

https://dspace.org/

Identification of Plasmodium Falciparum proteins interacting with the erythrocyte membrane skeleton protein spectrin

Lauterbach, Sonja Brigitte

2009-04-08T12:04:43Z
http://hdl.handle.net/10539/6878

Downloaded from a DSpace repository. https://dspace.org/

IDENTIFICATION OF *PLASMODIUM FALCIPARUM* PROTEINS INTERACTING WITH THE ERYTHROCYTE MEMBRANE SKELETON PROTEIN SPECTRIN

Sonja Brigitte Lauterbach

A thesis submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Doctor of Philosophy.

Johannesburg, October 2008.

DECLARATION

I declare this thesis to be my own, unaided work. It is being submitted to the University of the Witwatersrand, Johannesburg, for the degree of Doctor of Philosophy. It has not been submitted for any other degree or examination at any other University.



(Sonja Lauterbach)

3rd day of October 2008

RESEARCH OUTPUT

Publications:

Lauterbach, S.B. and Coetzer, T.L. (2008) The M18 aspartyl aminopeptidase of *Plasmodium falciparum* binds to human erythrocyte spectrin *in vitro*. *Malar J*, **7**, 161.

Lauterbach, S.B., Lanzillotti, R. and Coetzer, T.L. (2003) Construction and use of *Plasmodium falciparum* phage display libraries to identify host parasite interactions. *Malar J*, **2**, 47.

Conferences:

Analysis of PFI1570c, a *Plasmodium falciparum* aminopeptidase that interacts with human erythrocyte spectrin.
Coetzer, T.L. and Lauterbach, S.B.
Poster presented at:

- the Molecular Approaches to Malaria Conference, Lorne, Australia, February 2008

Characterisation of PFI1570c, a *Plasmodium falciparum* aminopeptidase.
Lauterbach, S.B. and Coetzer, T.L.
Poster presented at:

- the SASBMB XXth Conference, Pietermaritzburg, South Africa, July 2006 (best poster award)
- the Health Sciences Research Day, University of the Witwatersrand, Johannesburg, South Africa, August 2006
- the Molecular Cell Biology Group Symposium, Johannesburg, South Africa, October 2006

Plasmodium falciparum proteins that interact with human erythrocyte spectrin.
Lauterbach, S.B. and Coetzer, T.L.
Poster presented at:

- the SASBMB XIXth Conference, Stellenbosch, South Africa, January 2005

A *Plasmodium falciparum* aminopeptidase interacts with human erythrocyte spectrin.
Coetzer, T.L. and Lauterbach, S.B.
Poster presented at:

- the American Society of Hematology 46th Annual Meeting and Exposition, San Diego, USA, December 2004 (published in *Blood*, **104**, 441a)

Identification of *Plasmodium falciparum* proteins that bind to human erythrocyte spectrin.
Lauterbach, S.B. and Coetzer, T.L.
Poster presented at:

- the Health Sciences Research Day, University of the Witwatersrand, Johannesburg, South Africa, August 2004
- the 15th Molecular Parasitology Meeting, Woods Hole, USA, September 2004
- the Molecular Cell Biology Group Symposium, Johannesburg, South Africa, October 2004

Construction of a *Plasmodium falciparum* phage-display library.
Lauterbach, S.B. and Coetzer, T.L.
Poster presented at:

- the IUSMB/SASMB Special Meeting on the Biochemical and Molecular Basis of Disease, Cape Town, South Africa, November 2001
- the Health Sciences Research Day, University of the Witwatersrand, Johannesburg, South Africa, August 2002

ABSTRACT

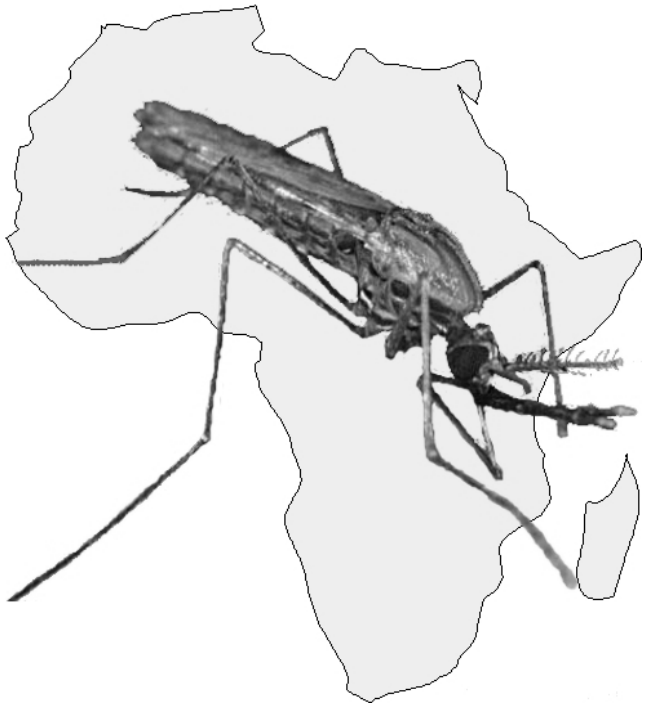
Malaria, which is caused by *Plasmodium* parasites, is responsible for the death of millions of humans every year in the tropical and subtropical regions of the world. Specifically *P. falciparum,* one of four malaria species infecting humans, is responsible for the greatest morbidity and mortality burden in African populations. The *Anopheles* mosquito transmits the parasite to the human host, where it infects and destroys human erythrocytes. The erythrocyte membrane therefore plays a vital role in all aspects of the pathogenic phase of the *P. falciparum* life cycle and protein-protein interactions between host and parasite are thus a key focus of research. The human erythrocyte maintains its shape with a structural network of proteins underneath the plasma membrane and the main protein component of this erythrocyte membrane skeleton is spectrin. To investigate host-parasite protein interactions, a novel application of phage display technology was developed, whereby purified human erythrocyte spectrin was biopanned against a *P. falciparum* phage-display library. The *P. falciparum* DNA inserts of interacting phage were compared to the PlasmoDB database and five interacting proteins were identified: a putative aminopeptidase (*Pf*M18AAP); a putative Ebl-1 like protein, which is proposed to participate in erythrocyte invasion; and three hypothetical proteins. The interaction of the hypothetical proteins with spectrin is the first information available on the function of these proteins. The five gene sequences were cloned into the pET-15b or pGEX-4T-2 expression vectors for purification of the recombinant proteins from *Escherichia coli*. Only the 6His-*Pf*M18AAP fusion protein was expressed in soluble form and purified by affinity selection. *Pf*M18AAP migrated as a 67 kDa peptide on SDS-PAGE and native gel analysis revealed multiple subunits of the enzyme, predominantly a tetramer and higher oligomers. Cleavage of the 6His-tag and subsequent IEF SDS-PAGE revealed three 65 kDa entities with pI ~6.6, ~6.7 and ~6.9. An *in vitro* coupled enzyme assay showed that *Pf*M18AAP cleaved an N-terminal aspartate from a peptide substrate with a maximum activity at pH 7.5 and 37 ºC. Inhibitor studies confirmed that the enzyme is a metalloprotease. Blot overlay assays with *Pf*M18AAP against spectrin and erythrocyte membrane proteins verified that

*Pf*M18AAP binds strongly to β-spectrin, as well as protein 4.1, protein 4.2, actin and glyceraldehyde-3-phosphate dehydrogenase. Comparison of the *Pf*M18AAP protein sequence to ten other M18 aminopeptidase sequences, including human and three other *Plasmodium* species, revealed that all the critical amino acids responsible for the binding of two catalytic metal ions, enzymatic catalysis and quaternary structure stabilisation are conserved. The peptide fragment, which initially bound to spectrin during phage display, is not found in other M18 aminopeptidases, suggesting that the presence of this fragment is an evolutionary development of *P. falciparum* that allows the protease to bind to human spectrin. Analysis of four M18 aminopeptidase crystal structures revealed that the spectrin-binding region forms an external loop on the protein and would thus be accessible to spectrin. Results from this study suggest that, apart from haemoglobin digestion, *Pf*M18AAP performs additional functions in the parasite and infected erythrocyte by cleaving spectrin and other erythrocyte membrane proteins. This would destabilise and disrupt the erythrocyte membrane skeleton to facilitate entry or exit from the host cell, or the insertion of parasite proteins into the host cell membrane. Further analysis and characterisation of *Pf*M18AAP and its interactions with the erythrocyte membrane proteins will shed more light on the multifunctional role of this parasite enzyme. Studies of this enzyme and the hypothetical proteins may also aid in the quest to discover new therapeutics to combat this killer disease.

This work is dedicated to four precious ladies in my family:

Christine Lauterbach (mom; 1943-2007),

Sheba (Fatty; 1989-1997), Nikita (Bigsy), and Killashandra (Shandy).



The men of experiment are like the ant, they only collect and use; the reasoners resemble spiders, who make cobwebs out of their own substance. But the bee takes the middle course: it gathers its materials from the flowers of the garden and field, but transforms and digests it by a power of its own. Not unlike this is the true business of philosophy (science); for it neither relies solely or chiefly on the powers of the mind, nor does it take the matter which it gathers from natural history and mechanical experiments and lay upon its memory of whole, as it finds it, but lays it up in the understanding altered and digested. Therefore, from a closer and purer league between these two faculties, the experimental and the rational (such as has ever been made), much may be hoped.

Francis Bacon (1561-1626)

ACKNOWLEDGEMENTS

I would like to acknowledge:

- Prof. Thérèsa Coetzer, who helped and encouraged me throughout my studies.
- The members of the *Plasmodium* Molecular Research Unit and the Department of Molecular Medicine and Haematology, who helped and guided me with my laboratory work.
- The National Research Foundation (South Africa), the National Health Laboratory Service (South Africa), the Stella and Paul Loewenstein Charitable and Educational Trust and the University of the Witwatersrand, who provided financial support.
- Cleo Kiley, Diane Kuhnert, Lara Noble, Chantal Drummond, Debbie Domeisen-Naude, Elizabeth Burger, Michelle Bowden and the members of the Liza-Michelle School of Dancing, who provided the emotional support required during the completion of my thesis.
- Diane Kuhnert, for her excellent tutorship in the use of DeepView Swiss-PdbViewer and Microsoft Word.

ETHICS CLEARANCE

Ethics clearance was granted by the University of the Witwatersrand Committee for research on human subjects (medical) in Protocol number M03-11-06.

TABLE OF CONTENTS

DECLARATION I
RESEARCH OUTPUT II
ABSTRACT IV
ACKNOWLEDGEMENTS VII
ETHICS CLEARANCE VII
TABLE OF CONTENTS VIII
LIST OF FIGURES XI
LIST OF TABLES XIV
ABBREVIATIONS XV

CHAPTER 1: INTRODUCTION 1

1.1 Malaria 1

1.1.1 The *P. falciparum* life cycle 2

1.2 The human erythrocyte 4

1.2.1 The erythrocyte membrane 4
1.2.2 Spectrin 6
1.2.3 The erythrocyte membrane skeleton 9
1.2.4 Integral membrane proteins 10
1.2.5 Membrane skeleton linkage to the lipid bilayer 11

1.3 The erythrocytic phase of the *P. falciparum* life cycle 12

1.3.1 Erythrocyte invasion 12
1.3.2 Erythrocyte remodelling 17
1.3.3 Haemoglobin digestion 21
1.3.4 Parasite escape 22

1.4 Parasite and erythrocyte membrane interactions 23

1.4.1 Protein interactions involved in erythrocyte invasion 25
1.4.2 Protein interactions facilitating parasite growth 26
1.4.3 Protein interactions mediating parasite escape 28

1.5 Erythrocyte protein abnormalities that protect against malaria infections 29

1.6 *P. falciparum* proteins 31

1.6.1 The *P. falciparum* genome and proteome 31
1.6.2 *P. falciparum* protein identification 32

1.7 Objectives 35

CHAPTER 2: PHAGE DISPLAY 36

2.1 Introduction 36

2.1.1 Identifying protein-protein interactions 36
2.1.2 Phage display technology 37
2.1.3 Phage display technology in malaria research 40
2.1.4 The T7Select® Phage Display System 42

2.2 Methods 45
2.2.1 Materials 45
2.2.2 Spectrin isolation and biotinylation 45
2.2.3 *P. falciparum* culturing 49
2.2.4 *P. falciparum* mRNA isolation 50
2.2.5 *P. falciparum* phage display library creation 52
2.2.6 *P. falciparum* phage display library biopanning 57
2.2.7 DNA sequencing of *P. falciparum* inserts 58
2.2.8 Identification and analysis of *P. falciparum* inserts 59
2.3 Results 59
2.3.1 Spectrin preparation 59
2.3.2 *P. falciparum* phage display library construction 63
2.3.3 Biopanning 68
2.3.4 Gene identification and peptide analysis 70
2.4 Discussion 74
2.4.1 Library creation and bacteriophage selection 75
2.4.2 Spectrin-binding bacteriophage 77
2.4.3 Identifying protein-protein interactions with phage display 79
CHAPTER 3: RECOMBINANT SPECTRIN-BINDING *P. FALCIPARUM* PROTEINS 81
3.1 Introduction 81
3.1.1 Expression of *P. falciparum* recombinant proteins 81
3.1.2 Chemical processing of inclusion bodies 86
3.2 Methods 87
3.2.1 The spectrin-binding *P. falciparum* proteins 87
3.2.2 Cloning 87
3.2.3 Protein expression 95
3.2.4 Recombinant *P. falciparum* aminopeptidase purification 96
3.3 Results 97
3.3.1 Cloning and expression of spectrin-binding *P. falciparum* proteins 97
3.3.2 Purification of the *P. falciparum* aminopeptidase 103
3.3.3 *P. falciparum* spectrin-binding protein features 106
3.4 Discussion 111
3.4.1 Cloning and *P. falciparum* protein expression 111
3.4.2 Characteristics of the *P. falciparum* spectrin-binding proteins 113
3.4.3 Possible functions of the spectrin-binding proteins in the erythrocytic phase of the *P. falciparum* life cycle 114

CHAPTER 4: CHARACTERISATION OF *P. FALCIPARUM* M18 ASPARTYL AMINOPEPTIDASE 116

4.1 Introduction 116

4.1.1 Enzymes and proteases 116
4.1.2 Metalloproteases 116
4.1.3 *P. falciparum* metalloproteases 118

4.2 Methods 119

4.2.1 Aminopeptidase sequence and structure analysis 119
4.2.2 Characterisation of recombinant *Pf*M18AAP 119
4.2.3 Protein overlays 125
4.2.4 Enzyme activity assay 127

4.3 Results 132

4.3.1 *Pf*M18AAP 132
4.3.2 M18 aminopeptidase family primary sequence features 132
4.3.3 *Pf*M18AAP secondary structure 138
4.3.4 Molecular weight and isoelectric point of the r*Pf*M18AAP monomer 141
4.3.5 The oligomeric state of native r*Pf*M18AAP 144
4.3.6 *Pf*M18AAP tertiary structure 148
4.3.7 r*Pf*M18AAP enzyme activity 154
4.3.8 Binding studies 159

4.4 Discussion 162

4.4.1 The *Pf*M18AAP active site 162
4.4.2 The *Pf*M18AAP spectrin-binding site 164
4.4.3 r*Pf*M18AAP protein and enzyme characteristics 165
4.4.4 r*Pf*M18AAP-erythrocyte membrane protein interactions 168
4.4.5 The function of *Pf*M18AAP in the infected erythrocyte 169
4.4.6 *Pf*M18AAP as a novel drug target 172

CHAPTER 5: CONCLUSION 173

APPENDICES 174
A1: Methods 174
A2: Vector maps 175
A3: Codon usage tables 178
A4: Automated sequencing results of PFI1570c 179

REFERENCES 184

LIST OF FIGURES

Figure 1: The global malaria distribution. 1
Figure 2: The *P. falciparum* life cycle. 3
Figure 3: The erythrocyte membrane. 5
Figure 4: Triple helical repeats. 7
Figure 5: The amino acid distribution of a spectrin repeat. 7
Figure 6: The spectrin tetramer. 8
Figure 7: The two-dimensional network of the red cell membrane skeleton. 10
Figure 8: Erythrocyte invasion by *P. falciparum*. 13
Figure 9: The *P. falciparum* merozoite. 14
Figure 10: The actin-myosin motor of *P. falciparum*. 16
Figure 11: *P. falciparum* protein transport to the erythrocyte cytosol. 19
Figure 12: *P. falciparum* protein trafficking to the erythrocyte surface. 21
Figure 13: Haemoglobin digestion. 22
Figure 14: *P. falciparum* escape from the erythrocyte. 23
Figure 15: *P. falciparum* and erythrocyte membrane protein interactions. 27
Figure 16: Biopanning of a phage display library. 39
Figure 17: The T7 bacteriophage. 43
Figure 18: The T7Select10-3 vector. 44
Figure 19: The biotinylation reaction. 47
Figure 20: Double digestion of the linker and *P. falciparum* insert. 52
Figure 21: Anchored 14-mer primers for *P. falciparum* cDNA synthesis. 53
Figure 22: Spectrin isolation from human erythrocytes. 60
Figure 23: Size exclusion chromatography of spectrin. 61
Figure 24: Biotinylated spectrin. 63
Figure 25: *P. falciparum* cultured in human erythrocytes. 64
Figure 26: *P. falciparum* total RNA. 65
Figure 27: Second strand synthesis from *P. falciparum* cDNA. 66
Figure 28: Gel filtration of *P. falciparum* DNA. 66
Figure 29: DNA inserts from the *P. falciparum* phage-display library. 67
Figure 30: *P. falciparum* DNA insert sequences. 70
Figure 31: Peptides isolated after fourth round biopanning. 73

Figure 32: Strategies to obtain soluble recombinant *P. falciparum* proteins from *E. coli*. 82
Figure 33: *P. falciparum* primers. 89
Figure 34: Vectors and *P. falciparum* amplicons before ligation. 98
Figure 35: pET-15b vector containing *P. falciparum* inserts. 99
Figure 36: pGEX-4T-2 vector containing *P. falciparum* inserts. 100
Figure 37: Expression of spectrin-binding *P. falciparum* proteins. 102
Figure 38: Induction and solubility study of recombinant *P. falciparum* aminopeptidase. 104
Figure 39: Sonication study of recombinant *P. falciparum* aminopeptidase. 105
Figure 40: Affinity purification of recombinant *P. falciparum* aminopeptidase. 106
Figure 41: Absolute mRNA expression levels for the spectrin-binding *P. falciparum* proteins. 110
Figure 42: Zinc binding in the active site of a metalloprotease. 117
Figure 43: Substrate cleavage by r*Pf*M18AAP, diazotisation and azo dye coupling reaction. 128
Figure 44: MEROPS database sequence alignment of M18 aminopeptidases. .. 136
Figure 45: Sequence alignment of the *Plasmodium Pf*M18AAP homologues. .. 138
Figure 46: PlasmoDB hydropathy plot, secondary structure and low complexity regions for PFI1570c (*Pf*M18AAP). 139
Figure 47: Jpred secondary structure prediction of *Pf*M18AAP and its aminopeptidase homologues. 141
Figure 48: Thrombin cleavage of r*Pf*M18AAP. 142
Figure 49: Two-dimensional gel electrophoresis of r*Pf*M18AAP. 144
Figure 50: Non-denaturing gel electrophoresis of r*Pf*M18AAP. 145
Figure 51: Ferguson plot of the r*Pf*M18AAP protein bands. 146
Figure 52: Non-denaturing gel electrophoresis standard curve. 146
Figure 53: Detection of r*Pf*M18AAP in non-denaturing gels. 148
Figure 54: Crystal structures of the monomers of the M18 aminopeptidases. 149
Figure 55: The *T. maritima* M18 aminopeptidase monomer and dimer. 152
Figure 56: The M18 aminopeptidase dodecamer. 153
Figure 57: Colour reaction of the r*Pf*M18AAP enzyme activity assay. 154
Figure 58: pH and temperature optima of r*Pf*M18AAP. 155
Figure 59: Variation in r*Pf*M18AAP enzymatic activity. 157
Figure 60: Michaelis-Menten and Lineweaver-Burk plots for r*Pf*M18AAP. 158
Figure 61: Spectrin dimers and tetramers overlaid with r*Pf*M18AAP. 159

Figure 62: Erythrocyte membrane proteins overlaid with r*Pf*M18AAP. 160
Figure 63: Spectrin tryptic digest overlaid with r*Pf*M18AAP. 161
Figure 64: The putative metal binding site of *Pf*M18AAP. 162
Figure 65: The putative substrate cleavage mechanism of *Pf*M18AAP. 164
Figure 66: The first 10 amino acids of the r*Pf*M18AAP-binding erythrocyte membrane proteins. 169

LIST OF TABLES

Table 1: *P. falciparum* proteins interacting with the erythrocyte membrane. 24
Table 2: Cycling parameters for T7 insert amplification. 56
Table 3: Spectrin dimer yields during spectrin processing. 62
Table 4: *P. falciparum* mRNA isolation. 64
Table 5: *P. falciparum* phage display library characteristics. 67
Table 6: Bacteriophage enrichment. 68
Table 7: *P. falciparum* genes identified after fourth round biopanning. 71
Table 8: Motifs and profiles of the *P. falciparum* spectrin-binding peptides. 74
Table 9: DNA fragments used for cloning. 87
Table 10: Annealing temperatures for *P. falciparum* gene specific primer sets. .. 90
Table 11: Cycling parameters for *P. falciparum* genomic DNA amplification. ... 90
Table 12: PlasmoDB data for the spectrin-binding *P. falciparum* proteins. 107
Table 13: Yeast two-hybrid data for the spectrin-binding *P. falciparum* proteins. 108
Table 14: Protein and mRNA expression data for the spectrin-binding *P. falciparum* proteins. 109
Table 15: Polyacrylamide percentages for the native gels. 124
Table 16: Sequence comparison of M18 aminopeptidases to *Pf*M18AAP. 133
Table 17: Approximate molecular weights of the r*Pf*M18AAP subunits. 147
Table 18: r*Pf*M18AAP relative activity in the presence of metal ions. 155
Table 19: r*Pf*M18AAP relative activity in the presence of protease inhibitors. . 156

ABBREVIATIONS

Å	angstrom
AMA-1	apical membrane antigen 1
atm	atmospheres
ATP	adenosine triphosphate
A_{260}	absorbance at 260nm
A_{280}	absorbance at 280nm
Big-1	bacterial immunoglobulin-like domain 1
bp	base pairs
BLAST	Basic Local Alignment Search Tool
BSA	bovine serum albumin
cAMP	cyclic adenosine monophosphate
cDNA	complementary deoxynucleic acid
cGMP	cyclic guanosine monophosphate
Ci	Curie
CSP	circumsporozoite protein
CTRP	TRAP-related protein
Da	dalton
DAP	human aspartyl aminopeptidase
dATP	deoxyadenosine triphosphate
dCTP	deoxycytosine triphosphate
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxyribonucleoside triphosphate
DPAP1	dipeptidyl peptidase I
DTT	1,4-dithiothreitol
EBA	erythrocyte binding antigens
E.C. number	Enzyme Commission number
EDTA	ethylenediaminetetraacetic acid
EST	expressed sequence tag
ER	endoplasmic reticulum
FEST	*falciparum*-exported serine/threonine kinase
FIRA	*falciparum* interspersed repeat antigen
GAP	glideosome-associated protein
GST	glutathione-S-transferase
G3PD	glyceraldehyde-3-phosphate dehydrogenase
G6PD	glucose-6-phosphate dehydrogenase

HAP	histo-aspartic protease
HE	hereditary elliptocytosis
HEPES	4-(2-hydroxyethyl)-piperazine-1-ethanesulphonic acid
HMW	high molecular weight
HPP	hereditary pyropoikilocytosis
HS	hereditary spherocytosis
IEF	isoelectric focusing
IMC	inner membrane complex
IPTG	isopropylthiogalactoside
IUBMB	International Union of Biochemistry & Molecular Biology
KAHRP	knob-associated histidine-rich protein
kb	kilo bases, 1000 base pairs
kDa	kilo dalton
K_m	Michaelis constant
K_R	retardation coefficient
LB	Luria broth
LMW	low molecular weight
LSA-1	*P. falciparum* liver stage antigen-1
mA	milli ampere
Mb	mega bases, 1000,000 base pairs
MBP	maltose binding protein
MESA	mature-parasite infected surface antigen
MMLV	Moloney Murine Leukemia Virus
MOI	multiplicity of infection
mRNA	messenger ribonucleic acid
MSP	merozoite surface proteins
MTIP	myosin tail interacting protein
MTRAP	merozoite TRAP homologue
MyoA	*Plasmodium* myosin
Na_2EDTA	disodium ethylenediaminetetraacetic acid
NHMec	4-methyl-7-coumarinylamide
NPL	native peptide library
NPP	New Permeability Pathways
NusA	N-utilising substance A
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PDB	Protein Data Bank
PEXEL motif	*Plasmodium* export element
PfA-M1	*P. falciparum* M1 aminopeptidase
PfEMP	*P. falciparum* erythrocyte membrane protein
*Pf*M17LAP	*P. falciparum* M17 leucyl aminopeptidase

*Pf*M18AAP	*P. falciparum* M18 aspartyl aminopeptidase
*Pfsbp*1	*P. falciparum* skeleton binding protein 1
PfSUB	*P. falciparum* subtilisin-like protease
PfROM	*P. falciparum* rhomboids
pfu	plaque forming units
PIPES	1,4-piperazinediethanesulphonic acid
PMSF	phenylmethanesulphonyl fluoride
RAMA	rhoptry associated membrane antigen
RAP	rhoptry associated protein
RESA	ring-infected erythrocyte surface antigen
R_m	relative mobility
RH	reticulocyte-binding homologues
RhopH	high molecular weight rhoptry protein
RIFINs	repetitive interspersed family
RIMA	ring membrane antigen
RNA	ribonucleic acid
RNase	ribonuclease
r*Pf*M18AAP	recombinant *P. falciparum* M18 aspartyl aminopeptidase
RPL	random peptide library
RPMI-1640 medium	Roswell Park Memorial Institute-1640 medium
SAO	Southeast Asian Ovalocytosis
SDS	sodium dodecylsulphate
SERA/SERP	serine-rich antigen/protein
SERPH	serine stretch protein homologue
SH3	*src* homology
SpD	spectrin dimer
SpT	spectrin tetramer
STEVOR	subtelomeric variant open reading frame family
TB	terrific broth
TBS	Tris buffered saline
TBST	Tris buffered saline with Tween®-20
TEA buffer	Tris EDTA acetic acid buffer
TE buffer	Tris EDTA buffer
TRAP	thrombospondin-related anonymous protein
Tris	tris(hydroxymethyl)amino-methane
tRNA	transfer ribonucleic acid
TVN	tubovesicular membrane network
V_{max}	maximum velocity
VTS	vacuolar transport signal
WHO	World Health Organisation

Chapter 1: Introduction

1.1 Malaria

Malaria, caused by an Apicomplexan of the genus *Plasmodium*, is a disease affecting humans in most tropical and subtropical regions of the world (Figure 1). Yearly 300-660 million people are diagnosed with clinical malaria and two thirds of the global incidence occurs in Africa (Snow *et al.*, 2005). The World Health Organisation estimates that annually 2.7 million deaths can be attributed to the disease (WHO, 2006). *P. vivax*, *P. ovale*, *P. malariae*, and *P. falciparum* are the four *Plasmodium* species infecting humans. *P. falciparum* causes the most severe and majority of malaria cases in Africa (Snow *et al.*, 2005) and often manifests itself as cerebral malaria in infected people. The disease can also induce anaemia and this results in the death of 10-20 % of infected children (WHO, 2006).

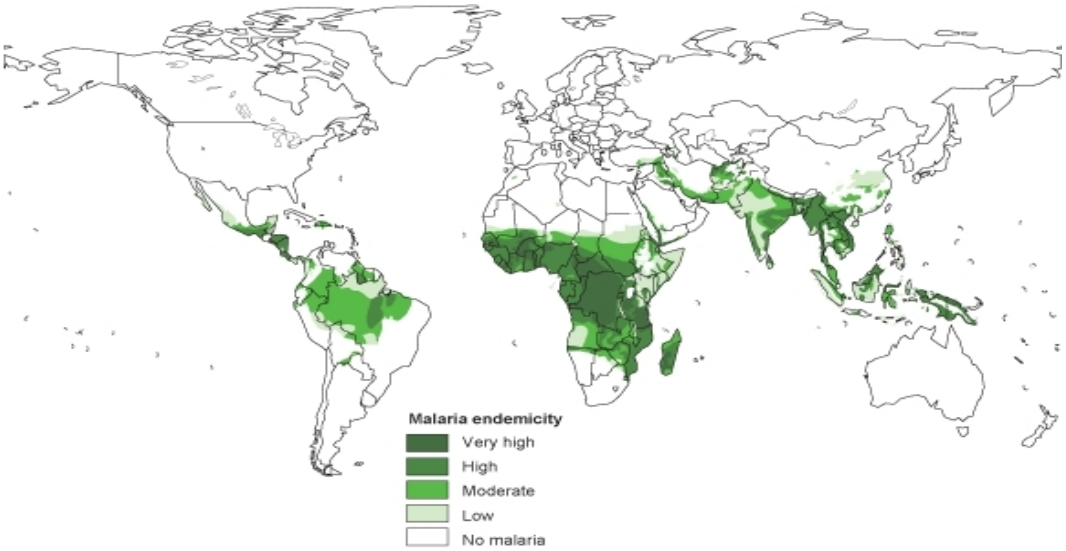


Figure 1: The global malaria distribution.
Map showing the endemicity of malaria. The severity of malaria throughout the tropical and subtropical regions of the world is indicated in green. The endemicity increases with increasing colour intensity (Roll Back Malaria, 2005).

The *Plasmodia* that cause malaria in humans, alternate their life cycle between the human and a mosquito host (reviewed in Miller *et al.*, 2002). The parasites use the mosquito as their primary host, where they complete their sexual replication and

undergo asexual replication, termed sporogony. The secondary human host is used for asexual replication, classified as schizogony. This phase includes the initial development in the liver known as pre-erythrocytic or exo-erythrocytic schizogony and the main asexual replication cycle that occurs in the human red cells, known as erythrocytic schizogony. It is during erythrocytic schizogony that the pathological and clinical symptoms of malaria occur, which include repeated attacks of fever, headache, vomiting and other flu-like symptoms. If the disease is not treated, or the parasites build up a resistance to the treatment, the parasite can kill its human host by infecting and destroying erythrocytes and by obstructing the capillaries that carry blood to the brain or other vital organs.

