerendsenPressureCompressibility    4.57E-5
BerendsenPressureRelaxationTime     20
```

```
BerendsenPressureFreq                   2
}

# Output
# outputName          $outputname

restartfreq             500     ;# 500steps = every 1ps
dcdfreq                 250
xstFreq                 250
outputEnergies          10
# outputPressure        100

#############################################################
## EXTRA PARAMETERS                                        ##
#############################################################

# Put here any custom parameters that are specific to
# this job (e.g., SMD, TclForces, etc...)

#############################################################
## EXECUTION SCRIPT                                        ##
#############################################################

# Minimisation
if {0} {
minimize                200
reinitvels              $temperature
}

run 3000
```

A6.3: Step 3, 20 ps molecular dynamics simulation

```
#############################################################
## JOB DESCRIPTION                                         ##
#############################################################
```

```
# MD Equilbruim Simulation 20 ps

# restart of a run
# PM9 in a Water Box
#
# to run SIngle CPU:  namd filename.conf > filename.log &
# to run multi CPU    charmrun +p4 /usr/local/bin/namd2
PM9_Equil_Sim_20ps_Step_3.namd > PM9_Equil_Sim_20ps_Step_3.log &

# charmrun +p4 /usr/local/bin/namd2
RM_MD_Equilibration_20ps_PM9_Step_3.namd >
RM_MD_Equilibration_20ps_PM9_Step_3.log &

#############################################################
## ADJUSTABLE PARAMETERS                                   ##
#############################################################

structure          Desmond_MD_frame304_NO_H_renum_autopsf.psf
coordinates        PM9_Phyre_model_H2O_Min.output.coor ;# this is text pdb file
defining the coords
outputname         PM9_Phyre_model_autopsf_equlib_MD_20ps

set inputname      RM_PM9_equil_temp_pressure_step2.out  ;# Base name of input
files for restarting jobs

# Continuing a job from the restart files
# if {0} is true

set rs             1 ;# use restart file? 1 = yes
if {$rs==1} {
binCoordinates     $inputname.restart.coor    ;# this is the binary file of the
coordinates from the last step of the previous simulation
#binVelocities     $inputname.restart.vel  ;# remove the "temperature" entry if you
use this!
```

```
extendedSystem	$inputname.restart.xsc	;# this is a text file of the cell paramters from the last step of the previous simulation
}

# Continuing a job but not from restart files
if {$rs==0} {
binCoordinates	$inputname.coor
# binVelocities	$inputname.vel	;# remove the "temperature" entry if you use this!
extendedSystem	$inputname.xsc
}

firsttimestep	0
set temperature	310

#####################################################################
## SIMULATION PARAMETERS                                           ##
#####################################################################

# Input
paraTypeCharmm	on
parameters	/Applications/VMD 1.8.7.app/Contents/vmd/plugins/noarch/tcl/readcharmmpar1.1/par_all27_prot_lipid_na.inp

# NOTE: Do not set the initial velocity temperature if you
# have also specified a .vel restart file!

temperature	$temperature

## To benefit from GPU acceleration you should set outputEnergies to 100 or higher in the simulation config file

outputEnergies 100

#keep temp if want to assign new velocitites
```

```
 # Periodic Boundary conditions
# NOTE: Do not set the periodic cell basis if you have also
# specified an .xsc restart file!
# Periodic Boundary conditions
# NOTE: Do not set the periodic cell basis if you have also
# specified an .xsc restart file!

## Periodic Boundary Conditions
##
## you get the info to make the following from:
## set sel [atomselect top all]
## cell basis vectors:
## set m [measure minmax $sel]
## foreach {j1 j2} $m {}
## foreach {x2 y2 z2} $j2 {}
## foreach {x1 y1 z1} $j1 {}
## expr $x2 - $x1
## expr $y2 - $y1
## expr $z2 - $z1
## cellOrigin:
## measure center $sel

#if {0} {
#cellBasisVector1      53.0    0.0   0.0
#cellBasisVector2      0.0   46.0   0.0
#cellBasisVector3       0.0    0.0  44.0
#cellOrigin            3.5   2.0  1.2
#}
wrapWater             on
wrapAll               on

# Force-Field Parameters
exclude               scaled1-4
1-4scaling            1.0
cutoff                12.0
```

```
switching               on
switchdist              10.0
pairlistdist            13.5

# Integrator Parameters
timestep                2.0  ;# 2fs/step
rigidBonds              all  ;# needed for 2fs steps
nonbondedFreq           1
fullElectFrequency      2
stepspercycle           10

#PME (for full-system periodic electrostatics)
if (Noedl, Se et al.) {
PME                     yes
PMEGridSizeX            64
PMEGridSizeY            64
PMEGridSizeZ            64

}

# Constant Temperature Control
langevin                on    ;# do langevin dynamics
langevinDamping         5     ;# damping coefficient (gamma) of 5/ps
langevinTemp            $temperature
langevinHydrogen        no    ;# don't couple langevin bath to hydrogens
# Constant Pressure Control (variable volume)
if (Noedl, Se et al.) {
useGroupPressure        yes ;# needed for 2fs steps
useFlexibleCell         no  ;# no for water box, yes for membrane
useConstantArea         no  ;# no for water box, yes for membrane

langevinPiston          on
langevinPistonTarget  1.01325 ;#  in bar -> 1 atm
langevinPistonPeriod  100.0
langevinPistonDecay   50.0
```

```
langevinPistonTemp    $temperature
}

# Output
# outputName          $outputname

restartfreq           500     ;# 500steps = every 1ps
dcdfreq               250
xstFreq               250
# outputPressure      100

# Fixed Atoms Constraint (set PDB beta-column to 1)
if {0} {
fixedAtoms            on
fixedAtomsFile        myfixedatoms.pdb
fixedAtomsCol         B
}

# IMD Settings (can view sim in VMD)
if {0} {
IMDon                 on
IMDport               3000    ;# port number (enter it in VMD)
IMDfreq               1       ;# send every 1 frame
IMDwait               no      ;# wait for VMD to connect before running?
}
###################################################################
## EXTRA PARAMETERS                                              ##
###################################################################

# Put here any custom parameters that are specific to
# this job (e.g., SMD, TclForces, etc...)

###################################################################
## EXECUTION SCRIPT                                              ##
###################################################################
```

```
# Minimisation
if {0} {
minimise            100
reinitvels          $temperature ;# the temperature statement above verforms vel reassignment
}

run 10000 ;# 20ps
```

A6.4: Step 4, 20 ns molecular dynamics simulation

```
#############################################################
## JOB DESCRIPTION                                         ##
#############################################################

# MD Equilbruim Simulation 20ns ps

# restart of a run
# PM9 in a Water Box
#
# to run SIngle CPU:  namd filename.conf > filename.log &
# to run multi CPU    charmrun ++local +p8 /usr/local/bin/namd2 +idlepoll MD_Equilibration_1ns_PM10_Step_4.namd > MD_Equilibration_1ns_PM10_Step_4.namd.log &

# charmrun +p4 /usr/local/bin/namd2  RM_MD_Equilibration_20ns_Step_4.namd > RM_MD_Equilibration_20ns_Step_4.namd.log &

#############################################################
## ADJUSTABLE PARAMETERS                                   ##
#############################################################

structure           Desmond_MD_frame304_NO_H_renum_autopsf.psf
coordinates         PM9_Phyre_model_H2O_Min.output.coor ;# this is text pdb file defining the coords
```

```
outputname          PM9_Phyre_model_autopsf_equlib_MD_20ns

set inputname       PM9_Phyre_model_autopsf_equlib_MD_20ps ;# Base name of input files for restarting jobs

# Continuing a job from the restart files
# if {0} is true

set rs              1 ;# use restart file? 1 = yes
if {$rs==1} {
binCoordinates      $inputname.restart.coor     ;# this is the binary file of the coordinates from the last step of the previous simulation
#binVelocities      $inputname.restart.vel  ;# remove the "temperature" entry if you use this!
extendedSystem      $inputname.restart.xsc    ;# this is a text file of the cell paramters from the last step of the previous simulation
}

# Continuing a job but not from restart files
if {$rs==0} {
binCoordinates      $inputname.coor
# binVelocities     $inputname.vel   ;# remove the "temperature" entry if you use this!
extendedSystem      $inputname.xsc
}

firsttimestep       0
set temperature     310

#####################################################################
## SIMULATION PARAMETERS                                           ##
#####################################################################

# Input
paraTypeCharmm      on
```

parameters /Applications/VMD 1.8.7.app/Contents/vmd/plugins/noarch/tcl/readcharmmpar1.1/par_all27_prot_lipid_na.inp

NOTE: Do not set the initial velocity temperature if you
have also specified a .vel restart file!

temperature $temperature

To benefit from GPU acceleration you should set outputEnergies to 100 or higher in the simulation config file

outputEnergies 100

#keep temp if want to assign new velocitites

Periodic Boundary conditions
NOTE: Do not set the periodic cell basis if you have also
specified an .xsc restart file!
Periodic Boundary conditions
NOTE: Do not set the periodic cell basis if you have also
specified an .xsc restart file!

Periodic Boundary Conditions
##
you get the info to make the following from:
set sel [atomselect top all]
cell basis vectors:
set m [measure minmax $sel]
foreach {j1 j2} $m {}
foreach {x2 y2 z2} $j2 {}
foreach {x1 y1 z1} $j1 {}
expr $x2 - $x1
expr $y2 - $y1
expr $z2 - $z1
cellOrigin:

```
## measure center $sel

#if {0} {
#cellBasisVector1     53.0    0.0   0.0
#cellBasisVector2     0.0   46.0   0.0
#cellBasisVector3     0.0    0.0  44.0
#cellOrigin           3.5   2.0  1.2
#}
wrapWater            on
wrapAll              on

# Force-Field Parameters
exclude              scaled1-4
1-4scaling           1.0
cutoff               12.0
switching            on
switchdist           10.0
pairlistdist         13.5

# Integrator Parameters
timestep             2.0  ;# 2fs/step
rigidBonds           all  ;# needed for 2fs steps
nonbondedFreq        1
fullElectFrequency   2
stepspercycle        10

#PME (for full-system periodic electrostatics)
if (Noedl, Se et al.) {
PME                  yes
PMEGridSizeX         64
PMEGridSizeY         64
PMEGridSizeZ         64

}
```

```
# Constant Temperature Control
langevin                on    ;# do langevin dynamics
langevinDamping         5     ;# damping coefficient (gamma) of 5/ps
langevinTemp            $temperature
langevinHydrogen        no    ;# don't couple langevin bath to hydrogens

# Constant Pressure Control (variable volume)
if (Noedl, Se et al.) {
useGroupPressure        yes ;# needed for 2fs steps
useFlexibleCell         no  ;# no for water box, yes for membrane
useConstantArea         no  ;# no for water box, yes for membrane

langevinPiston          on
langevinPistonTarget  1.01325 ;#  in bar -> 1 atm
langevinPistonPeriod  100.0
langevinPistonDecay   50.0
langevinPistonTemp    $temperature
}

# Output
# outputName            $outputname

restartfreq             500     ;# 500steps = every 1ps
dcdfreq                 250
xstFreq                 250
# outputPressure        100

# Fixed Atoms Constraint (set PDB beta-column to 1)
if {0} {
fixedAtoms              on
fixedAtomsFile          myfixedatoms.pdb
fixedAtomsCol           B
}

# IMD Settings (can view sim in VMD)
```

```
if {0} {
IMDon           on
IMDport         3000    ;# port number (enter it in VMD)
IMDfreq         1       ;# send every 1 frame
IMDwait         no      ;# wait for VMD to connect before running?
}

#############################################################
## EXTRA PARAMETERS                                        ##
#############################################################

# Put here any custom parameters that are specific to
# this job (e.g., SMD, TclForces, etc...)

#############################################################
## EXECUTION SCRIPT                                        ##
#############################################################

# Minimisation
if {0} {
minimise                100
reinitvels              $temperature ;# the temperature statement above verforms vel reassignment
}