1.1.1 The *P. falciparum* life cycle

The intricate *P. falciparum* life cycle (Figure 2) (reviewed in Tuteja, 2007) commences in the human host when a female *Anopheles* mosquito injects a small amount of anticoagulant-containing saliva with approximately 20 haploid sporozoites into human subcutaneous tissue, or less commonly directly into the bloodstream. Once parasites have entered the bloodstream they travel to the liver, where they traverse several host cells before a single sporozoite comes to rest in a hepatocyte and develops asexually to produce 30,000-50,000 merozoites in about 5 to 7 days. Upon maturation the merozoites are released from the liver cells into the bloodstream, where they attach themselves to erythrocytes, which they subsequently invade.

Inside the erythrocyte, the parasite begins to enlarge as a uninucleate ring form, and continues to develop into a young feeding stage, known as the ring or early trophozoite. Trophozoites continue to divide asexually to produce a schizont that has approximately 16-32 nuclei. Further parasite maturation results in mononucleated merozoites that rupture the erythrocytes, thereby releasing themselves into the bloodstream. Merozoites reinvade healthy erythrocytes, resulting in a replication cycle that repeats itself approximately every 48 hours during a *P. falciparum* infection.

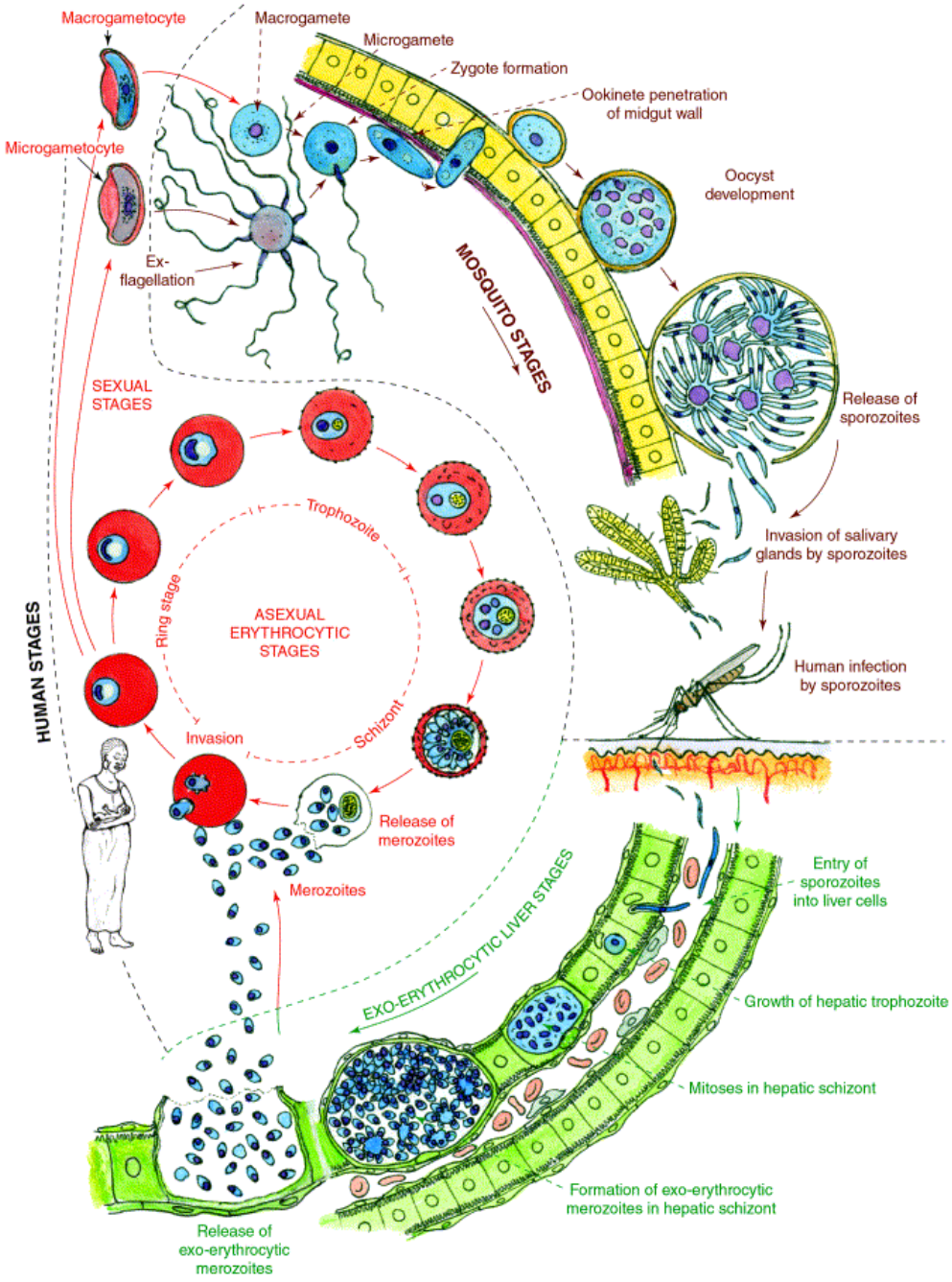


Figure 2: The *P. falciparum* life cycle.
Diagram showing the three distinct stages and the different parasite forms during the *P. falciparum* life cycle. The mosquito stage is labelled in purple; the exo-erythrocytic liver stage is labelled in green; and the asexual erythrocytic stage is labelled in red (Bannister and Mitchell, 2003).

Not all merozoites commit themselves to erythrocytic schizogony. Instead, parasites differentiate into macrogametocytes and microgametocytes (Fujioka and Aikawa, 1999) that are ingested by an *Anopheles* mosquito when it bites an

infected human host. The gametocytes develop into female (macro) and male (micro) gametes (Fujioka and Aikawa, 1999), which escape from the erythrocytes once they have entered the mosquito's gut and fuse to form a diploid zygote called the ookinete. The ookinete penetrates the gut wall and forms an oocyst on its outer surface. The diploid oocyst undergoes meiosis and eventually ~8,000 haploid sporozoites erupt from the oocyst and migrate to the salivary glands of the mosquito (Sinden and Gilles, 2002). The sporozoites remain in the salivary glands until the mosquito bites a new human host and thereby releases the parasites into the human bloodstream.

1.2 The human erythrocyte

The human erythrocyte functions as the main oxygen carrier in the body. Varying from 7.5 to 7.8 μm in diameter, the cells contain no nucleus and lack intracellular organelles (Bull and Breton-Gorius, 1995). Haemoglobin makes up 90 % of the cellular composition and it is the haem moiety contained within the haemoglobin that binds oxygen. Due to a lack of mitochondria the cell generates its energy through direct glycolysis of glucose and the hexose monophosphate shunt.

1.2.1 The erythrocyte membrane

Erythrocytes need to be supple and at the same time robust to allow them to withstand large deformations during their passage through the capillaries and spleen. These characteristics are provided by the erythrocyte membrane which contains two structural elements made up of lipids and proteins (Figure 3).

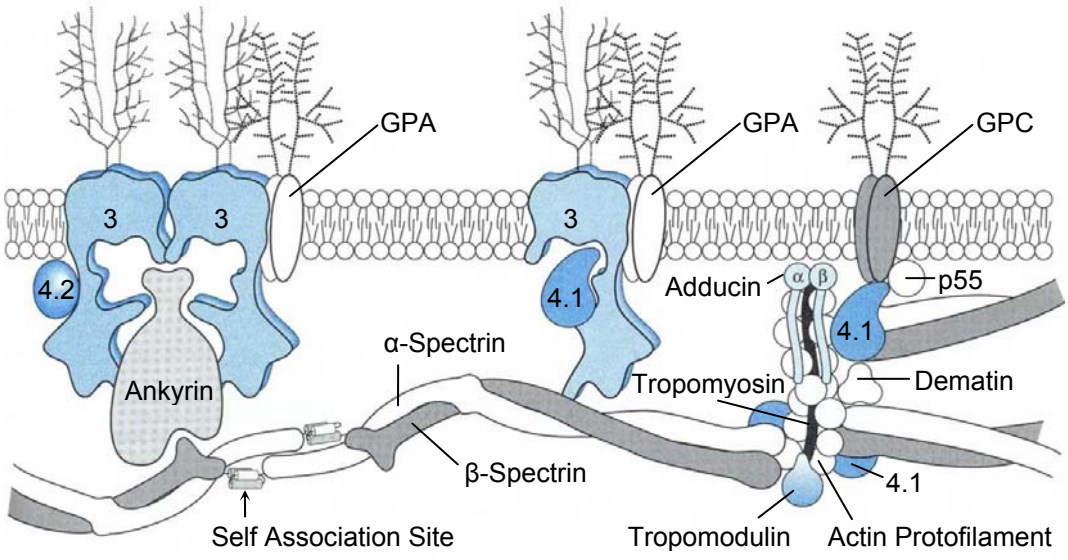


Figure 3: The erythrocyte membrane.
Diagram showing the structure of the erythrocyte membrane. The erythrocyte membrane consists of two structural components, the phospholipid bilayer and the membrane skeleton. The skeleton is attached to the integral proteins (band 3 and GPC) in the lipid bilayer via ankyrin and protein 4.1. GPA = glycophorin A; GPC = glycophorin C; 3 = band 3; 4.1 = protein 4.1; 4.2 = protein 4.2 (Walensky *et al.*, 2003).

The first structural element is the lipid bilayer that forms the permeability barrier between the internal and external environment of the erythrocyte. It is composed of two phospholipid layers that are arranged in such a way, that the hydrophobic fatty acid chains form the core of the membrane and the polar heads face outwards towards the blood plasma (outer leaflet) and the erythrocyte cytoplasm (inner leaflet). The outer leaflet contains phosphatidylcholine and sphingomyelin and the inner leaflet consists of phosphatidylserine and phosphatidylethanolamine, thus showing an asymmetry in the phospholipid arrangement. The phospholipid bilayer is very rich in nonesterified cholesterol, the majority of which is associated with sphingolipids in structures called rafts. These rafts form detergent-resistant plasma membrane microdomains and function in sorting and signalling processes of the cell (Brown and London, 1998). Embedded within the erythrocyte lipid bilayer are the integral membrane proteins, the anion exchanger (band 3) and the glycophorins (reviewed in Walensky *et al.*, 2003).

The second structural element is the membrane skeleton consisting of a protein scaffold that lies about 10 nm below the inner side of the lipid bilayer and it is this membrane skeleton that gives the erythrocyte its structural integrity. The erythrocyte membrane skeleton is anchored to the integral membrane proteins by linker proteins such as ankyrin and protein 4.1 (reviewed in Walensky *et al.*, 2003).

1.2.2 Spectrin

Making up 50-75 % of the total skeletal protein mass (Yu *et al.*, 1973), spectrin is the predominant protein in the membrane skeleton. Spectrin occurs as an αβ-heterodimer consisting of a 240 kDa alpha and 220 kDa beta chain (Marchesi, 1979).

Each chain is composed of multiple homologous 106 amino acid residue motifs (termed repeats) which fold into triple α-helical bundles with the first and third helices in parallel and the intervening second helix in an antiparallel arrangement (Speicher and Marchesi, 1984) (Figure 4). The amino acids in each α-helical bundle are arranged as a heptad motif with a seven-residue symmetry (Figure 5), where the first and fourth positions (a & d) are filled by hydrophobic and the fifth and seventh positions (e & g) are occupied by charged amino acids. The charged amino acids link the three bundles together and embed the hydrophobic amino acids in the repeat (Yan *et al.*, 1993). Formation of the repeat leaves the outer polar uncharged, positively and negatively charged amino acids (b, c & f) of each α-helical bundle available for hydrogen and electrostatic bond formation with other molecules.

The spectrin α and β chains self-associate in an antiparallel fashion starting at the dimer nucleation site, which forms at the amino terminus of β-spectrin and the carboxyl terminus of α-spectrin (Figure 6) (Ursitti *et al.*, 1996). The rest of the αβ-heterodimer then proceeds to associate side-by-side in a zipper-like fashion (Speicher *et al.*, 1992), forming a supercoiled rope-like structure.

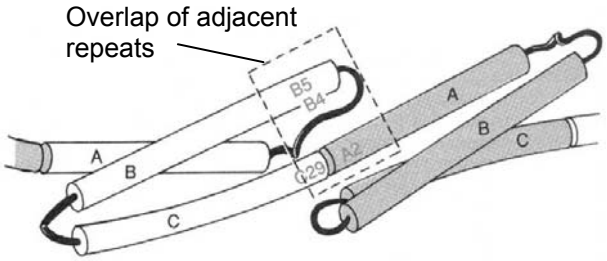


Figure 4: Triple helical repeats.
Diagram showing two spectrin repeats. Three α-helical bundles, containing 106 amino acids, fold with the first and third helices (A and C) in parallel and the intervening second helix (B) in an antiparallel arrangement to form a single spectrin repeat. Two repeats are stabilised by the interaction of the second helix of the first repeat (B4 and B5) with the third helix of the first repeat (C29) and the first helix of the second repeat (A2) (Walensky *et al.*, 2003).

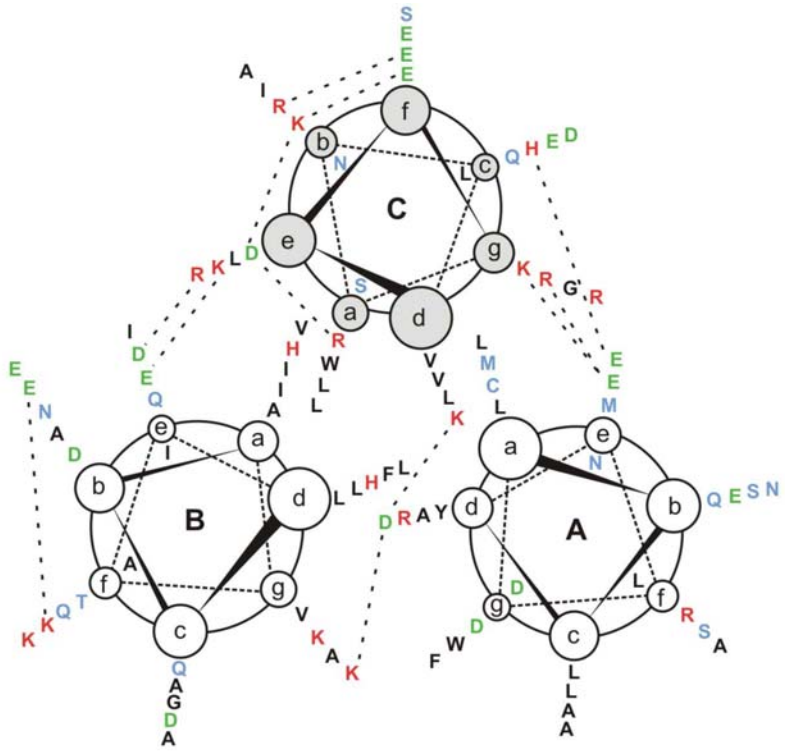


Figure 5: The amino acid distribution of a spectrin repeat.
Cross-section of a *Drosophila* α-spectrin repeat. Three heptad motif containing α-helical bundles align in a parallel (A & C) and anti-parallel (B) fashion to form the repeat. Hydrophobic (a & d) and charged (e & g) amino acids stabilise repeat formation by ionic interactions. The remaining polar uncharged, negatively and positively charged (b, c & f) amino acids are available for interactions with the spectrin-binding peptides, via hydrogen bonds and ionic interactions. Ionic interactions are represented as dashed lines and amino acids are coloured according to class: blue = polar uncharged; black = nonpolar aliphatic and aromatic; red = positively charged; green = negatively charged (adapted from Yan *et al.*, 1993).

The spectrin tetramer is the predominant form of spectrin found in the erythrocyte membrane skeleton (Ungewickell and Gratzer, 1978). A tetramer is formed when two dimers self-associate at the tetramerisation site (self-association site), which is at the opposite end to the dimer nucleation site in the spectrin molecule (Figure 6). At this tetramerisation site one N-terminal α-helix from the α-chain and two C-terminal α-helices from the β-chain link together to form a complete triple helix repeat (Tse *et al.*, 1990). Thus, spectrin tetramers are long flexible molecules arranged in a lattice framework, which lends stability and flexibility to the membrane skeleton.

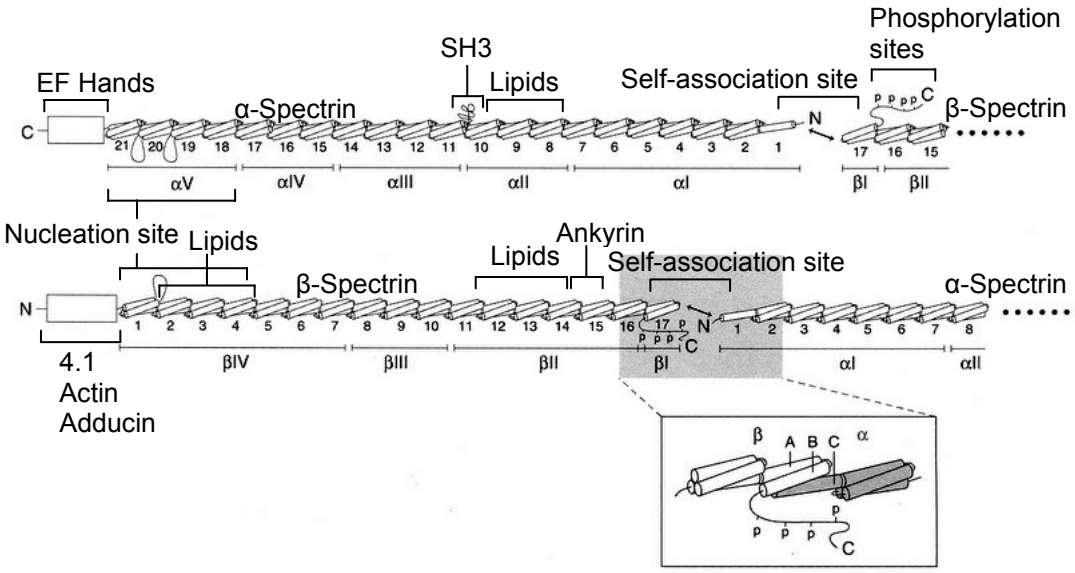


Figure 6: The spectrin tetramer.
Diagram showing the association of α- and β-spectrin triple helical repeats into dimers and a tetramer. The spectrin dimer, consisting of an α- and β-spectrin monomer, forms at the nucleation site. The tetramer is created when two spectrin dimers join at the self-association sites (inset). One N-terminal α-helix from α-spectrin (C) and two C-terminal α-helices from β-spectrin (A and B) link together to form a complete repeat. Features of the α- and β-spectrin monomers are marked: Alpha-spectrin contains a *src* homology domain (SH3) and EF hands. Beta-spectrin contains 4.1, actin, adducin and ankyrin binding sites and the phosphorylation sites. Repeats involved in phosphatidylserine binding are labelled as 'Lipids' and the five α- and four β-spectrin tryptic digest domains are labelled 'αI-V' and 'βI-IV' respectively (adapted from Walensky *et al.*, 2003).

Each spectrin monomer has additional functional domains (Figure 6). Alpha spectrin contains a *src* homology (SH3) domain in the 10th repeat, and two or three calcium-binding EF hands at the C-terminal end of the molecule. The SH3 domain functions as a protein attachment site and it has been shown that tyrosine

kinases interact with this domain (Ziemnicka-Kotula *et al.*, 1998). The N-terminus of β-spectrin contains a region of 272 amino acids that binds protein 4.1, actin and adducin (Li and Bennett, 1996) and the 15th repeat binds ankyrin, which is part of the band 3 complex (Kennedy *et al.*, 1991). The C-terminus of β-spectrin has an additional stretch of 52 amino acids that contains phosphorylated serines and threonines (Harris and Lux, 1980). Beta-spectrin also binds calmodulin, which, in the presence of calcium, regulates the interactions between spectrin, protein 4.1 and actin (Anderson and Morrow, 1987). In addition, both spectrin molecules contain repeats that interact with the lipid bilayer (An *et al.*, 2004). Finally, a limited tryptic digest of the αβ-heterodimer liberates five α-spectrin (αI = 80 kDa, αII = 46 kDa, αIII = 52 kDa, αIV = 41 kDa and αV = 41 kDa) and four β-spectrin (β I = 17 kDa, βII = 65 kDa, βIII = 33 kDa and βIV = 74 kDa) peptide domains (Speicher *et al.*, 1992).

1.2.3 The erythrocyte membrane skeleton

The lattice framework of the membrane skeleton consists of a two-dimensional network of spectrin tetramers and junctional complexes (Figure 7A). The junctional complexes are found at the distal end of the spectrin tetramer. The N-termini of six β-spectrin chains (Byers and Branton, 1985, Liu *et al.*, 1987) link to two actin protofilaments (Brenner and Korn, 1979) via protein 4.1 (Ungewickell *et al.*, 1979) and dematin (protein 4.9) (Siegel and Branton, 1985) (Figure 7b). The two actin filaments are stabilised and their length regulated by two tropomyosin molecules, that each bind six actin monomers (Fowler, 1996). The barbed ends of the actin filaments are capped by adducin and the tropomyosin molecules are capped at the opposite end by tropomodulin (Fowler, 1996).

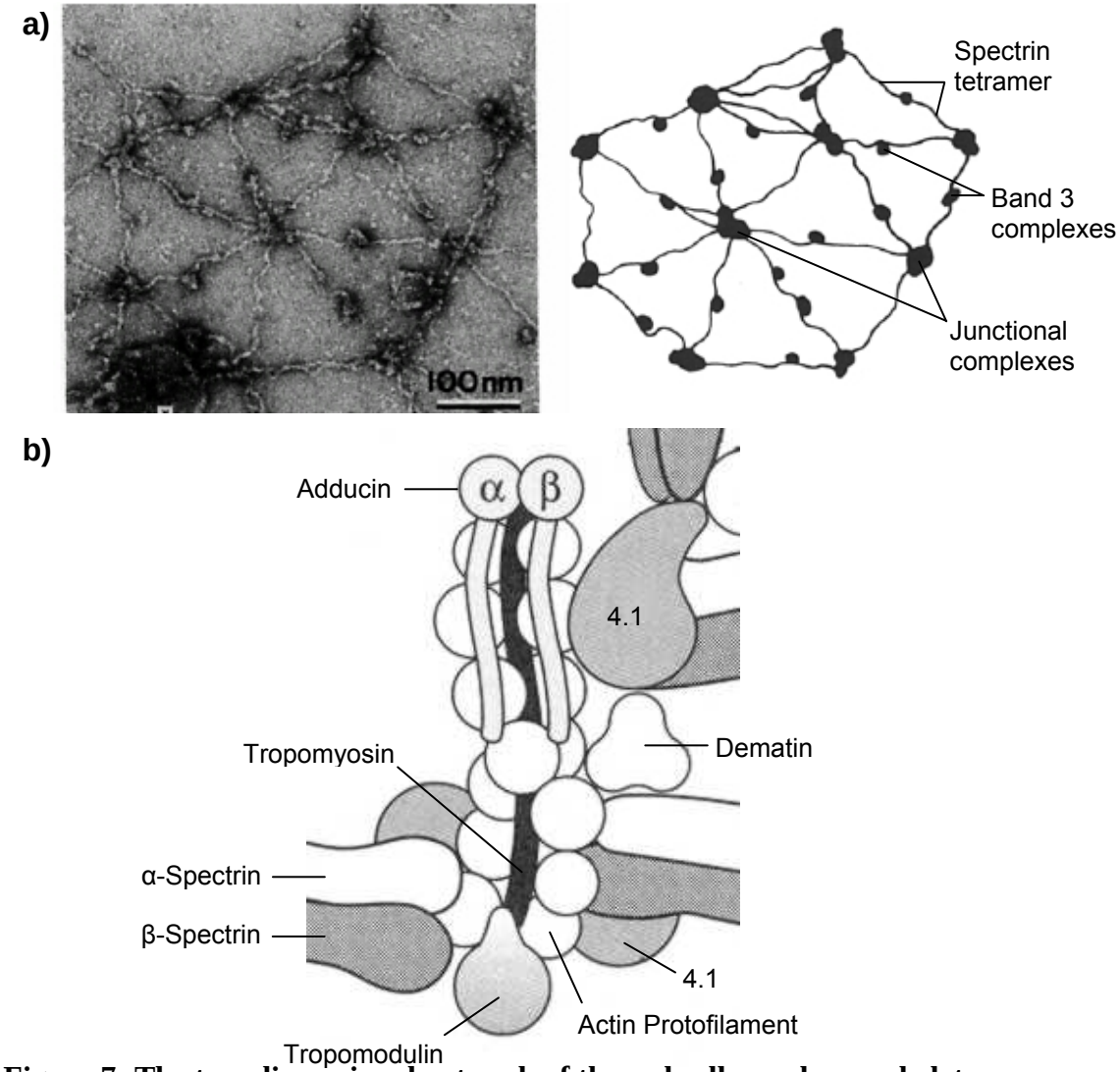


Figure 7: The two-dimensional network of the red cell membrane skeleton. Electron micrograph and diagram (a) depicting the two-dimensional network of the membrane skeleton and diagram (b) showing the junctional complex. Junctional complexes are found at the ends of the spectrin tetramer and band 3 complexes are located along the spectrin tetramer. The spectrin tetramers are attached to actin protofilaments via protein 4.1 and dematin. The actin filaments are stabilised by tropomyosin molecules and capped by adducin. The tropomyosin molecules are capped by tropomodulin ((a) Liu *et al.*, 1987, (b) adapted from Walensky *et al.*, 2003).

1.2.4 Integral membrane proteins

The erythrocyte integral membrane proteins, namely band 3 and the glycophorins, penetrate or traverse the erythrocyte membrane by interacting with the hydrophobic lipid core of the phospholipid bilayer (Figure 3).

Band 3 is the major integral membrane protein of the red cell. It forms a 52 kDa transmembrane channel, which functions in the exchange of HCO_3^- and Cl^- and therefore actively participates in the exchange of CO_2. The 43 kDa N-terminal cytoplasmic domain serves as an anchor point for the membrane skeleton by interacting with ankyrin, protein 4.1, and protein 4.2 (Bennett and Stenbuck, 1980). Two complexes associate with band 3. The first is the Rh complex, which is linked to ankyrin and protein 4.2 via the Rh polypeptide and CD47 respectively (Bruce *et al.*, 2003). The second is a glycolytic enzyme complex containing glyceraldehyde-3-phosphate dehydrogenase (Kant and Steck, 1973), phosphofructokinase and aldolase (Jenkins *et al.*, 1984), haemoglobin (Walder *et al.*, 1984), lactate dehydrogenase and pyruvate kinase (Campanella *et al.*, 2005). This complex interacts with the cytoplasmic domain of band 3.

The glycophorins have sialic acid moieties bound to domains that are exposed on the erythrocyte surface. These sialic acid chains code for blood group variants and also give the erythrocyte surface a net negative charge, which prevents the erythrocytes from adhering to each other and the vessel walls (reviewed in Chasis and Mohandas, 1992).

1.2.5 Membrane skeleton linkage to the lipid bilayer

The membrane skeleton is linked to the integral proteins of the lipid bilayer via primary and secondary attachment sites. Primary attachment occurs through the band 3 complex, where band 3 binds to the C-terminal end of β-spectrin via ankyrin (Bennett and Stenbuck, 1979, Byers and Branton, 1985). Protein 4.2 further strengthens this membrane interaction by linking ankyrin and band 3 (Korsgren and Cohen, 1988). In the secondary attachment site, protein 4.1 links the actin filaments in the junctional complex to glycophorin C (Anderson and Lovrien, 1984) and this interaction is stabilised by p55 (Figure 3) (Alloisio *et al.*, 1993).

The membrane skeleton linkage to the lipid bilayer is also strengthened by the direct interaction of the skeleton proteins with the phospholipids of the erythrocyte membrane. Alpha- and beta-spectrin have specific repeats (Figure 6) that bind phosphatidylserines (An *et al.*, 2004) and the 30 kDa basic domain of protein 4.1 also interacts with phosphatidylserines (Cohen *et al.*, 1988, Rybicki *et al.*, 1988).

1.3 The erythrocytic phase of the *P. falciparum* life cycle

The malaria parasite spends the main part of its asexual life cycle in the human erythrocyte. To ensure its survival, the parasite has to enter an erythrocyte by disrupting the organisation of the erythrocyte membrane, and once inside, the site of invasion has to be repaired. During development and maturation the parasite depletes the nutrient resources available in the erythrocyte and therefore the parasite has to modify the host cell and host cell membrane to obtain additional nutrients and eliminate waste products. The parasite also needs to remain undetected by the host's immune system and this is achieved by further modifications to the host cell and membrane. Finally, the parasite has to disrupt the erythrocyte membrane to ensure its release from the erythrocyte.

1.3.1 Erythrocyte invasion

The host cell invasion model suggests that the parasite merozoite goes through three phases to enter the erythrocyte (Figure 8) (reviewed in Topolska *et al.*, 2004). These phases are: 1) an initial low-affinity interaction between any part of the merozoite and the host cell surface, followed by the reorientation of the merozoite to allow its apical prominence to contact the host cell; 2) irreversible attachment and formation of an electron-dense junction between the apical prominence and the host cell and finally; 3) invasion, in which the parasite propels itself forward into a membrane enclosed compartment called the parasitophorous vacuole, which eventually seals behind the parasite (Dvorak *et al.*, 1975).

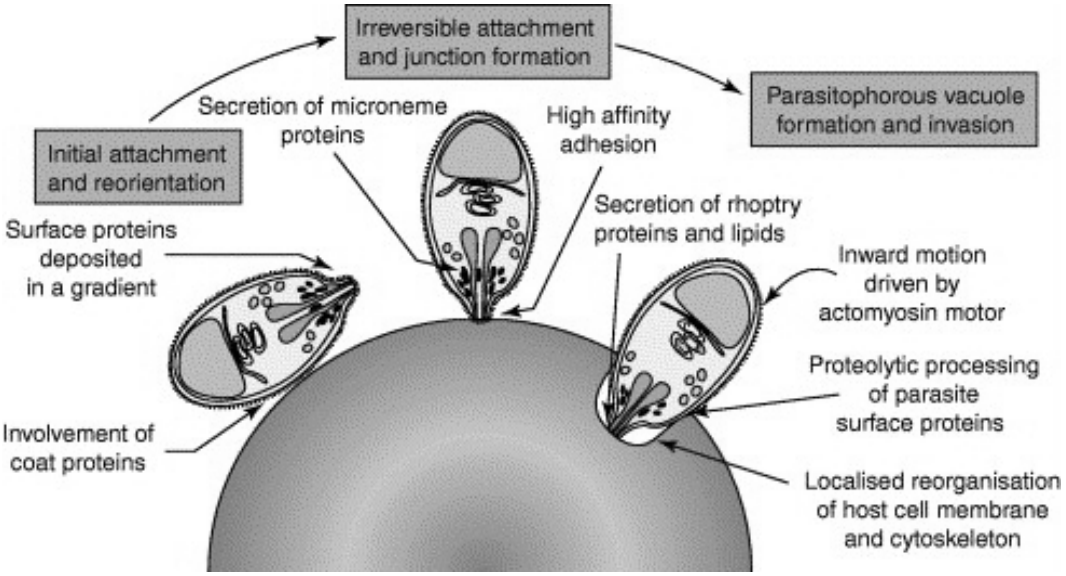


Figure 8: Erythrocyte invasion by *P. falciparum*.
Diagram showing the three phases of invasion. The merozoite undergoes initial attachment and reorientation; followed by irreversible attachment and junction formation; and parasitophorous vacuole formation and invasion (Kats *et al.*, 2006, originally adapted from Chitnis and Blackman, 2000).

During each phase a specific set of proteins located either on the surface of, or in specific organelles and compartments of the merozoite, are used to recognise and disrupt the erythrocyte (Figure 9). The organelles involved in invasion are the rhoptries, which are made up of two, large, tear-shaped organelles terminating at the apex of the parasite, and the micronemes, which are small tube-shaped structures associated with the rhoptries. Additional proteins are released from the dense granules, which are intermediate in size, spherical, and denser than the micronemes and rhoptries (Langreth *et al.*, 1978).

The first phase of invasion occurs when the merozoite comes into contact with an uninfected erythrocyte. This contact is accidental in nature and results from the recognition of red cell receptors by the coat filaments of the merozoite. These filaments are made up of parasite proteins such as the merozoite surface proteins (MSP-1-10), serine-rich antigen (SERA/SERP) and S antigen. This initial interaction between parasite and host is weak, non-specific and reversible, and merozoites are often seen to dissociate from erythrocytes.

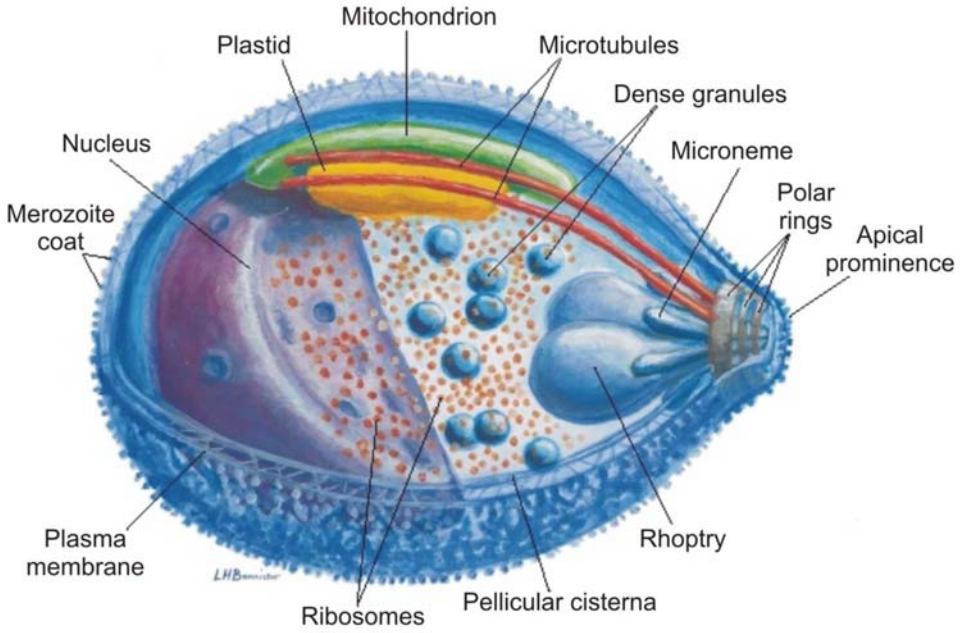


Figure 9: The *P. falciparum* merozoite.
Diagram showing the structure and organelles of the *P. falciparum* merozoite. The merozoite contains secretory organelles involved in invasion, namely the micronemes, rhoptries and dense granules (Bannister *et al.*, 2000).

Invasion continues with the parasite reorientating itself so that its apical end comes into contact with the erythrocyte surface. This reorientation strengthens the interaction between the merozoite surface and the red cell surface and eventually becomes irreversible, and the parasite commits itself to erythrocyte invasion. After contact, intracellular stores of Ca^{2+} are released within the parasite (Lovett and Sibley, 2003) and this activates the discharge of parasite proteins, such as the erythrocyte binding antigens (EBA-175, EBA-140/BAEBL, EBA-181/JESEBL) and the apical membrane antigen 1 (AMA-1) from the micronemes. The erythrocyte binding antigens form a junction with the erythrocyte membrane receptors, such as Glycophorin A, B, and C/D (Hadley *et al.*, 1987) and these interactions are to some degree redundant and interchangeable (Mayer *et al.*, 2002). The tight junction remains during invasion and is seen as a ring that moves backwards over the parasite surface as the parasite actively invades the erythrocyte (Aikawa *et al.*, 1978). Shortly after junction formation the rhoptries release their parasite proteins, namely the rhoptry associated proteins (RAP-1, RAP-2, RAP-3), high molecular weight rhoptry proteins (RhopH-1, RhopH-2,

RhopH-3), and proteases (gp76), which aid in the invasion process. The rhoptries also discharge parasite lipids (Etzion *et al.*, 1991). These rhoptry proteins and lipids, together with proteins and lipids from the host cell membrane, are used to form the parasitophorous vacuole membrane (Stewart *et al.*, 1986). This membrane surrounds the parasitophorous vacuole which is a sack-like structure that provides a secure environment for the parasite to reside and grow in. The dense granules are the last organelles to discharge their protein content, such as the ring-infected erythrocyte surface antigen (RESA), ring membrane antigen (RIMA), and the subtilisin-like proteases (PfSUB-1, PfSUB-2). These proteins either integrate themselves into the parasitophorous vacuole or cross the parasitophorous vacuole membrane to interact with the erythrocyte membrane lipid bilayer and erythrocyte membrane skeleton proteins. PfSUB-2 has been implicated in the shedding of MSP-1 and AMA-1 from the parasite surface as it invades the erythrocyte (Howell *et al.*, 2003). Another set of proteases, the rhomboids (PfROM-1 and PfROM-4), cleave AMA-1, thrombospondin-related anonymous protein (TRAP), circumsporozoite protein (CSP) and TRAP-related protein (CTRP), merozoite TRAP homologue (MTRAP), PFF0800c, EBA-175, EBA-140/BAEBL, EBA-181/JESEBL, MAEBL, and the reticulocyte-binding homologues (RH-1, RH-2a, RH-2b, and RH-4), and could therefore be involved in the shedding process (Baker *et al.*, 2006).

Invasion requires the use of an actin/myosin motor complex termed the glideosome (Figure 10) (Keeley and Soldati, 2004, reviewed in Cowman and Crabb, 2006). This complex is located in the parasite's pellicle, between the parasite plasma membrane and the inner membrane complex (IMC). The IMC is constructed from a single large flattened vesicle that is joined at a single suture line running the length of the long axis of the parasite (Morrissette and Sibley, 2002). The IMC is not found at the apical prominence (Bannister *et al.*, 2000), thereby allowing the micronemes, rhoptries and dense granules to discharge their contents to the external environment. Within the glideosome, *Plasmodium* myosin (MyoA) (Pinder *et al.*, 1998) attaches to the Myosin Tail Interacting Protein (MTIP) via its light chain and the other end of myosin interacts with actin (Webb

et al., 1996). MTIP in turn interacts with glideosome-associated protein 45 (GAP45) and glideosome-associated protein 50 (GAP50) to link the glideosome to the IMC (Jones *et al.*, 2006). The glideosome is also present in other invasive stages of *Plasmodium*, and in sporozoites it has been shown that the glideosome connects to the parasite plasma membrane via aldolase and the adhesin TRAP (Buscaglia *et al.*, 2003). It is therefore thought that the link to the merozoite plasma membrane is formed by the secreted microneme or rhoptry ligands and possibly aldolase. The parasite gains entry into the host cell using the force that is generated by the actin-myosin motor. Because the motor is anchored to the IMC, the generated movement is highly directional and causes the parasite to be propelled forwards into the host cell. During invasion the adhesin-aldolase-actin complex moves towards the posterior end of the parasite and this complex is released from the host cell by proteases such as the subtilisin-like proteases and rhomboids mentioned previously.

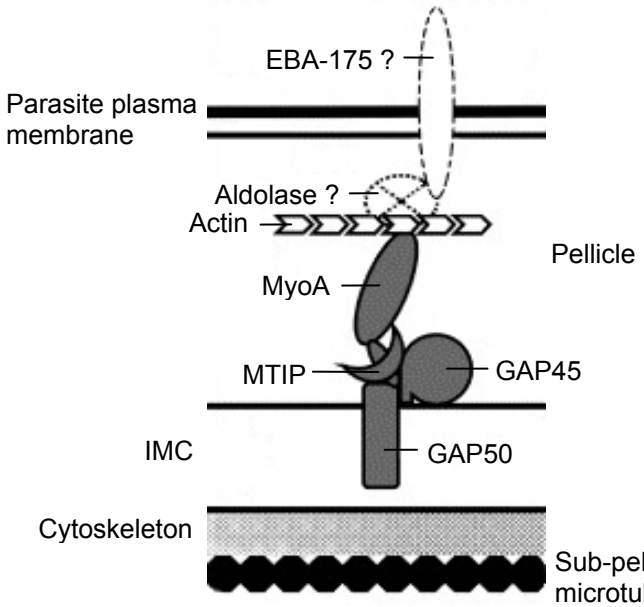


Figure 10: The actin-myosin motor of *P. falciparum*.
Diagram showing the actin-myosin motor found in the *P. falciparum* glideosome. The glideosome consists of glideosome-associated protein 50 (Gap50), Myosin Tail Interacting Protein (MTIP), glideosome-associated protein 45 (GAP45) and *Plasmodium* myosin (MyoA). It attaches to the inner membrane complex (IMC) via GAP50 and interacts with the actin filaments via MyoA. Aldolase and erythrocyte binding antigen-175 (EBA-175) could form the link between actin and the parasite plasma membrane. Myosin moves along the actin filaments and thereby moves the parasite into the host cell (adapted from Jones *et al.*, 2006).

1.3.2 Erythrocyte remodelling

Enhancing the permeability of the erythrocyte membrane

Once the parasite has invaded the erythrocyte the ring stage parasite seems to be 'dormant' within the parasitophorous vacuole until approximately 12-16 hours later when it develops into the trophozoite stage. This maturation is concomitant with a rapid increase in metabolic and biosynthetic activity, requiring an energy source for glycolysis (Roth, 1990) and other molecules for protein, RNA and DNA synthesis. The parasite's main source of amino acids for protein synthesis comes from the degradation of haemoglobin in the food vacuole (Sherman, 1977), but the acquisition of the other essential components required for survival poses a problem for the parasite, because it resides in a host cell that is devoid of and lacks the ability to synthesise macromolecules. In addition to acquiring nutrients from the plasma external to the erythrocyte, the parasite needs to eliminate the excess amino acids from the digestion of haemoglobin (Krugliak *et al.*, 2002), lactate from glycolysis (Kanaani and Ginsburg, 1991) and other toxic waste products produced during parasite growth. The parasite also needs to regulate the influx and efflux of ions and water as it matures within the erythrocyte. The parasite overcomes these obstacles by upregulating and activating transporters and channels within the erythrocyte plasma membrane via phosphorylation (Decherf *et al.*, 2004). The parasite also incorporates its own transporters and channels into the erythrocyte membrane to obtain essential nutrients, such as the vitamin pantothenic acid (Saliba *et al.*, 1998). All these transporters and channels have been collectively termed New Permeability Pathways (NPP) (Kirk, 2001).

Modifying the erythrocyte to accommodate cellular trafficking

As the parasite progresses to the late trophozoite stage it further modifies the erythrocyte to satisfy its own survival and growth requirements. The parasite shuttles parasite proteins, required for cytoadherence and rosetting, to the erythrocyte surface and incorporates them into structures termed 'knobs' (Atkinson and Aikawa, 1990). However, the transport of proteins to the

erythrocyte surface is not a straightforward procedure because the parasite resides within the parasitophorous vacuole that is surrounded by a membrane. This membrane does not contain any transporters or channels and is often referred to as a molecular sieve that allows the free exchange of nutrients and metabolites, but prevents the movement of proteins to and from the erythrocyte cytosol (Desai and Rosenberg, 1997). Underlying the transport of proteins from the parasite through the parasitophorous vacuole to the erythrocyte membrane and the transport of lipids from the erythrocyte membrane to the parasitophorous vacuole membrane are several transport mechanisms.

Protein transport in eukaryotic cells is normally initiated through the classical secretory pathway and this is also the case in the parasite (Das *et al.*, 1994). In this pathway proteins that contain a hydrophobic N-terminal signal sequence are transported into the endoplasmic reticulum (ER) after leaving the ribosome. From the ER the proteins move into the Golgi complex where they are sorted and subsequently transported in vesicles to the cellular organelles or to the plasma membrane of the cell. A classical Golgi complex containing cisternae has however not been visualised in the parasite (Aikawa, 1971), but components of the Golgi apparatus have been found and it has therefore been suggested that the parasite contains a highly modified 'unstacked' Golgi form, where the Golgi compartments are not located adjacent to one another (Van Wye *et al.*, 1996). It should also be noted that not all exported parasite proteins contain an N-terminal signal sequence that determines transport via the ER and thus it seems that the parasite also relies on an ER-Golgi-independent secretory pathway (Mattei *et al.*, 1999) that has yet to be defined.

Once the parasite proteins have been processed via the ER they have to traverse the parasitophorous vacuole. Parasite proteins can be either soluble or membrane-bound and this defines if the protein is transported freely through the parasitophorous vacuole and erythrocyte cytosol, or if the protein has to be transported via vesicles either in a soluble form or attached to the membrane of the transport vesicles (Lingelbach and Przyborski, 2006). Vesicle transport is

thought to occur via a one-step or a two-step process (Figure 11) for membrane-bound proteins.

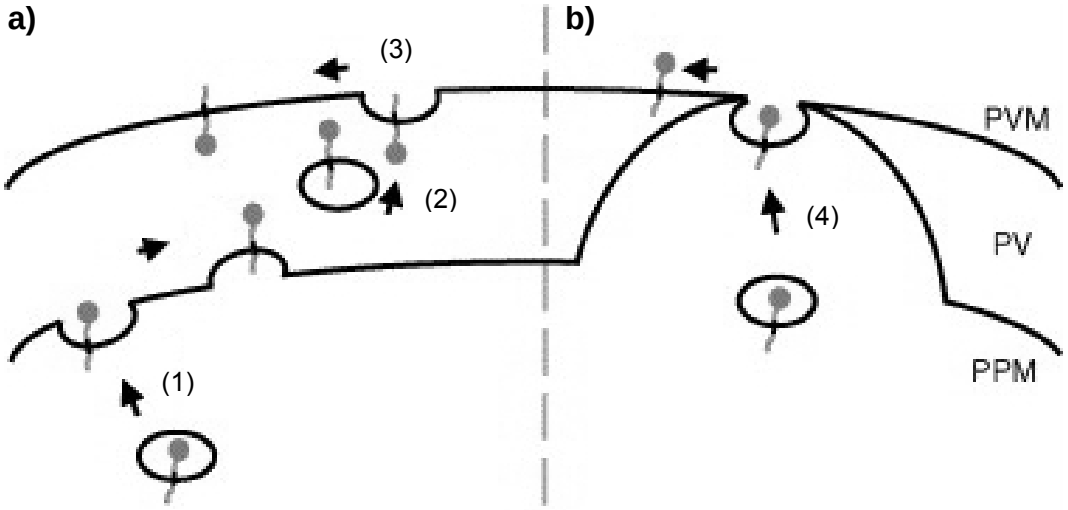


Figure 11: *P. falciparum* protein transport to the erythrocyte cytosol.
Diagram showing the transport of membrane-bound proteins across the parasitophorous vacuole via the two-step (a) or one-step (b) process. In the two-step process the secretory vesicle containing the exported protein (grey dot) releases its contents into the parasitophorous vacuole via the parasite plasma membrane (1). The protein is transported in another vesicle across the parasitophorous vacuole (2) and released into the erythrocyte cytosol via the parasitophorous vacuole membrane (3). In the one-step process the secretory vesicle containing the exported protein joins to the parasite plasma membrane where it is in close contact with the parasitophorous vacuole membrane and releases the protein directly into the erythrocyte cytosol (4) (Lingelbach and Przyborski, 2006).

In the one-step process secretory vesicles within the parasite join the parasite plasma membrane where it is in close contact with the parasitophorous vacuole membrane. The contents of these vesicles are thus released directly into the erythrocyte cytosol (Elmendorf and Haldar, 1993). The two-step process entails parasite secretory vesicles releasing their contents into the parasitophorous vacuole. The proteins are then transported across the vacuolar space in another secretory vesicle and are finally released into the erythrocyte cytosol through the parasitophorous vacuole membrane either by protein translocation complexes or by vesicle budding (Ansorge *et al.*, 1996, Lingelbach and Przyborski, 2006). The two-step process seems to be the predominant form of trafficking utilised by the

parasite (Ansorge *et al.*, 1996) and transport across the parasitophorous vacuolar membrane requires the presence of transport signal. A vacuolar transport signal (VTS) (Hiller *et al.*, 2004) and a *Plasmodium* export element (PEXEL motif) (Marti *et al.*, 2004) have been characterised to date, and are located downstream from the hydrophobic N-terminal signal sequence recognised by the endoplasmic reticulum.

After the parasite proteins have reached the erythrocyte cytosol, they need to be transported to the erythrocyte membrane. This transport occurs via two proposed transport models (Figure 12). The first is the vesicular transport model (Taraschi *et al.*, 2001), in which parasite proteins are transported in vesicles across the erythrocyte cytosol to the Maurer's clefts. Maurer's clefts are 'secretory organelles' (reviewed in Przyborski *et al.*, 2003) that consist of stacks of flattened lamellae and are located just below the erythrocyte plasma membrane (Langreth *et al.*, 1978). In the second model, termed the lateral diffusion model, the tubovesicular membrane network (TVN) is utilised to transport proteins to the Maurer's clefts (Elmendorf and Haldar, 1994, Wickert *et al.*, 2004). The TVN extends from the parasitophorous vacuolar membrane into the erythrocyte cytosol and consists of a collection of membranous structures. It is interesting to note that the *P. falciparum* Golgi proteins that have been identified to date, have been found to be associated with the TVN (Elmendorf and Haldar, 1994, Van Wye *et al.*, 1996).

The final transportation step in both models is the release of the parasite proteins from the Maurer's clefts into the erythrocyte membrane and the incorporation of these proteins into the knobs, which are often seen to be closely associated with the Maurer's clefts (Aikawa *et al.*, 1986).

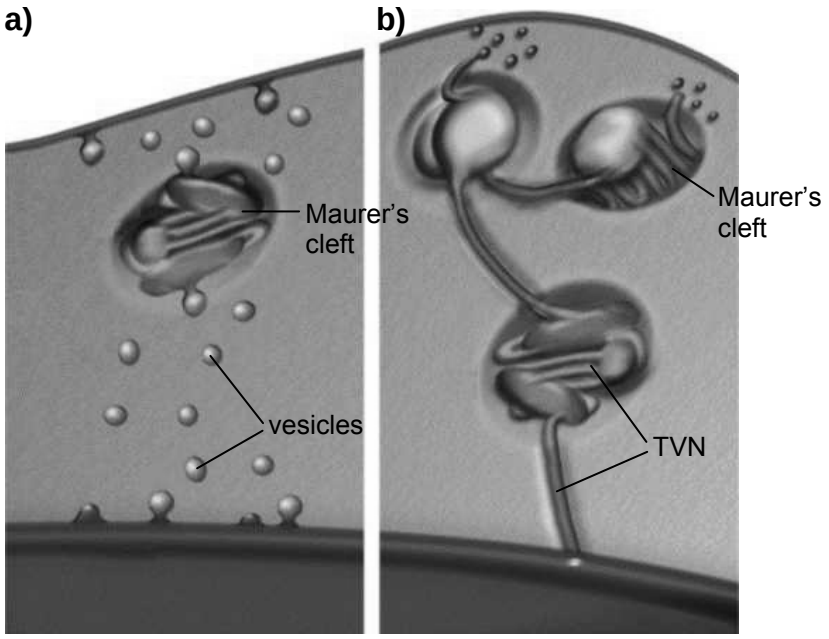


Figure 12: *P. falciparum* protein trafficking to the erythrocyte surface.
Diagram showing the vesicular transport (a) and the lateral diffusion model (b). In the vesicular transport model parasite proteins are transported in vesicles across the erythrocyte cytosol to the Maurer's clefts. In the lateral diffusion model proteins are transported via the tubovesicular membrane network (TVN) to the Maurer's clefts. In both models the proteins are eventually transported in vesicles to the erythrocyte membrane and surface (adapted from Przyborski *et al.*, 2003).

1.3.3 Haemoglobin digestion

Several *P. falciparum* proteases are involved in the degradation of haemoglobin inside the food vacuole (Figure 13). Digestion of haemoglobin commences at the hinge region of the α-globin chain and is mediated by the aspartic proteases plasmepsins-I, -II, and -IV (Gluzman *et al.*, 1994, Banerjee *et al.*, 2002, Wyatt and Berry, 2002). Further degradation of the globin chains is facilitated by the same enzymes, as well the histo-aspartic protease (HAP) (Banerjee *et al.*, 2002), and three cysteine proteases, falcipain-2, -2', and -3 (Shenai *et al.*, 2000, Sijwali and Rosenthal, 2004, Sijwali *et al.*, 2001). The M16 metalloprotease falcilysin converts short globin polypeptides into oligopeptides consisting of 2-10 amino acids (Eggleson *et al.*, 1999). Finally, the calpain-like dipeptidyl peptidase I (DPAP1) hydrolyses the oligopeptides into 2-10 amino acid residues (Klemba *et al.*, 2004) which are exported to the cytoplasm where they are cleaved into single amino acid units by the cytosolic neutral aminopeptidases (Curley *et al.*, 1994).

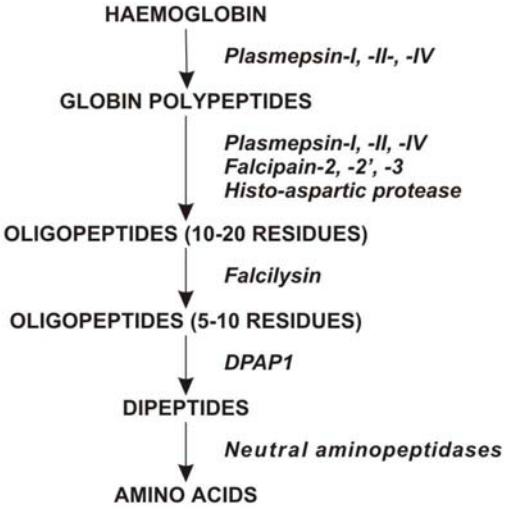


Figure 13: Haemoglobin digestion.
Haemoglobin digestion in the food vacuole is initiated by the aspartic proteases, plasmepsin -I, -II, and -IV. Further processing into dipeptides is performed by the cysteine proteases falcipain-2, -2', and -3, followed by the metalloprotease falcilysin and finally the dipeptidyl peptidase DPAP1. Dipeptides are transported into the cytosol where they are digested by neutral aminopeptidases (adapted from Klemba *et al.*, 2004).

1.3.4 Parasite escape

Approximately 40 hours after invasion the trophozoite develops into a schizont containing several daughter nuclei. These mature into merozoites which escape the host cell at approximately 48 hours after the parasite entered the erythrocyte. The two-step parasite exit process is initiated at the schizont stage when the parasite secretes proteases, probably a cysteine protease (Salmon *et al.*, 2001), into the parasitophorous vacuole. Micrographs of *P. falciparum* schizont-infected erythrocytes have shown that the material on each side of the parasitophorous vacuole membrane has an equal density and structure, indicating that the parasitophorous vacuole membrane has become permeable to host cell proteins (Langreth *et al.*, 1978). The proteases implicated in the second step of the two-step parasite exit process are an aspartic protease and a cysteine or serine protease that degrade the membrane skeleton of the erythrocyte (Wickham *et al.*, 2003).

The degradation of the membranes and the conversion of the schizont into merozoites, give rise to the irregular schizont (Figure 14). Just before escape the

parasites rearrange themselves into a short-lived distinctive ‘flower’ shape (Trager, 1956, Glushakova *et al.*, 2005). This rearrangement is concomitant with an influx of fluid into the erythrocyte due to the excess colloid-osmotic pressure of intracellular proteins since most of the haemoglobin has been digested by the parasite (Lew, 2005). The build-up of osmotic pressure causes the parasitophorous vacuole membrane and the erythrocyte plasma membrane to blow apart, thereby scattering the merozoites (Trager, 1956, Glushakova *et al.*, 2005). The force of the merozoite release is strong enough to bring the merozoites into contact with new erythrocytes even in places where the parasites have been sequestered. The parasite therefore does not need to be motile and only becomes motile when a host cell is encountered and the parasite has committed itself to invasion.

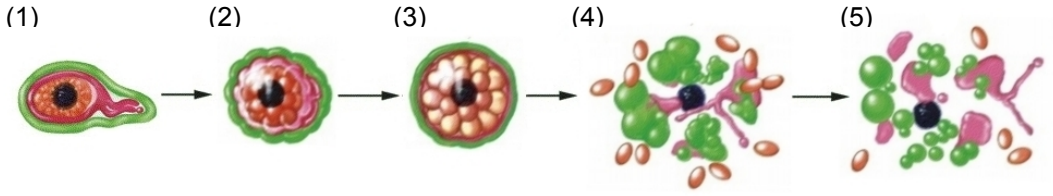


Figure 14: *P. falciparum* escape from the erythrocyte.
Diagram showing the morphological changes parasite and erythrocyte undergo during *P. falciparum* escape. The irregular schizont surrounding the food vacuole (black), initiates the two-step escape process by first weakening the parasitophorous vacuole membrane (pink) and then the erythrocyte membrane (green) by enzymatic proteolysis (1). The schizont continues to develop into merozoites (orange) (2) and the influx of fluid through the damaged erythrocyte membrane results in the distinctive ‘flower’ shape (3). The build-up of osmotic pressure causes the parasitophorous vacuole membrane and erythrocyte membrane to rupture, ‘blowing’ the cell apart and scattering the mature merozoites (4). Vesiculation of the erythrocyte and parasitophorous vacuole membranes results from the violent and sudden rupture of the erythrocyte (5) (Glushakova *et al.*, 2005).

1.4 Parasite and erythrocyte membrane interactions

Malaria proteins interact with the erythrocyte surface and the membrane skeleton proteins at several points during the erythrocytic life cycle. A summary of these interactions is listed in Table 1.

Table 1: *P. falciparum* proteins interacting with the erythrocyte membrane.

Parasite protein	EM protein	Stage	Possible function
MSP-1	band 3; spectrin	invasion	initial erythrocyte contact and contact during active invasion
MSP-9/ABRA	band 3	invasion	cleavage of band 3
EBA-175	glycophorin A	invasion	contact during active invasion
BAEBL/ EBA-140	glycophorin C/D	invasion	contact during active invasion
RH-1	receptor Y	invasion	contact during active invasion
RH-2b	receptor Z	invasion	contact during active invasion
JESEBL/ EBA-181	receptor E	invasion	contact during active invasion
SERA	erythrocyte and PVM phospholipids	invasion; escape	disrupts phospholipids
RhopH complex	erythrocyte phospholipids	invasion	disrupts phospholipids
gp76	band 3, glycophorin A	invasion	membrane skeleton destabilisation by band 3 and glycophorin A cleavage; PV formation
RESA	β-spectrin	invasion; growth	chaperone in membrane skeleton repair; thermal stability; inhibits subsequent parasite invasion
FEST	EM skeleton	growth	possible role in phosphorylating protein 4.1 and spectrin
MESA/ PfEMP-2	protein 4.1	growth	anchors knobs; membrane skeleton destabilisation by binding to protein 4.1 where glycophorin C and p55 bind
PfEMP-1	spectrin-actin junction	growth	anchors knobs
PfEMP-3	EM skeleton	growth	transports PfEMP-1 from Maurer's clefts to the erythrocyte surface
*Pf*332	EM skeleton	growth	-
KAHRP/ HRP-I	α-spectrin; ankyrin	growth	anchors knobs by interacting with spectrin and PfEMP-1
*Pfsbp*1	EM skeleton	growth	attaches Maurer's clefts to membrane skeleton network
RIFINs	erythrocyte phospholipids	growth	rosetting and antigenic variation
STEVOR	erythrocyte phospholipids	growth	rosetting and antigenic variation
plasmepsin-II	α-spectrin; protein 4.1; actin	escape	disrupts membrane skeleton
plasmepsin-IV	spectrin	escape	disrupts membrane skeleton
37 kDa acidic protease	β-spectrin; protein 4.1	escape	disrupts membrane skeleton
falcipain-2	ankyrin; protein 4.1	escape	disrupts membrane skeleton

Abbreviations: ABRA = acidic-basic repeat antigen; EBA= erythrocyte binding antigen; EM = erythrocyte membrane; FEST = *falciparum*-exported serine/threonine kinase; HRP = histidine-rich protein; KAHRP = knob-associated histidine-rich protein; MESA = mature parasite-infected surface antigen; MSP = merozoite surface protein; PfEMP = *P. falciparum* erythrocyte membrane protein; *Pfsbp*1 = *P. falciparum* skeleton binding protein-1; PV = parasitophorous vacuole; PVM = parasitophorous vacuole membrane; RESA = ring-infected erythrocyte surface antigen; RH = reticulocyte-binding homologues; RhopH = high molecular weight rhoptry protein; RIFINs = repetitive interspersed family; SERA = serine-rich antigen; STEVOR = subtelomeric variant open reading frame family.
References for each protein-protein interaction are given in the text.