run 10000000 ;# 20000ps 20ns

#run 5000000 ;# 1000ps 1ns
```

Appendix 7: Top docking hits from *in silico* screens

Table A7.1: **Top docking hits from the GlaxoSmithKline library.** Inhibitors are ranked by Glide docking score and are displayed with their respective ChEMBL identifier and corresponding structure number.

Compound	Schrödinger 'Glide' Docking Score	Structure number
TCMDC-131544	-11.13	1
TCMDC-139334	-10.65	2
TCMDC-140588	-10.58	3
TCMDC-135362	-10.38	4
TCMDC-141769	-10.19	5
TCMDC-141484	-10.18	6
TCMDC-137201	-10.16	7
TCMDC-140556	-10.06	8
TCMDC-139797	-9.96	9
TCMDC-141567	-9.96	10
TCMDC-136458	-9.95	11
TCMDC-139016	-9.94	12
TCMDC-141514	-9.92	13
TCMDC-141101	-9.89	14

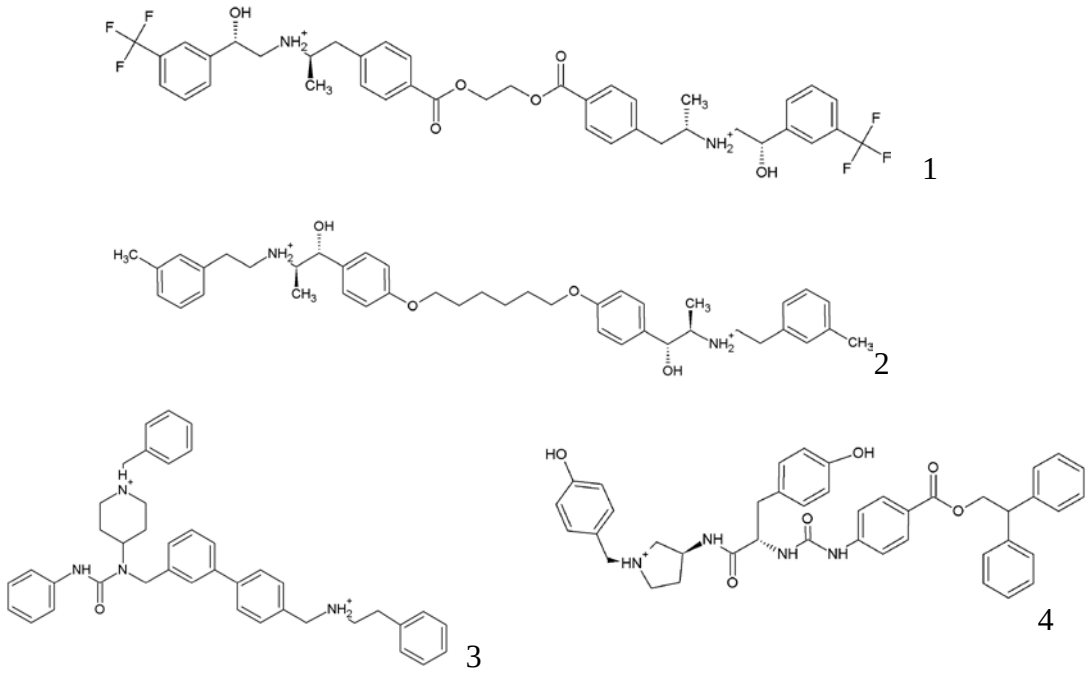
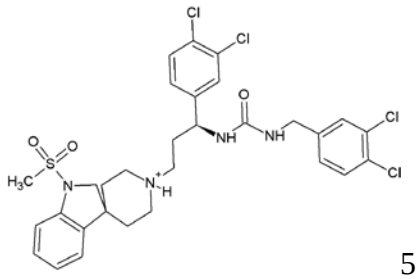
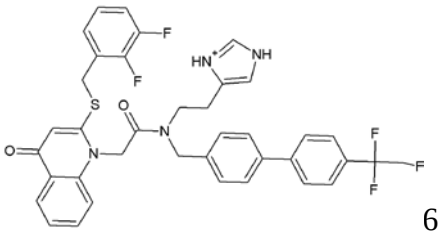


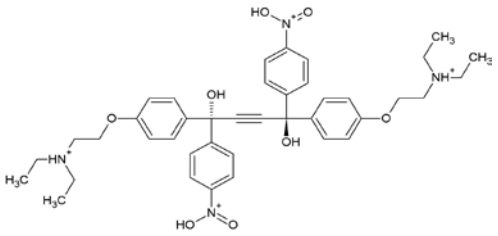
Figure A7.1a: **Structures of top binding GlaxoSmithKline inhibitors.** Numbers correspond to compounds in Table A7.1.



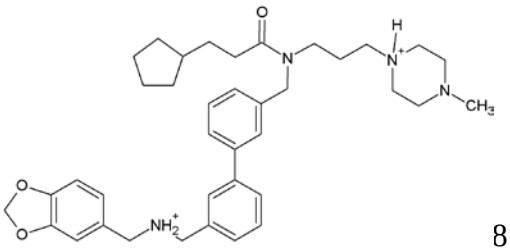
5



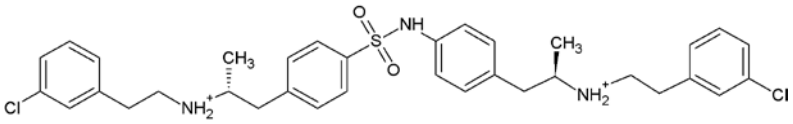
6



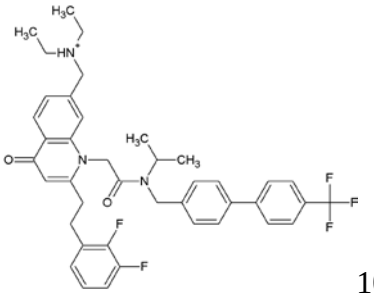
7



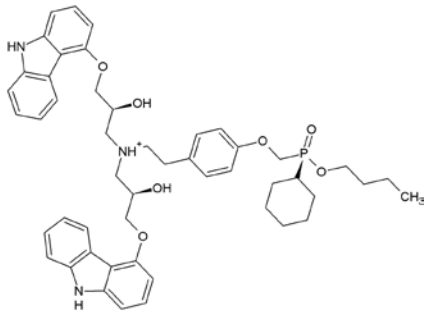
8



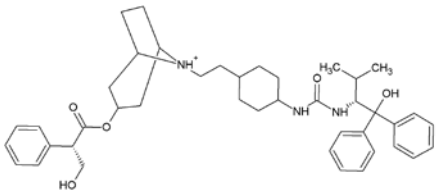
9



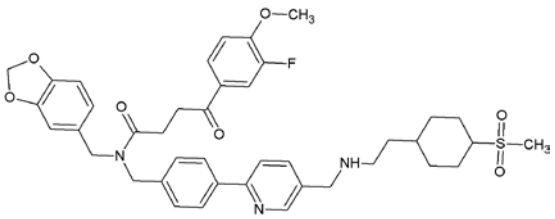
10



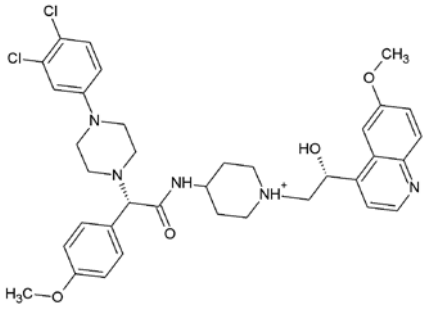
12



11



13

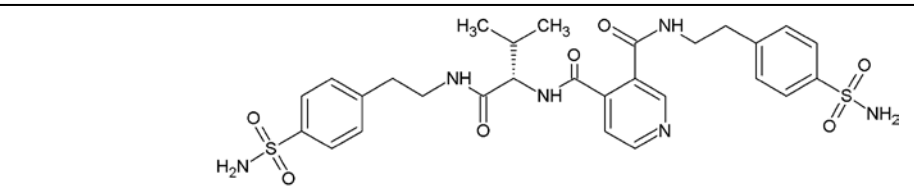


14

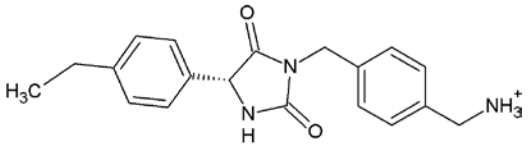
Figure A7.1b: **Structures of top binding GlaxoSmithKline inhibitors.** Numbers correspond to compounds in Table A7.1.

Table A7.2: **Top docking hits from an *in silico* screen of the Ambinter library.** The top 11 hits are presented with corresponding Glide docking scores and structure number.

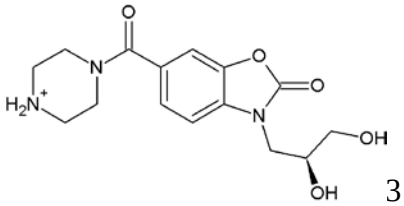
Compound code	Glide docking score	Structure number
Amb17039013	-11.25	1
Amb17456959	-10.72	2
Amb18502744	-10.66	3
Amb17151146	-10.64	4
Amb17320578	-10.38	5
Amb16943777	-10.25	6
Amb17308586	-10.09	7
Amb17228674	-10.07	8
Amb16940839	-10.05	9
Amb16943683	-10.04	10
Amb17092319	-10.03	11



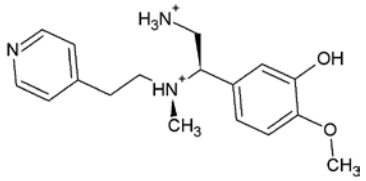
1



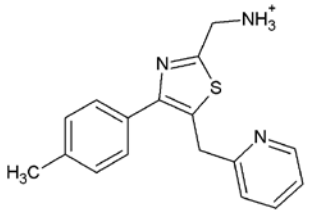
2



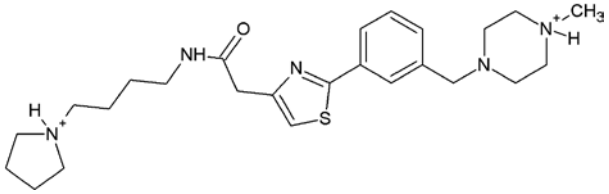
3



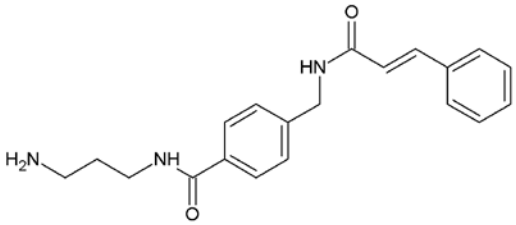
4



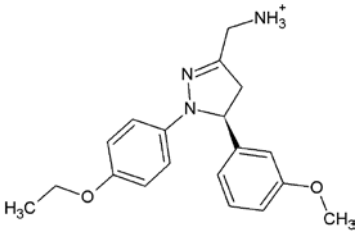
5



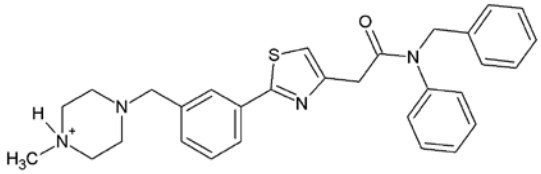
6



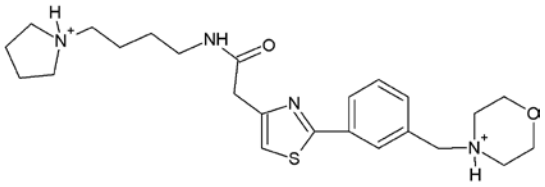
7



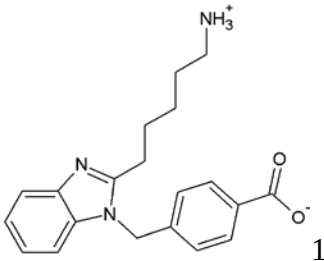
8



9



10



11

Figure A7.2: **Structures of top binding Ambinter compounds.** Numbers correspond to compounds in Table A7.2.

Appendix 8: Phylogenetic tree

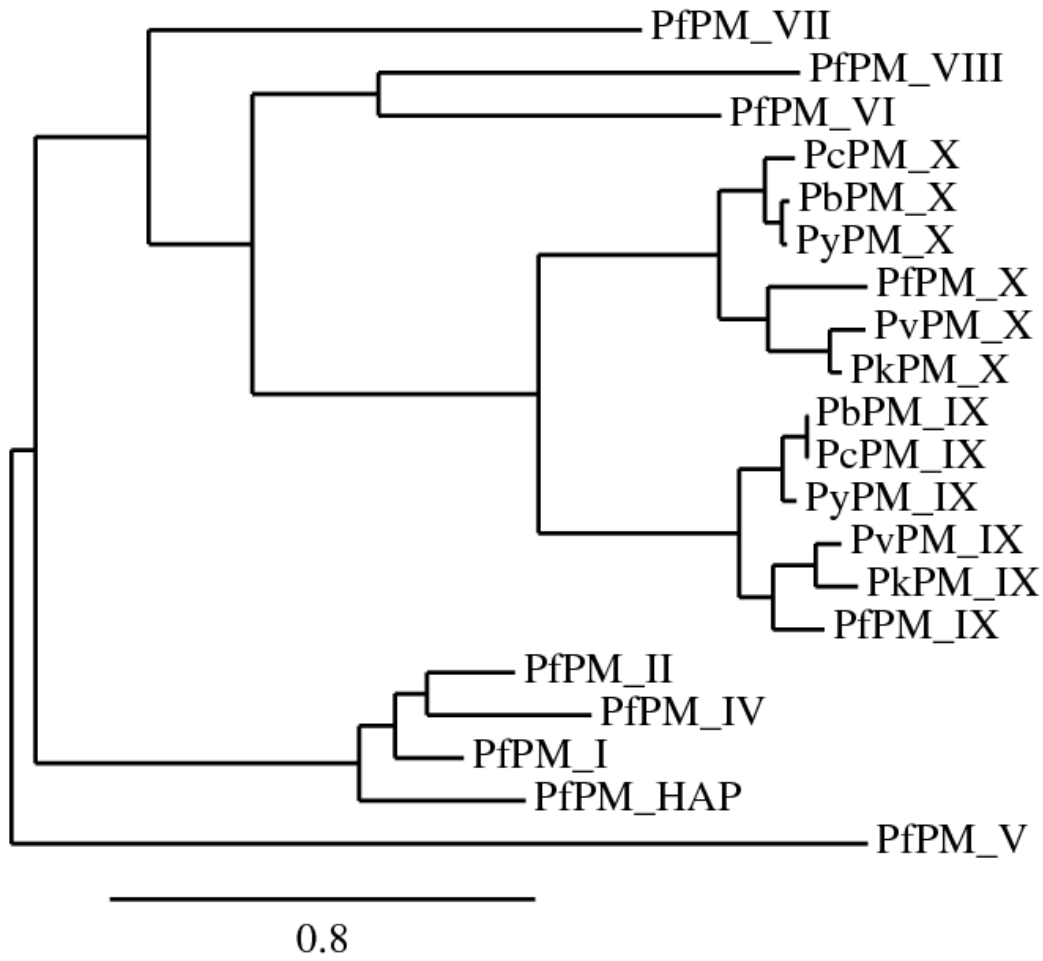


Figure A8.1: **Phylogenetic tree of selected plasmepsins.** Tree generated with the Phylongeny.fr online server. Pf = *Plasmodium falciparum.* Pc = *Plasmodium chabaudi.* Pv = *Plasmodium vivax.* Pk = *Plasmodium knowlesi.* Pb = *Plasmodium bergei.* Py = *Plasmodium yoellii.* (Dereeper, Guignon et al., 2008)

Bibliography

1 Abu Bakar, N., Klonis, N., Hanssen, E., Chan, C. and Tilley, L. (2010) Digestive-vacuole genesis and endocytic processes in the early intraerythrocytic stages of *Plasmodium falciparum*, *J Cell Sci*, 123(Pt 3), pp. 441-50.

2 Achan, J., Kakuru, A., Ikilezi, G., Ruel, T., Clark, T. D., Nsanzabana, C., Charlebois, E., Aweeka, F., *et al.* (2012) Antiretroviral agents and prevention of malaria in HIV-infected Ugandan children, *N Engl J Med*, 367(22), pp. 2110-8.

3 Adjuik, M., Agnamey, P., Babiker, A., Borrmann, S., Brasseur, P., Cisse, M., Cobelens, F., Diallo, S., *et al.* (2002) Amodiaquine-artesunate versus amodiaquine for uncomplicated *Plasmodium falciparum* malaria in African children: a randomised, multicentre trial, *Lancet*, 359(9315), pp. 1365-72.

4 Agnandji, S. T. and Lell, B. and Soulanoudjingar, S. S. and Fernandes, J. F. and Abossolo, B. P. and Conzelmann, C. and Methogo, B. G. and Doucka, Y., *et al.* (2011) First results of phase 3 trial of RTS,S/AS01 malaria vaccine in African children, *N Engl J Med*, 365(20), pp. 1863-75.

5 Ahn, D. G., Lee, W., Choi, J. K., Kim, S. J., Plant, E. P., Almazan, F., Taylor, D. R., Enjuanes, L., *et al.* (2011) Interference of ribosomal frameshifting by antisense peptide nucleic acids suppresses SARS coronavirus replication, *Antiviral Res*, 91(1), pp. 1-10.

6 Ahn, D. G., Shim, S. B., Moon, J. E., Kim, J. H., Kim, S. J. and Oh, J. W. (2011) Interference of hepatitis C virus replication in cell culture by antisense peptide nucleic acids targeting the X-RNA, *J Viral Hepat*, 18(7), pp. e298-306.

7 Aikawa, M., Hepler, P. K., Huff, C. G. and Sprinz, H. (1966) The feeding mechanism of avian malarial parasites, *J Cell Biol*, 28(2), pp. 355-73.

8 Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990) Basic local alignment search tool, *J Mol Biol*, 215(3), pp. 403-10.

9 Ambinter (2014) *Ambinter. A supplier of advanced chemicals to advanced users.* France. Available at: www.ambinter.com 2014).

10 Andrews, K. T., Fairlie, D. P., Madala, P. K., Ray, J., Wyatt, D. M., Hilton, P. M., Melville, L. A., Beattie, L., *et al.* (2006) Potencies of human immunodeficiency virus protease inhibitors *in vitro* against *Plasmodium falciparum* and *in vivo* against murine malaria, *Antimicrob Agents Chemother*, 50(2), pp. 639-48.

11 Angov, E., Hillier, C. J., Kincaid, R. L. and Lyon, J. A. (2008) Heterologous protein expression is enhanced by harmonizing the codon usage frequencies of the target gene with those of the expression host, *PLoS One*, 3(5), pp. e2189.

12 Anthonsen, H. W., Baptista, A., Drablos, F., Martel, P. and Petersen, S. B. (1994) The blind watchmaker and rational protein engineering, *J Biotechnol*, 36(3), pp. 185-220.

13 Aravind, L., Iyer, L. M., Wellems, T. E. and Miller, L. H. (2003) Plasmodium biology: genomic gleanings, *Cell*, 115(7), pp. 771-85.

14 Armstrong, C. M. and Goldberg, D. E. (2007) An FKBP destabilization domain modulates protein levels in *Plasmodium falciparum*, *Nat Methods*, 4(12), pp. 1007-9.

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

16 Asojo, O. A., Afonina, E., Gulnik, S. V., Yu, B., Erickson, J. W., Randad, R., Medjahed, D. and Silva, A. M. (2002) Structures of Ser205 mutant plasmepsin II from *Plasmodium falciparum* at 1.8 A in complex with the inhibitors rs367 and rs370, *Acta Crystallogr D Biol Crystallogr*, 58(Pt 12), pp. 2001-8.

17 Asojo, O. A., Gulnik, S. V., Afonina, E., Yu, B., Ellman, J. A., Haque, T. S. and Silva, A. M. (2003) Novel uncomplexed and complexed structures of plasmepsin II, an aspartic protease from *Plasmodium falciparum*, *J Mol Biol*, 327(1), pp. 173-81.

18 Augagneur, Y., Wesolowski, D., Tae, H. S., Altman, S. and Ben Mamoun, C. (2012) Gene selective mRNA cleavage inhibits the development of *Plasmodium falciparum*, *Proc Natl Acad Sci U S A*, 109(16), pp. 6235-40.

19 Banerjee, R., Liu, J., Beatty, W., Pelosof, L., Klemba, M. and Goldberg, D. E. (2002) Four plasmepsins are active in the *Plasmodium falciparum* food vacuole, including a protease with an active-site histidine, *Proc Natl Acad Sci U S A*, 99(2), pp. 990-5.

20 Barr, J. (2011) A short history of dapsone, or an alternative model of drug development, *J Hist Med Allied Sci*, 66(4), pp. 425-67.

21 Bates, M. D., Meshnick, S. R., Sigler, C. I., Leland, P. and Hollingdale, M. R. (1990) *In vitro* effects of primaquine and primaquine metabolites on exoerythrocytic stages of *Plasmodium berghei*, *Am J Trop Med Hyg*, 42(6), pp. 532-7.

22 Bates, P. A., Kelley, L. A., MacCallum, R. M. and Sternberg, M. J. (2001) Enhancement of protein modeling by human intervention in applying the automatic programs 3D-JIGSAW and 3D-PSSM, *Proteins,* Suppl 5, pp. 39-46.

23 Beeson, J. G., Amin, N., Kanjala, M. and Rogerson, S. J. (2002) Selective accumulation of mature asexual stages of *Plasmodium falciparum*-infected erythrocytes in the placenta, *Infect Immun,* 70(10), pp. 5412-5.

24 Beeson, J. G. and Duffy, P. E. (2005) The immunology and pathogenesis of malaria during pregnancy, *Curr Top Microbiol Immunol,* 297, pp. 187-227.

25 Bejon, P., Lusingu, J., Olotu, A., Leach, A., Lievens, M., Vekemans, J., Mshamu, S., Lang, T., *et al.* (2008) Efficacy of RTS,S/AS01E vaccine against malaria in children 5 to 17 months of age, *N Engl J Med,* 359(24), pp. 2521-32.

26 Bejon, P., Mwacharo, J., Kai, O., Mwangi, T., Milligan, P., Todryk, S., Keating, S., Lang, T., *et al.* (2006) A phase 2b randomised trial of the candidate malaria vaccines FP9 ME-TRAP and MVA ME-TRAP among children in Kenya, *PLoS Clin Trials,* 1(6), pp. e29.

27 Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. and Bourne, P. E. (2000) The Protein Data Bank, *Nucleic Acids Res,* 28(1), pp. 235-42.

28 Bernstein, N. K., Cherney, M. M., Loetscher, H., Ridley, R. G. and James, M. N. (1999) Crystal structure of the novel aspartic proteinase zymogen proplasmepsin II from *Plasmodium falciparum*, *Nat Struct Biol,* 6(1), pp. 32-7.

29 Beyer, B. B., Goldfarb, N. E. and Dunn, B. M. (2004) Expression, purification, and characterization of aspartic endopeptidases: Plasmodium plasmepsins and "short" recombinant human pseudocathepsin, *Curr Protoc Protein Sci*, Chapter 21, pp. Unit 21 14.

30 Bhattacharjee, A. K. and Karle, J. M. (1998) Functional correlation of molecular electronic properties with potency of synthetic carbinolamine antimalarial agents, *Bioorg Med Chem,* 6(10), pp. 1927-33.

31 Bhaumik, P., Horimoto, Y., Xiao, H., Miura, T., Hidaka, K., Kiso, Y., Wlodawer, A., Yada, R. Y., *et al.* (2011) Crystal structures of the free and inhibited forms of plasmepsin I (PMI) from *Plasmodium falciparum*, *J Struct Biol,* 175(1), pp. 73-84.

32 Bindschedler, M., Lefevre, G., Ezzet, F., Schaeffer, N., Meyer, I. and Thomsen, M. S. (2000) Cardiac effects of co-artemether (artemether/lumefantrine) and

mefloquine given alone or in combination to healthy volunteers, *Eur J Clin Pharmacol,* 56(5), pp. 375-81.

33 Birkholtz, L. M., Blatch, G., Coetzer, T. L., Hoppe, H. C., Human, E., Morris, E. J., Ngcete, Z., Oldfield, L., *et al.* (2008) Heterologous expression of plasmodial proteins for structural studies and functional annotation, *Malar J,* 7, pp. 197.

34 Bleicher, K. H., Bohm, H. J., Muller, K. and Alanine, A. I. (2003) Hit and lead generation: beyond high-throughput screening, *Nat Rev Drug Discov,* 2(5), pp. 369-78.