1.4.1 Protein interactions involved in erythrocyte invasion

To initiate and commit to invasion several parasite ligands interact with the glycophorin moieties and band 3 on the erythrocyte surface. MSP-1 and MSP-9/ABRA interact with band 3 (Goel *et al.*, 2003, Kushwaha *et al.*, 2002) and MSP-1 also interacts with spectrin (Herrera *et al.*, 1993). EBA-175 binds to glycophorin A (Sim *et al.*, 1994) and BAEBL/EBA-140 associates with glycophorin C/D (Mayer *et al.*, 2001). Receptors Y, Z, and E are utilized by RH-1, RH-2b and JESEBL/EBA-181 respectively (Rayner *et al.*, 2001, Duraisingh *et al.*, 2003, Gilberger *et al.*, 2003) and the ligands of glycophorin B and receptor X have yet to be identified (Dolan *et al.*, 1994). During invasion several proteases and phospholipases are released by the micronemes and rhoptries. These enzymes induce structural changes in the erythrocyte membrane, such as the stripping of the underlying erythrocyte membrane skeleton proteins from the area where the merozoite has attached itself to the erythrocyte surface. Interestingly Dluzewski *et al*. (1983) showed that immunological and chemical crosslinking of spectrin inhibited parasite invasion in resealed erythrocytes. The cysteine protease SERA and the rhoptry proteins of the RhopH complex associate with the phospholipids of the inner leaflet of the erythrocyte membrane and could thus facilitate the entry of the parasites into the host cell (Perkins and Ziefer, 1994). The protease, gp76, has been suggested to be involved in invasion by degrading band 3 and glycophorin A (Roggwiller *et al.*, 1996). In addition, an erythrocyte cyclic adenosine monophosphate-independent kinase that phosphorylates spectrin is necessary for invasion (Rangachari *et al.*, 1986) because phosphorylation of β-spectrin decreases the rigidity and stability of the erythrocyte membrane (Manno *et al.*, 1995). Finally RESA, which interacts with spectrin, is thought to act as a chaperone during the repair of the membrane skeleton after invasion has been completed (Foley and Tilley, 1995). The protein binds to β-spectrin, close to the self-association site where it stabilises the spectrin tetramer, thereby increasing the thermostability of the erythrocyte and preventing further invasion by other malaria parasites (Pei *et al.*, 2007).

1.4.2 Protein interactions facilitating parasite growth

As the parasite proliferates within the erythrocyte it causes striking structural and morphological changes to its host cell. These include the loss of the typical erythrocyte shape, alterations in the mechanical properties of the cell, and modifications in the phosphorylation state of erythrocyte membrane skeleton proteins. Phosphorylation of ankyrin in uninfected erythrocytes destabilises the interaction between ankyrin and spectrin tetramers (Lu *et al.*, 1985) and the phosphorylation of protein 4.1 disrupts the interaction between protein 4.1, spectrin and actin (Ling *et al.*, 1988), and destabilises the interaction between band 3 and the erythrocyte membrane skeleton (Manno *et al.*, 2005). Thus an increased phosphorylation of protein 4.1 which is observed in *P. falciparum* infected erythrocytes (Chishti *et al.*, 1994) could reduce the mechanical stability of the membrane skeleton and contribute significantly to the morphological changes observed in infected erythrocytes. To date the enzyme responsible for the phosphorylation of protein 4.1 has not been identified, but a *falciparum*-exported serine/threonine kinase (FEST) (Kun *et al.*, 1997) could be one such candidate.

Another striking change in the malaria-infected erythrocyte is the appearance of knobs on the host cell surface (Langreth *et al.*, 1978). Parasite proteins cluster amongst host proteins in these knobs (Figure 15) and are responsible for altering the adhesive properties of the erythrocyte. Several of the parasite proteins, for example *P. falciparum* erythrocyte membrane protein 3 (PfEMP-3), RESA, mature parasite-infected surface antigen (MESA/PfEMP-2), FEST, and *falciparum* interspersed repeat antigen (FIRA), are distributed throughout the membrane skeleton, while others, such as the knob-associated histidine-rich protein (KAHRP/HRP-I) and PfEMP-1, are found with erythrocyte membrane skeleton proteins associated with the membrane knobs. Studies have shown that MESA/PfEMP-2 binds to protein 4.1 (Waller *et al.*, 2003) and RESA interacts with spectrin (Foley *et al.*, 1991). KAHRP/HRP-I interacts with repeat 4 on α-spectrin (Pei *et al.*, 2005), a phosphorylated region of ankyrin (Magowan *et al.*, 2000), and the cytoplasmic tail of PfEMP-1 (Waller *et al.*, 1999). The cytoplasmic

domain of PfEMP-1 has also been shown to bind to the spectrin-actin junction of the membrane skeleton (Oh *et al.*, 2000). The Maurer's clefts are connected to the erythrocyte membrane skeleton network via *P. falciparum* skeleton binding protein 1 (*Pfsbp*1) (Blisnick *et al.*, 2000).

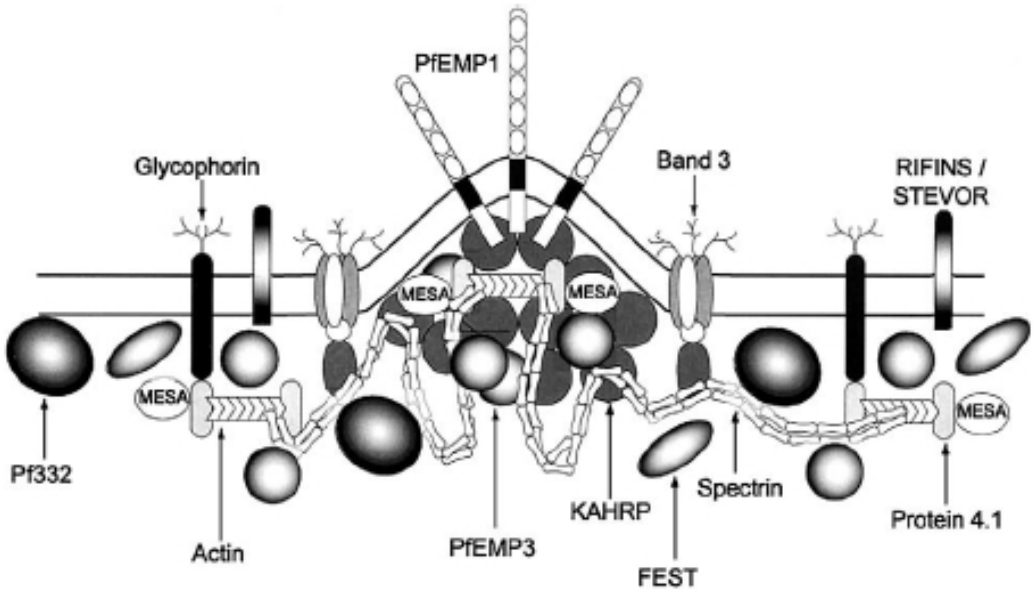


Figure 15: *P. falciparum* and erythrocyte membrane protein interactions.
Diagram showing the *P. falciparum* proteins that interact with the erythrocyte plasma membrane and membrane skeleton proteins. Interactions between parasite proteins and membrane skeleton proteins disrupt the host protein interactions that stabilise the membrane skeleton. These interactions also facilitate the display of malaria proteins on the cell surface in structures called knobs. MESA: mature parasite-infected surface antigen; PfEMP: *Plasmodium falciparum* erythrocyte membrane protein; KAHRP: knob-associated histidine-rich protein; FEST: *falciparum*-exported serine/threonine kinase; RIFINs: products of the repetitive interspersed family; STEVOR: products of the subtelomeric variant open reading frame family (Cooke *et al.*, 2004).

Although the binding of malaria proteins to the membrane skeleton ensures and strengthens the display of malaria proteins on the cell surface, the binding also disrupts the host protein interactions that stabilise the membrane skeleton network. For example, the binding site of KAHRP/HRP-I on ankyrin coincides with the binding domains of band 3 and spectrin and the parasite therefore disrupts the connections in the band 3 complex (Magowan *et al.*, 2000). MESA/PfEMP-2 binds to protein 4.1 at the same domain as glycophorin C and p55. This could

influence the binding capacity of p55 (Waller *et al.*, 2003), which would again destabilise the integrity of the membrane skeleton network. The precise function of MESA/PfEMP-2 in the erythrocyte skeleton is still unknown, but accumulation of MESA/PfEMP-2 in the cytosol of erythrocytes deficient in protein 4.1 leads to intracellular death of the parasite (Magowan *et al.*, 1995).

Parasite growth has also been shown to decrease the mobility of band 3 and glycophorin (Parker *et al.*, 2004), as well as causing band 3 to cluster in the erythrocyte membrane (Winograd and Sherman, 2004). These alterations contribute to the decreased deformability of the erythrocyte membrane and have been observed during the ring stage of growth, before the parasite introduces the knobs into the erythrocyte surface.

1.4.3 Protein interactions mediating parasite escape

Sodium dodecylsulphate (SDS) polyacrylamide gel electrophoresis (PAGE) analysis and microscopic analysis of erythrocyte membranes revealed a decrease of α- and β-spectrin, as well as band 3 in *Plasmodium*-infected erythrocytes (Weidekamm *et al.*, 1973, Konigk and Mirtsch, 1977, Sherman and Jones, 1979, Garcia *et al.*, 1997) indicating that the parasite degrades the erythrocyte skeleton proteins during growth within the erythrocyte. The decrease of membrane skeleton proteins was seen at the end of the erythrocytic life cycle when late trophozoites and schizonts were present in the host cell. The parasite could therefore be weakening the erythrocyte membrane to enable it to escape. Wickham *et al*. (2003) showed that the use of protease inhibitors prevented the release of parasites from their erythrocyte host and proposed that parasite exit involved two stages of membrane lysis. In the first stage, the parasitophorous vacuole membrane is lysed and in the second stage the red cell membrane skeleton proteins are broken down. The first lysis stage is prevented by addition of E64 (Wickham *et al.*, 2003), a cysteine protease inhibitor and the second lysis stage is inhibited by the addition of leupeptin and pepstatin (Lyon and Haynes, 1986), which are cysteine/serine and aspartic protease inhibitors respectively. A

set of related papain-like proteases (papain is a cysteine protease), namely the SERAs and the serine stretch protein homologue (SERPH) have been implicated in the breakdown of the parasitophorous vacuole membrane, because inhibition of SERA prevented the release of schizonts from their erythrocyte host (Li *et al.*, 2002). Several parasite enzymes have been suggested to play role in the breakdown of the membrane skeleton. A 37 kDa acidic protease cleaves β-spectrin and protein 4.1 (Deguercy *et al.*, 1990). Although the primary function of the aspartic protease plasmepsin-II is to digest haemoglobin in the acidic parasite food vacuole, it is also the main enzyme found in a parasitic extract that cleaves spectrin, specifically the SH3 domain of α-spectrin, at a neutral pH. Plasmepsin-II also interacts with actin and protein 4.1 (Le Bonniec *et al.*, 1999) and another aspartic protease, plasmepsin-IV, has been shown to degrade spectrin (Wyatt and Berry, 2002). Another food vacuole enzyme, the cysteine protease falcipain-2, cleaves ankyrin and protein 4.1 within the spectrin-actin-binding domain (Dua *et al.*, 2001) and has been implicated in parasite release (Dhawan *et al.*, 2003).

1.5 Erythrocyte protein abnormalities that protect against malaria infections

Several erythrocyte mutations and polymorphisms, that have arisen due to selective pressure of the malaria parasite (reviewed in Williams, 2006), protect against infection by either preventing invasion or hampering the growth of the parasite within the erythrocyte.

The haemoglobinopathies, which are caused by a decrease in or incorrect folding of globin, were the first erythrocyte abnormalities to be implicated in malaria protection. They include the structural variants of haemoglobin (haemoglobins S, C or E) and the thalassaemias (α- and β-thalassaemia). Mutations in these globin coding genes are often homozygous lethal and tolerable only in the heterozygous state, when they protect against malaria (reviewed in Min-Oo and Gros, 2005). The haemoglobin S mutation has arisen several times in Africa (Cheboulne *et al.*, 1988), indicating that the mutation was caused by the same selective pressure.

The enzymopathies, namely glucose-6-phosphate dehydrogenase (G6PD) and pyruvate kinase deficiency, are caused by mutations in the genes that code for these erythrocyte enzymes. The protective mechanisms of the enzyme deficiencies against malaria have not been well characterised, but it is thought that G6PD-deficient erythrocytes are more susceptible to haemolysis due to the oxidative stress caused by the parasite (reviewed in Weatherall, 2008). Protection against malaria due to pyruvate kinase deficiency has been shown in the mouse model (Min-Oo *et al.*, 2003) and recently also in humans in an *in vitro* study (Durand and Coetzer, 2008, Ayi *et al.*, 2008).

Mutations in the erythrocyte membrane skeleton proteins are responsible for hereditary spherocytosis (HS) and the spectrum of hereditary elliptocytosis (HE) disorders. In HS the attachment of the membrane skeleton to the phospholipid bilayer is disrupted due to mutations in spectrin, or band 3, or in the linker proteins, namely ankyrin and protein 4.2. HS cells with a decreased spectrin content inhibit malaria parasite growth *in vitro*, and the extent of growth inhibition is related to the extent of spectrin deficiency (Schulman *et al.*, 1990).

Hereditary pyropoikilocytosis (HPP) and some cases of HE are caused by spectrin mutations that prevent the self association of spectrin dimers into tetramers (Coetzer *et al.*, 1990). *In vitro* studies demonstrated that these abnormal erythrocytes inhibited invasion by *P. falciparum* (Facer, 1989). Other cases of HE are caused by protein 4.1 or glycophorin C deficiency and these erythrocytes also reduce parasite invasion, as well as retard the growth of parasites *in vitro* (Chishti *et al.*, 1996, Schulman *et al.*, 1990).

Southeast Asian Ovalocytosis (SAO) is the only red cell membrane disorder that has been shown to have developed due to the malaria burden (Foo *et al.*, 1992). It protects against the disease by preventing the development of cerebral malaria (Allen *et al.*, 1999), as well as decreasing the parasitaemia by preventing erythrocyte invasion by the parasite (Cortes *et al.*, 2004). SAO is caused by an in-frame 27 nucleotide deletion in the band 3 gene (Jarolim *et al.*, 1991). The mutant

protein loses its anion exchange function and binds more tightly to ankyrin, resulting in a rigid membrane (Mohandas *et al.*, 1992), that presumably prevents the parasite from invading the red cell *in vitro* (Kidson *et al.*, 1981). A rapid decline of adenosine triphosphate (ATP) levels within SAO erythrocytes has also been shown to prevent invasion by the parasite *in vitro* (Dluzewski *et al.*, 1992).

Finally, the polymorphic red cell receptors displayed on the erythrocyte surface have also been implicated in malaria resistance. *P. falciparum* uses several erythrocyte receptors during invasion and the lack of one receptor, such as Glycophorin A or B (Hadley *et al.*, 1987) or the Gerbich blood group system (Maier *et al.*, 2003), does therefore not provide complete protection against *P. falciparum* invasion. In contrast, *P. vivax* only uses the Duffy antigen as a receptor and the absence of this antigen in some West African populations has resulted in protection against *P. vivax* infections (Miller *et al.*, 1976).

1.6 *P. falciparum* proteins

1.6.1 The *P. falciparum* genome and proteome

The complete *P. falciparum* 3D7 nuclear genome sequence was published in 2002 (Gardner *et al.*, 2002) and contains 14 chromosomes, ranging in size from 0.643 Mb (chromosome 1) to 3.290 Mb (chromosome 14). Two non-nuclear genomes, consisting of the 6 kb mitochondrial genome and the 35 kb apicoplast genome, are also present. The 22,853,764 bp nuclear genome encodes 5,268 predicted genes, implying an average frequency of one gene every 4,338 bp. Of these protein-coding genes, 3,208 (60.9 %) have no assigned function or homology to genes of other organisms with known genome sequence and have been designated as 'hypothetical proteins'. Introns are found in 53.9 % of the genes and their AT-composition rises to 90 % when compared to the overall AT-content of the whole genome, which is 80.6 %. *P. falciparum* has the highest AT-content of all the organisms sequenced to date. The *Homo sapiens* genome, for example, only has a 59 % AT-content (Lander *et al.*, 2001).

The central chromosome regions encode 'housekeeping' genes and the highly variable end regions encode polymorphic antigens and surface-related molecules. The genome of this intracellular parasite encodes fewer enzymes and transporters when compared to the genomes of free-living eukaryotic microbes. Additionally, a large subset of the genes is responsible for immune evasion and host-parasite interactions (Gardner *et al.*, 2002).

The mean length of *P. falciparum* coding genes (excluding introns) is 2.28 kb. This is larger than in other organisms, for example the 1.34 kb mean length of the *H. sapiens* coding genes (Lander *et al.*, 2001), and this can be attributed to the presence of numerous codons encoding low complexity regions found in *Plasmodium* proteins. These low complexity regions may be over several hundred amino acids in length and contain stretches of one or a few amino acids, particularly asparagine and lysine (Pizzi and Frontali, 2001). Low complexity regions generally occur between domains of a protein, but in *P. falciparum* proteins they are also found within the domains. They tend to be located on the outer surface of the folded protein, where they form unstructured soluble stretches that do not interfere with the functions of the domains or the rest of the protein (Aravind *et al.*, 2003). No physical function has thus far been assigned to low complexity regions, but it has been thought that they may be employed as decoys for the human immune system (Anders, 1986).

1.6.2 *P. falciparum* protein identification

Before the 1970s, *P. falciparum* proteins were isolated directly from malaria infected patients. This research however had limited success and thus the establishment of parasites in continuous culture by Trager and Jensen (1976) provided a new source of parasite proteins in the laboratory. Purification of proteins directly from the parasite by conventional methods was difficult due to the presence of erythrocyte protein contaminants and insufficient quantities of parasite protein being isolated. Combinations of other methods were however successfully employed to isolate and identify parasite proteins.

Several parasite surface proteins, for example MSP-1 (Holder and Freeman, 1982), were isolated by immunoprecipitation from parasite culture supernatants with human sera containing parasite-specific antibodies. Hybridoma technology was also applied to produce antibodies that recognise specific parasite antigens, for example CSP (Nardin *et al.*, 1982). The antibodies were subsequently used to purify parasite proteins by affinity chromatography and for immunolocalisation experiments to determine the exact location of the proteins, such as AMA-1 (Peterson *et al.*, 1989), in the parasite.

Standard methods, such as SDS-PAGE, chromatography and centrifugation, were also applied to isolate *P. falciparum* proteins. For example, Kilejian (1979) compared the erythrocyte membrane protein composition of cultured parasitised and non-parasitised erythrocytes by SDS-PAGE and isolated KHARP/HRP-I, and Perkins (1984) used immobilised erythrocyte glycophorins to isolate RESA and glycophorin-binding protein from lysed parasites. Goldberg *et al*. (1990) isolated *P. falciparum* food vacuoles by applying a combination of saponin lysis and Percoll gradient centrifugation and this group subsequently used anion exchange chromatography and gel filtration to isolate plasmepsin-I and -II, falcipain-I and falcilysin from these food vacuoles (Gluzman *et al.*, 1994, Eggleson *et al.*, 1999). Sanders *et al*. (2005) used sucrose density centrifugation to isolate erythrocyte detergent-resistant membranes from cells that were infected with parasites. The amino acid sequence of the N-termini of some isolated proteins, for example falcipain-2 (Shenai *et al.*, 2000) and MSP-6 (Trucco *et al.*, 2001) were determined by Edman sequencing and the amino acid sequence of peptide fragments of other proteins, such as RhopH1 (Kaneko *et al.*, 2001), were determined by tandem mass spectroscopy.

The genes coding for several isolated malaria proteins, for example MSP-2 and the rhoptry associated membrane antigen (RAMA) (Smythe *et al.*, 1988), were identified by screening a *P. falciparum* bacteriophage λ cDNA library with MSP-2 and RAMA-specific antibodies. Some parasite genes, for example MAEBL (Blair *et al.*, 2002) and TRAP (Robson *et al.*, 1988), were isolated by

hybridisation and polymerase chain reaction (PCR) experiments using specific or degenerate oligonucleotide primers designed from the DNA and amino acid sequences of previously identified related malaria proteins or other conserved protein motifs. Primers were also used to screen mRNA, genomic DNA, expressed sequence tag (EST) libraries and cDNA libraries to ultimately obtain the complete DNA sequences coding for malaria proteins.

After the publication of the *P. falciparum* genome sequence in 2002 and the subsequent availability of the annotated genome at PlasmoDB (www.plasmodb.org) (Bahl *et al.*, 2003), researchers probed the genomic and EST databases with conserved regions of known malaria proteins and motifs of non-malaria proteins to identify several malaria protein families, for example the plasmepsins (Banerjee *et al.*, 2002) and the rhomboids (Baker *et al.*, 2006).

While the above mentioned techniques are adequate to isolate and identify parasite proteins, other techniques need to be employed to study protein-protein interactions. Some of these include the yeast two-hybrid (LaCount *et al.*, 2005) and the phage display system (Smith, 1985). Phage display has only been used to study antibody-antigen interactions, for example to identify epitopes that interact with AMA-1 antibodies (Coley *et al.*, 2001). This study will use a novel application of phage display to identify malaria proteins that interact with human erythrocyte spectrin.

1.7 Objectives

During the erythrocytic stage of the *P. falciparum* life cycle the parasite has to enter, grow within and escape from the erythrocyte. All three processes utilise protein-protein interactions between the parasite and the erythrocyte membrane skeleton. Particularly protein-protein interactions with erythrocyte spectrin are of interest because spectrin is the main component of the membrane skeleton. Therefore, the aim of this project is to use phage display to identify *P. falciparum* proteins that interact with human erythrocyte spectrin.

This will be achieved by:

- Constructing a *P. falciparum* phage-display library by inserting *P. falciparum* cDNA into the T7 bacteriophage genome.
- Biopanning the *P. falciparum* phage-display library against immobilised erythrocyte spectrin.
- Sequencing the *P. falciparum* cDNA inserts found in the spectrin-binding bacteriophage and identifying the *P. falciparum* genes that match the cDNA inserts in the PlasmoDB database.
- Cloning the identified *P. falciparum* genes into expression vectors to produce recombinant proteins.
- Studying the protein-protein interaction between spectrin and the *P. falciparum* proteins with blot overlays.
- Performing functional and structural studies to further characterise the *P. falciparum* proteins that interact with spectrin.

Chapter 2: Phage display

2.1 Introduction

2.1.1 Identifying protein-protein interactions

Many processes occurring within the cell depend on protein-protein interactions, and therefore the majority of modern life-science research is concerned with identifying natural interactions and then studying these interactions in finer detail. Protein-protein interactions can be identified by screening a protein against multiple compounds or cell and tissue extracts to identify an interacting protein or ligand.

Protein-protein interactions rely on the contact between certain amino acids of each protein and these interacting amino acid segments can be arranged in two ways. Amino acid residues are either located next to each other as a stretch of amino acids in their primary protein sequence, or they can be located far apart from each other and are brought together by protein folding. An amino acid stretch can also form part of a binding site that has to fold into the correct conformation for a protein-protein interaction to occur. Binding sites are characterised according to their structural and functional characteristics. Structural characteristics are studied with x-ray crystallography and relate to the actual position of the interacting amino acids within the complete protein structure. Functional characteristics are studied by site-directed mutagenesis to determine the type of amino acid residues that are important for the protein-protein interactions. These amino acids can normally not be replaced by other amino acids within the protein sequence. Another type of binding site that can only be discovered *in vitro* has also been characterised. This site does not contain the same amino acid sequence as the native protein but still binds to ligands, since it mimics the native binding amino acid sequence (Sidhu *et al.*, 2003).

Over the years there has been an increase and improvement in the methods applied in elucidating protein-protein interactions. New methods, for example the

yeast two-hybrid system (Fields and Song, 1989), have been used, but they are very labour-intensive and often not suitable for high-throughput evaluation of multiple gene products. Phage display on the other hand, can be used to screen multiple molecules.

2.1.2 Phage display technology

Phage display is a combinatorial technology that was introduced to the scientific community by Smith in 1985, who inserted segments of the *Eco*RI endonuclease gene between the N-terminal and C-terminal segments of the pIII capsid gene of the filamentous bacteriophage f1 (Smith, 1985). Smith was able to isolate bacteriophage expressing the endonuclease protein fused to the bacteriophage capsid protein by affinity selection with an antibody specific to the endonuclease (Smith, 1985). Since then the technique has for example been used to identify ligands of peptide receptors (Cesareni, 1992), enzyme substrates (Maenaka *et al.*, 1996), antibodies against specific antigens (Huse *et al.*, 1992) and to improve or modify the affinity of proteins for their binding partners (Lowman *et al.*, 1991).

The novelty of the technique relies on the expression of foreign peptides or proteins fused to one of the coat proteins of a bacteriophage, such as f1 (Smith, 1985, de la Cruz *et al.*, 1988), fd (Parmley and Smith, 1988), M13 (Makowski, 1993, Devlin *et al.*, 1990), MS2 (Mastico *et al.*, 1993), P4 (Lindqvist and Naderi, 1995), λ (Sternberg and Hoess, 1995), T4 (Jiang *et al.*, 1997) or T7 (Houshmand *et al.*, 1999). The protein fusions are created by inserting a DNA fragment into the bacteriophage genome, adjacent to, or in, the DNA sequence of one of the capsid proteins of the bacteriophage. Subsequent infection of a bacterial host with the modified phage results in the expression of the fusion peptide or protein on the capsid surface. If random sequences are inserted, it is possible to obtain a large library with 10^6-10^9 different peptide or protein sequences, each displayed on a different phage (Hoess, 1993, Parmley and Smith, 1988). The phenotype of the displayed peptide is physically linked to the DNA sequence inserted in the phage

genome and because the bacteriophage is infective it can be propagated either individually or as a library.

To select for phage displaying peptide sequences, a procedure termed biopanning is used (Parmley and Smith, 1988). This process of affinity selection involves incubating the phage library with a target peptide or protein that has been immobilised on a solid support (Figure 16). Phage displaying the appropriate peptide or protein domain bind to the immobilised target protein, while non-binding phage are removed by washing. Bound phage are eluted from the target protein and amplified in bacteria specific for the virus. Amplified viruses can be used for further rounds of biopanning, usually three to four times, to obtain phage with the strongest binding ability. The DNA of the phage from the final round of selection can be sequenced and inserted into an expression system to further characterise the peptide.

The advantages of the phage display system are: (1) the creation of a library that contains millions of peptides, instead of having to clone each peptide individually; (2) the selection of high-affinity phage from a library, even if they are present at low levels; (3) the ability to amplify bacteriophage in bacteria, resulting in the enrichment of rare binding phage; and (4) the linkage of the displayed peptide with its encoding DNA sequence, which allows one to identify and further process binding sequences (Hoess, 1993, Sidhu *et al.*, 2003).

Many phage display libraries have been created over the last twenty years and can be classified into two categories based on the type of DNA inserted into the library. The first is the random peptide library (RPL). RPLs are created by inserting synthetic random degenerate oligonucleotides into the phage genome (Devlin *et al.*, 1990). These types of libraries display peptides of a predetermined length on the bacteriophage surface and because the libraries are universal in nature, they can be used to select peptides that react with several different target proteins. The main disadvantage of RPLs is that one can isolate peptide mimics which have binding abilities similar to that of the native protein, but do not

contain the native amino acid sequence. Mimics can however be very useful in drug discovery.

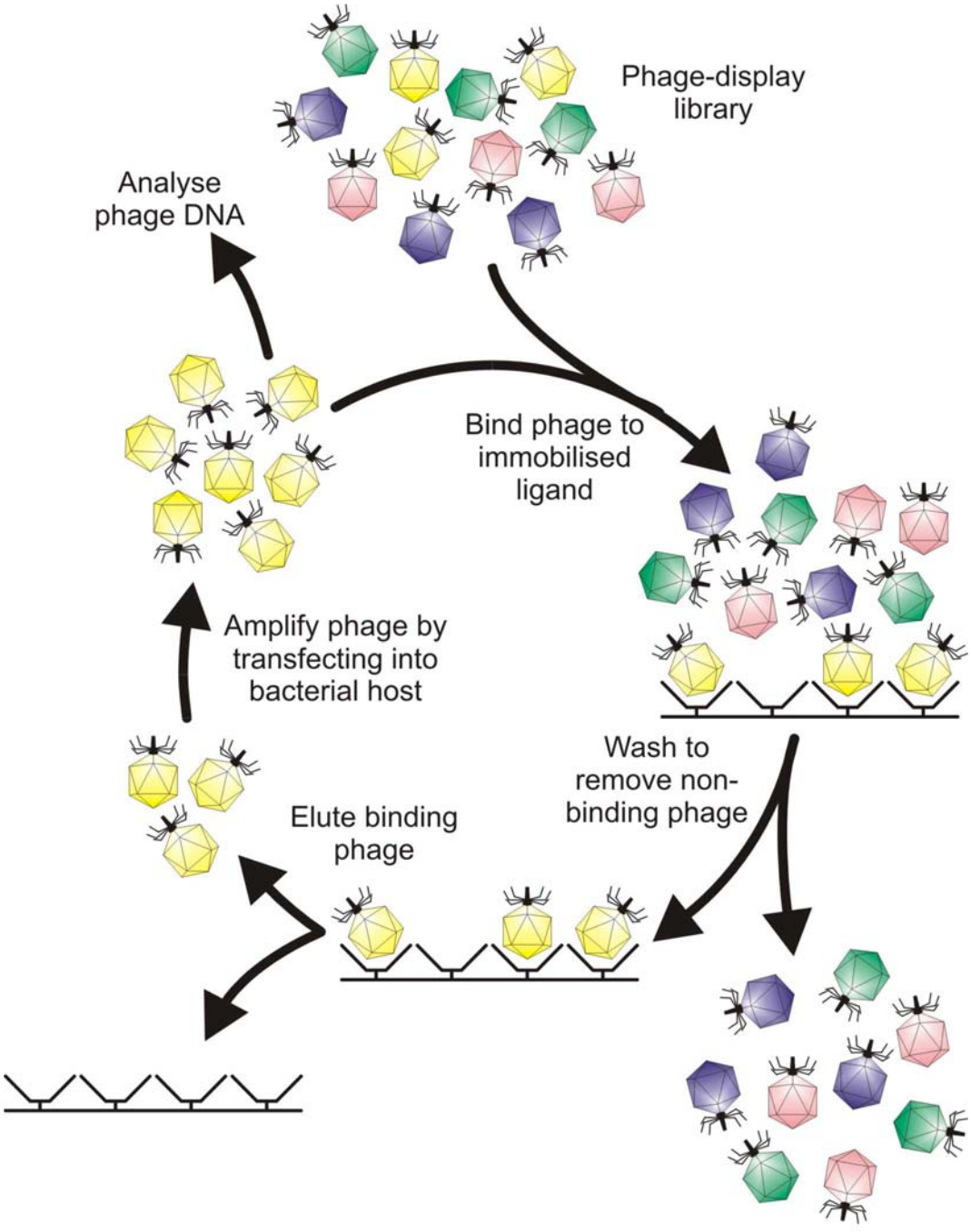


Figure 16: Biopanning of a phage display library.
Diagram showing the biopanning process. A T7 phage display library is added to the protein of interest that has been immobilised via the streptavidin-biotin interaction on a solid surface. Non-binding phage are removed by washing. Binding phage are eluted and used for amplification in the bacterial host. After amplification the phage are isolated and used for another round of biopanning.

The second type of library is the native peptide library (NPL), which is constructed from randomly generated fragments of a specific genome (Jacobsson and Frykberg, 1998) or reverse transcribed mRNA (Jespers *et al.*, 1995), thereby making the insert sizes in NPLs less uniform than those in a RPL. NPLs are gene or genome specific and therefore a new library needs to be created every time a new organism is studied. The major advantages of NPLs are that native proteins that are highly specific to an interaction can be isolated from a library, even if the library is relatively small.

Successful selection of a peptide-binding bacteriophage from a phage display library depends on the affinity of the displayed peptide for the immobilised target protein. Peptide binding in phage display can be influenced by the number of peptides displayed on the bacteriophage surface (i.e. the copy number). A large copy number (multivalent or polyvalent display) is used to isolate low and high affinity binders from a library, whereas a single bacteriophage displayed peptide (monovalent display) is used to isolate only high affinity binders (Hoess, 2001). Multivalent display is therefore more appropriate for an NPL, because this type of library is likely to contain only a few high-affinity binders and a large number of low affinity peptides that interact with the target. RPLs on the other hand are often used to detect high affinity peptides and therefore monovalent display would be more applicable.

2.1.3 Phage display technology in malaria research

Phage display technology has been applied to malaria research, but only in the context of studying interactions between antigens and antibodies. These antigen-antibody studies were performed in three ways.

The first type of experimental approach used phage display to express fragments of specific *P. falciparum* proteins on the surface of the bacteriophage. The repeat region of CSP was expressed on the surface of f1 phage either linked to gene III or gene VIII and the recombinant phage were used to test the immunogenicity of

the displayed peptide fragments in rabbits (de la Cruz *et al.*, 1988, Greenwood *et al.*, 1991). Heal and co-workers used the same approach to test the immunogenicity of the *P. falciparum* liver stage antigen-1 (LSA-1) in mice (Heal *et al.*, 1999).

The second type of experimental approach used RPLs or fragments of a *P. falciparum* protein displayed on the bacteriophage surface to identify epitopes that interacted with specific *P. falciparum* antibodies. Adda *et al*. (1999) panned a random 15-mer and 17-mer peptide library against a monoclonal antibody which was raised against the C-terminal recombinant fragment of *P. falciparum* RESA. The group identified mimotopes which are similar to the peptide motif found in the 5' and 3' repeat sequences of RESA. Coley and co-workers displayed AMA-1 fragments on the surface of M13 and panned them against two antibodies known to interact with AMA-1. The binding fragments were analysed and mapped to specific regions within the AMA-1 protein sequence. Further studies using mutant AMA-1 fragments and a random 15-mer fd library were used to identify the exact epitope that interacted with the AMA-1 antibody (Coley *et al.*, 2001). Casey *et al*. used a 20-mer RPL to find peptide mimics that interact with a known AMA-1 antibody. The peptide fragments were subsequently used to isolate other AMA-1 antibodies from human sera and to create new AMA-1 antibodies in rabbits that inhibited parasite invasion (Casey *et al.*, 2004).

The third type of experimental approach used antibody display libraries to find antibody fragments that interact with specific *P. falciparum* proteins. Lundquist and co-workers created an antibody phage display library by inserting cDNA from human peripheral blood leukocytes from immune individuals into the M13 genome. The library was biopanned against *P. falciparum* MSP-3 to isolate three antibody fragments that interacted with the parasite protein (Lundquist *et al.*, 2006). Similar approaches were used to identify human antibodies interacting with Pfs48/45 (Roeffen *et al.*, 2001) and MSP-1 (Sowa *et al.*, 2001).

RPL phage display was also used to study and identify peptides that interact with host cell surfaces. Eda *et al*. (2004) panned a 7-mer disulfide bond-constrained RPL against *P*. *falciparum*-infected human erythrocytes. The isolated peptide was coupled to a compound with moderate haemolytic activity and added to parasite-infected erythrocytes. The peptide bound to the surface of the infected erythrocytes, which were subsequently lysed by the haemolytic compound, thereby inhibiting parasite growth. Ghosh and co-workers injected or fed a 12-mer random peptide library to mosquitoes and dissected them to obtain bacteriophage that had bound to the salivary glands and the luminal side of the midgut. The group isolated one peptide that specifically interacted with both these surfaces and inhibited both sporozoite invasion of the salivary glands as well as ookinete invasion of the midgut (Ghosh *et al.*, 2001).

This study used phage display to express random *P*. *falciparum* protein fragments on the surface of bacteriophage T7 by inserting *P*. *falciparum* cDNA into the bacteriophage genome using the T7Select® Phage Display System. This approach is novel because it created an NPL that is based on the *P*. *falciparum* genome, which was utilised to investigate the interaction of human erythrocyte spectrin with *P*. *falciparum* peptides. The library can also be used to identify malaria proteins involved in any protein-protein interactions.

2.1.4 The T7Select® Phage Display System

Filamentous bacteriophage are non-lytic viruses and assembly of the infective viral particle involves the secretion of the capsid protein through the bacterial plasma membrane (Smith and Petrenko, 1997). This secretion limits the type of proteins that can be expressed on the bacteriophage capsid because a signal sequence is required to direct the protein to the plasma membrane of the host bacteria. The use of lytic bacteriophage such as P4, T4, T7 and λ that assemble the viral particle within the cytoplasm of bacteria and subsequently lyse their host cell, has vastly increased the type and number of proteins expressed on phage capsids. The T7Select® System created by Novagen (Novagen, Inc., Madison, USA)

utilises the double-stranded DNA bacteriophage, T7, to display peptides on the bacteriophage surface (Rosenberg *et al.*, 1996).

The 40 kb T7 bacteriophage genome codes for approximately 50 genes. These include the early genes, the genes involved in DNA metabolism and the genes responsible for the assembly of the bacteriophage (genes 9-18) (van Regenmortel *et al.*, 2000). The T7 capsid consists of 415 copies of the T7 capsid protein, which are arranged as 60 hexamers on the faces and 11 pentamers at the vertices of the shell to make up an icosahedron. The last vertex contains the head-tail connector and a short conical tail, that contains six tail fibres, which completes the rest of the phage particle (Figure 17) (Swanson, 1999). Phage assembly takes place within the *Escherichia coli* cytoplasm by the formation of a procapsid that contains scaffolding protein, capsid protein, the head-tail connector, and an internal protein structure. After the bacteriophage DNA has entered the procapsid, the scaffolding proteins are released and conformational changes lead to the formation of the mature bacteriophage particle to which the tail and tail fibres attach at the head-tail connector (Cerritelli and Studier, 1996).

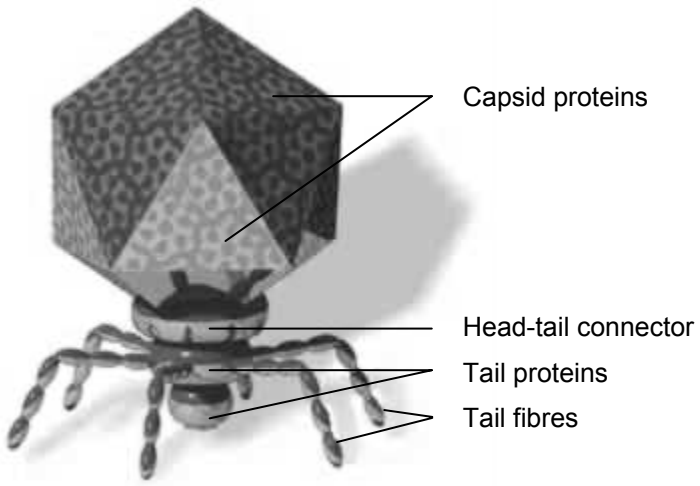


Figure 17: The T7 bacteriophage.
Schematic representation of the structure of the T7 bacteriophage. The bacteriophage capsid contains 415 copies of the capsid protein and is connected to the tail proteins by the head-tail connector. Six tail fibres extend from the bacteriophage tail (Swanson, 1999).

The capsid of wildtype T7 is made up of either one form or a mixture of two forms of the capsid protein (Rosenberg *et al.*, 1996), protein 10A consisting of 344 amino acids and protein 10B consisting of 397 amino acids. The 10B protein is produced because of a translational frameshift at codon 341 of the capsid gene and represents ~10 % of the capsid proteins (Condron *et al.*, 1991). The T7Select® System utilises the 10B gene to express peptides on the surface of T7. The T7Select10-3 vector was created by removing the frameshift at codon 341 and introducing a multiple cloning cassette at codon 348 of the 10B gene (Rosenberg *et al.*, 1996) (Figure 18). This modification leads to the production of only the 10B protein with a fused peptide, and because the size and sequence of the attached peptide could have a negative effect on viral assembly, the virus particles are amplified in *E. coli* cells containing an additional plasmid expressing multiple copies of the 10A protein. The viral particles created in this system thus only express 5-15 fused capsid proteins with peptides of up to 1200 amino acids on their surface making it an ideal system to display larger *P. falciparum* peptide fragments with a medium copy number on the bacteriophage surface.

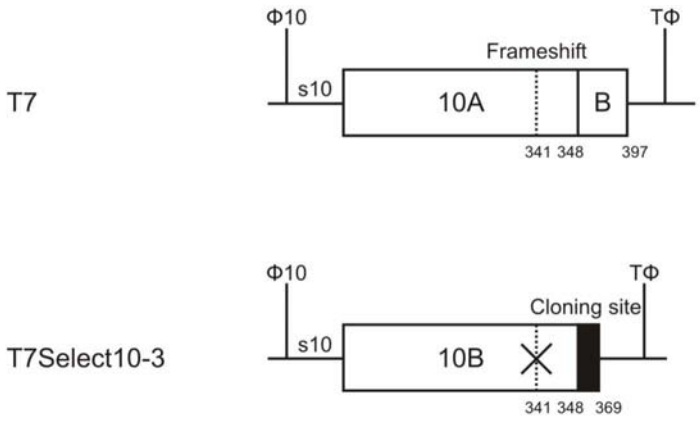


Figure 18: The T7Select10-3 vector.
Diagram showing the creation of the T7Select10-3 vector. The frameshift (codon 341) responsible for the translation of the 10B capsid protein was removed and the nucleotide sequence (codon 348-397) coding for the additional amino acids of the 10B capsid protein was replaced by a multiple cloning site (codon 348-369) to enable the insertion of foreign cDNA (adapted from Rosenberg *et al.*, 1996).

T7 has several propagation properties that make it an ideal phage display system. Bacteriophage plaques can be produced within 3 hours at 37 ºC and cultures

generally lyse 1-2 hours after infection, which decreases the turnaround time between several rounds of panning. The phage particles are extremely robust and can withstand the harsh experimental conditions employed in the panning and eluting procedure. The viral DNA is easy to isolate and therefore restriction and DNA sequence analysis of clones are simple to perform (Rosenberg *et al.*, 1996).

2.2 Methods

2.2.1 Materials

P. falciparum strain FCR-3 was a gift from the Department of Pharmacy and Pharmacology (University of the Witwatersrand, Johannesburg, South Africa). All chemical reagents were of analytical or molecular biology grade. Milli-Q water prepared with the Milli-Q^{TM} Water System (Millipore Corporation, Bedford, USA) was used for all experimental procedures.

2.2.2 Spectrin isolation and biotinylation

Isolation of human spectrin from erythrocytes

Erythrocyte ghosts were obtained from whole human blood by rupturing the erythrocytes by hypotonic lysis (Dodge *et al.*, 1963). Twelve 6ml ACD tubes (BD Vacutainer Systems, Plymouth, UK) containing human blood donated by volunteers, were centrifuged at 800 *g* for 10 minutes at 4 ºC using a Jouan BR 3.11 centrifuge (Jouan Inc., Winchester, USA). The plasma and buffy coat were removed and the erythrocytes washed three times in cold 0.9 % sodium chloride. Erythrocytes were lysed with 30 volumes of cold erythrocyte lysis buffer (3 mM sodium phosphate buffer, pH 8, 0.1 mM Na_2EDTA, 0.1 mM phenylmethanesulphonyl fluoride (PMSF; see Appendix A1 for stock solution; Roche Diagnostics Gmbh, Mannheim, Germany) and centrifuged at 25,000 *g* for 15 minutes at 4 ºC using a Beckman® J2-21 centrifuge (Beckman Coulter, Inc., Fullerton, USA). The supernatant and the residual white cell pellet were removed and the ghosts washed a further three times with erythrocyte lysis buffer. Spectrin

was extracted from the erythrocyte ghosts by low ionic strength extraction (Coetzer and Palek, 1986). This was achieved by adding 30 volumes of fresh spectrin extraction buffer (0.1 mM sodium phosphate, pH 8, 0.1 mM Na_2EDTA, and 0.1 mM 1,4-dithiothreitol (DTT; Roche Diagnostics Gmbh, Mannheim, Germany) and centrifuging at 30,000 *g* for 30 minutes at 4 ºC. DTT prevents disulphide bond formation and thereby inhibits the aggregation of spectrin. The supernatant was aspirated until there was a 1:½ ratio of erythrocyte ghosts to buffer remaining in the tube. The suspension was incubated overnight at 4 ºC. Pefabloc SC (an irreversible protease inhibitor) (Roche Diagnostics Gmbh, Mannheim, Germany) was added to a concentration of 0.1 mM. The sample was centrifuged at 150,000 *g* for 30 minutes at 4 ºC, using a Beckman® L8-70M ultracentrifuge (Beckman Coulter, Inc., Fullerton, USA) and the supernatant containing the crude spectrin extract collected and pooled.

Size exclusion chromatography and concentration of spectrin

Spectrin tetramers and dimers were removed from high molecular weight complexes and band 4.1 by size exclusion chromatography (Zail and Coetzer, 1984). The crude spectrin sample was incubated at 37 ºC for 30 minutes to dimerise the spectrin and loaded onto a 3 cm x 100 cm Sepharose® (CL) 4B column (kept at 4 ºC) (Pharmacia Fine Chemicals AB, Uppsala, Sweden) equilibrated with spectrin column buffer (10 mM sodium phosphate buffer, pH 7.6, 130 mM KCl, 20 mM NaCl, 1 mM Na_2EDTA, 1 mM DTT) at a flow rate of 20 ml/hr. The elution profile of the collected fractions (2.5ml per tube) was plotted using a LKB Bromma 2238 Uvicord S II UV reader (Pharmacia Fine Chemicals AB, Uppsala, Sweden) (settings: 280 nm, absorbance 0.5) and a LKB Bromma 2210 Recorder (settings: 0.5mm/min, 500mV). Fractions containing only the spectrin dimers were pooled and the concentration of a 10 µl sample determined with the Coomassie Plus™ Protein Assay Reagent (Pierce Biotechnology Incorporated, Rockford, USA). Spectrin tetramers were not used for the biotinylation procedure because they eluted just after the high molecular weight complexes and contamination with other erythrocyte membrane proteins

was possible. The spectrin dimers were concentrated to 2.5 ml using an Amicon Stirred Ultrafiltration Cell kept on ice (Model 8050; Millipore Corporation, Bedford, USA) that contained an Ultracel Amicon YM 10 Disc Membrane (molecular weight exclusion limit of 100,000 Da; Millipore Corporation, Bedford, USA). The concentration procedure was performed with nitrogen (African Oxygen Limited, Johannesburg, South Africa) set at 0.7 atm. The sample was further concentrated to ~5 mg/ml with a 0.5-3 ml Slide-A-Lyzer® Dialysis cassette (Pierce Biotechnology Incorporated, Rockford, USA) and Polyethylene glycol 2000 (PEG 2000) (Sigma-Aldrich Corporation, St. Louis, USA).

Biotinylation of purified spectrin

The concentrated spectrin sample was dialysed against phosphate buffered saline (PBS; 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.4 mM KH_2PO_4, pH 7.2) for 2-3 hours at 4 ºC using a 0.5-3 ml Slide-A-Lyzer® Dialysis cassette. The spectrin was biotinylated for 2-4 hrs at room temperature with continuous stirring in a reaction that contained a molar ratio of 1:10 of D-biotin-N-hydroxysuccinimide ester (Roche Diagnostics Gmbh, Mannheim, Germany) to spectrin dimers (MW 460 kDa) (Figure 19).

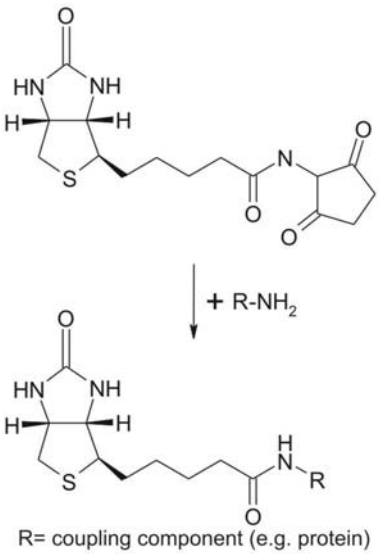


Figure 19: The biotinylation reaction.
(adapted from Roche Applied Science, 1999)

The sample was dialysed overnight at 4 ºC against PBS to remove free biotin and centrifuged at 8,500 *g* for 10 minutes at room temperature using a Sorvall® RMC-14 Refrigerated Microcentrifuge (Sorvall Products, L.P., Newtown, USA). The supernatant was diluted to a final concentration of 1 mg/ml with PBS. One hundred microlitre aliquots were stored at -20 ºC.

Analysis of spectrin

The biotinylated spectrin was electrophoresed on a Laemmli SDS-polyacrylamide gel (Laemmli, 1970) and a gradient Fairbanks SDS-polyacrylamide gel (Coetzer *et al.*, 1988, Fairbanks *et al.*, 1971) to analyse its purity.

Ten micrograms spectrin and 45 µg erythrocyte membrane proteins were prepared by adding 5 x Suspension Solution (50 mM Tris-HCl, 5 mM Na_2EDTA, 5 % SDS (w/v), 25 % sucrose (w/v), pH 8), 40 x sucrose/dye (2.5 % sucrose (w/v), 0.5 % bromophenol blue (w/v)) and 2 % β-mercaptoethanol (v/v) and heating the sample for 2 minutes in boiling water. A 14 cm x 16 cm x 1.5 mm Laemmli SDS-polyacrylamide gel containing a stacking gel composed of 4 % polyacrylamide (29.2 % acrylamide/0.8 % bisacrylamide (w/v) stock solution) in lower gel buffer (375 mM Tris-HCl, pH 8.8) and 0.1 % SDS (w/v) and a resolving gel composed of 12 % polyacrylamide in upper gel buffer (125 mM Tris-HCl, pH 6.8) and 0.1 % SDS (w/v), was cast and the spectrin electrophoresed against the erythrocyte membrane proteins at 75 V for 17 hours with Laemmli running buffer (25 mM Tris, 192 mM glycine, 0.1 % SDS (w/v). The protein bands were stained with Coomassie Blue stain (0.05 % Coomassie Blue R-250 (w/v) (BDH Laboratory Supplies, Poole, UK), 25 % isopropanol (v/v), 10 % glacial acetic acid (v/v)).

A 14 cm x 16 cm x 1.5 mm Fairbanks SDS-polyacrylamide gel was set up with an exponential polyacrylamide gradient of 3.5-17.5 % (40 % acrylamide/1.5 % bisacrylamide (w/v) stock solution) in Tris acetate buffer (40 mM Tris, 20 mM sodium acetate, 1 mM Na_2EDTA, pH 7.4) and 0.2 % SDS (w/v). Protein samples were solubilised the same way as for the Laemmli SDS-polyacrylamide gel. Five

micrograms spectrin was electrophoresed against 20 µg erythrocyte membrane proteins at 45 V for 17 hours using Tris acetate buffer containing 0.1 % SDS (w/v). The protein bands were stained with Coomassie Blue stain.

To test successful biotinylation of spectrin, one well from a Streptawell plate (transparent, 12x8-well strips; Roche Diagnostics Gmbh, Mannheim, Germany) was coated with 0.5 µg of biotinylated spectrin in 200 µl Tris buffered saline (TBS; 10 mM Tris-HCl, 150 mM NaCl, pH 8), by allowing the protein to bind for 1 hour at 4 ºC. A second control well was incubated with 200 µl TBS. The wells were washed 3 times with 300 µl TBS-T (TBS with 0.12 % Tween®-20 (v/v) (Calbiochem®, San Diego, USA), blocked with 200 µl 3 % Bovine Serum Albumin Fraction V (w/v) (Roche Diagnostics Gmbh, Mannheim, Germany) in TBS and analysed with a 1:100 dilution of a rabbit anti-α-spectrin and rabbit anti-β-spectrin primary polyclonal antibody mix (prepared at St. Elizabeth's Medical Center, Boston, USA), followed by a 1:1000 dilution of goat anti-rabbit IgG peroxidase (Roche Diagnostics Gmbh, Mannheim, Germany). The secondary antibody was detected with 0.5 mg/ml 4-chloro-1-napthol (Sigma-Aldrich Corporation, St. Louis, USA) and 0.1 % hydrogen peroxide solution (Sigma-Aldrich Corporation, St. Louis, USA).

2.2.3 *P. falciparum* culturing

P. falciparum cultures (~5 % parasitaemia and ~5 % haematocrit) were maintained in 175 cm^3 culture flasks (Nunc™, Roskilde, Denmark) with 30 ml RPMI-1640 medium containing 25 mM HEPES (Invitrogen Ltd, Renfrewshire, UK) supplemented with 20 mM glucose (Saarchem (Pty) Ltd., Wadeville, South Africa), 44 mg hypoxanthine (Sigma-Aldrich Corporation, St. Louis, USA), 0.05 mg/ml gentamycin (Sigma-Aldrich Corporation, St. Louis, USA), 0.2 % sodium bicarbonate (w/v) (adapted from Trager and Jensen, 1976). Before use, the filter sterilised medium (Sterivex™-GS 0.2 µm Filter unit, Millipore Corporation, Bedford, USA) was gassed with 100 % carbon dioxide (African Oxygen Ltd., Johannesburg, South Africa) until the pH indicator changed to a orange-yellow

colour, and 10 % heat inactivated AB plasma (v/v) was added to the medium. The cultures were examined daily by staining a blood smear with the Rapid Haematology Stain (Diagnostic Media Products, NHLS, Sandringham, South Africa). The culture medium was changed daily and the parasitaemia and haematocrit levels were controlled by removing a portion of parasitised cells from the flask and adding fresh human erythrocytes (donated by volunteers) that had been washed with sterile PBS. Cultures were grown in a 37 ºC Nuaire™ IR Autoflow incubator (Nuaire™, Plymouth, USA).

2.2.4 *P. falciparum* mRNA isolation

Total RNA isolation

Total RNA was obtained from the parasites using the guanidinium isothiocyanate isolation method described by Chomczynski and Sacchi (1987). All RNA isolation experiments were performed in an RNase-free environment (as described in Ausubel *et al.*, 1994). Parasite suspensions from four 175 cm^3 culture flasks (mixed cultures with a 10-20 % parasitaemia) were centrifuged for 5 minutes at 900 *g* using a Jouan BR 3.11 centrifuge. Thirty millilitres 0.05 % saponin (w/v, dissolved in PBS; USB Corporation, Cleveland, USA) was added to each tube to lyse the parasite-infected erythrocytes. The samples were incubated for 5 minutes at room temperature, followed by centrifugation at 900 *g* for 10 minutes. Parasite pellets were washed three times in PBS. The 100-200 µl parasite pellets were lysed with 5 ml 4 M guanidinium isothiocyanate (Sigma-Aldrich Corporation, St. Louis, USA) solution containing 25 mM sodium citrate, pH 7, 0.5 % lauryl sarcosine (w/v) (Sigma-Aldrich Corporation, St. Louis, USA), and 0.1 M β-mercaptoethanol (Merck KGaA, Darmstadt, Germany). The total RNA was removed from the cellular debris by adding 1/10 volume 2 M sodium acetate, pH 4, an equal volume phenol pH 4.2 (MP Biomedicals, LLC, Aurora, USA), and 1/5 volume chloroform and centrifuging the mixed sample at 1,500 *g* for 15 minutes at 4 ºC. An equal volume of chloroform was added to the aqueous phase and the centrifugation step repeated. One millilitre aliquots of the supernatant were pipetted into RNaseZap® (Ambion Inc., Austin, USA) treated 2.0 ml Eppendorf

tubes (Eppendorf AG, Hamburg, Germany) and an equal volume of 100 % cold isopropanol added. The samples were incubated at -20 ºC for 1-2 hours and centrifuged at 16,000 *g* for 10 minutes at 4 ºC in a Sorvall® RMC-14 Refrigerated Microcentrifuge. The pellets were air dried for 1 hour at room temperature and resuspended in a total volume of 120 µl nuclease-free water (Promega Corporation, Madison, USA). A total of 10 µl of the sample was kept aside for spectrophotometric analysis at 260 and 280 nm (as described in Ausubel *et al.*, 1994) and visualisation on a 1 % agarose gel (containing 0.5 µg/ml ethidium bromide (Sigma-Aldrich Corporation, St. Louis, USA)) in Tris EDTA acetic acid buffer (TEA buffer; 40 mM Tris, 1 mM Na_2EDTA, pH 8, 0.1 % glacial acetic acid (v/v)).

mRNA isolation

mRNA was separated from the total RNA using the Dynabeads® mRNA Direct™ Kit (Dynal A.S, Oslo, Norway), which uses magnetic Oligo $(dT)_{25}$ beads that bind mRNA poly-A tails to extract mRNA directly from cells. The protocol was therefore modified to extract mRNA from total RNA. The 110 µl total RNA sample was made up to 1 ml with Binding/Lysis Buffer and added to the magnetic particles (250 µl Dynabeads® Oligo $(dT)_{25}$ bead suspension from the stock, washed with 250 µl Binding/Lysis Buffer). The mRNA was allowed to bind to the beads for 5 minutes at room temperature while continuously inverting the tube by hand. The Binding/Lysis Buffer was removed by placing the tube on the Magnetic Particle Separator (Roche Diagnostics Gmbh, Mannheim, Germany) and aspirating the supernatant. Beads were washed twice by continuously inverting the sample for 5 minutes at room temperature using 1 ml Washing Buffer A and once with 1 ml Washing Buffer B. mRNA was eluted from the beads in 25 µl 10 mM Tris-HCl, pH 7.5 by heating the sample to 65 ºC for 2 minutes. The concentration and purity of the mRNA was determined by spectrophotometric analysis at 260 and 280 nm and visualisation on a 1 % agarose gel in TEA buffer.

2.2.5 *P. falciparum* phage display library creation

P. falciparum cDNA synthesis

The OrientExpress™ cDNA Kit (Novagen, Inc., Madison, USA) was used to create the *P. falciparum* phage display library according to the protocol of the manufacturer. *P. falciparum* mRNA was reverse transcribed in the presence of 5-methyl dCTP to protect internal *Eco*RI and *Hind*III restriction sites. The *P. falciparum* mRNA strand was digested by RNase H and double stranded DNA synthesised using the reverse transcribed DNA strand as the template. A single directional linker was added to both sides of the dsDNA (Figure 20). The linker has an internal *Eco*RI site and a *Hind*III site is created at the 3' end of the construct by a combination of the poly-A tail and part of the linker. Digestion with *Eco*RI and *Hind*III thus facilitated directional cloning of the dsDNA into the T7Select®10-3b Vector (Figure 20 and vector map in Appendix A2). The protocol in the OrientExpress™ cDNA Manual (Novagen, 1999a) was modified by precipitating the DNA with ethanol instead of isopropanol. The library creation procedure was monitored by incorporating 5 µCi [α-^{32}P]dATP (specific activity = 3000 Ci/mmol; Amersham Biosciences, Ltd., Buckinghamshire, UK) into the first and second strand synthesis reaction and electrophoresing the *P. falciparum* dsDNA against size markers (Appendix A1) on a 8 % denaturing polyacrylamide gel.

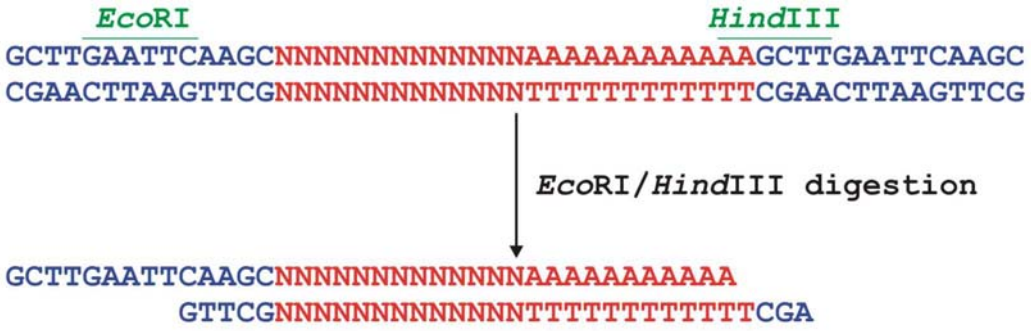


Figure 20: Double digestion of the linker and *P. falciparum* insert.

Diagram showing the creation of the *Eco*RI/*Hind*III *P. falciparum* insert after addition and double digestion of a single linker. Red nucleotides = *P. falciparum* DNA; blue nucleotides = directional linker; green lines = *Eco*RI and *Hind*III restriction sites.

The *P. falciparum* library was created using 4 µg *P. falciparum* mRNA. The mRNA and 2 µg of a two-base anchored primer cocktail (P129, P130, P131, P132) (Integrated DNA Technologies, Inc., Coralville, USA) (reaction volume = 20 µl) (Figure 21) were heated at 70 ºC for 10 minutes to reduce RNA secondary structure. The sample was cooled on ice and the mRNA subjected to first strand synthesis with 800 units MMLV Reverse Transcriptase in the presence of 1 x First Strand Buffer, 10 mM DTT and 0.5 x Methylated dNTP Mix for 1 hour at 37 ºC (reaction volume = 50 µl). The MMLV enzyme was subsequently inactivated by heating to 70 ºC for 10 minutes. Second strand synthesis entailed digesting the mRNA template with 1.6 units RNase H and completing the DNA synthesis with 50 units DNA polymerase I in the presence of 1 x Second Strand Buffer, 2.4 mM DTT and 0.08 x Methylated dNTP Mix for 90 minutes at 15 ºC (reaction volume = 250 µl).

P129
5' TTTTTTTTTTTTVA 3'
P130
5' TTTTTTTTTTTTVg 3'
P131
5' TTTTTTTTTTTTVC 3'
P132
5' TTTTTTTTTTTTVT 3'
V= A,g or C

Figure 21: Anchored 14-mer primers for *P. falciparum* cDNA synthesis.