35 Bloland, P. (2001) *Drug resistance in malaria*, Chamberlee: World Health Organisation.

36 Blumberg, L., Lee, R. P., Lipman, J. and Beards, S. (1996) Predictors of mortality in severe malaria: a two year experience in a non-endemic area, *Anaesth Intensive Care*, 24(2), pp. 217-23.

37 Boddey, J. A., Hodder, A. N., Gunther, S., Gilson, P. R., Patsiouras, H., Kapp, E. A., Pearce, J. A., de Koning-Ward, T. F., *et al.* (2010) An aspartyl protease directs malaria effector proteins to the host cell, *Nature*, 463(7281), pp. 627-31.

38 Bonilla, J. A., Bonilla, T. D., Yowell, C. A., Fujioka, H. and Dame, J. B. (2007) Critical roles for the digestive vacuole plasmepsins of *Plasmodium falciparum* in vacuolar function, *Mol Microbiol,* 65(1), pp. 64-75.

39 Borrmann, S., Sasi, P., Mwai, L., Bashraheil, M., Abdallah, A., Muriithi, S., Fruhauf, H., Schaub, B., *et al.* (2011) Declining responsiveness of *Plasmodium falciparum* infections to artemisinin-based combination treatments on the Kenyan coast, *PLoS One,* 6(11), pp. e26005.

40 Bouillon, A., Giganti, D., Benedet, C., Gorgette, O., Petres, S., Crublet, E., Girard-Blanc, C., Witkowski, B., *et al.* (2013) *In Silico* screening on the three-dimensional model of the *Plasmodium vivax* SUB1 protease leads to the validation of a novel anti-parasite compound, *J Biol Chem,* 288(25), pp. 18561-73.

41 Bousema, T., Okell, L., Shekalaghe, S., Griffin, J. T., Omar, S., Sawa, P., Sutherland, C., Sauerwein, R., *et al.* (2010) Revisiting the circulation time of *Plasmodium falciparum* gametocytes: molecular detection methods to estimate the duration of gametocyte carriage and the effect of gametocytocidal drugs, *Malar J,* 9, pp. 136.

42 Bowers, K., Chow, E., Xu, H., Dror, R. O., Eastwood, M. P., Gregersen, B. A., Klepeis, J. L., Kolossvary, I., *et al.* 'Scalable Algorithms for Molecular

Dynamics Simulations on Commodity Clusters'. *Proceedings of the ACM/IEEE Conference on Supercomputing*, Tampa, Florida, 11-17 November.

43 Bozdech, Z., Llinas, M., Pulliam, B. L., Wong, E. D., Zhu, J. and DeRisi, J. L. (2003) The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*, *PLoS Biol*, 1(1), pp. E5.

44 Bray, P. G., Mungthin, M., Ridley, R. G. and Ward, S. A. (1998) Access to hematin: the basis of chloroquine resistance, *Mol Pharmacol*, 54(1), pp. 170-9.

45 Briolant, S., Henry, M., Oeuvray, C., Amalvict, R., Baret, E., Didillon, E., Rogier, C. and Pradines, B. (2010) Absence of association between piperaquine *in vitro* responses and polymorphisms in the *pfcrt*, *pfmdr1*, *pfmrp*, and *pfnhe* genes in *Plasmodium falciparum*, *Antimicrob Agents Chemother*, 54(9), pp. 3537-44.

46 Brueckner, R., Ohrt, C., Baird, J. K. and Milhous, W. (2001) 8-aminoquinolines, in Rosenthal, P. (ed.) *Antimalarial Chemotherapy: Mechanisms of action, resistance, and new directions in drug discovery*: Humana Press, pp. 123-148.

47 Bruneel, F., Hocqueloux, L., Alberti, C., Wolff, M., Chevret, S., Bedos, J. P., Durand, R., Le Bras, J., *et al.* (2003) The clinical spectrum of severe imported falciparum malaria in the intensive care unit: report of 188 cases in adults, *Am J Respir Crit Care Med*, 167(5), pp. 684-9.

48 Buonsenso, D. and Cataldi, L. (2010) Watch out for malaria: still a leading cause of child death worldwide, *Ital J Pediatr*, 36, pp. 58.

49 Butterworth, A. S., Skinner-Adams, T. S., Gardiner, D. L. and Trenholme, K. R. (2013) *Plasmodium falciparum* gametocytes: with a view to a kill, *Parasitology*, 140(14), pp. 1718-34.

50 Castel, D. A. (2003) Antimalarial Agents, in J., A.D. (ed.) *Burger's Medicinal Chemistry and Drug Discovery.* 6th ed. New York: Wiley, pp. 943.

51 Cawley, N. X., Olsen, V., Zhang, C. F., Chen, H. C., Tan, M. and Loh, Y. P. (1998) Activation and processing of non-anchored yapsin 1 (Yap3p), *J Biol Chem*, 273(1), pp. 584-91.

52 Charman, S. A., Arbe-Barnes, S., Bathurst, I. C., Brun, R., Campbell, M., Charman, W. N., Chiu, F. C., Chollet, J., *et al.* (2011) Synthetic ozonide drug candidate OZ439 offers new hope for a single-dose cure of uncomplicated malaria, *Proc Natl Acad Sci U S A*, 108(11), pp. 4400-5.

53 Choi, Y. J., Kim, H. S., Lee, S. H., Park, J. S., Nam, H. S., Kim, H. J., Kim, C. J., Jeong, D. J., *et al.* (2011) Evaluation of peptide nucleic acid array for the

detection of hepatitis B virus mutations associated with antiviral resistance, *Arch Virol,* 156(9), pp. 1517-24.

54 Clemente, J. C., Govindasamy, L., Madabushi, A., Fisher, S. Z., Moose, R. E., Yowell, C. A., Hidaka, K., Kimura, T., *et al.* (2006) Structure of the aspartic protease plasmepsin 4 from the malarial parasite *Plasmodium malariae* bound to an allophenylnorstatine-based inhibitor, *Acta Crystallogr D Biol Crystallogr,* 62(Pt 3), pp. 246-52.

55 Coombs, G. H., Goldberg, D. E., Klemba, M., Berry, C., Kay, J. and Mottram, J. C. (2001) Aspartic proteases of *Plasmodium falciparum* and other parasitic protozoa as drug targets, *Trends Parasitol,* 17(11), pp. 532-7.

56 Cooper, J. B. (2002) Aspartic proteinases in disease: a structural perspective, *Curr Drug Targets,* 3(2), pp. 155-73.

57 Correa, A. and Oppezzo, P. (2011) Tuning different expression parameters to achieve soluble recombinant proteins in *E. coli*: advantages of high-throughput screening, *Biotechnol J,* 6(6), pp. 715-30.

58 Dalal, S. and Klemba, M. (2007) Roles for two aminopeptidases in vacuolar hemoglobin catabolism in *Plasmodium falciparum*, *J Biol Chem,* 282(49), pp. 35978-87.

59 Dame, J. B., Yowell, C. A., Omara-Opyene, L., Carlton, J. M., Cooper, R. A. and Li, T. (2003) Plasmepsin 4, the food vacuole aspartic proteinase found in all *Plasmodium spp*. infecting man, *Mol Biochem Parasitol,* 130(1), pp. 1-12.

60 Daubenberger, C. A., Tisdale, E. J., Curcic, M., Diaz, D., Silvie, O., Mazier, D., Eling, W., Bohrmann, B., *et al.* (2003) The N'-terminal domain of glyceraldehyde-3-phosphate dehydrogenase of the apicomplexan *Plasmodium falciparum* mediates GTPase Rab2-dependent recruitment to membranes, *Biol Chem,* 384(8), pp. 1227-37.

61 Davis, T. M., Karunajeewa, H. A. and Ilett, K. F. (2005) Artemisinin-based combination therapies for uncomplicated malaria, *Med J Aust,* 182(4), pp. 181-5.

62 De Clercq, E. (2009) The history of antiretrovirals: key discoveries over the past 25 years, *Rev Med Virol,* 19(5), pp. 287-99.

63 De Vries, P. J. and Dien, T. (1996) Clinical pharmacology and therapeutic potential of artemisinin and its derivatives in the treatment of malaria., *Drugs,* 52, pp. 818-836.

64 Degliesposti, G., Kasam, V., Da Costa, A., Kang, H. K., Kim, N., Kim, D. W., Breton, V., Kim, D., *et al.* (2009) Design and discovery of plasmepsin II inhibitors using an automated workflow on large-scale grids, *ChemMedChem*, 4(7), pp. 1164-73.

65 Dell'Agli, M., Parapini, S., Galli, G., Vaiana, N., Taramelli, D., Sparatore, A., Liu, P., Dunn, B. M., *et al.* (2006) High antiplasmodial activity of novel plasmepsins I and II inhibitors, *J Med Chem*, 49(25), pp. 7440-9.

66 DePristo, M. A., Zilversmit, M. M. and Hartl, D. L. (2006) On the abundance, amino acid composition, and evolutionary dynamics of low-complexity regions in proteins, *Gene*, 378, pp. 19-30.

67 Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., Dufayard, J. F., Guindon, S., *et al.* (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist, *Nucleic Acids Res*, 36(Web Server issue), pp. W465-9.

68 Dixon, S. L., Smondyrev, A. M., Knoll, E. H., Rao, S. N., Shaw, D. E. and Friesner, R. A. (2006a) PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results, *J Comput Aided Mol Des*, 20(10-11), pp. 647-71.

69 Dixon, S. L., Smondyrev, A. M. and Rao, S. N. (2006b) PHASE: a novel approach to pharmacophore modeling and 3D database searching, *Chem Biol Drug Des*, 67(5), pp. 370-2.

70 Djogbenou, L., Pasteur, N., Akogbeto, M., Weill, M. and Chandre, F. (2010) Insecticide resistance in the *Anopheles gambiae* complex in Benin: a nationwide survey, *Med Vet Entomol*.

71 Dondorp, A. M., Nosten, F., Yi, P., Das, D., Phyo, A. P., Tarning, J., Lwin, K. M., Ariey, F., *et al.* (2009) Artemisinin resistance in *Plasmodium falciparum* malaria, *N Engl J Med*, 361(5), pp. 455-67.

72 Doolittle, R. F. (1992) Stein and Moore Award address. Reconstructing history with amino acid sequences, *Protein Sci*, 1(2), pp. 191-200.

73 Draper, S. J., Goodman, A. L., Biswas, S., Forbes, E. K., Moore, A. C., Gilbert, S. C. and Hill, A. V. (2009) Recombinant viral vaccines expressing merozoite surface protein-1 induce antibody- and T cell-mediated multistage protection against malaria, *Cell Host Microbe*, 5(1), pp. 95-105.

74 Duffy, P. E. and Fried, M. (2005) Malaria in the pregnant woman, *Curr Top Microbiol Immunol*, 295, pp. 169-200.

75 Dunn, B. M. (2002) Structure and mechanism of the pepsin-like family of aspartic peptidases, *Chem Rev,* 102(12), pp. 4431-58.

76 Durrand, V., Berry, A., Sem, R., Glaziou, P., Beaudou, J. and Fandeur, T. (2004) Variations in the sequence and expression of the *Plasmodium falciparum* chloroquine resistance transporter (Pfcrt) and their relationship to chloroquine resistance *in vitro*, *Mol Biochem Parasitol,* 136(2), pp. 273-85.

77 Eastman, R. T. and Fidock, D. A. (2009) Artemisinin-based combination therapies: a vital tool in efforts to eliminate malaria, *Nat Rev Microbiol,* 7(12), pp. 864-74.

78 Egan, T. J., Combrinck, J. M., Egan, J., Hearne, G. R., Marques, H. M., Ntenteni, S., Sewell, B. T., Smith, P. J., *et al.* (2002) Fate of haem iron in the malaria parasite *Plasmodium falciparum*, *Biochem J,* 365(Pt 2), pp. 343-7.

79 Eisele, T. P., Larsen, D. and Steketee, R. W. (2010) Protective efficacy of interventions for preventing malaria mortality in children in *Plasmodium falciparum* endemic areas, *Int J Epidemiol,* 39 Suppl 1, pp. i88-101.

80 Eisele, T. P., Miller, J. M., Moonga, H. B., Hamainza, B., Hutchinson, P. and Keating, J. (2011) Malaria infection and anemia prevalence in Zambia's Luangwa District: an area of near-universal insecticide-treated mosquito net coverage, *Am J Trop Med Hyg,* 84(1), pp. 152-7.

81 Eksi, S., Czesny, B., van Gemert, G. J., Sauerwein, R. W., Eling, W. and Williamson, K. C. (2007) Inhibition of *Plasmodium falciparum* oocyst production by membrane-permeant cysteine protease inhibitor E64d, *Antimicrob Agents Chemother,* 51(3), pp. 1064-70.