The dsDNA was purified by phenol (pH 7.9)/chloroform/ isoamyl alcohol (25:24:1) extraction. The dsDNA was precipitated for 20 minutes at -70 ºC with 10 µg Glycogen, an equal volume of 4 M Ammonium Acetate, pH 4.8 and two and a half volumes 100 % cold ethanol, followed by centrifugation at 16,000 *g* for 30 minutes at 4 ºC. The pellet was washed with 70 % ethanol and air dried for 20 minutes before being resuspended in 20 µl TE Buffer. DNA ends were blunted with 1.5 units T4 DNA Polymerase, 1 x Flush Buffer, 5 mM DTT, and 3 mM dNTP Mix for 20 minutes at 11 ºC (reaction volume = 30 µl), followed by phenol-chloroform extraction and ethanol precipitation as above. The pellet was resuspended in 10 µl TE Buffer. One hundred picomoles Directional *Eco*RI/*Hind*III Linkers and the DNA were phosphorylated in the presence of 5 units T4 Polynucleotide Kinase, 1 x Ligation Buffer, 0.1 mM ATP, and 10 mM

DTT for 5 minutes at 37 ºC and then ligated to each other in the presence of 6-8 Weiss units T4 DNA Ligase for 20 hours at 16 ºC (reaction volume = 20 µl). The ligase was heat inactivated at 70 ºC for 10 minutes, before the DNA was subjected to restriction digestion with 100 units *Hind*III in 1 x *Hind*III Buffer for 2 hours at 37 ºC (reaction volume = 100 µl), followed by restriction digestion with 100 units *Eco*RI in 1 x *Eco*RI Adjustment Buffer for 4 hours at 37 ºC (reaction volume = 115 µl). The DNA was purified with phenol-chloroform as above and fragments smaller than 300 bp removed by size exclusion chromatography through 1 ml Gel Filtration Resin (Sepharose® 4B) equilibrated with Column Buffer. The 250 µl eluate was subjected to ethanol precipitation as above with 1 µg Glycogen but no addition of 4 M Ammonium Acetate and resuspended in 20 µl TE Buffer. The concentration and purity of vector-ready DNA was determined by spectrophotometric analysis at 260 and 280 nm.

Ligation and viral packaging

P. falciparum DNA was inserted into the T7 vector using the T7Select® System (Novagen, Inc., Madison, USA) according to the instructions of the manufacturer. A molar ratio of 1500 bp DNA insert (estimation of the average length of the inserts after DNA synthesis) to vector of 1:1 to 3:1 should be used for the ligation reaction (Novagen, 2000b). Therefore, ~0.05 pmol *P. falciparum* DNA was ligated to 0.02 pmol T7Select® Vector Arms in the presence of 0.4-0.6 Weiss units T4 DNA Ligase, 1 mM ATP, 1 x Ligase Buffer, and 10 mM DTT for 16 hours at 16 ºC (reaction volume = 5 µl). The 5 µl ligation reaction was packaged in 12.5 µl T7Select® Packaging Extract for 2 hours at room temperature and the reaction terminated with 132.5 µl sterile Luria broth (LB; 1 % Bacto™ Tryptone (w/v) (Becton, Dickinson and Co., Sparks, USA), 0.5 % Yeast Extract (w/v) (Oxoid Ltd., Basingstoke, UK), 1 % NaCl (w/v), 1 % 10 mM Tris-HCl, pH 7.5 (v/v)) supplemented with 50 µg/ml ampicillin (Roche Diagnostics Gmbh, Mannheim, Germany). Ten microlitres of the packaged sample was kept aside for a 10^3-10^6 dilution assay. The remaining packaging mixture was added to a 50 ml log phase BLT5403 *E. coli* culture (bacterial cells provided in kit; Novagen, Inc.,

Madison, USA) grown in M9LB medium (LB medium supplemented with 0.4 % glucose (w/v), 1 mM $MgSO_4$, 1 x M9 salts (18.7 mM NH_4Cl, 22 mM KH_2PO_4, 22.5 mM Na_2HPO_4)) containing 50 µg/ml ampicillin and incubated at 37 ºC while shaking at 225 rpm in a Labotec® Orbital Shaker (Labotec, Midrand, South Africa) until lysis was observed (3-4 hours). Sodium chloride was added to a final concentration of 0.5 M and the suspension centrifuged at 8,000 *g* for 10 minutes using a Beckman® J2-21 centrifuge. Ten microlitres of the bacteriophage-containing supernatant was kept aside for a 10^6-10^9 dilution assay to determine the bacteriophage concentration. The remainder of the supernatant was kept at 4 ºC for daily use, or stored long term at –70 ºC with an equal volume of 80 % glycerol (v/v) (Saarchem (Pty) Ltd., Wadeville, South Africa).

Dilution assay

The packaging efficiency and total plaque forming units were determined by a 10^3-10^6 dilution assay. The bacteriophage concentration of an amplified culture was determined from a 10^6-10^9 dilution assay.

The assay involved setting up the above mentioned serial dilutions in LB medium. One hundred microlitres of each dilution was added to 250 µl log phase BLT5403 *E. coli* cells and 3 ml warm top agarose (1 % tryptone (w/v), 0.5 % yeast extract (w/v), 0.5 % NaCl (w/v), 0.6 % SeaKem LE agarose (w/v) (FMC® BioProducts, Rockland, USA)). The mixture was poured onto 20 ml agar plates (LB medium, 1.5 % agar (w/v) (Oxoid Ltd., Basingstoke, UK) containing 50 µg/ml ampicillin. Inverted plates were incubated at 37 ºC until plaques formed (3-4 hours). The plaque forming units (pfu) per ml (phage titre), total number of pfu (library size), and packaging efficiency and concentration were determined from the following formulae (Novagen, 2000b):

phage titre (pfu/ml) = (number of plaques on plate) x dilution x 10

where 10 takes into account the 0.1 ml of the dilution plated

library size (pfu) = phage titer x total sample volume

where total sample volume is the 0.15 ml final volume used during packaging reaction

packaging efficiency (pfu/µg) = library size / amount vector DNA

where 0.5 µg vector DNA was used in packaging reaction

DNA insert analysis by polymerase chain reaction (PCR)

Plaques were removed from agar plates and placed into 50 µl 10 mM Na_2EDTA, pH 8. The samples were vortexed and incubated at 65 ºC for 10 minutes. After cooling, the tubes were centrifuged at 16,000 *g* for 3 minutes at 4 ºC using a Sorvall® RMC-14 Refrigerated Microcentrifuge. One microlitre phage lysate was mixed with 0.5 µl T7SelectUP Primer (5 pmol/µl) (Novagen, Inc., Madison, USA), 0.5 µl T7SelectDOWN Primer (5 pmol/µl) (Novagen, Inc., Madison, USA), 12.5 µl PCR Master Mix (Promega Corporation, Madison, USA), and 10.5 µl H_2O. The sample was placed in the Eppendorf Mastercycler® Gradient machine (Eppendorf AG, Hamburg, Germany) under the cycling conditions shown in Table 2 (Novagen, 2000b). Ten microlitres of the PCR product was visualised against the 100bp DNA Ladder (Promega Corporation, Madison, USA) on a 1% agarose gel in TEA buffer.

Table 2: Cycling parameters for T7 insert amplification.

Segment	Cycles	Temperature	Time
1	1	80 ºC	2 min
2	1-35	94 ºC	50 sec
		50 ºC	1 min
		72 ºC	1 min
3	1	72 ºC	6 min

2.2.6 *P. falciparum* phage display library biopanning

Binding spectrin to Streptavidin Magnetic Particles

Hundred microlitres biotinylated spectrin was bound to 50 µl Streptavidin Magnetic Particles (Roche Diagnostics Gmbh, Mannheim, Germany) at room temperature for 1 hour by continuous inversion by hand to prevent the beads from settling at the bottom of the tube. Unbound spectrin was removed by placing the beads on the Magnetic Particle Separator. The beads were washed four times with 500 µl PBS containing 0.05 % Tween®-20 (v/v) and kept in 500 µl PBS at 4 ºC until further use. The PBS was removed before the use of the beads.

Biopanning the P. falciparum library against spectrin

The biopanning process entailed incubating the *P. falciparum* phage display library with spectrin bound to Streptavidin Magnetic Particles. Bacteriophage that did not interact with spectrin were washed away and bound bacteriophage eluted from the beads and amplified in *E. coli*. The biopanning was performed four times to ensure specific enrichment of bacteriophage interacting with spectrin.

The multiplicity of infection (MOI) for the biopanning procedure can be calculated from the following formula:

MOI = volume of phage used (ml) x phage titre (pfu/ml) / library size (pfu)

Five hundred microlitres of the *P. falciparum* phage display library (1 x 10^7 pfu/ml diluted in PBS) was bound to 100 µl Streptavidin Magnetic Particles for 1 hour at room temperature by continuous inversion by hand. This background selection step was performed to eliminate bacteriophage that could bind to the streptavidin-coated beads. The suspension was placed on the Magnetic Particle Separator and the unbound bacteriophage removed and added to the spectrin-coated beads. Bacteriophage were allowed to bind to the beads for 1 hour at room temperature while continuously inverting the tube by hand. Unbound bacteriophage were removed and the beads transferred to a 15 ml tube (Nunc™,

Roskilde, Denmark) and washed three times with 10 ml PBS containing 0.05 % Tween®-20 (v/v), for 10 minutes at room temperature. A final wash was performed with 10 ml PBS. Bound bacteriophage were eluted from the beads by incubating them in 250 µl PBS containing 1 % SDS (w/v) for 15 minutes at room temperature. The supernatant was added to a 50 ml log phase BLT5403 *E. coli* culture grown in M9LB medium supplemented with 50 µg/ml ampicillin and incubated overnight at 37 ºC using a Labotec® Orbital Shaker (225 rpm). Sodium chloride was added to a final concentration of 0.5 M and the suspension centrifuged at 8,000 *g* for 10 minutes using a Beckman® J2-21 centrifuge. A 10^6-10^9 dilution assay was performed to determine the bacteriophage concentration and the biopanning procedure repeated three times.

2.2.7 DNA sequencing of *P. falciparum* inserts

The nucleotide sequence of the *P. falciparum* DNA fragments coding for the peptides that interact with spectrin was analysed by Sanger sequencing (Sanger *et al.*, 1977). Sequencing was accomplished by first isolating the bacteriophage DNA from a plaque and then amplifying the *P. falciparum* DNA fragment by PCR using the T7SelectUP and T7SelectDOWN Primers. The remainder of the plaque was excised from the agar plate and placed into 500 µl phage extraction buffer (20 mM Tris-HCl, 100 mM NaCl, 6 mM $MgSO_4$, pH 8) and stored at 4 ºC. The double stranded DNA product was prepared for the sequencing reaction with ExoSap-IT® (USB Corporation, Cleveland, USA). This enzyme cocktail contains exonuclease I, which digests residual primers, and shrimp alkaline phosphatase, which dephosphorylates unincorporated nucleotides. The prepared DNA was sequenced using the T7 Sequenase version 2.0 DNA sequencing kit (Amersham Biosciences, Ltd., Buckinghamshire, UK) in the presence of [α-^{32}P]dATP and the T7SelectUP Primer.

2.2.8 Identification and analysis of *P. falciparum* inserts

P. falciparum database gene search

The sequence of each insert was compared to the PlasmoDB (Version 4.4, available at http://plasmodb.org) (Bahl *et al.*, 2003) *P. falciparum* genomic sequence (nt) database and the *P. falciparum* annotated genes (nt) databases using the WashU-BLASTN 2.0 search algorithm (gapped alignment using a Blosum62 matrix). *P. falciparum* genes showing a 95 % or higher identity to the query sequence were further analysed by establishing the reading frame of the DNA in the T7 bacteriophage. The nucleotide sequence was translated into the amino acid sequence using Gene Runner (version 3.05) (Hastings Software Inc., Hastings, USA) and compared to the PlasmoDB annotated protein database using the WashU-BLASTP 2.0 algorithm (using a gapped Blosum62 matrix). Identical results for the best matches from both queries, as well as a +1 reading frame in the amino acid sequence query, indicated that these *P. falciparum* genes code for proteins that interact with spectrin.

Peptide alignment and motif scan

The amino acid sequences of the peptides displayed on the bacteriophage surface were compared to each other with ClustalW (http://www.ebi.ac.uk/clustalw/) (Chenna *et al.*, 2003) and analysed with Motif Scan (http://myhits.isb-sib.ch/cgi-bin/motif_scan) (Pagni *et al.*, 2004) to determine if there were any amino acid or motif similarities between all the spectrin-binding peptides. Motif Scan compared the peptide sequence to the Prosite and the Pfam peptide motif databases.

2.3 Results

2.3.1 Spectrin preparation

SDS-PAGE analysis of the crude spectrin extract (Figure 22) showed that small amounts of protein 4.1, actin, and protein 4.9 had been co-extracted from the

erythrocyte ghosts, but no ankyrin contamination was evident. A Fairbanks SDS-polyacrylamide gel was used to verify the absence of ankyrin, because ankyrin co-migrates with β-spectrin on a Laemmli SDS-polyacrylamide gel.

Size exclusion chromatography (Figure 23) was used to purify the spectrin dimers and tetramers from the contaminating erythrocyte membrane proteins. The small protein peak before the spectrin tetramers consisted of high molecular weight (HMW) spectrin/protein 4.1/actin complexes. The protein peak after the spectrin dimers contained the other contaminating membrane proteins seen in Figure 22.

Purified spectrin dimers were concentrated and used for the biotinylation reaction. Tetramers were excluded because there was a risk of contamination with the HMW complexes.

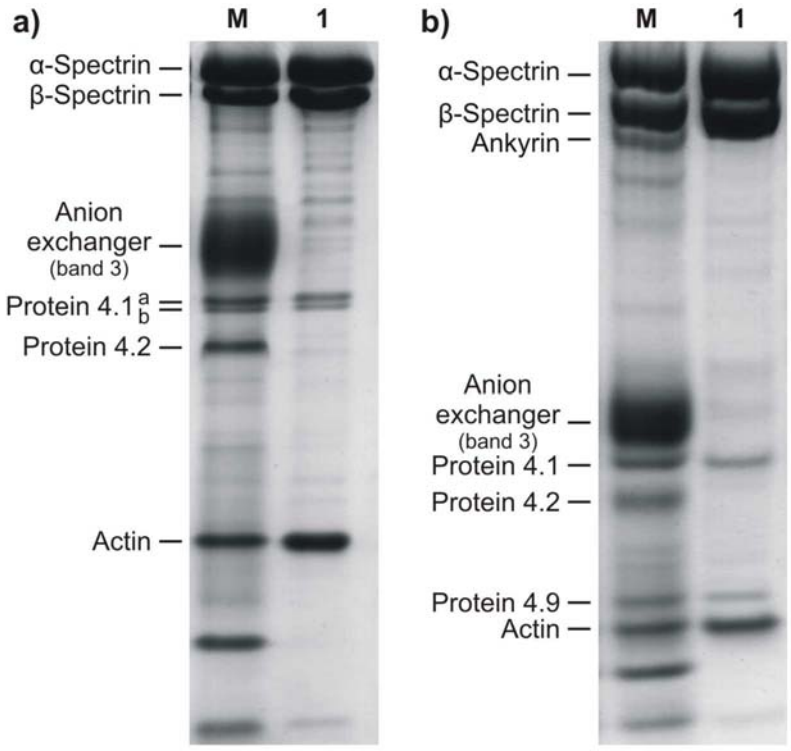


Figure 22: Spectrin isolation from human erythrocytes.
Photographs of a Laemmli (a) and Fairbanks (b) gel showing spectrin and additional proteins isolated from erythrocytes after hypotonic lysis and low ionic strength extraction. Lane M contains 45 µg (a) and 20 µg (b) erythrocyte membrane proteins respectively and lane 1 contains 10 µg (a) and 5 µg (b) spectrin respectively.

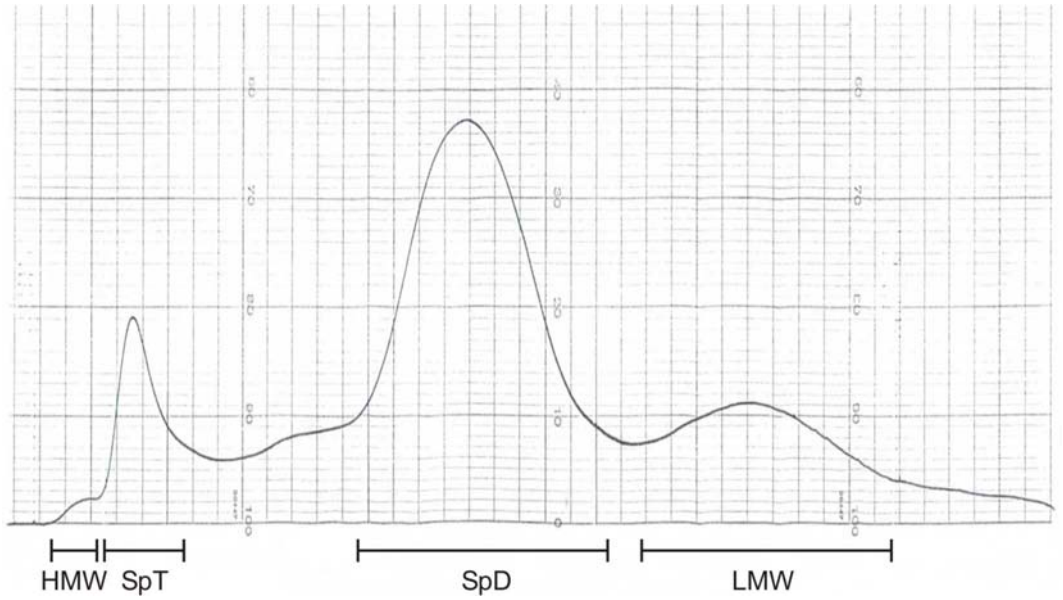


Figure 23: Size exclusion chromatography of spectrin.
Elution profile of crude spectrin passed through a Sepharose® (CL) 4B column. Spectrin tetramers (SpT; 920 kDa) elute before spectrin dimers (SpD; 460 kDa). The first small peak contains high molecular weight complexes (HMW: spectrin/protein 4.1/actin) and the last peak contains the low molecular weight proteins (LMW: actin and protein 4.1).

The total amount and concentration of protein, as well as the percentage yield of spectrin dimers after each step of the procedure are listed in Table 3. The spectrin dimers obtained after biotinylation were only 23 % of the crude spectrin extract because the high molecular weight complexes, actin and protein 4.1 were removed during size exclusion and spectrin tetramers were also not included in the concentration and biotinylation procedure. Furthermore, only 39 % of the spectrin dimers (after size exclusion) were recovered at the end of the biotinylation procedure because a large amount of protein was lost during the concentration and particularly in the dialysis steps.

Table 3: Spectrin dimer yields during spectrin processing.

	Total amount protein (mg)	Protein concentration (mg/ml)	Yield from total spectrin (%)	Yield from spectrin dimers (%)
Crude spectrin extract	18.5	1.85	100	-
Spectrin tetramers after size exclusion	1.3	0.06	7.0	-
Spectrin dimers after size exclusion	10.9	0.15	59	100
Spectrin dimers after concentration	7.7	3.06	42	70
Spectrin dimers after biotinylation	4.3	2.27	23	39

Analysis of the biotinylated spectrin dimers on Laemmli and Fairbanks SDS-polyacrylamide gels showed that the spectrin samples were >90 % pure (Figure 24) and that the visible contaminant was actin. These gels seem to contain less spectrin than Figure 22 even though similar amounts were loaded and Bradford analysis was used to determine the spectrin concentration in the samples. However, the Bradford reagent contains Coomassie G-250 Dye which binds to basic and aromatic groups in amino acid side chains (Pierce, 2005). Biotin contains two NH-groups in its structure, which could bind additional Coomassie G-250 Dye, thereby overestimating the amount of protein present in the sample. The percentage yield of spectrin dimers after biotinylation would therefore also have been lower.

The spectrin biotinylation was checked by coating a Streptawell with the biotinylated spectrin and then detecting the spectrin with α- and β- spectrin antibodies. The well containing spectrin turned light purple with 4-chloro-1-napthol, while the well containing no spectrin remained clear (results not shown). This indicates that the spectrin was biotinylated successfully and would therefore bind to streptavidin.

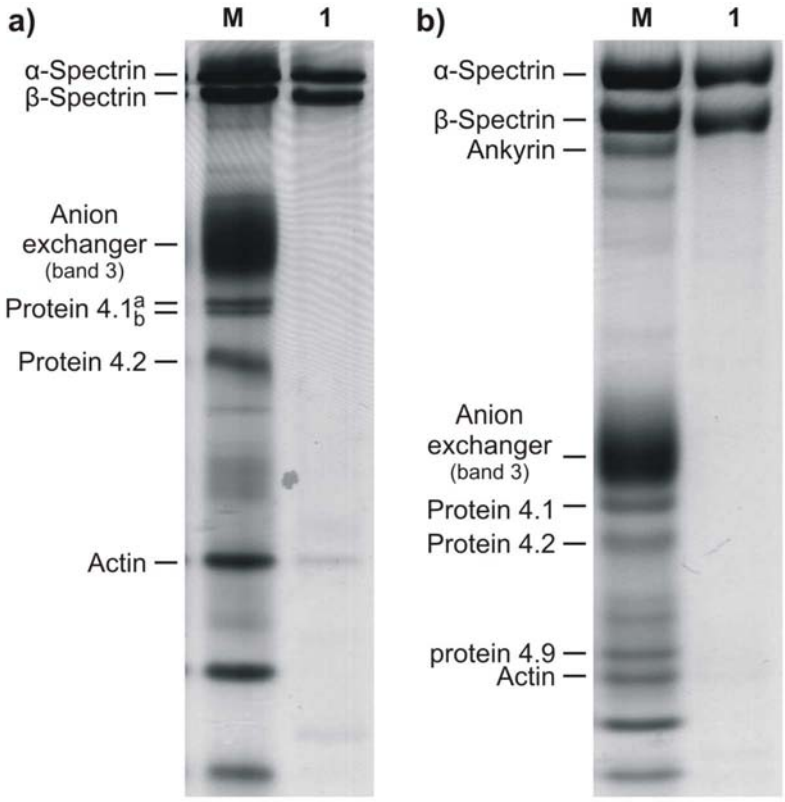


Figure 24: Biotinylated spectrin.
Photographs of a Laemmli (a) and Fairbanks (b) gel showing biotinylated spectrin. Spectrin was purified by size exclusion chromatography and subsequently biotinylated. Lane M contains 45 µg (a) and 20 µg (b) erythrocyte membrane proteins respectively and lane 1 contains 10 µg (a) and 5 µg (b) biotinylated spectrin respectively.

2.3.2 *P. falciparum* phage display library construction

Two mRNA isolations were performed from mixed *P. falciparum* cultures, depicted in Figure 25 and listed in Table 4. Four 175 cm^3 culture flasks were used for each isolation.

The isolated mRNA had a lower A_{260}/A_{280} ratio than that of pure mRNA which is 2 (as described in Ausubel *et al.*, 1994) and was thus either contaminated with proteins, or the Dynabeads® Oligo $(dT)_{25}$ beads that had not been completely removed from the mRNA sample (Dynal®, 2000).

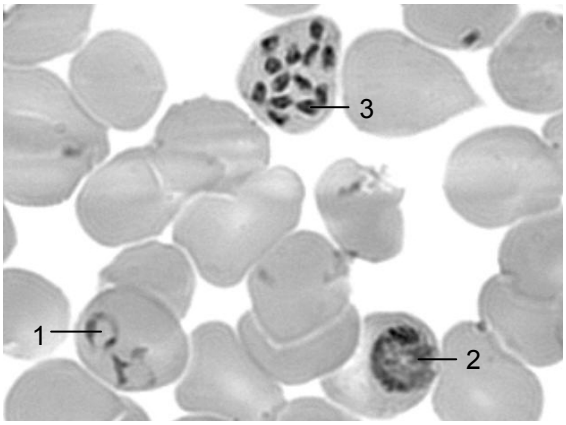


Figure 25: *P. falciparum* cultured in human erythrocytes.
Photograph showing a *P. falciparum* culture at 1000x magnification. The parasite's nuclear material is stained by the methylene blue component of the Rapid Haematology Stain. Early trophozoites (rings) (1) have the shape of a ring within the erythrocyte, the trophozoite (2) appears as a large circular body within the erythrocyte, and schizonts (3) are small ovoid bodies inside the erythrocytes.

Table 4: *P. falciparum* mRNA isolation.

mRNA isolation	Culture parasitaemia and composition	Total RNA	mRNA	Percentage mRNA isolated from total RNA
1	20 %; ~2/3 R* ~1/3 T/S#	403 µg (3.7 µg/µl)	7.7 µg (0.31 µg/µl) A_{260}/A_{280} = 1.5	1.9
2	10 %; ~2/3 R ~1/3 T/S	169 µg (1.5 µg/µl)	3.8 µg (0.15 µg/µl) A_{260}/A_{280} = 1.75	2.3

* R = rings
# T/S = trophozoites and schizonts

Figure 26 shows the *P. falciparum* total RNA containing the 28s (4104 bp), 18s (1384 bp) and 5s ribosomal RNA (rRNA) subunits (Daily *et al.*, 2004), as well as mRNA. Visualisation of the *P. falciparum* mRNA on a 1 % agarose gel (result not shown) revealed mRNA as a faint smear with only traces of 18s and 28s rRNA subunits present. rRNA was a contaminant because the *P. falciparum* genome is AT-rich and Oligo $(dT)_{25}$ beads were used for the isolation procedure.

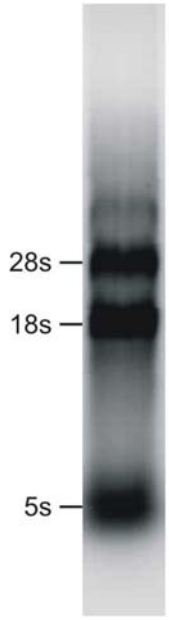


Figure 26: *P. falciparum* total RNA.
Agarose gel showing total RNA isolated from *P. falciparum* cultures. The 28s (4104 bp), 18s (1384 bp), and 5s rRNA subunits are the predominant forms of RNA. The faint mRNA smear can be seen in the region of the 28s and 18s rRNA.

The phage display library was created with 4 µg *P. falciparum* mRNA from the second mRNA isolation. DNA synthesis was monitored by the addition of [α-^{32}P]dATP in the first and second strand synthesis reactions. The *P. falciparum* DNA generated during second strand synthesis is shown in Figure 27 as a smear. The majority of the *P. falciparum* DNA was greater than 722 bp.

The [α-^{32}P]dATP labelled *P. falciparum* DNA was also used to monitor the gel filtration procedure (Figure 28). The gel filtration column retains DNA smaller than 300 bp and therefore the initial fractions eluted from the column should contain the large DNA fragments. Fraction 3 contained the DNA that was greater than 722 bp and was used for the ligation reaction. This fraction contained 963 ng DNA at a concentration of 3.85 ng/µl with an A_{260}/A_{280} absorbance ratio of 1.4. This ratio is lower than 1.8-1.9 which is the A_{260}/A_{280} absorbance ratio of pure DNA (as described in Ausubel *et al.*, 1994), indicating that the sample was contaminated with protein or aromatic substances, such as phenol. The size markers in Figure 28 show up as multiple bands due to non-specific amplification during the PCR reaction and the sensitivity of the radioisotope.

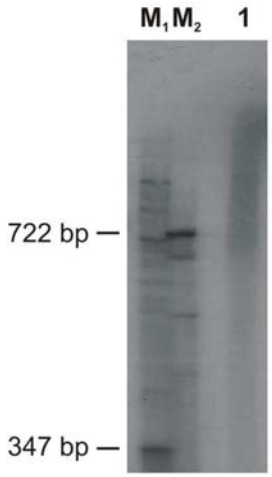


Figure 27: Second strand synthesis from *P. falciparum* cDNA.
Autoradiograph of a denaturing polyacrylamide gel showing *P. falciparum* DNA synthesised in the presence of [α-^{32}P]dATP. Lanes M_1 and M_2 contain 1 µl molecular weight markers, synthesized by PCR utilising primers specific for human erythrocyte α-spectrin (exon 2; 347 bp) and band 3 (exon 18 and 19; 722 bp) respectively. Lane 1 contains 2 µl *P. falciparum* dsDNA, which migrates as a smear of >722 bp.

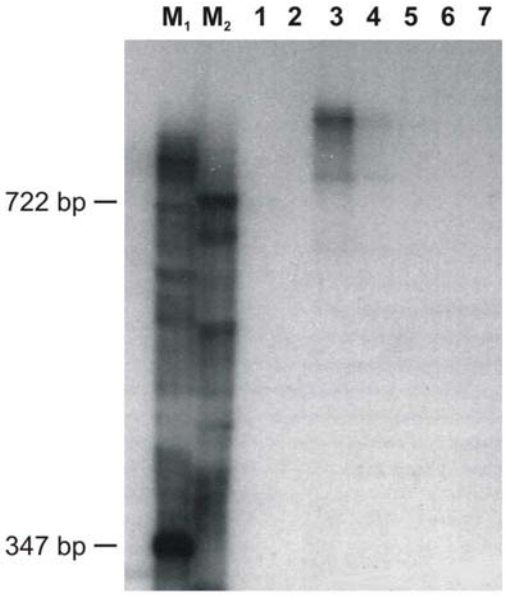


Figure 28: Gel filtration of *P. falciparum* DNA.
Autoradiograph of a denaturing polyacrylamide gel showing elution fractions of *P. falciparum* DNA after gel filtration. Lanes M_1 and M_2 contain 1 µl molecular weight markers, which were synthesized by PCR utilising primers specific for human erythrocyte α-spectrin (exon 2; 347 bp) and band 3 (exon 18 and 19; 722 bp) respectively. Lanes 1-7 represent the following elution fractions: fraction 1 = 0.1 ml; fraction 2 = 0.2 ml; fraction 3 = 0.25ml; and fractions 4-7 = 0.1ml each. Two microlitres of each DNA fraction was electrophoresed after it had been precipitated with ethanol and resuspended in 20 µl TE buffer.

One microlitre of this sample, containing 0.05 pmol of ~1500 bp DNA (size estimation based on OrientExpress™ cDNA Manual), was used for the ligation reaction. This gave an insert to vector ratio of 2.5:1, which is well within the recommended ratio range for the ligation reaction (Novagen, 2000b). The ligated fragments were packaged into viral particles to create the *P. falciparum* phage-display library. The characteristics of this library were compared to those of a library created with the T7Select® Packaging Control DNA (Novagen, Inc., Madison, USA) (Table 5). The *P. falciparum* library was much smaller than the control library and had a much lower packaging efficiency.

Table 5: *P. falciparum* phage display library characteristics.

Library	Phage titre (pfu/ml)	Library size (pfu)	Packaging efficiency (pfu/µg)
P. falciparum	3.01×10^7	4.52×10^6	9.03×10^6
T7Select® Packaging Control	4.53×10^9	6.80×10^8	1.36×10^9

Amplification of randomly selected plaques from the *P. falciparum* phage-display library showed DNA inserts of mainly 50-200 bp in length (Figure 29) and the largest DNA fragment found to date was 1200 bp (result not shown).