82 English, M., Waruiru, C., Amukoye, E., Murphy, S., Crawley, J., Mwangi, I., Peshu, N. and Marsh, K. (1996) Deep breathing in children with severe malaria: indicator of metabolic acidosis and poor outcome, *Am J Trop Med Hyg,* 55(5), pp. 521-4.

83 Ersmark, K., Nervall, M., Gutierrez-de-Teran, H., Hamelink, E., Janka, L. K., Clemente, J. C., Dunn, B. M., Gogoll, A., *et al.* (2006) Macrocyclic inhibitors of the malarial aspartic proteases plasmepsin I, II, and IV, *Bioorg Med Chem,* 14(7), pp. 2197-208.

84 Ersmark, K., Samuelsson, B. and Hallberg, A. (2006) Plasmepsins as potential targets for new antimalarial therapy, *Med Res Rev,* 26(5), pp. 626-66.

85 Esposito, D. and Chatterjee, D. K. (2006) Enhancement of soluble protein expression through the use of fusion tags, *Curr Opin Biotechnol,* 17(4), pp. 353-8.

86 Ezzet, F., van Vugt, M., Nosten, F., Looareesuwan, S. and White, N. J. (2000) Pharmacokinetics and pharmacodynamics of lumefantrine (benflumetol) in acute falciparum malaria, *Antimicrob Agents Chemother,* 44(3), pp. 697-704.

87 Famin, O., Krugliak, M. and Ginsburg, H. (1999) Kinetics of inhibition of glutathione-mediated degradation of ferriprotoporphyrin IX by antimalarial drugs, *Biochem Pharmacol,* 58(1), pp. 59-68.

88 Ferrari, V. and Cutler, D. J. (1991) Kinetics and thermodynamics of chloroquine and hydroxychloroquine transport across the human erythrocyte membrane, *Biochem Pharmacol,* 41(1), pp. 23-30.

89 Fidock, D. A. and Wellems, T. E. (1997) Transformation with human dihydrofolate reductase renders malaria parasites insensitive to WR99210 but does not affect the intrinsic activity of proguanil, *Proc Natl Acad Sci U S A,* 94(20), pp. 10931-6.

90 Fitch, C. D. (2004) Ferriprotoporphyrin IX, phospholipids, and the antimalarial actions of quinoline drugs, *Life Sci,* 74(16), pp. 1957-72.

91 Flick, K., Ahuja, S., Chene, A., Bejarano, M. T. and Chen, Q. (2004) Optimized expression of *Plasmodium falciparum* erythrocyte membrane protein 1 domains in *Escherichia coli*, *Malar J,* 3, pp. 50.

92 Flick, K. and Chen, Q. (2004) var genes, PfEMP1 and the human host, *Mol Biochem Parasitol,* 134(1), pp. 3-9.

93 Florens, L., Washburn, M. P., Raine, J. D., Anthony, R. M., Grainger, M., Haynes, J. D., Moch, J. K., Muster, N., *et al.* (2002) A proteomic view of the *Plasmodium falciparum* life cycle, *Nature,* 419(6906), pp. 520-6.

94 Francesconi, P., Fabiani, M., Dente, M. G., Lukwiya, M., Okwey, R., Ouma, J., Ochakachon, R., Cian, F., *et al.* (2001) HIV, malaria parasites, and acute febrile episodes in Ugandan adults: a case-control study, *AIDS,* 15(18), pp. 2445-50.

95 Francis, S. E., Gluzman, I. Y., Oksman, A., Knickerbocker, A., Mueller, R., Bryant, M. L., Sherman, D. R., Russell, D. G., *et al.* (1994) Molecular characterization and inhibition of a *Plasmodium falciparum* aspartic hemoglobinase, *EMBO J,* 13(2), pp. 306-17.

96 Francis, S. E., Sullivan, D. J., Jr. and Goldberg, D. E. (1997) Hemoglobin metabolism in the malaria parasite *Plasmodium falciparum*, *Annu Rev Microbiol*, 51, pp. 97-123.

97 Franke, M. F., Spiegelman, D., Ezeamama, A., Aboud, S., Msamanga, G. I., Mehta, S. and Fawzi, W. W. (2010) Malaria parasitemia and CD4 T cell count, viral load, and adverse HIV outcomes among HIV-infected pregnant women in Tanzania, *Am J Trop Med Hyg*, 82(4), pp. 556-62.

98 Frederich, M., Dogne, J. M., Angenot, L. and De Mol, P. (2002) New trends in anti-malarial agents, *Curr Med Chem*, 9(15), pp. 1435-56.

99 Freville, A., Landrieu, I., Garcia-Gimeno, M. A., Vicogne, J., Montbarbon, M., Bertin, B., Verger, A., Kalamou, H., *et al.* (2012) *Plasmodium falciparum* inhibitor-3 homolog increases protein phosphatase type 1 activity and is essential for parasitic survival, *J Biol Chem*, 287(2), pp. 1306-21.

100 Fried, M. and Duffy, P. E. (1998) Maternal malaria and parasite adhesion, *J Mol Med*, 76(3-4), pp. 162-71.

101 Fried, M., Nosten, F., Brockman, A., Brabin, B. J. and Duffy, P. E. (1998) Maternal antibodies block malaria, *Nature*, 395(6705), pp. 851-2.

102 Friedman, T. C., Gordon, V. M., Leppla, S. H., Klimpel, K. R., Birch, N. P. and Loh, Y. P. (1995) *In vitro* processing of anthrax toxin protective antigen by recombinant PC1 (SPC3) and bovine intermediate lobe secretory vesicle membranes, *Arch Biochem Biophys*, 316(1), pp. 5-13.

103 Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., Repasky, M. P., Knoll, E. H., *et al.* (2004) Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy, *J Med Chem*, 47(7), pp. 1739-49.

104 Friesner, R. A., Murphy, R. B., Repasky, M. P., Frye, L. L., Greenwood, J. R., Halgren, T. A., Sanschagrin, P. C. and Mainz, D. T. (2006) Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes, *J Med Chem*, 49(21), pp. 6177-96.

105 Gagaring, K., Borboa, R., Francek, C., Chen, Z., Buenviaje, J., Plouffe, D., Winzeler, E., Brinker, A., *et al.* (2013) Novartis-GNF Malaria Box (Accessed.

106 Gamo, F. J., Sanz, L. M., Vidal, J., de Cozar, C., Alvarez, E., Lavandera, J. L., Vanderwall, D. E., Green, D. V., *et al.* (2010) Thousands of chemical starting points for antimalarial lead identification, *Nature*, 465(7296), pp. 305-10.

107 Gardiner, D. L., Holt, D. C., Thomas, E. A., Kemp, D. J. and Trenholme, K. R. (2000) Inhibition of *Plasmodium falciparum clag9* gene function by antisense RNA, *Mol Biochem Parasitol,* 110(1), pp. 33-41.

108 Gardiner, D. L., Trenholme, K. R., Skinner-Adams, T. S., Stack, C. M. and Dalton, J. P. (2006) Overexpression of leucyl aminopeptidase in *Plasmodium falciparum* parasites. Target for the antimalarial activity of bestatin, *J Biol Chem,* 281(3), pp. 1741-5.

109 Gardner, M. J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R. W., Carlton, J. M., Pain, A., *et al.* (2002) Genome sequence of the human malaria parasite *Plasmodium falciparum*, *Nature,* 419(6906), pp. 498-511.

110 Gassmann, M., Grenacher, B., Rohde, B. and Vogel, J. (2009) Quantifying Western blots: pitfalls of densitometry, *Electrophoresis,* 30(11), pp. 1845-55.

111 Genton, B. and D'Acremont (2001) Clinical features of malaria in returning travelers and migrants., in SP, H. (ed.) *Travelers' malaria*: BC Decker, pp. 371-392.

112 Ghose, A. K., Viswanadhan, V. N. and Wendoloski, J. J. (1999) A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases, *J Comb Chem,* 1(1), pp. 55-68.

113 Ghose, A. K., Viswanadhan, V.N. and Wendoloski, J.J. (1998) Prediction of Hydrophobic (Lipophilic) Properties of Small Organic Molecules Using Fragment Methods: An Analysis of AlogP and CLogP Methods, *Journal of Physical Chemistry A,* 102, pp. 3762-3772.

114 Gildberg, A., Olsen, R. L. and Bjarnason, J. B. (1990) Catalytic properties and chemical composition of pepsins from Atlantic cod (*Gadus morhua*), *Comp Biochem Physiol B,* 96(2), pp. 323-30.

115 GlaxoSmithKline 2008. Update on GSK's malaria treatments: Dacart and Lapdap. London, UK

Geneva, Switzerland: Brentford: GlaxoSmithKline.

116 Gligorijevic, B., McAllister, R., Urbach, J. S. and Roepe, P. D. (2006) Spinning disk confocal microscopy of live, intraerythrocytic malarial parasites. 1. Quantification of hemozoin development for drug sensitive versus resistant malaria, *Biochemistry,* 45(41), pp. 12400-10.

117 Glushakova, S., Mazar, J., Hohmann-Marriott, M. F., Hama, E. and Zimmerberg, J. (2009) Irreversible effect of cysteine protease inhibitors on the release of malaria parasites from infected erythrocytes, *Cell Microbiol,* 11(1), pp. 95-105.

118 Goldberg, D. E., Slater, A. F., Beavis, R., Chait, B., Cerami, A. and Henderson, G. B. (1991) Hemoglobin degradation in the human malaria pathogen *Plasmodium falciparum*: a catabolic pathway initiated by a specific aspartic protease, *J Exp Med,* 173(4), pp. 961-9.

119 Goldberg, D. E., Slater, A. F., Cerami, A. and Henderson, G. B. (1990) Hemoglobin degradation in the malaria parasite *Plasmodium falciparum*: an ordered process in a unique organelle, *Proc Natl Acad Sci U S A,* 87(8), pp. 2931-5.

120 Graslund, S., Nordlund, P., Weigelt, J., Hallberg, B. M., Bray, J., Gileadi, O., Knapp, S., Oppermann, U., *et al.* (2008) Protein production and purification, *Nat Methods,* 5(2), pp. 135-46.

121 Greenwood, J. R., Calkins, D., Sullivan, A. P. and Shelley, J. C. (2010) Towards the comprehensive, rapid, and accurate prediction of the favorable tautomeric states of drug-like molecules in aqueous solution, *J Comput Aided Mol Des,* 24(6-7), pp. 591-604.

122 Grimwade, K., French, N., Mbatha, D. D., Zungu, D. D., Dedicoat, M. and Gilks, C. F. (2004) HIV infection as a cofactor for severe falciparum malaria in adults living in a region of unstable malaria transmission in South Africa, *AIDS,* 18(3), pp. 547-54.

123 Groom, C. R., Allen, F. H. and Henderson, S. (2013) The Cambridge Structural Database (CSD), *Reference Module in Chemistry, Molecular Sciences and Chemical Engineering*: Elsevier.

124 Guiguemde, W. A., Shelat, A. A., Bouck, D., Duffy, S., Crowther, G. J., Davis, P. H., Smithson, D. C., Connelly, M., *et al.* (2010) Chemical genetics of *Plasmodium falciparum*, *Nature,* 465(7296), pp. 311-5.

125 Gupta, S., Thapar, M. M., Mariga, S. T., Wernsdorfer, W. H. and Bjorkman, A. (2002) *Plasmodium falciparum: in vitro* interactions of artemisinin with amodiaquine, pyronaridine, and chloroquine, *Exp Parasitol,* 100(1), pp. 28-35.

126 Guruprasad, L., Tanneeru, K. and Guruprasad, K. (2011) Structural rationale for the recognition of arginine at P(3) in PEXEL motif containing proteins of *Plasmodium falciparum* by plasmepsin V, *Protein Pept Lett,* 18(6), pp. 634-41.

127 Gutierrez-de-Teran, H., Nervall, M., Ersmark, K., Liu, P., Janka, L. K., Dunn, B., Hallberg, A. and Aqvist, J. (2006) Inhibitor binding to the plasmepsin IV aspartic protease from *Plasmodium falciparum*, *Biochemistry*, 45(35), pp. 10529-41.

128 Haase, S. and de Koning-Ward, T. F. (2010) New insights into protein export in malaria parasites, *Cell Microbiol*, 12(5), pp. 580-7.

129 Halgren, T. A., Murphy, R. B., Friesner, R. A., Beard, H. S., Frye, L. L., Pollard, W. T. and Banks, J. L. (2004) Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening, *J Med Chem*, 47(7), pp. 1750-9.

130 Hassan Alin, M., Bjorkman, A. and Wernsdorfer, W. H. (1999) Synergism of benflumetol and artemether in *Plasmodium falciparum*, *Am J Trop Med Hyg*, 61(3), pp. 439-45.

131 Hawley, S. R., Bray, P. G., Mungthin, M., Atkinson, J. D., O'Neill, P. M. and Ward, S. A. (1998) Relationship between antimalarial drug activity, accumulation, and inhibition of heme polymerization in *Plasmodium falciparum in vitro*, *Antimicrob Agents Chemother*, 42(3), pp. 682-6.

132 Hawthorne, P. L., Trenholme, K. R., Skinner-Adams, T. S., Spielmann, T., Fischer, K., Dixon, M. W., Ortega, M. R., Anderson, K. L., *et al.* (2004) A novel *Plasmodium falciparum* ring stage protein, REX, is located in Maurer's clefts, *Mol Biochem Parasitol*, 136(2), pp. 181-9.

133 Hay, S. I., Guerra, C. A., Tatem, A. J., Noor, A. M. and Snow, R. W. (2004) The global distribution and population at risk of malaria: past, present, and future, *Lancet Infect Dis*, 4(6), pp. 327-36.

134 Hien, T. T. and White, N. J. (1993) Qinghaosu, *Lancet*, 341(8845), pp. 603-8.

135 Hill, J., Tyas, L., Phylip, L. H., Kay, J., Dunn, B. M. and Berry, C. (1994) High level expression and characterisation of Plasmepsin II, an aspartic proteinase from *Plasmodium falciparum*, *FEBS Lett*, 352(2), pp. 155-8.

136 Hobbs, C. V., Tanaka, T. Q., Muratova, O., Van Vliet, J., Borkowsky, W., Williamson, K. C. and Duffy, P. E. (2013) HIV treatments have malaria gametocyte killing and transmission blocking activity, *J Infect Dis*, 208(1), pp. 139-48.

137 Hobbs, C. V., Voza, T., Coppi, A., Kirmse, B., Marsh, K., Borkowsky, W. and Sinnis, P. (2009) HIV protease inhibitors inhibit the development of preerythrocytic-stage plasmodium parasites, *J Infect Dis*, 199(1), pp. 134-41.

138 Hochman, S. and Kim, K. (2012) The Impact of HIV Coinfection on Cerebral Malaria Pathogenesis, *J Neuroparasitology*, 3.

139 Hoppe, H. C., van Schalkwyk, D. A., Wiehart, U. I., Meredith, S. A., Egan, J. and Weber, B. W. (2004) Antimalarial quinolines and artemisinin inhibit endocytosis in *Plasmodium falciparum*, *Antimicrob Agents Chemother*, 48(7), pp. 2370-8.

140 Hu, G., Cabrera, A., Kono, M., Mok, S., Chaal, B. K., Haase, S., Engelberg, K., Cheemadan, S., *et al.* (2010) Transcriptional profiling of growth perturbations of the human malaria parasite *Plasmodium falciparum*, *Nat Biotechnol*, 28(1), pp. 91-8.

141 Humphrey, W., Dalke, A. and Schulten, K. (1996) VMD: visual molecular dynamics, *J Mol Graph*, 14(1), pp. 33-8, 27-8.

142 Humphries, D., Mosites, E., Otchere, J., Twum, W. A., Woo, L., Jones-Sanpei, H., Harrison, L. M., Bungiro, R. D., *et al.* (2011) Epidemiology of hookworm infection in Kintampo North Municipality, Ghana: patterns of malaria coinfection, anemia, and albendazole treatment failure, *Am J Trop Med Hyg*, 84(5), pp. 792-800.

143 Hyde, J. E. (2005) Drug-resistant malaria, *Trends Parasitol*, 21(11), pp. 494-8.

144 Ibrahim, M. L., Steenkeste, N., Khim, N., Adam, H. H., Konate, L., Coppee, J. Y., Ariey, F. and Duchemin, J. B. (2009) Field-based evidence of fast and global increase of *Plasmodium falciparum* drug-resistance by DNA-microarrays and PCR/RFLP in Niger, *Malar J*, 8, pp. 32.

145 Imani, P. D., Musoke, P., Byarugaba, J. and Tumwine, J. K. (2011) Human immunodeficiency virus infection and cerebral malaria in children in Uganda: a case-control study, *BMC Pediatr*, 11, pp. 5.

146 Jackson, K. E., Spielmann, T., Hanssen, E., Adisa, A., Separovic, F., Dixon, M. W., Trenholme, K. R., Hawthorne, P. L., *et al.* (2007) Selective permeabilization of the host cell membrane of *Plasmodium falciparum*-infected red blood cells with streptolysin O and equinatoxin II, *Biochem J*, 403(1), pp. 167-75.

147 Jani, D., Nagarkatti, R., Beatty, W., Angel, R., Slebodnick, C., Andersen, J., Kumar, S. and Rathore, D. (2008) HDP-a novel heme detoxification protein from the malaria parasite, *PLoS Pathog*, 4(4), pp. e1000053.

148 Jones, C. M., Machin, C., Mohammed, K., Majambere, S., Ali, A. S., Khatib, B. O., McHa, J., Ranson, H., *et al.* (2012) Insecticide resistance in *Culex*

quinquefasciatus from Zanzibar: implications for vector control programmes, *Parasit Vectors*, 5, pp. 78.

149 Kamya, M. R., Gasasira, A. F., Yeka, A., Bakyaita, N., Nsobya, S. L., Francis, D., Rosenthal, P. J., Dorsey, G., *et al.* (2006) Effect of HIV-1 infection on antimalarial treatment outcomes in Uganda: a population-based study, *J Infect Dis*, 193(1), pp. 9-15.

150 Karlsen, S., Hough, E. and Olsen, R. L. (1998) Structure and proposed amino-acid sequence of a pepsin from atlantic cod (*Gadus morhua*), *Acta Crystallogr D Biol Crystallogr*, 54(Pt 1), pp. 32-46.

151 Kasam, V., Salzemann, J., Botha, M., Dacosta, A., Degliesposti, G., Isea, R., Kim, D., Maass, A., *et al.* (2009) WISDOM-II: screening against multiple targets implicated in malaria using computational grid infrastructures, *Malar J*, 8, pp. 88.

152 Kelley, L. A. and Sternberg, M. J. (2009) Protein structure prediction on the Web: a case study using the Phyre server, *Nat Protoc*, 4(3), pp. 363-71.

153 Kim, Y., Babnigg, G., Jedrzejczak, R., Eschenfeldt, W. H., Li, H., Maltseva, N., Hatzos-Skintges, C., Gu, M., *et al.* (2011) High-throughput protein purification and quality assessment for crystallization, *Methods*, 55(1), pp. 12-28.

154 Kim, Y. M., Lee, M. H., Piao, T. G., Lee, J. W., Kim, J. H., Lee, S., Choi, K. M., Jiang, J. H., *et al.* (2006) Prodomain processing of recombinant plasmepsin II and IV, the aspartic proteases of *Plasmodium falciparum*, is auto- and trans-catalytic, *J Biochem*, 139(2), pp. 189-95.

155 Kirkpatrick, D. L., Watson, S. and Ulhaq, S. (1999) Structure-based drug design: combinatorial chemistry and molecular modeling, *Comb Chem High Throughput Screen*, 2(4), pp. 211-21.

156 Klemba, M., Beatty, W., Gluzman, I. and Goldberg, D. E. (2004) Trafficking of plasmepsin II to the food vacuole of the malaria parasite *Plasmodium falciparum*, *J Cell Biol*, 164(1), pp. 47-56.

157 Klemba, M., Gluzman, I. and Goldberg, D. E. (2004) A *Plasmodium falciparum* dipeptidyl aminopeptidase I participates in vacuolar hemoglobin degradation, *J Biol Chem*, 279(41), pp. 43000-7.

158 Klemba, M. and Goldberg, D. E. (2005) Characterization of plasmepsin V, a membrane-bound aspartic protease homolog in the endoplasmic reticulum of *Plasmodium falciparum*, *Mol Biochem Parasitol*, 143(2), pp. 183-91.

159 Kolevzon, N., Nasereddin, A., Naik, S., Yavin, E. and Dzikowski, R. (2014) Use of Peptide Nucleic Acids to Manipulate Gene Expression in the Malaria Parasite *Plasmodium falciparum*, *PLoS One*, 9(1), pp. e86802.

160 Kolinski, A., Betancourt, M. R., Kihara, D., Rotkiewicz, P. and Skolnick, J. (2001) Generalized comparative modeling (GENECOMP): a combination of sequence comparison, threading, and lattice modeling for protein structure prediction and refinement, *Proteins*, 44(2), pp. 133-49.

161 Kolinski, A., Rotkiewicz, P., Ilkowski, B. and Skolnick, J. (1999) A method for the improvement of threading-based protein models, *Proteins*, 37(4), pp. 592-610.

162 Krudsood, S., Tangpukdee, N., Thanchatwet, V., Wilairatana, P., Srivilairit, S., Pothipak, N., Jianping, S., Guoqiao, L., *et al.* (2007) Dose ranging studies of new artemisinin-piperaquine fixed combinations compared to standard regimens of artemisisnin combination therapies for acute uncomplicated falciparum malaria, *Southeast Asian J Trop Med Public Health*, 38(6), pp. 971-8.

163 Kuhn, H., Sahu, B., Rapireddy, S., Ly, D. H. and Frank-Kamenetskii, M. D. (2010) Sequence specificity at targeting double-stranded DNA with a gamma-PNA oligomer modified with guanidinium G-clamp nucleobases, *Artif DNA PNA XNA*, 1(1), pp. 45-53.

164 Kulzer, S., Charnaud, S., Dagan, T., Riedel, J., Mandal, P., Pesce, E. R., Blatch, G. L., Crabb, B. S., *et al.* (2012) *Plasmodium falciparum*-encoded exported hsp70/hsp40 chaperone/co-chaperone complexes within the host erythrocyte, *Cell Microbiol*, 14(11), pp. 1784-95.

165 Kulzer, S., Rug, M., Brinkmann, K., Cannon, P., Cowman, A., Lingelbach, K., Blatch, G. L., Maier, A. G., *et al.* (2010) Parasite-encoded Hsp40 proteins define novel mobile structures in the cytosol of the *P. falciparum*-infected erythrocyte, *Cell Microbiol*, 12(10), pp. 1398-420.

166 Lacaze-Dufaure, C., Najjar, F. and Andre-Barres, C. (2010) First computational evidence of a competitive stepwise and concerted mechanism for the reduction of antimalarial endoperoxides, *J Phys Chem B*, 114(30), pp. 9848-53.

167 LaCount, D. J., Vignali, M., Chettier, R., Phansalkar, A., Bell, R., Hesselberth, J. R., Schoenfeld, L. W., Ota, I., *et al.* (2005) A protein interaction network of the malaria parasite *Plasmodium falciparum*, *Nature*, 438(7064), pp. 103-7.

168 Langreth, S. G., Jensen, J. B., Reese, R. T. and Trager, W. (1978) Fine structure of human malaria *in vitro*, *J Protozool*, 25(4), pp. 443-52.

169 Lanners, H. N. (1991) Effect of the 8-aminoquinoline primaquine on culture-derived gametocytes of the malaria parasite *Plasmodium falciparum*, *Parasitol Res*, 77(6), pp. 478-81.

170 Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., *et al.* (2007) Clustal W and Clustal X version 2.0, *Bioinformatics*, 23(21), pp. 2947-8.

171 Laskowski, R. M., M.; Moss, D.; Thornton, J. (1993) PROCHECK - a program to check the stereochemical quality of protein structures, *J Applied Crystallography*, 26, pp. 283-291.

172 Latek, D., Ekonomiuk, D. and Kolinski, A. (2007) Protein structure prediction: combining *de novo* modeling with sparse experimental data, *J Comput Chem*, 28(10), pp. 1668-76.

173 Leach, A. R. (2001) *Molecular Modelling: Principles and Applications.* 2nd edn. Dorchester: Pearson Education Limited.

174 Lefevre, G., Bindschedler, M., Ezzet, F., Schaeffer, N., Meyer, I. and Thomsen, M. S. (2000) Pharmacokinetic interaction trial between co-artemether and mefloquine, *Eur J Pharm Sci*, 10(2), pp. 141-51.

175 Lek-Uthai, U., Suwanarusk, R., Ruengweerayut, R., Skinner-Adams, T. S., Nosten, F., Gardiner, D. L., Boonma, P., Piera, K. A., *et al.* (2008) Stronger activity of human immunodeficiency virus type 1 protease inhibitors against clinical isolates of *Plasmodium vivax* than against those of *P. falciparum*, *Antimicrob Agents Chemother*, 52(7), pp. 2435-41.