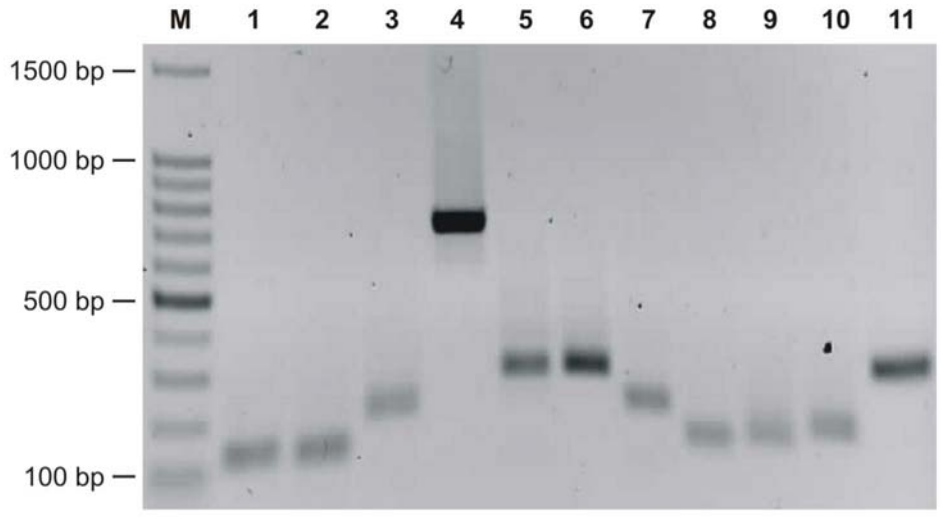


Figure 29: DNA inserts from the *P. falciparum* phage-display library.
Agarose gel showing the PCR product of *P. falciparum* inserts in T7 bacteriophage plaques. All fragments include 107 bp of the T7Select10-3b vector sequence. Lane M contains the Promega 100 bp DNA Ladder. Lanes 1-11 contain 10 µl of the PCR product synthesised with the T7Up and T7Down Primers.

2.3.3 Biopanning

Biopanning was initially performed in Streptawell plates before the method with the Streptavidin Magnetic Particles was established. One hundred microlitres of a *P. falciparum* phage display library that had a concentration of 4.5 x 10^7 pfu/ml was added to two wells containing immobilised spectrin. This gave the binding reaction a multiplicity of infection (MOI) of 1, which means that only one bacteriophage copy of each unique *P. falciparum* DNA insert was present in the sample. A titering assay was performed on the binding phage in the first well to monitor enrichment and the second well was used for phage amplification for the next round of biopanning. Table 6 shows that the enrichment in the first two rounds of biopanning was within the 10^3-10^5 pfu range. The third and fourth round of biopanning however showed a decrease in binding indicating that sufficient enrichment had already been achieved in the second round of biopanning.

Table 6: Bacteriophage enrichment.

Biopanning round	Number of binding phage (pfu)
1st	$4.5x10^4$
2nd	$4.5x10^5$
3rd	$2.1x10^5$
4th	$1.1x10^5$

Once the method with the Streptavidin Magnetic Particles was established a mixture of two *P. falciparum* libraries was biopanned against purified spectrin. The first library was discussed in section 2.3.2 and the second library was provided by Roberto Lanzillotti and had a library size of 6.2 x 10^5 pfu. The new combined library had a size of 2.6 x 10^6 pfu and was used at a MOI of 2 for biopanning. This means that at least two copies of each phage were present in the 500 µl sample of 1 x 10^7 pfu/ml bacteriophage used in the first round of biopanning. Biopanning with the Streptavidin Magnetic Particles resulted in a phage enrichment similar to that observed during biopanning with the Streptawell plates (results not shown).

After the fourth round of biopanning 80 plaques were selected for PCR amplification with the T7SelectUP and T7SelectDOWN Primers. Of these only 44 were sequenced because the remaining PCR products were barely visible on a 1% agarose gel. The PCR amplification could have been poor for numerous reasons, such as impurities in the DNA template that prevented the *Taq* DNA polymerase from functioning properly, or insufficient DNA template or secondary structures within the DNA sequences that prevented the primers and *Taq* DNA polymerase from binding to the DNA template. The aim of this project was however not to isolate all possible parasite peptides that bind to spectrin and therefore the failed PCR amplifications were not repeated. Thirteen of the sequenced DNA inserts consisted of A-stretches (Figure 30a) and the remaining 31 DNA fragments showed 14 different DNA sequences rich in adenosines and thymidines which is typical for the *P. falciparum* genome.

Four of the 14 *P. falciparum* DNA fragments did not have the linker region between the T7 vector DNA and the *P. falciparum* DNA (Figure 30b), indicating that the *Eco*RI enzyme had cleaved the *P. falciparum* DNA internally. First and second strand synthesis were performed in the presence of 5-methyl dCTP to protect the DNA from the restriction enzymes that are used during the linker digestion procedure, and *Eco*RI should not cleave methyl-protected DNA. It has however been shown that *Eco*RI sites that are flanked by a guanosine or cytidine can be cleaved by the enzyme when the DNA is fully and hemi-methylated (Tasseron-de Jong *et al.*, 1988).

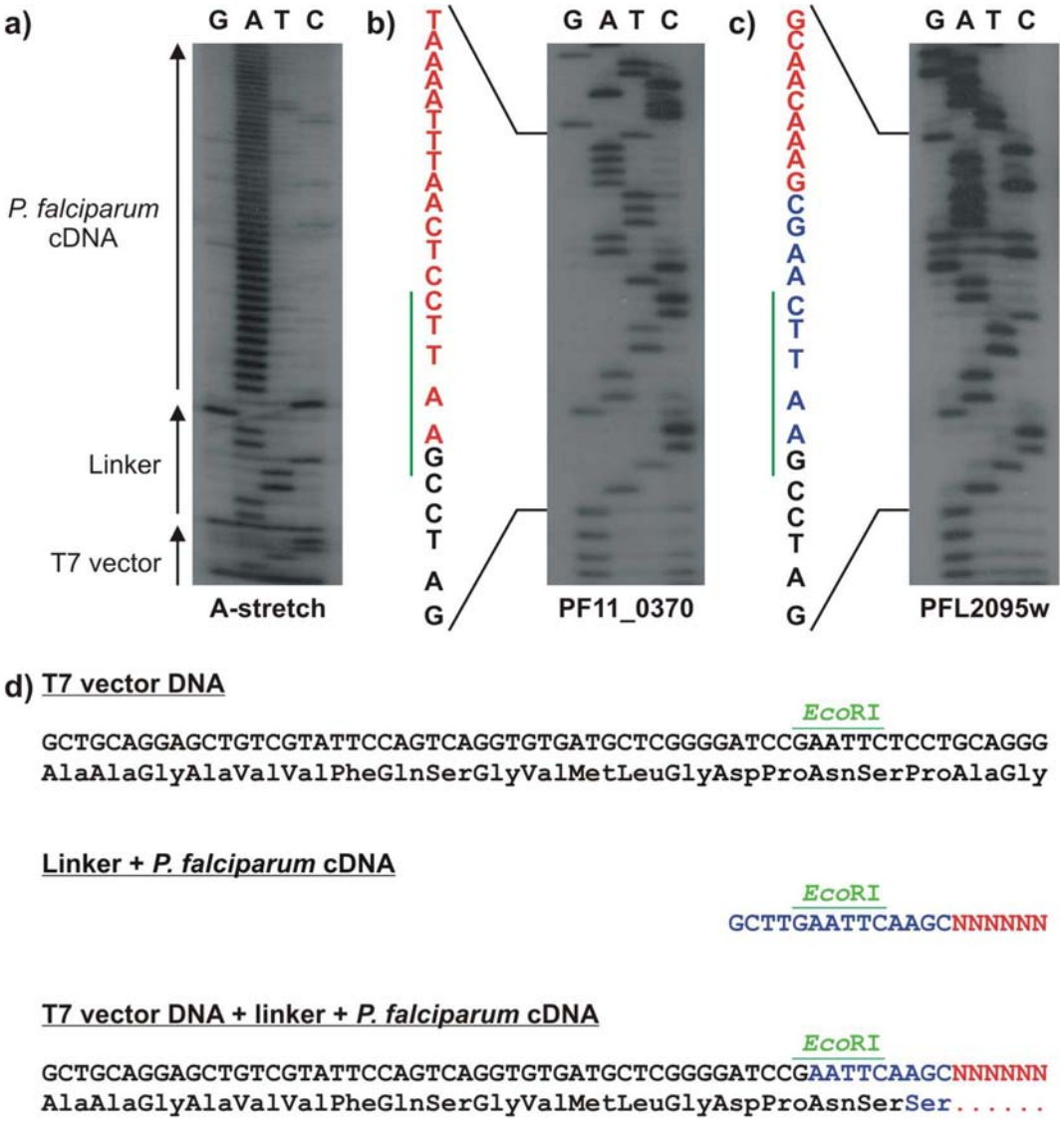


Figure 30: *P. falciparum* DNA insert sequences.
Autoradiographs (a, b & c) showing DNA sequences that were isolated from plaques and diagram (d) showing the cloning procedure facilitated by directional linkers. The autoradiographs represent 8 % denaturing polyacrylamide gels of a *P. falciparum* A-stretch (a) and *P. falciparum* DNA sequences without (b) and with (c) the linker region. The linker region is absent in (b) because *Eco*RI cleaved the methyl-protected *P. falciparum* DNA internally. Red nucleotides = *P. falciparum* DNA; blue nucleotides = directional linker; black nucleotides = vector nucleotides; green line = *Eco*RI restriction site.

2.3.4 Gene identification and peptide analysis

All 14 sequences found after fourth round biopanning were identified in the PlasmoDB genomic sequence (nt) database (Table 7). However, only 12 of these

sequences were found in the annotated gene section of the database. Five of the genes were expressed in the correct frame (they had a +1 frame when compared to the PlasmoDB annotated protein database). One fragment was inserted backwards (-2 frame). This is surprising because the directional linker used for the cloning procedure, should only allow the fragment to be inserted into the +1, +2, or the +3 frame.

Table 7: *P. falciparum* genes identified after fourth round biopanning.

PlasmoDB gene name or Chromosome number	Insert size (bp)	Percentage nucleotide identity	Reading frame	Special features	Predicted PlasmoDB protein
PFA0125c	~150	96	+1	-	putative Ebl-1 like protein
PFE0980c	~250	97	+1	-	hypothetical protein
MAL8P1.104	~400	100	+1	no 5' *Eco*RI/*Hind* III linker	hypothetical protein
MAL8P1.151	~60	96	+1	-	hypothetical protein
PFI1570c	~100	97	+1	no 5' *Eco*RI/*Hind* III linker	putative amino-peptidase
PF10_0232	~1200	90	+2	-	hypothetical protein
MAL13P1.249	~200	100	+2	-	hypothetical protein
PFC0135c	~100	97	+3	no 5' *Eco*RI/*Hind* III linker	conserved protein
PF08_0131	~300	100	+3	-	1-cys peroxidoxin
PFL2095w	~200	100	+3	-	translation initiation factor SUI1, putative
PFF0125c	~350	98	-2	inserted backwards	hypothetical protein
MAL7_5.8s	~100	98	-	no translation	5.8s rRNA
Chromosome 11	~300	99	+2	intron of PF_0370	-
Chromosome 13	~200	97	-	intergenic	-

When the nucleotide sequences of the +2, +3 and -2 frame DNA inserts were translated with Gene Runner, it was found that the majority of the inserts were expressed as small peptides with four to eight amino acids since premature stop codons terminated translation. The +1 frame sequences were expressed as peptides with 20 or more amino acids (Figure 31).

The majority of the expressed peptides contained numerous positively (mainly lysine) and negatively (mainly aspartate) charged amino acids (Figure 31), but no consistent charge pattern was observed. Two of the in-frame sequences (PFE0980c and MAL8P1.104) also contained stretches of polar uncharged asparagine residues and PFE0980c had two repetitive amino acid sequences, namely KDN and KENNNNKGN.

A ClustalW alignment with all the peptide fragments showed no sequence similarities (results not shown), and motif and profile searches did not reveal any putative motif that was present in every sequence (Table 8). Two sequences (PFF0125c & MAL7_5.8s) contained no motif and the other sequences contained a mixture of N-glycosylation and N-myristoylation sites, as well as protein kinase C, casein kinase II, and cyclic adenosine monophosphate- and cyclic guanosine monophosphate-dependent protein kinase phosphorylation sites. These are all post-translational modification sites that could be used by the parasite to regulate its proteins. A putative bacterial immunoglobulin-like domain 1 (Big-1) was found in several of the shorter out-of-frame sequences. This domain is generally associated with several other domains in bacteria and phage surface proteins that mediate bacteria host-cell interactions (Finn *et al.*, 2006). The complete domain is normally over 50 amino acids in length and only segments thereof were identified in the ≤8 amino acid spectrin-binding sequences.

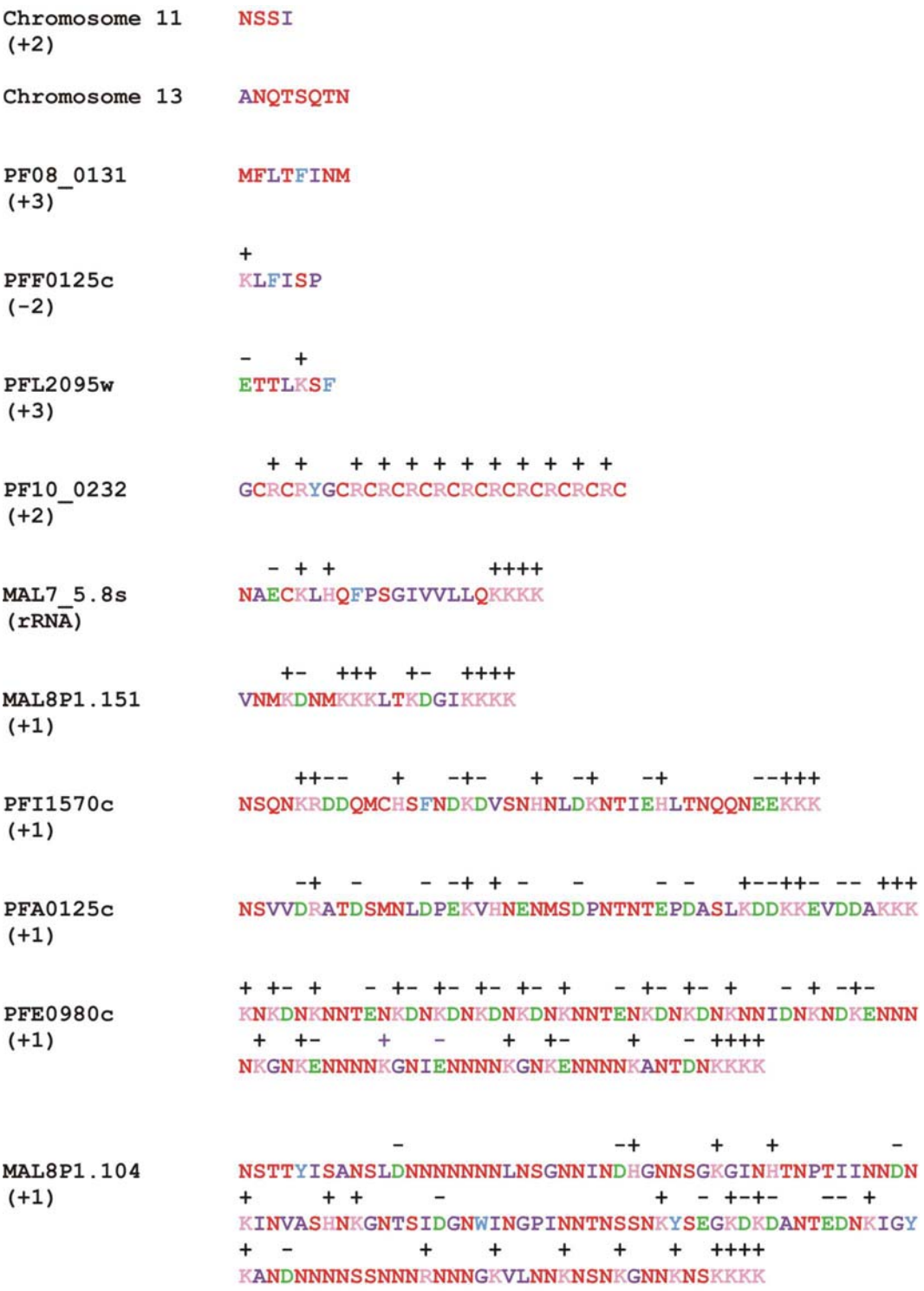


Figure 31: Peptides isolated after fourth round biopanning.
Diagram showing the amino acid sequences of the peptides isolated after fourth round biopanning against spectrin. The gene name as well as the expression frame is given. The amino acids are colour coded according to class and charges are shown above the amino acids. Red = polar uncharged; purple = nonpolar aliphatic; blue = aromatic; pink = positively charged; green = negatively charged.

Table 8: Motifs and profiles of the *P. falciparum* spectrin-binding peptides.

Gene name	Motif/Profile (amino acid sequence)
Chromosome 11	1 N-glycosylation site (NSSI)
	Big-1* domain profile (NSSI)
Chromosome 13	1 N-glycosylation site (NQTS)
	Big-1 domain profile (ANQT--SQTN)
PF08_0131	Big-1 domain profile (MFLTFI)
PFF0125c	-
PFL0295w	1 Protein kinase C phosphorylation site (TLK)
	Big-1 domain profile (ETTLK-SF)
PF10_0232	Arginine-rich region profile
	Cysteine-rich region profile
MAL7_5.8s	-
MAL8P1.151	1 cAMP**- and cGMP***-dependent protein kinase phosphorylation site (KK T)
PFI1570c	1 Casein kinase II phosphorylation site (SXXD)
PFA0125c	1 N-glycosylation site (NMSD)
	3 Casein kinase II phosphorylation sites (SVVD; TNTE; SXXD)
	1 Protein kinase C phosphorylation site (SLK)
PFE0980c	2 N-glycosylation sites (NNTE)
	3 N-myristoylation sites (GNKENN; GNIENN)
	Asparagine-rich region profile
	Lysine-rich region profile
MAL8P1.104	6 N-glycosylation sites (NSTT; NNSG; NHTN; NTSI; NNTN; NNSS)
	1 Casein kinase II phosphorylation site (TSID)
	5 N-myristoylation sites (GNNIND; GNNSGK; GINHTN; GNWING; GNNKNS)
	4 Protein kinase C phosphorylation sites (SGK; SNK; SKK)
	Asparagine-rich region profile

* Big-1 = bacterial immunoglobulin-like domain 1
** cAMP = cyclic adenosine monophosphate
*** cGMP = cyclic guanosine monophosphate

2.4 Discussion

A *P. falciparum* phage display library was created by inserting DNA fragments generated from *P. falciparum* mRNA into the T7Select10-3 vector. The library was used to isolate several *P. falciparum* peptides that interact with human erythrocyte spectrin.

2.4.1 Library creation and bacteriophage selection

Characteristics of the P. falciparum phage display library

The phage display library was created from mRNA that had been isolated from mixed *P. falciparum* cultures to ensure that transcripts from all the parasite stages of the erythrocytic phase were present. This *P. falciparum* phage display library was however not representative of all the *P. falciparum* mRNA transcripts. Several mRNA transcripts and especially rare mRNAs, which transcribe as low copy numbers in the parasite, were probably not present in the library because only a small amount of the isolated mRNA was used for reverse transcription and only a portion of the *P. falciparum* DNA was used for the ligation and packaging reaction. Some DNA was also lost during the multiple processing steps that involved phenol-chloroform extraction and ethanol precipitation.

First strand synthesis was performed with anchored poly-T primers that should have bound to the poly-A tail at the 3' end of each mRNA strand. However, most of the *P. falciparum* DNA inserts that were isolated after fourth round biopanning coded for an internal region of the respective *P. falciparum* gene. The primers would have bound within the mRNA sequences because the *P. falciparum* genome is 80% AT-rich (Gardner *et al.*, 2002) and parasite genes contain adenosine-stretches that translate into the low complexity regions commonly found in *P. falciparum* proteins (Pizzi and Frontali, 2001).

The size of the *P. falciparum* inserts of the phage display library was checked by PCR and it was found that most of the amplified inserts were between 50 and 200 bp in length. This was surprising because the gel filtration step should have removed DNA fragments smaller than 300 bp prior to ligation to the bacteriophage DNA. This seemed to be the case since α-^{32}P labelled dsDNA was >722 bp after gel filtration (Figure 28), although the incorporation of less [α-^{32}P]dATP into the smaller DNA fragments would have made them less visible. The *Eco*RI digestion prior to ligation could also have cleaved the *P. falciparum* DNA internally, thereby creating smaller DNA fragments, even though the *Eco*RI

sites should have been protected by methylation. This was confirmed by the isolation of several bacteriophage that contained *P. falciparum* DNA inserts without a 5' linker region. Insert lengths could have been increased by reverse transcribing the mRNA at a higher temperature, such as 55 ºC. Elevated temperatures remove hairpin loops from the mRNA template, which can be responsible for early termination of the transcription reaction. However, higher temperatures can degrade the mRNA template, but this could be prevented with the addition of a RNA chaperone such as T4 bacteriophage gene 32 protein in the reaction, thereby stabilising the mRNA template (Piche and Schernthaner, 2005). Phage display insert sizes may also be increased by using a gel filtration resin with a higher molecular weight cut-off. The collection of smaller gel filtration fractions and subsequently using only those fractions with large cDNA could also improve the insert composition of the phage display library.

The presence of DNA fragments smaller than 300 bp could have been responsible for the small library size and low packaging efficiency. The molar ratio of insert to vector that was used for the ligation reaction (2.5:1) was calculated on the assumption that the prepared *P. falciparum* DNA was approximately 1.5 kb in length (Novagen, 2000b). If most of the fragments were smaller, then the molar ratio would have been much higher.

Biopanning against spectrin

The biopanning procedure was initiated with a background binding step to eliminate bacteriophage binding to streptavidin. None of the spectrin-binding peptides contained the streptavidin binding motif HPQ or any of the weaker binding motifs HPN, HPM, and HPL (Devlin *et al.*, 1990), indicating that the background binding step was sufficient to prevent any non-specific interactions.

Biopanning with the first *P. falciparum* phage display library was performed with a MOI of 1 which should be sufficient to isolate high affinity bacteriophage (Novagen, 2000b). When the second *P. falciparum* phage display library was used

for biopanning, the MOI was raised to 2, to ensure that all the primary recombinants had a chance to bind to spectrin. Some bacteriophage replicate at a slower rate than others, because the DNA insert present in the bacteriophage interferes with replication (Wang and Yu, 2004) and therefore the other bacteriophage would outgrow the slower growing phage during each round of biopanning. Increasing the multiplicity of infection during biopanning ensures that the slow growing bacteriophage can also bind to spectrin.

The number of phage recovered from the first and second round of biopanning should be in the range of 10^3–10^5 pfu, increasing to $>10^6$ pfu in the third and fourth rounds. When no further increase in phage number is observed, enrichment is unlikely to occur with additional rounds of biopanning (Novagen, 2000b). However, bacteriophage enrichment was only observed during the first two rounds of biopanning. This could be due to the small library size (small number of primary recombinants) of the *P. falciparum* phage display library, which enables most of the different spectrin binding bacteriophage to bind within the first two rounds of biopanning.

Bacteriophage displaying small peptides were isolated more frequently during fourth round biopanning than those displaying larger peptides. The peptide-binding regions on spectrin could have been more accessible to smaller peptides than to the larger ones. Identical small peptides were also isolated several times indicating that enrichment did occur during biopanning.

2.4.2 Spectrin-binding bacteriophage

Characteristics of the P. falciparum DNA inserts

All the sequenced DNA inserts were identified in the PlasmoDB database. Two of the sequences did not match any of the annotated genes, one of which was located in the intron of PF11_0370 and the other between two genes on chromosome 13. Another sequence was identified as part of the 5.8s rRNA. This was surprising because *P. falciparum* mRNA was used to create the library and this should have

excluded introns, intergenic regions and rRNA. An explanation for isolating introns and intergenic regions is that the DNA fragments could be part of *P. falciparum* genes that have either not been correctly annotated, or belong to a gene that still has to be annotated in the PlasmoDB database. The other explanation for the presence of these three isolated sequences is that genomic DNA could have contaminated the total RNA isolated from the parasites, but this is highly unlikely since genomic DNA is double stranded and only single stranded DNA can bind to the beads. The most likely scenario is that rRNA and pre-mRNA containing introns were isolated from the total RNA with the Dynabeads® Oligo $(dT)_{25}$ beads due to the AT-richness of the genome.

The insert that was identified as part of PFF0125c was found in the -2 frame. The linkers that were used to insert the *P. falciparum* DNA into the T7 vector only allow the inserts to be cloned in the +1, +2 or +3 frame, thus indicating that the insert is part of a gene that is transcribed from the complementary strand of PFF0125c. PlasmoDB lists another gene (PFF0120w) on the complementary strand approximately 1000 nucleotides downstream from PFF0125c, so it is very unlikely that the insert is part of that gene. Therefore this *P. falciparum* insert could again be part of a gene that still has to be annotated in the PlasmoDB database.

Characteristics of the bacteriophage-displayed peptides and the P. falciparum proteins that interact with spectrin

The *P. falciparum* peptides that bound to spectrin showed no single common motif or profile when compared to each other. The peptides were composed of a variety of positively and negatively charged and polar uncharged amino acids, the most prevalent being lysine, aspartate and asparagine. The codons for these amino acids are AAA (lysine), GAU (aspartate) and AAU (asparagine) and are often found in *P. falciparum* proteins due to the A+T richness of the *P. falciparum* genome (Gardner *et al.*, 2002). Asparagine and lysine are also very common in

the hydrophilic low complexity regions of *P. falciparum* proteins (Pizzi and Frontali, 2001) and will be discussed further in the following chapter.

Three peptides contained only polar uncharged amino acids, while the rest of the peptides contained a variety of polar uncharged and charged amino acids. The polar uncharged amino acids presumably bound to the spectrin surface via hydrogen bonds and the charged amino acids bound to spectrin via hydrogen bonds and electrostatic interactions, since the outer amino acids of each triple helical bundle of spectrin are mainly positively and negatively charged (Figure 5). The biopanning procedure yielded five DNA inserts out of twelve that expressed the peptides in the +1 frame. This correlates well with the fact that only one in every three inserts will be in the +1 frame when an expression library is created with directional linkers (compared to one in six in-frame inserts when the library is created without directional linkers). Even though the other peptides were not in the correct frame, they could be peptide mimics. The study of these mimics could provide further information on the interactions between *P. falciparum* proteins and spectrin.

The five in-frame peptides isolated after fourth round biopanning are part of *P. falciparum* proteins identified in the PlasmoDB database. Three of the proteins have been described as hypothetical and the other two as a putative Ebl-1 like protein and a putative aminopeptidase. Approximately 60.9 % of the *P. falciparum* genome contains genes that have no assigned function or homology to genes of other organisms (Gardner *et al.*, 2002), which matches with three of the five spectrin-binding proteins being described as hypothetical proteins.

2.4.3 Identifying protein-protein interactions with phage display

The *P. falciparum* phage display library and the techniques employed here were successfully used to isolate peptides that interact with spectrin. The library creation was time-consuming and expensive, but once it was established, the biopanning procedure was rapid and cost effective.

A disadvantage of phage-display is that more than one library needs to be created to ensure a full representation of mRNA transcripts. This was however not a critical issue in this study because the aim was not to isolate all the *P. falciparum* spectrin-binding peptides. The chance of isolating more peptides was however increased by combining two *P. falciparum* phage display libraries for the biopanning procedure.

An advantage specific to the T7Select® Phage Display System is the use of directional linkers during the cloning procedure. Directional linkers ensure that cDNA can only be inserted into the vector in one of three frames in contrast to all six frames when directional linkers are omitted from the cloning procedure. This decreases the risk of isolating false positives in a NPL, because at least one third of the isolated sequences should be expressed as in-frame peptides.

Another major advantage of the phage display system is that once the libraries have been created, they can be used to identify peptides that interact with any immobilised ligand, thus making phage display a very powerful tool in the elucidation of parasite-host protein-protein interactions.

In conclusion, five genes, namely PFI1570c, MAL8P1.151, MAL8P1.104, PFE0980c and PFA0125c, that code for *P. falciparum* proteins that interact with erythrocyte spectrin, were chosen for cloning and recombinant protein expression to further analyse and characterise these parasite proteins.

Chapter 3: Recombinant spectrin-binding *P. falciparum* proteins

3.1 Introduction

3.1.1 Expression of *P. falciparum* recombinant proteins

The successful production of biologically active recombinant proteins depends on the expression micro-environment and the compatibility of the native gene sequence with the expression host. Over the years, *E. coli* has been the preferred expression host, because the bacterium is easy to manipulate genetically and very cost-effective to maintain in culture. Successful expression of proteins within *E. coli* is influenced by the structural features of the inserted gene, the transcription efficiency and stability of the transcribed mRNA, the correct and efficient translation, correct protein folding (i.e. solubility), toxicity of the protein and the degradation of the protein by proteases (Schumann and Ferreira, 2004).

Recently, two large scale *P. falciparum* recombinant protein production studies attempted to define some of the factors that affect the expression and solubility of recombinant parasite proteins. In the first study Mehlin *et al*. (2006) attempted to express recombinant proteins that spanned several *P. falciparum* protein classes. Out of the 1000 genes, 337 recombinant proteins were expressed and only 63 of these were soluble. In the second study Vedadi *et al*. (2007) selected a set of 468 *P. falciparum* genes, which coded for putative haemoglobin metabolism, biosynthesis, hypothetical, exported and apicoplast-targeted proteins. The group increased their sample size to 1008 genes, by selecting orthologues from six other Apicomplexan species, namely *P. yoelii yoelii*, *P. vivax*, *P. berghei*, *P. knowlesi*, *Cryptosporidium parvum* and *Toxoplasma gondii*. Only 304 proteins were expressed in soluble form and only 36 of these were successfully crystallised. The results from both studies highlight the difficulties and the low success rates that are often encountered when cloning and expressing *P. falciparum* recombinant proteins. They also indicate that a trial and error based approach, utilising several strategies, is generally necessary to express soluble parasite proteins (Figure 32).

These strategies reduce the problems associated with the physical characteristics of the recombinant protein, the gene and mRNA that code for the protein and the micro-environment of the expression host.