176 Lengeler, C. (2004) Insecticide-treated bed nets and curtains for preventing malaria, *Cochrane Database Syst Rev*, (2), pp. CD000363.

177 Li, T., Yowell, C. A., Beyer, B. B., Hung, S. H., Westling, J., Lam, M. T., Dunn, B. M. and Dame, J. B. (2004) Recombinant expression and enzymatic subsite characterization of plasmepsin 4 from the four Plasmodium species infecting man, *Mol Biochem Parasitol*, 135(1), pp. 101-9.

178 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(2), pp. 1432-7.

179 Liu, J., Istvan, E. S., Gluzman, I. Y., Gross, J. and Goldberg, D. E. (2006) *Plasmodium falciparum* ensures its amino acid supply with multiple acquisition pathways and redundant proteolytic enzyme systems, *Proc Natl Acad Sci U S A*, 103(23), pp. 8840-5.

180 Liu, P., Marzahn, M. R., Robbins, A. H., Gutierrez-de-Teran, H., Rodriguez, D., McClung, S. H., Stevens, S. M., Jr., Yowell, C. A., *et al.* (2009) Recombinant plasmepsin 1 from the human malaria parasite *plasmodium falciparum*: enzymatic characterization, active site inhibitor design, and structural analysis, *Biochemistry*, 48(19), pp. 4086-99.

181 Livak, K. J. and Schmittgen, T. D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method, *Methods*, 25(4), pp. 402-8.

182 Ljunggren, E. (2006) *Identification of the target of Antiretroviral agents in P.falciparum.* Master in Pharmacy Masters, Göteborgs University, Sweden, Brisbane.

183 Loria, P., Miller, S., Foley, M. and Tilley, L. (1999) Inhibition of the peroxidative degradation of haem as the basis of action of chloroquine and other quinoline antimalarials, *Biochem J*, 339 (Pt 2), pp. 363-70.

184 Lozano, R. and Naghavi, M. and Foreman, K. and Lim, S. and Shibuya, K. and Aboyans, V. and Abraham, J. and Adair, T., *et al.* (2012) Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010, *Lancet*, 380(9859), pp. 2095-128.

185 Macadangdang, B., Zhang, N., Lund, P. E., Marple, A. H., Okabe, M., Gottesman, M. M., Appella, D. H. and Kimchi-Sarfaty, C. (2011) Inhibition of multidrug resistance by SV40 pseudovirion delivery of an antigene peptide nucleic acid (PNA) in cultured cells, *PLoS One*, 6(3), pp. e17981.

186 Mackintosh, C. L., Beeson, J. G. and Marsh, K. (2004) Clinical features and pathogenesis of severe malaria, *Trends Parasitol*, 20(12), pp. 597-603.

187 MacPherson, G. G., Warrell, M. J., White, N. J., Looareesuwan, S. and Warrell, D. A. (1985) Human cerebral malaria. A quantitative ultrastructural analysis of parasitized erythrocyte sequestration, *Am J Pathol*, 119(3), pp. 385-401.

188 Author (2010) *Desmond Molecylar Dynamics System* (Version 2.4).

189 Magowan, C., Brown, J. T., Liang, J., Heck, J., Coppel, R. L., Mohandas, N. and Meyer-Ilse, W. (1997) Intracellular structures of normal and aberrant *Plasmodium falciparum* malaria parasites imaged by soft x-ray microscopy, *Proc Natl Acad Sci U S A*, 94(12), pp. 6222-7.

190 Maier, A. G., Braks, J. A., Waters, A. P. and Cowman, A. F. (2006) Negative selection using yeast cytosine deaminase/uracil phosphoribosyl transferase in

Plasmodium falciparum for targeted gene deletion by double crossover recombination, *Mol Biochem Parasitol,* 150(1), pp. 118-21.

191 Makanga, M., Bassat, Q., Falade, C. O., Premji, Z. G., Krudsood, S., Hunt, P., Walter, V., Beck, H. P., *et al.* (2011) Efficacy and safety of artemether-lumefantrine in the treatment of acute, uncomplicated *Plasmodium falciparum* malaria: a pooled analysis, *Am J Trop Med Hyg,* 85(5), pp. 793-804.

192 Makanga, M. and Krudsood, S. (2009) The clinical efficacy of artemether/lumefantrine (Coartem), *Malar J,* 8 Suppl 1, pp. S5.

193 Malkin, E., Long, C. A., Stowers, A. W., Zou, L., Singh, S., MacDonald, N. J., Narum, D. L., Miles, A. P., *et al.* (2007) Phase 1 study of two merozoite surface protein 1 (MSP1(42)) vaccines for *Plasmodium falciparum* malaria, *PLoS Clin Trials,* 2(4), pp. e12.

194 Marks, F., von Kalckreuth, V., Kobbe, R., Adjei, S., Adjei, O., Horstmann, R. D., Meyer, C. G. and May, J. (2005) Parasitological rebound effect and emergence of pyrimethamine resistance in *Plasmodium falciparum* after single-dose sulfadoxine-pyrimethamine, *J Infect Dis,* 192(11), pp. 1962-5.

195 Marsh, K., Forster, D., Waruiru, C., Mwangi, I., Winstanley, M., Marsh, V., Newton, C., Winstanley, P., *et al.* (1995) Indicators of life-threatening malaria in African children, *N Engl J Med,* 332(21), pp. 1399-404.

196 Mayxay, M., Thongpraseuth, V., Khanthavong, M., Lindegardh, N., Barends, M., Keola, S., Pongvongsa, T., Phompida, S., *et al.* (2006) An open, randomized comparison of artesunate plus mefloquine vs. dihydroartemisinin-piperaquine for the treatment of uncomplicated *Plasmodium falciparum* malaria in the Lao People's Democratic Republic (Laos), *Trop Med Int Health,* 11(8), pp. 1157-65.

197 McGeorge, R. (2009) *Examining the Antimalarial Activity of the HIV Protease Inhibitors.* Bachelor of Science with Honours, Griffith University, Brisbane, Australia.

198 McGowan, S., Porter, C. J., Lowther, J., Stack, C. M., Golding, S. J., Skinner-Adams, T. S., Trenholme, K. R., Teuscher, F., *et al.* (2009) Structural basis for the inhibition of the essential *Plasmodium falciparum* M1 neutral aminopeptidase, *Proc Natl Acad Sci U S A,* 106(8), pp. 2537-42.

199 Mehlin, C., Boni, E., Buckner, F. S., Engel, L., Feist, T., Gelb, M. H., Haji, L., Kim, D., *et al.* (2006) Heterologous expression of proteins from *Plasmodium*

falciparum: results from 1000 genes, *Mol Biochem Parasitol*, 148(2), pp. 144-60.

200 Modjarrad, K. and Vermund, S. H. (2010) Effect of treating co-infections on HIV-1 viral load: a systematic review, *Lancet Infect Dis*, 10(7), pp. 455-63.

201 Moon, R. P., Tyas, L., Certa, U., Rupp, K., Bur, D., Jacquet, C., Matile, H., Loetscher, H., *et al.* (1997) Expression and characterisation of plasmepsin I from *Plasmodium falciparum*, *Eur J Biochem*, 244(2), pp. 552-60.

202 Moore, J. M., Ayisi, J., Nahlen, B. L., Misore, A., Lal, A. A. and Udhayakumar, V. (2000) Immunity to placental malaria. II. Placental antigen-specific cytokine responses are impaired in human immunodeficiency virus-infected women, *J Infect Dis*, 182(3), pp. 960-4.

203 Mount, D. W. (2007) Using the Basic Local Alignment Search Tool (BLAST), *CSH Protoc*, 2007, pp. pdb top17.

204 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(12), pp. 4968-78.

205 Mueller, I., Zimmerman, P. A. and Reeder, J. C. (2007) *Plasmodium malariae* and *Plasmodium ovale*--the "bashful" malaria parasites, *Trends Parasitol*, 23(6), pp. 278-83.

206 Murphy, G. S. and Oldfield, E. C., 3rd (1996) Falciparum malaria, *Infect Dis Clin North Am*, 10(4), pp. 747-75.

207 Murray, C. J., Rosenfeld, L. C., Lim, S. S., Andrews, K. G., Foreman, K. J., Haring, D., Fullman, N., Naghavi, M., *et al.* (2012) Global malaria mortality between 1980 and 2010: a systematic analysis, *Lancet*, 379(9814), pp. 413-31.

208 Murray, C. J. and Vos, T. and Lozano, R. and Naghavi, M. and Flaxman, A. D. and Michaud, C. and Ezzati, M. and Shibuya, K., *et al.* (2012) Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010, *Lancet*, 380(9859), pp. 2197-223.

209 N'Guessan, R., Corbel, V., Akogbeto, M. and Rowland, M. (2007) Reduced efficacy of insecticide-treated nets and indoor residual spraying for malaria control in pyrethroid resistance area, Benin, *Emerg Infect Dis*, 13(2), pp. 199-206.

210 Nathoo, S., Serghides, L. and Kain, K. C. (2003) Effect of HIV-1 antiretroviral drugs on cytoadherence and phagocytic clearance of *Plasmodium falciparum*-parasitised erythrocytes, *Lancet,* 362(9389), pp. 1039-41.

211 Nielsen, M., Lundegaard, C., Lund, O. and Petersen, T. N. (2010) CPHmodels-3.0--remote homology modeling using structure-guided sequence profiles, *Nucleic Acids Res,* 38(Web Server issue), pp. W576-81.

212 Nkhoma, E. T., Poole, C., Vannappagari, V., Hall, S. A. and Beutler, E. (2009) The global prevalence of glucose-6-phosphate dehydrogenase deficiency: a systematic review and meta-analysis, *Blood Cells Mol Dis,* 42(3), pp. 267-78.

213 Noedl, H., Se, Y., Schaecher, K., Smith, B. L., Socheat, D., Fukuda, M. M. and Artemisinin Resistance in Cambodia 1 Study, C. (2008) Evidence of artemisinin-resistant malaria in western Cambodia, *N Engl J Med,* 359(24), pp. 2619-20.

214 Noedl, H., Socheat, D. and Satimai, W. (2009) Artemisinin-resistant malaria in Asia, *N Engl J Med,* 361(5), pp. 540-1.

215 Nosten, F., Rogerson, S. J., Beeson, J. G., McGready, R., Mutabingwa, T. K. and Brabin, B. (2004) Malaria in pregnancy and the endemicity spectrum: what can we learn?, *Trends Parasitol,* 20(9), pp. 425-32.

216 Nzila, A., Ochong, E., Nduati, E., Gilbert, K., Winstanley, P., Ward, S. and Marsh, K. (2005) Why has the dihydrofolate reductase 164 mutation not consistently been found in Africa yet?, *Trans R Soc Trop Med Hyg,* 99(5), pp. 341-6.

217 O'Neill, P. M., Willock, D. J., Hawley, S. R., Bray, P. G., Storr, R. C., Ward, S. A. and Park, B. K. (1997) Synthesis, antimalarial activity, and molecular modeling of tebuquine analogues, *J Med Chem,* 40(4), pp. 437-48.

218 Olajuyigbe, F. M. (2013) Preliminary Studies on the Expression and Purification of Functionally Active Recombinant Plasmepsin 9 from *Plasmodium falciparum*, *International Journal of Biochemistry Research and Review,* 3(2), pp. 107-118.

219 Olotu, A., Fegan, G., Wambua, J., Nyangweso, G., Awuondo, K. O., Leach, A., Lievens, M., Leboulleux, D., *et al.* (2013) Four-year efficacy of RTS,S/AS01E and its interaction with malaria exposure, *N Engl J Med,* 368(12), pp. 1111-20.

220 Ortiz, A. R., Kolinski, A., Rotkiewicz, P., Ilkowski, B. and Skolnick, J. (1999) *Ab initio* folding of proteins using restraints derived from evolutionary information, *Proteins,* Suppl 3, pp. 177-85.

221 Ostermann, N., Gerhartz, B., Worpenberg, S., Trappe, J. and Eder, J. (2004) Crystal structure of an activation intermediate of cathepsin E, *J Mol Biol*, 342(3), pp. 889-99.

222 Pachlatko, E., Rusch, S., Muller, A., Hemphill, A., Tilley, L., Hanssen, E. and Beck, H. P. (2010) MAHRP2, an exported protein of *Plasmodium falciparum*, is an essential component of Maurer's cleft tethers, *Mol Microbiol*, 77(5), pp. 1136-52.

223 Papadopoulos, J. S. and Agarwala, R. (2007) COBALT: constraint-based alignment tool for multiple protein sequences, *Bioinformatics*, 23(9), pp. 1073-9.

224 Parikh, S., Gut, J., Istvan, E., Goldberg, D. E., Havlir, D. V. and Rosenthal, P. J. (2005) Antimalarial activity of human immunodeficiency virus type 1 protease inhibitors, *Antimicrob Agents Chemother*, 49(7), pp. 2983-5.

225 Parikh, S., Liu, J., Sijwali, P., Gut, J., Goldberg, D. E. and Rosenthal, P. J. (2006) Antimalarial effects of human immunodeficiency virus type 1 protease inhibitors differ from those of the aspartic protease inhibitor pepstatin, *Antimicrob Agents Chemother*, 50(6), pp. 2207-9.

226 Peatey, C. L., Andrews, K. T., Eickel, N., MacDonald, T., Butterworth, A. S., Trenholme, K. R., Gardiner, D. L., McCarthy, J. S., *et al.* (2010) Antimalarial asexual stage-specific and gametocytocidal activities of HIV protease inhibitors, *Antimicrob Agents Chemother*, 54(3), pp. 1334-7.

227 Peatey, C. L., Skinner-Adams, T. S., Dixon, M. W., McCarthy, J. S., Gardiner, D. L. and Trenholme, K. R. (2009) Effect of antimalarial drugs on *Plasmodium falciparum* gametocytes, *J Infect Dis*, 200(10), pp. 1518-21.

228 Penny, M. A., Maire, N., Studer, A., Schapira, A. and Smith, T. A. (2008) What should vaccine developers ask? Simulation of the effectiveness of malaria vaccines, *PLoS One*, 3(9), pp. e3193.

229 Peters, P. J., Thigpen, M. C., Parise, M. E. and Newman, R. D. (2007) Safety and toxicity of sulfadoxine/pyrimethamine: implications for malaria prevention in pregnancy using intermittent preventive treatment, *Drug Saf*, 30(6), pp. 481-501.

230 Peti, W. and Page, R. (2007) Strategies to maximize heterologous protein expression in *Escherichia coli* with minimal cost, *Protein Expr Purif*, 51(1), pp. 1-10.

231 Phillips, J., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R., *et al.* (2005) Scalable molecular dynamics with NAMD, *Journal of Computational Chemistry*, 26(16), pp. 1781-1802.

232 PlasmoDB (2009) PlasmoDB: The Plasmodium Genome Resource. Available at: www.plasmodb.org

233 Porter, K. A., Cole, S. R., Eron, J. J., Jr., Zheng, Y., Hughes, M. D., Lockman, S., Poole, C., Skinner-Adams, T. S., *et al.* (2012) HIV-1 protease inhibitors and clinical malaria: a secondary analysis of the AIDS Clinical Trials Group A5208 study, *Antimicrob Agents Chemother*, 56(2), pp. 995-1000.

234 Prade, L., Jones, A. F., Boss, C., Richard-Bildstein, S., Meyer, S., Binkert, C. and Bur, D. (2005) X-ray structure of plasmepsin II complexed with a potent achiral inhibitor, *J Biol Chem*, 280(25), pp. 23837-43.

235 Price, R. N., Uhlemann, A. C., Brockman, A., McGready, R., Ashley, E., Phaipun, L., Patel, R., Laing, K., *et al.* (2004) Mefloquine resistance in *Plasmodium falciparum* and increased pfmdr1 gene copy number, *Lancet*, 364(9432), pp. 438-47.

236 Pukrittayakamee, S., Vanijanonta, S., Chantra, A., Clemens, R. and White, N. J. (1994) Blood stage antimalarial efficacy of primaquine in *Plasmodium vivax* malaria, *J Infect Dis*, 169(4), pp. 932-5.

237 Quashie, N. B., Duah, N. O., Abuaku, B., Quaye, L., Ayanful-Torgby, R., Akwoviah, G. A., Kweku, M., Johnson, J. D., *et al.* (2013) A SYBR Green 1-based *in vitro* test of susceptibility of Ghanaian *Plasmodium falciparum* clinical isolates to a panel of anti-malarial drugs, *Malar J*, 12, pp. 450.

238 Ragheb, D., Bompiani, K., Dalal, S. and Klemba, M. (2009) Evidence for catalytic roles for *Plasmodium falciparum* aminopeptidase P in the food vacuole and cytosol, *J Biol Chem*, 284(37), pp. 24806-15.

239 Ranson, H., N'Guessan, R., Lines, J., Moiroux, N., Nkuni, Z. and Corbel, V. (2011) Pyrethroid resistance in African anopheline mosquitoes: what are the implications for malaria control?, *Trends Parasitol*, 27(2), pp. 91-8.

240 Raynes, K. (1999) Bisquinoline antimalarials: their role in malaria chemotherapy, *Int J Parasitol*, 29(3), pp. 367-79.

241 Reddy, G. S., Ali, A., Nalam, M. N., Anjum, S. G., Cao, H., Nathans, R. S., Schiffer, C. A. and Rana, T. M. (2007) Design and synthesis of HIV-1 protease inhibitors incorporating oxazolidinones as P2/P2' ligands in pseudosymmetric dipeptide isosteres, *J Med Chem*, 50(18), pp. 4316-28.

242 Redmond, A. M., Skinner-Adams, T., Andrews, K. T., Gardiner, D. L., Ray, J., Kelly, M. and McCarthy, J. S. (2007) Antimalarial activity of sera from subjects taking HIV protease inhibitors, *AIDS,* 21(6), pp. 763-5.

243 Rijken, M. J., McGready, R., Phyo, A. P., Lindegardh, N., Tarning, J., Laochan, N., Than, H. H., Mu, O., *et al.* (2011) Pharmacokinetics of dihydroartemisinin and piperaquine in pregnant and nonpregnant women with uncomplicated falciparum malaria, *Antimicrob Agents Chemother,* 55(12), pp. 5500-6.

244 Ringwald, P., Eboumbou, E. C., Bickii, J. and Basco, L. K. (1999) *In vitro* activities of pyronaridine, alone and in combination with other antimalarial drugs, against *Plasmodium falciparum*, *Antimicrob Agents Chemother,* 43(6), pp. 1525-7.

245 Ripka, A. S., Satyshur, K. A., Bohacek, R. S. and Rich, D. H. (2001) Aspartic protease inhibitors designed from computer-generated templates bind as predicted, *Org Lett,* 3(15), pp. 2309-12.

246 Rogers, W. O., Sem, R., Tero, T., Chim, P., Lim, P., Muth, S., Socheat, D., Ariey, F., *et al.* (2009) Failure of artesunate-mefloquine combination therapy for uncomplicated *Plasmodium falciparum* malaria in southern Cambodia, *Malar J,* 8, pp. 10.

247 Rogerson, S. (2003) HIV-1, antiretroviral therapy, and malaria, *Lancet,* 362(9389), pp. 1008-9.

248 Roseler, A., Prieto, J. H., Iozef, R., Hecker, B., Schirmer, R. H., Kulzer, S., Przyborski, J., Rahlfs, S., *et al.* (2012) Insight into the selenoproteome of the malaria parasite *Plasmodium falciparum*, *Antioxid Redox Signal,* 17(4), pp. 534-43.

249 Rosenthal, P. J. (2002) Hydrolysis of erythrocyte proteins by proteases of malaria parasites, *Curr Opin Hematol,* 9(2), pp. 140-5.

250 Rosenthal, P. J. (2004) Cysteine proteases of malaria parasites, *Int J Parasitol,* 34(13-14), pp. 1489-99.

251 Russell, R. B. and Barton, G. J. (1992) Multiple protein sequence alignment from tertiary structure comparison: assignment of global and residue confidence levels, *Proteins,* 14(2), pp. 309-23.

252 Russo, I., Babbitt, S., Muralidharan, V., Butler, T., Oksman, A. and Goldberg, D. E. (2010) Plasmepsin V licenses Plasmodium proteins for export into the host erythrocyte, *Nature,* 463(7281), pp. 632-6.

253 Schlitzer, M. (2007) Malaria chemotherapeutics part I: History of antimalarial drug development, currently used therapeutics, and drugs in clinical development, *ChemMedChem*, 2(7), pp. 944-86.

254 Schrodinger 2009. Maestro 9.0: User Manual. Schrodinger Press.

255 Author (2011) *Schrodinger Suite 2011 Protein Preparation Wizard* (Version Epik 2.2, Impact 5.7, Prime 3.0).

256 Author (2013) *Schrodinger Release 2013_3: LigPrep* (Version 2.8).

257 Schwartz, E., Sadetzki, S., Murad, H. and Raveh, D. (2001) Age as a risk factor for severe *Plasmodium falciparum* malaria in nonimmune patients, *Clin Infect Dis*, 33(10), pp. 1774-7.

258 Schwede, T., Kopp, J., Guex, N. and Peitsch, M. C. (2003) SWISS-MODEL: An automated protein homology-modeling server, *Nucleic Acids Res*, 31(13), pp. 3381-5.

259 Seidens, T. (2010) *Characterising Plasmodium falciparum plasmepsin X*. Diploma of Biotechnology, Hochschule Weihenstephan-Triesdorf University of Applied Sciences.

260 Shea, M., Jakle, U., Liu, Q., Berry, C., Joiner, K. A. and Soldati-Favre, D. (2007) A family of aspartic proteases and a novel, dynamic and cell-cycle-dependent protease localization in the secretory pathway of *Toxoplasma gondii*, *Traffic*, 8(8), pp. 1018-34.

261 Shelley, J. C., Cholleti, A., Frye, L. L., Greenwood, J. R., Timlin, M. R. and Uchimaya, M. (2007) Epik: a software program for pK(a) prediction and protonation state generation for drug-like molecules, *J Comput Aided Mol Des*, 21(12), pp. 681-91.

262 Sherman, W., Day, T., Jacobson, M. P., Friesner, R. A. and Farid, R. (2006) Novel procedure for modeling ligand/receptor induced fit effects, *J Med Chem*, 49(2), pp. 534-53.

263 Shi, J., Blundell, T. L. and Mizuguchi, K. (2001) FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties, *J Mol Biol*, 310(1), pp. 243-57.

264 Shivakumar, D., Williams, J., Wu, Y., Damm, W., Shelley, J. and Sherman, W. (2010) Prediction of Absolute Solvation Free Energies using Molecular Dynamics Free Energy Perturbation and the OPLS Force Field, *J. Chem. Theory Comput.*, 6, pp. 1509-1519.

265 Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega, *Mol Syst Biol,* 7, pp. 539.

266 Sijwali, P. S., Kato, K., Seydel, K. B., Gut, J., Lehman, J., Klemba, M., Goldberg, D. E., Miller, L. H., *et al.* (2004) *Plasmodium falciparum* cysteine protease falcipain-1 is not essential in erythrocytic stage malaria parasites, *Proc Natl Acad Sci U S A,* 101(23), pp. 8721-6.

267 Sijwali, P. S., Koo, J., Singh, N. and Rosenthal, P. J. (2006) Gene disruptions demonstrate independent roles for the four falcipain cysteine proteases of *Plasmodium falciparum*, *Mol Biochem Parasitol,* 150(1), pp. 96-106.

268 Sijwali, P. S. and Rosenthal, P. J. (2004) Gene disruption confirms a critical role for the cysteine protease falcipain-2 in hemoglobin hydrolysis by *Plasmodium falciparum*, *Proc Natl Acad Sci U S A,* 101(13), pp. 4384-9.

269 Silamut, K., Phu, N. H., Whitty, C., Turner, G. D., Louwrier, K., Mai, N. T., Simpson, J. A., Hien, T. T., *et al.* (1999) A quantitative analysis of the microvascular sequestration of malaria parasites in the human brain, *Am J Pathol,* 155(2), pp. 395-410.

270 Singh, B., Kim Sung, L., Matusop, A., Radhakrishnan, A., Shamsul, S. S., Cox-Singh, J., Thomas, A. and Conway, D. J. (2004) A large focus of naturally acquired *Plasmodium knowlesi* infections in human beings, *Lancet,* 363(9414), pp. 1017-24.

271 Skinner-Adams, T. S., Butterworth, A. S., Porter, K. A., D'Amico, R., Sawe, F., Shaffer, D., Siika, A., Hosseinipour, M. C., *et al.* (2012) The frequency of malaria is similar among women receiving either lopinavir/ritonavir or nevirapine-based antiretroviral treatment, *PLoS One,* 7(4), pp. e34399.

272 Skinner-Adams, T. S., Lawrie, P. M., Hawthorne, P. L., Gardiner, D. L. and Trenholme, K. R. (2003) Comparison of *Plasmodium falciparum* transfection methods, *Malar J,* 2, pp. 19.

273 Skinner-Adams, T. S., McCarthy, J. S., Gardiner, D. L. and Andrews, K. T. (2008) HIV and malaria co-infection: interactions and consequences of chemotherapy, *Trends Parasitol,* 24(6), pp. 264-71.

274 Skinner-Adams, T. S., McCarthy, J. S., Gardiner, D. L., Hilton, P. M. and Andrews, K. T. (2004) Antiretrovirals as antimalarial agents, *J Infect Dis,* 190(11), pp. 1998-2000.

275 Skinner-Adams, T. S., Stack, C. M., Trenholme, K. R., Brown, C. L., Grembecka, J., Lowther, J., Mucha, A., Drag, M., *et al.* (2010) *Plasmodium falciparum* neutral aminopeptidases: new targets for anti-malarials, *Trends Biochem Sci,* 35(1), pp. 53-61.

276 Skinner, T. S., Manning, L. S., Johnston, W. A. and Davis, T. M. (1996) *In vitro* stage-specific sensitivity of *Plasmodium falciparum* to quinine and artemisinin drugs, *Int J Parasitol,* 26(5), pp. 519-25.

277 Skolnick, J., Kolinski, A., Kihara, D., Betancourt, M., Rotkiewicz, P. and Boniecki, M. (2001) *Ab initio* protein structure prediction via a combination of threading, lattice folding, clustering, and structure refinement, *Proteins,* Suppl 5, pp. 149-56.

278 Slater, A. F., Swiggard, W. J., Orton, B. R., Flitter, W. D., Goldberg, D. E., Cerami, A. and Henderson, G. B. (1991) An iron-carboxylate bond links the heme units of malaria pigment, *Proc Natl Acad Sci U S A,* 88(2), pp. 325-9.

279 Soldati-Favre, D. 19/11/2013 2013. *RE: Conditional PM IX knock-outs.* Type to McGeorge, R.

280 Soldati, D., Kim, K., Kampmeier, J., Dubremetz, J. F. and Boothroyd, J. C. (1995) Complementation of a *Toxoplasma gondii* ROP1 knock-out mutant using phleomycin selection, *Mol Biochem Parasitol,* 74(1), pp. 87-97.

281 Song, J., Socheat, D., Tan, B., Seila, S., Xu, Y., Ou, F., Sokunthea, S., Sophorn, L., *et al.* (2011) Randomized trials of artemisinin-piperaquine, dihydroartemisinin-piperaquine phosphate and artemether-lumefantrine for the treatment of multi-drug resistant falciparum malaria in Cambodia-Thailand border area, *Malar J,* 10, pp. 231.

282 Sowunmi, A., Fateye, B. A., Adedeji, A. A., Fehintola, F. A., Bamgboye, A. E., Babalola, C. P., Happi, T. C. and Gbotosho, G. O. (2005) Effects of antifolates--co-trimoxazole and pyrimethamine-sulfadoxine--on gametocytes in children with acute, symptomatic, uncomplicated, *Plasmodium falciparum* malaria, *Mem Inst Oswaldo Cruz,* 100(4), pp. 451-5.

283 Spaccapelo, R., Janse, C. J., Caterbi, S., Franke-Fayard, B., Bonilla, J. A., Syphard, L. M., Di Cristina, M., Dottorini, T., *et al.* (2010) Plasmepsin 4-deficient *Plasmodium berghei* are virulence attenuated and induce protective immunity against experimental malaria, *Am J Pathol,* 176(1), pp. 205-17.

284 Spangenberg, T., Burrows, J. N., Kowalczyk, P., McDonald, S., Wells, T. N. and Willis, P. (2013) The open access malaria box: a drug discovery catalyst for neglected diseases, *PLoS One,* 8(6), pp. e62906.

285 Stack, C. M., Lowther, J., Cunningham, E., Donnelly, S., Gardiner, D. L., Trenholme, K. R., Skinner-Adams, T. S., Teuscher, F., *et al.* (2007) Characterization of the *Plasmodium falciparum* M17 leucyl aminopeptidase. A protease involved in amino acid regulation with potential for antimalarial drug development, *J Biol Chem,* 282(3), pp. 2069-80.

286 Sutherland, C. J., Tanomsing, N., Nolder, D., Oguike, M., Jennison, C., Pukrittayakamee, S., Dolecek, C., Hien, T. T., *et al.* (2010) Two nonrecombining sympatric forms of the human malaria parasite *Plasmodium ovale* occur globally, *J Infect Dis,* 201(10), pp. 1544-50.

287 Svenson, J. E., MacLean, J. D., Gyorkos, T. W. and Keystone, J. (1995) Imported malaria. Clinical presentation and examination of symptomatic travelers, *Arch Intern Med,* 155(8), pp. 861-8.

288 Sweetman, S. E. (2005) Martindale: The Complete Drug Reference, in Sweetman, S. (ed.). 34th ed. London: Pharmaceutical Press; Royal Pharmaceutical Society of Great Britain, pp. 453-456.

289 Talman, A. M., Blagborough, A. M. and Sinden, R. E. (2010) A *Plasmodium falciparum* strain expressing GFP throughout the parasite's life-cycle, *PLoS One,* 5(2), pp. e9156.

290 Targett, G. A. and Greenwood, B. M. (2008) Malaria vaccines and their potential role in the elimination of malaria, *Malar J,* 7 Suppl 1, pp. S10.

291 Taylor, T. E., Borgstein, A. and Molyneux, M. E. (1993) Acid-base status in paediatric *Plasmodium falciparum* malaria, *Q J Med,* 86(2), pp. 99-109.

292 Taylor, W. R. and White, N. J. (2004) Antimalarial drug toxicity: a review, *Drug Saf,* 27(1), pp. 25-61.

293 ter Kuile, F. O., Parise, M. E., Verhoeff, F. H., Udhayakumar, V., Newman, R. D., van Eijk, A. M., Rogerson, S. J. and Steketee, R. W. (2004) The burden of co-infection with human immunodeficiency virus type 1 and malaria in pregnant women in sub-saharan Africa, *Am J Trop Med Hyg,* 71(2 Suppl), pp. 41-54.

294 Teuscher, F., Lowther, J., Skinner-Adams, T. S., Spielmann, T., Dixon, M. W., Stack, C. M., Donnelly, S., Mucha, A., *et al.* (2007) The M18 aspartyl aminopeptidase of the human malaria parasite *Plasmodium falciparum, J Biol Chem,* 282(42), pp. 30817-26.

295 Thanaraj, T. A. and Argos, P. (1996) Ribosome-mediated translational pause and protein domain organization, *Protein Sci*, 5(8), pp. 1594-612.

296 Tonkin, C. J., van Dooren, G. G., Spurck, T. P., Struck, N. S., Good, R. T., Handman, E., Cowman, A. F. and McFadden, G. I. (2004) Localization of organellar proteins in *Plasmodium falciparum* using a novel set of transfection vectors and a new immunofluorescence fixation method, *Mol Biochem Parasitol*, 137(1), pp. 13-21.

297 Trager, W. and Jensen, J. B. (1976) Human malaria parasites in continuous culture, *Science*, 193(4254), pp. 673-5.

298 Trager, W., Rudzinska, M. A. and Bradbury, P. C. (1966) The fine structure of *Plasmodium falciparum* and its host erythrocytes in natural malarial infections in man, *Bull World Health Organ*, 35(6), pp. 883-5.

299 Trampuz, A., Jereb, M., Muzlovic, I. and Prabhu, R. M. (2003) Clinical review: Severe malaria, *Crit Care*, 7(4), pp. 315-23.

300 Triglia, T., Healer, J., Caruana, S. R., Hodder, A. N., Anders, R. F., Crabb, B. S. and Cowman, A. F. (2000) Apical membrane antigen 1 plays a central role in erythrocyte invasion by Plasmodium species, *Mol Microbiol*, 38(4), pp. 706-18.

301 Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R. and Leunissen, J. A. (2007) Primer3Plus, an enhanced web interface to Primer3, *Nucleic Acids Res*, 35(Web Server issue), pp. W71-4.

302 Vedadi, M., Lew, J., Artz, J., Amani, M., Zhao, Y., Dong, A., Wasney, G. A., Gao, M., *et al.* (2007) Genome-scale protein expression and structural biology of *Plasmodium falciparum* and related Apicomplexan organisms, *Mol Biochem Parasitol*, 151(1), pp. 100-10.

303 Verkhivker, G., Tiana, G., Camilloni, C., Provasi, D. and Broglia, R. A. (2008) Atomistic simulations of the HIV-1 protease folding inhibition, *Biophys J*, 95(2), pp. 550-62.

304 Vivas, L., Rattray, L., Stewart, L. B., Robinson, B. L., Fugmann, B., Haynes, R. K., Peters, W. and Croft, S. L. (2007) Antimalarial efficacy and drug interactions of the novel semi-synthetic endoperoxide artemisone *in vitro* and *in vivo*, *J Antimicrob Chemother*, 59(4), pp. 658-65.

305 Warren, G. L., Do, T. D., Kelley, B. P., Nicholls, A. and Warren, S. D. (2012) Essential considerations for using protein-ligand structures in drug discovery, *Drug Discov Today*, 17(23-24), pp. 1270-81.

306 Waugh, A. and Grant, A. (2007) *Anatomy and Physiology in Health and Illness* 10 edn. Churchhill Livingstone: Elsevier.

307 White, N. J. (2008) *Plasmodium knowlesi*: the fifth human malaria parasite, *Clin Infect Dis*, 46(2), pp. 172-3.

308 Whitworth, J., Morgan, D., Quigley, M., Smith, A., Mayanja, B., Eotu, H., Omoding, N., Okongo, M., *et al.* (2000) Effect of HIV-1 and increasing immunosuppression on malaria parasitaemia and clinical episodes in adults in rural Uganda: a cohort study, *Lancet*, 356(9235), pp. 1051-6.

309 WHO (1990) Severe and complicated malaria. World Health Organization, Division of Control of Tropical Diseases, *Trans R Soc Trop Med Hyg*, 84 Suppl 2, pp. 1-65.

310 WHO (2000) Severe falciparum malaria. World Health Organization, Communicable Diseases Cluster, *Trans R Soc Trop Med Hyg*, 94 Suppl 1, pp. S1-90.

311 WHO (2008) *World Malaria Report*, Geneva: World Health Organisation.

312 WHO 2012. Malaria Fact Sheet No. 94. Geneva: World Health Organisation.

313 WHO (2013) *World Malaria Report*, Geneva, Switzerland: World Health Organisation. Available at: http://who.int/malaria/publications/world_malaria_report_2013/en/.

314 WHO Global Malaria Programme, M. f. M. V. 2014. Global Malaria Mapper. *In:* Medicines for Malaria Venture, W.G.M.P. (ed.).

315 Wootton J., F. S. (1993) Statistics of local complexity in amino acid sequences and sequence databases., *Comput Chem*, 17, pp. 149-163.

316 Wootton, J. C. (1994) Non-globular domains in protein sequences: automated segmentation using complexity measures, *Comput Chem*, 18(3), pp. 269-85.

317 Wyatt, D. M. and Berry, C. (2002) Activity and inhibition of plasmepsin IV, a new aspartic proteinase from the malaria parasite, *Plasmodium falciparum*, *FEBS Lett*, 513(2-3), pp. 159-62.

318 Xiao, H., Briere, L. A., Dunn, S. D. and Yada, R. Y. (2010) Characterization of the monomer-dimer equilibrium of recombinant histo-aspartic protease from *Plasmodium falciparum*, *Mol Biochem Parasitol*, 173(1), pp. 17-24.

319 Xiao, H., Tanaka, T., Ogawa, M. and Yada, R. Y. (2007) Expression and enzymatic characterization of the soluble recombinant plasmepsin I from *Plasmodium falciparum*, *Protein Eng Des Sel*, 20(12), pp. 625-33.

320 Yasuda, Y., Kageyama, T., Akamine, A., Shibata, M., Kominami, E., Uchiyama, Y. and Yamamoto, K. (1999) Characterization of new fluorogenic substrates for the rapid and sensitive assay of cathepsin E and cathepsin D, *J Biochem*, 125(6), pp. 1137-43.

321 Zhou, Z., Schnake, P., Xiao, L. and Lal, A. A. (2004) Enhanced expression of a recombinant malaria candidate vaccine in *Escherichia coli* by codon optimization, *Protein Expr Purif*, 34(1), pp. 87-94.

322 Zilversmit, M. M., Volkman, S. K., DePristo, M. A., Wirth, D. F., Awadalla, P. and Hartl, D. L. (2010) Low-complexity regions in *Plasmodium falciparum*: missing links in the evolution of an extreme genome, *Mol Biol Evol*, 27(9), pp. 2198-209.