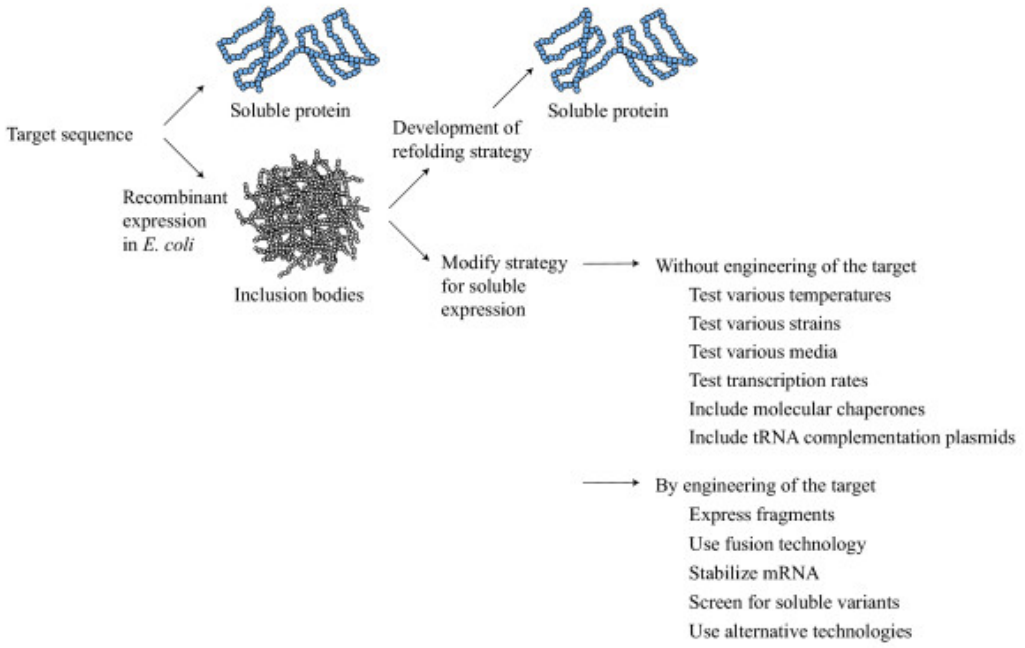


Figure 32: Strategies to obtain soluble recombinant *P. falciparum* proteins from *E. coli.*
Diagram showing the downstream applications that can be employed to obtain soluble recombinant proteins. Two strategies can be followed once a protein is produced in inclusion bodies. The first is to denature and subsequently refold the protein. The other strategy includes the modification of culture conditions and/or the protein, to aid the bacterial cell to produce soluble protein (Sorensen and Mortensen, 2005b).

Protein modifications

P. falciparum proteins are generally larger than their respective homologues in other species and often possess long, disordered loops within their structures (reviewed in Aravind *et al.*, 2003). These features, as well as the presence of long repetitive amino acid stretches within the protein sequence (Singh *et al.*, 2004), contribute to the insolubility of *P. falciparum* proteins when they are expressed in heterologous systems. The removal of these structures as well as the omission of transmembrane regions, signal peptides, transit peptides and export motifs from

the protein of interest should enhance the successful expression of soluble recombinant *P. falciparum* proteins (Vedadi *et al.*, 2007). Mehlin *et al*. (2006) also found that protein size and pI were good predictors for *P. falciparum* protein expression and solubility. Larger proteins were more difficult to express, a high pI caused insolubility and a low pI decreased expression but increased solubility. In addition, proteins with a relatively high homology to *E. coli* proteins were more likely to express in soluble form. Vedadi *et al*. (2007) however found that protein size and pI could not be used to predict the successful expression of soluble Apicomplexan proteins.

The solubility of some proteins can be enhanced by modifying the recombinant protein by tag-technology or by site-directed mutagenesis of protease-specific sites or other regions that destabilise the protein structure. Tag-technology attaches a reporter protein, for example the green fluorescent protein (Pedelacq *et al.*, 2006), or a solubility tag, such as the maltose binding protein (MBP) (Kapust and Waugh, 1999), the N-utilising substance A (NusA) (Zheng *et al.*, 2003), or glutathione-S-transferase (GST), to the recombinant protein. These tags may be used to purify the soluble fusion protein and it is assumed that a correctly folded tag suggests a correctly folded recombinant protein.

DNA and mRNA modifications

Fifty-four percent of *P. falciparum* genes contain introns (Gardner *et al.*, 2002), requiring the use of cDNA for cloning the full length gene. Due to the 80.6 % AT-richness of the genome (Gardner *et al.*, 2002), parasite genes are characterised by long, continuous stretches of adenosines and thymidines, making error-free PCR amplification of the genes problematic. The use of a high fidelity DNA polymerase mix, which contains *Tgo* and *Taq* DNA polymerase can decrease the introduction of errors during amplification, because the *Tgo* DNA polymerase has a 3'-5' exonuclease proofreading activity that increases the fidelity of DNA synthesis ~3-fold (Roche Applied Science, 2004) when compared to the fidelity of

Taq DNA polymerase, which has a total error rate of 1 x 10^{-4} to 2 x 10^{-5} errors per base pair (Tindall and Kunkel, 1988).

The AT-richness of the *P. falciparum* genome also interferes with protein expression in *E. coli* in other ways. For example, AT-richness results in a different codon usage in the parasite when compared to the expression host (Appendix A3). Certain *E. coli* tRNAs, that correspond to amino acids that are abundant in *P. falciparum* proteins, may therefore only be available in limited amounts. This can cause ribosomal stalling, which leads to translational errors such as amino acid substitutions, frameshifts, and premature translation termination (Sorensen and Mortensen, 2005a). A high content of AT-repeats within the mRNA template can also cause the ribosomal unit to stall (Flick *et al.*, 2004). *E. coli* sporadically substitutes amino acids located in the low-complexity regions of the *P. falciparum* sequence, resulting in expressed proteins which appear homogenous by SDS-PAGE, but some molecules have the incorrect amino acid sequence (Schneider *et al.*, 2005). However, both *P. falciparum* protein expression studies mentioned previously (Mehlin *et al.*, 2006, Vedadi *et al.*, 2007) stated that codon usage and AT-content had minimal effect on protein expression.

Other factors may also complicate protein production. There is evidence that *P. falciparum* proteins bind their own mRNA (Zhang and Rathod, 2002), thereby inhibiting protein expression. Parasite genes also often contain cryptic *E. coli* start sites (Turgut-Balik *et al.*, 2001) that cause the expression of truncated proteins.

Several strategies can be used to overcome codon bias and ribosomal stalling. One strategy, termed codon optimisation, uses gene synthesis to create a *P. falciparum* sequence with codons reflecting the tRNA pool of the host *E. coli* system. This approach should theoretically increase expression levels and alleviate the occurrence of translational errors (Calderone *et al.*, 1996). It can however also influence the stability of the mRNA, thereby reducing transcription levels (Wu *et al.*, 2004). Mehlin *et al.* (2006) successfully used codon optimisation to express 12 *P. falciparum* proteins that did not express in *E. coli* from their native genes,

but all the proteins were insoluble. In contrast, codon harmonisation relies on the modification of the *P. falciparum* sequence by site-directed mutagenesis to match the codon usage of *E. coli*, but instead of replacing all the codons, only codons that interfere with protein translation (i.e. where the ribosomal unit stalls) and folding are replaced (Hillier *et al.*, 2005). Substitution of only a few codons may enable translation in *E. coli* to occur in the same manner as translation in the parasite. Codon optimisation and harmonisation can be very time-consuming and expensive and therefore the use of other strategies is often preferred. For example, an *E. coli* host that has been co-transformed with a plasmid that contains the genes encoding the rare *E. coli* tRNAs can be used (Calderone *et al.*, 1996).

Manipulation of the expression host

Expressed proteins will only be soluble when the correct structure is formed during the post-translational folding process. Newly synthesised polypeptides remain in an intermediate folded form within the bacterial cytoplasm until chaperones fold the proteins into their functional forms (reviewed in Mogk *et al.*, 2002). Incorrectly folded proteins accumulate as aggregates within the cell and tend to be toxic. To cope with this toxicity, *E. coli* stores the aggregates as confined structures termed inclusion bodies (reviewed in Mogk *et al.*, 2002). In general, eukaryotic proteins are more prone to inclusion body formation within *E. coli*, because the bacterium does not possess the appropriate redox environment to introduce the correct disulphide bonds and post-translational modifications into the recombinant proteins (Weickert *et al.*, 1996).

Protein solubility within the *E. coli* host can be enhanced by co-expressing the protein with a plasmid encoding additional *E. coli* chaperones, some of which drive folding attempts and others that prevent protein aggregation (reviewed in Mogk *et al.*, 2002). Inclusion body formation can also be inhibited or minimised by controlling the temperature and the expression rate of the host cell. Growing *E. coli* at 37 ºC has been shown to increase the formation of inclusion bodies, because the hydrophobic interactions causing protein aggregation are favoured at

these temperatures (Kiefhaber *et al.*, 1991). Lowering the temperature on the other hand, favours the slower and correct folding of the protein, resulting in the formation of fewer inclusion bodies. The expression rate of the host cell can be controlled with promoters that tightly regulate transcription and prevent 'leaky' expression of the protein (Weickert *et al.*, 1996). Gradually increasing the level of inducer, using a natural inducing agent, inducing the cultures at post-log phase (Flick *et al.*, 2004), or slowly increasing the plasmid copy number during culturing (Trepod and Mott, 2002) can also be employed to control expression rate. These methods decrease the induction level, which in turn reduces the protein concentration within the cell, which favours correct protein folding to produce higher amounts of soluble protein (Weickert *et al.*, 1996).

However, if soluble recombinant protein expression in *E. coli* is not successful, another host, for example the yeast *Pichia pistoria* (Yokoyama, 2003), can be used. Expression from a baculovirus vector in insect or mammalian cells (Kost *et al.*, 2005) provides another alternative. Mehlin *et al*. (2006), for example, used 17 *P. falciparum* proteins, which expressed as insoluble proteins in *E. coli* and expressed them in the baculovirus-insect cell system. All these proteins were expressed as soluble recombinant proteins, but only one showed a high yield. Cell-free protein synthesis, based on the coupled transcription-translation system of *E. coli* or wheat embryos (Yokoyama, 2003), could also be used in an attempt to express soluble recombinant *P. falciparum* proteins.

3.1.2 Chemical processing of inclusion bodies

If none of the above mentioned strategies yield soluble recombinant *P. falciparum* proteins, the inclusion bodies can be processed directly to obtain soluble proteins. This method involves solubilising inclusion bodies with urea or guanidinium hydrochloride and subsequently refolding the denatured protein *in vitro* by dilution, dialysis or on-column refolding methods (reviewed in Mogk *et al.*, 2002). Protein refolding is however time consuming and there is no guarantee that the protein has been refolded into its native state. In addition, protein refolding is not

always conducive to high product yields, and in the case of refolded enzymes the activity may also be lost.

3.2 Methods

3.2.1 The spectrin-binding *P. falciparum* proteins

Table 9 lists the nucleotide regions of MAL8P1.104, MAL8P1.151, PFA0125c and PFE0980c that were used for cloning. The spectrin binding region lies in the middle of these ~1.5 kb fragments and because these genes are large, only one segment of each gene (excluding introns) was cloned. In contrast, most of the PFI1570c gene (lacks 2 & 8 amino acids at the N- & C-terminus respectively) was used for the cloning procedure, because it is fairly small and consists of one exon.

Table 9: DNA fragments used for cloning.

PlasmoDB gene name	Size of gene (bp)	Introns present in gene	Spectrin binding region	Cloning region
PFI1570c	1713	no	646-771	7-1689
MAL8P1.151	4092	yes	2751-2811	2578-4056
PFA0125c	4704	yes	2914-3060	2182-3750
MAL8P1.104	5325	no	3151-3561	2449-4155
PFE0980c	7467	yes	1261-1521	1156-2670

3.2.2 Cloning

Isolation of P. falciparum genomic DNA

DNA was isolated from *P. falciparum* FCR-3 cultures by a method adapted from Hang *et al*. (1995), which entails lysing the erythrocytes and subsequently rupturing the parasites by boiling. Two hundred microlitres whole parasite suspension (Section 2.2.3; removed after the parasites were resuspended in fresh medium) were combined with 400 µl lysis buffer (0.2 % NaCl (w/v), 1 % Triton® X-100 (v/v) (BDH Laboratory Supplies, Poole, UK), 1 mM EDTA) and vortexed to ensure complete lysis of the erythrocytes. The sample was centrifuged at 13,000 *g* for 10 minutes at 4 ºC in an Eppendorf 5415 R Centrifuge (Eppendorf

AG, Hamburg, Germany). The supernatant was removed and the pellet washed twice by resuspension in 100 µl wash buffer (10 mM Tris-HCl, 50 mM KCl, pH 8.3), brief vortexing and centrifugation at 16,000 *g* for 5 minutes at 4 ºC. One hundred microlitres of PCR buffer (10 mM Tris-HCl, 50 mM KCl, 1.5 mM $MgCl_2$, pH 8.3) was added to the pellet and the sample boiled for 10 minutes. The concentration of the DNA was determined by spectrophotometric analysis at 260 nm (as described in Ausubel *et al.*, 1994).

Amplification of P. falciparum gene fragments

P. falciparum gene segments were amplified using primers containing appropriate restriction enzyme sites to facilitate cloning of the genes into the pET-15b (Novagen, Inc., Madison, USA) or pGEX-4T-2 (Amersham Biosciences, Ltd., Buckinghamshire, UK) expression vectors (vector maps in Appendix A2). The primers (Figure 33) contained 19-24 nucleotides matching the *P. falciparum* gene, 6 nucleotides at the 5' end matching the *Bam*HI, *Nde*I or *Xho*I restriction sites, followed by 5-6 nucleotides 5' to the *Bam*HI and *Xho*I or ~11 nucleotides 5' to the *Nde*I restriction site to ensure complete digestion by the restriction enzymes. All the downstream primers, except PFS16, contained an additional stop codon at the end of the *P. falciparum* sequence. PFI1570c was amplified with primers for both vectors, while MAL8P1.104 and PFA0125c were amplified with primers for pET-15b, and MAL8P1.151 and PFE0980c were amplified with primers for pGEX-4-T2. The four latter sequences could not be inserted into both vectors because they contained either *Xho*I or *Nde*I restriction sites within their gene sequence.

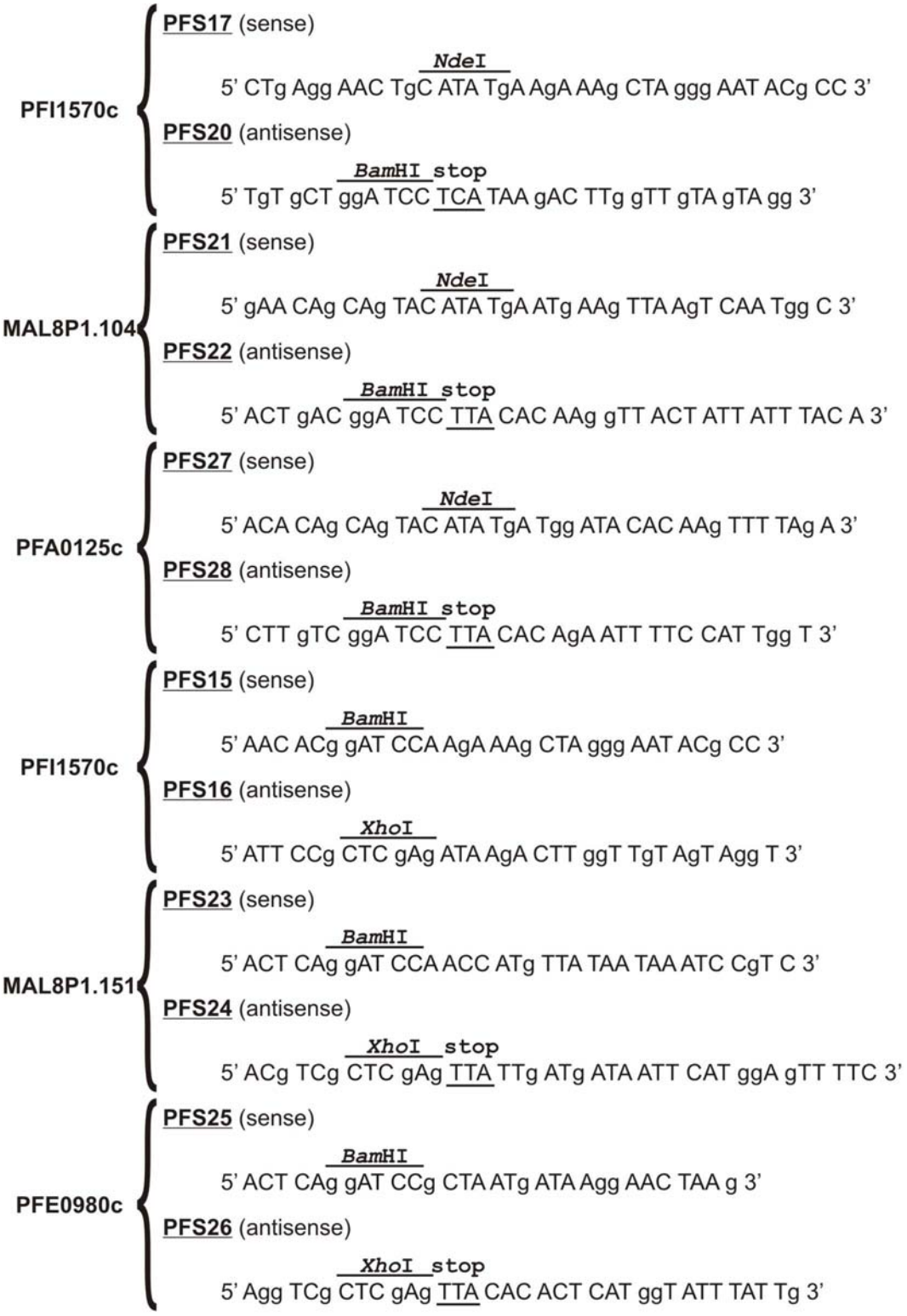


Figure 33: *P. falciparum* primers.
Diagram showing the primers used to amplify the *P. falciparum* genes. The restriction sites and stop codon sites are marked. *Nde*I and *Bam*HI sites were used to clone the DNA into pET-15b (primers PFS17/PFS20, PFS21/PFS22, PFS27/PFS28) and *Bam*HI and *Xho*I were used to clone the DNA into pGEX-4T-2 (primers PFS15/PFS16, PFS23/PFS24, PFS25/PFS26).

Approximately 500 ng *P. falciparum* DNA was amplified with 300 nM of each primer (Inqaba Biotechnical Industries, Pretoria, South Africa) in the presence 1 x High Fidelity PCR Master (Roche Diagnostics Gmbh, Mannheim, Germany) under the conditions listed in Table 10 and Table 11 using the Eppendorf Mastercycler® Gradient machine (reaction volume = 50 µl). The amplification reaction was divided into two segments to allow initial annealing of the primer to the *P. falciparum* specific region in the first five cycles, followed by 30 cycles with a higher annealing temperature, which was determined from the lower Tm (given by the oligonucleotide manufacturer) of each primer pair. The extension step was performed at 68 ºC instead of 72 ºC to slow down the activity of the *Taq* DNA polymerase so that it can cope with the AT-richness of the *P. falciparum* gene sequences. Three microlitres of the amplicons were visualised against a 1 kb DNA Ladder (Promega Corporation, Madison, USA) on a 1 % agarose gel in TEA buffer. The remaining sample was subjected to phenol-chloroform extraction and ethanol precipitation (as described in Ausubel *et al.*, 1994) and resuspended in 45 µl nuclease-free water.

Table 10: Annealing temperatures for *P. falciparum* gene specific primer sets.

P. falciparum gene	Primer set	1st Annealing temperature	2nd Annealing temperature	Vector
PFI1570c	PFS15/PFS16	57 ºC	68 ºC	pGEX-4T-2
PFI1570c	PFS17/PFS20	57 ºC	68 ºC	pET-15b
MAL8P1.104	PFS21/PFS22	46 ºC	65 ºC	pET-15b
MAL8P1.151	PFS23/PFS24	58 ºC	65 ºC	pGEX-4T-2
PFE0980c	PFS25/PFS26	54 ºC	65 ºC	pGEX-4T-2
PFA0125c	PFS27/PFS28	47 ºC	65 ºC	pET-15b

Table 11: Cycling parameters for *P. falciparum* genomic DNA amplification.

Segment	Cycles	Temperature	Time
1	1	94 ºC	2 min
2	1-5	94 ºC	45 sec
		1st annealing temp	1 min
		68 ºC	2 min
3	1-30	94 ºC	45 sec
		2nd annealing temp	1 min
		68 ºC	2 min
4	1	68 ºC	7 min

Restriction enzyme digestion of the PCR products

The 45 µl DNA samples were digested overnight at 37 ºC in the presence of 20 units *Nde*I (Promega Corporation, Madison, USA) (for ligation into pET-15b) or *Xho*I (Promega Corporation, Madison, USA) (for ligation into pGEX-4T-2) with 1 x Buffer D (Promega Corporation, Madison, USA) and 10 µg acetylated Bovine Serum Albumin (Promega Corporation, Madison, USA). The sample was subjected to phenol-chloroform extraction followed by ethanol precipitation and resuspension in 45 µl nuclease-free water. The sample was digested at 37 ºC for 4 hours with 20 units *Bam*HI (Promega Corporation, Madison, USA) in the presence of 1 x Buffer E (Promega Corporation, Madison, USA) and 10 µg acetylated bovine serum albumin (BSA). Phenol-chloroform extraction and ethanol precipitation were repeated and the DNA resuspended in 5 µl nuclease-free water.

Isolation and restriction enzyme digestion of the plasmid vector

Plasmids were isolated from *E. coli* by alkaline lysis (as described in Ausubel *et al.*, 1994). *E. coli* containing the pET-15b vector or pGEX-4T-2 vector were grown in 10 ml LB medium, containing 50 µg/ml ampicillin, overnight at 37 ºC in a Labotec® Orbital Shaker (225 rpm). Two millilitre aliquots of the cells were centrifuged at 16,000 *g* for 2 minutes at room temperature in an Eppendorf 5415 R Centrifuge. The pellets were resuspended in 100 µl cold 50 mM glucose, 10 mM Na_2EDTA, 25 mM Tris-HCl, pH 8 and incubated at room temperature for 5 minutes. The EDTA chelates divalent metal ions, primarily magnesium and calcium. Removal of these cations destabilises the cell membrane and also inhibits DNases. Glucose maintains the osmolarity and prevents the buffer from bursting the cells. Two hundred microlitres 0.2 M NaOH, 1 % SDS (w/v) (prepared on the day of use) was added to the suspension and mixed by carefully tapping the tube. The SDS detergent ruptures the *E. coli* plasma membrane, while the NaOH loosens the cell walls and denatures the bacterial and plasmid DNA. The bacterial DNA becomes linearised and the circular plasmid DNA remains topologically

constrained. One hundred and fifty microlitres 3.3 M acetic acid, 1.6 M potassium acetate (prepared on the day of use) was added, the sample vortexed and placed on ice for 15 minutes, followed by centrifugation at 16,000 *g* for 5 minutes at 4 ºC. Potassium acetate renatures the circular plasmid DNA, while the linear single stranded bacterial DNA precipitates in the high salt environment. SDS forms a precipitate with the potassium acetate, allowing it to be removed from the plasmid DNA. After phenol-chloroform extraction and ethanol precipitation (no addition of sodium acetate), the pellet was resuspended in 100 µl Tris EDTA buffer (TE buffer; 1 mM Na_2EDTA, 10 mM Tris-HCl, pH 7.5) containing 5 µg DNase-free Ribonuclease A (Fermentas International Inc., Burlington, Canada) and incubated at 37 ºC for half an hour before addition of 40 µl 3.3 M acetic acid, 1.6 M potassium acetate. Ribonuclease A digests any remaining RNA in the plasmid DNA sample. The sample was subjected to phenol-chloroform extraction and ethanol precipitation after which the DNA pellets were pooled and resuspended in 175 µl nuclease-free water. The pET-15b vector was digested with *Nde*I and *Bam*HI and the pGEX-4T-2 vector was digested with *Xho*I and *Bam*HI. A 1 hour incubation at 37 ºC in the presence of 5 units calf intestinal alkaline phosphatase (Roche Diagnostics Gmbh, Mannheim, Germany) was performed before the final phenol-chloroform extraction step to prevent religation of the vector. Both vectors were stored in a final volume of 200 µl nuclease-free water.

Ligation of the PCR fragments into the expression vector

The amplified DNA was inserted into the vector with the Rapid Ligation Kit (Roche Diagnostics Gmbh, Mannheim, Germany). One microlitre DNA and 0.5 µl vector were visualised against 10 µl MassRuler™ DNA Ladder, Mix (Fermentas International Inc., Burlington, Canada) on a 1 % agarose gel in TEA buffer to determine the concentration of both samples. A 1:4 molar ratio of vector to insert was diluted in 1 x DNA Dilution Buffer (Roche Diagnostics Gmbh, Mannheim, Germany) (volume = 10 µl; total amount of DNA ≤200 ng). Five units T4 DNA Ligase and 1 x DNA Ligation Buffer were added to the DNA and the mixture incubated at 16 ºC for 30 minutes in the Eppendorf Mastercycler® Gradient

machine (reaction volume = 21µl). A second ligation reaction without amplified DNA was set up as a control reaction.

Transformation of DH5α™ E. coli or BL21-Codon Plus® (DE3)-RIL E. coli cells

DH5α™ *E. coli* cells (Invitrogen Ltd, Renfrewshire, UK) and BL21-Codon Plus® (DE3)-RIL *E. coli* cells (Stratagene, La Jolla, USA) were made competent with PIPES buffer. Ten millilitres LB medium (containing 50 µg/ml chloramphenicol (Sigma-Aldrich Corporation, St. Louis, USA) for the BL-21 cells) was inoculated with either DH5α™ or BL21-Codon Plus® (DE3)-RIL *E. coli* cells and grown overnight at 37 ºC in a Labotec® Orbital Shaker (225 rpm). The 10 ml culture was used to inoculate 100 ml LB medium (containing 50 µg/ml chloramphenicol for the BL-21 cells) and grown for another further 90 minutes at 37 ºC with shaking. The cells were centrifuged at 4,000 *g* for 15 minutes at 4 ºC using the Beckman® J2-21 centrifuge and the spent medium removed. The pellet was resuspended in 10 ml cold sterile PIPES buffer (10 mM 1,4-piperazinediethanesulphonic acid (Sigma-Aldrich Corporation, St. Louis, USA), 100 mM $CaCl_2$, 15 % glycerol (v/v), pH 7) and incubated on ice for 20 minutes. The cells were centrifuged at 4,000 *g* for 10 minutes at 4 ºC and the supernatant discarded. The pellet was resuspended in 2 ml PIPES buffer and the cells stored at -70 ºC as 100 µl aliquots. Competent DH5α™ *E. coli* cells were transformed with the ligation reaction. Plasmids with the correct DNA inserts were subsequently used to transform competent BL21-Codon Plus® (DE3)-RIL *E. coli* cells for the expression studies. This BL21 (DE3) strain was chosen because it contains the ColE1-compatible, pACYC-based plasmid with extra copies of the *argU*, *ileY*, and *leuW* tRNA genes (Stratagene, 2005).

Five microlitres ligation reaction or 3 µl purified plasmid was added to 25 µl competent *E. coli* cells and mixed gently by tapping the tube. The sample was incubated on ice for 30 minutes and subsequently placed in a 42 ºC waterbath for 1½ minutes, followed by a 5 minute incubation on ice. The cell suspension was

spread on agar plates containing either 50 µg/ml ampicillin for the DH5α™ *E. coli* cells, or 50 µg/ml ampicillin and 50 µg/ml chloramphenicol for the BL21-Codon Plus® (DE3)-RIL *E. coli* cells. The plates were inverted and incubated overnight at 37 ºC in a Heraeus incubator (Heraeus Instruments, Hanau, Germany).

Restriction enzyme analysis of the plasmid constructs

DH5α *E. coli* colonies containing plasmids were placed into 14 ml Falcon® tubes (Becton Dickinson Labware, Franklin Lakes, USA) containing 2 ml LB medium, supplemented with 50 µg/ml ampicillin, and grown at 37 ºC overnight in a Labotec® Orbital Shaker (225 rpm). The plasmids were extracted from the DH5α *E. coli* cells with the FastPlasmid™ Mini Kit (Eppendorf AG, Hamburg, Germany). This kit is based on the alkaline lysis method and isolates ~20 µg high-copy plasmid DNA from a 1.5 ml *E. coli* culture. A single solution lyses bacterial cells, denatures and solubilises cellular components, degrades RNA, and traps DNA on a matrix. The resuspended cell mixture is transferred to a mini spin column that binds the plasmid DNA, which is subsequently washed with an alcohol-based solution to remove any remaining impurities. The plasmid DNA is eluted from the column with 50 µl low-salt elution buffer. Ten microlitres of the pET-15b-*P. falciparum* gene construct was digested with 10 units *Sty*I (Fermentas International Inc., Burlington, Canada) and 10 units *Eco*RI (Fermentas International Inc., Burlington, Canada) in 2 x Tango™ Buffer (Fermentas International Inc., Burlington, Canada) for 1 hour at 37 ºC. Double digestion of 10 µl pGEX-4T-2-*P. falciparum* was performed with 10 units *Eco*RV (Fermentas International Inc., Burlington, Canada) and 10 units *Eco*RI with the identical buffer and incubation conditions. Ten microlitres of the reactions were electrophoresed against the Promega 1 kb DNA Ladder and the Promega 100 bp Ladder on a 1 % agarose gel in TEA buffer. Constructs showing the correct digestion pattern were sequenced by Inqaba Biotechnical Industries (Pretoria, South Africa) and used for transformation of BL21-Codon Plus® (DE3)-RIL cells.

3.2.3 Protein expression

Expression in pET-15b and pGEX-4T-2 is controlled by different promoters, namely the tac promoter in pGEX-4T-2 and the T7 promoter in pET-15b. Expression of the *P. falciparum* proteins from the pET-15b plasmid can therefore only occur in *E. coli* cells that have the T7 RNA polymerase gene inserted into their genome. BL21-Codon Plus® (DE3)-RIL *E. coli* cells containing the *P. falciparum* constructs were grown in 10 ml LB medium containing 200 µg/ml ampicillin and 200 µg/ml chloramphenicol overnight at 37 ºC using a Labotec® Orbital Shaker prior to induction with isopropylthiogalactoside (IPTG, a nonmetabolisable inducing agent; Promega Corporation, Madison, USA) or Overnight Express™ Instant TB Medium (Novagen, Inc., Madison, USA). Overnight Express™ medium is a modified terrific broth that allows the bacteria to autoinduce in the presence of lactose (secondary energy source) once the glucose (primary energy source) from the medium has been depleted by the bacteria. Induction with IPTG was performed by inoculating 48 ml LB broth with 2 ml of the *P. falciparum* construct-containing culture and growing the cells for 4 hours at 37 ºC with shaking. The culture was induced with 1 mM IPTG and grown for 15 hours at 37 ºC with shaking. Induction with Overnight Express™ Instant TB Medium was performed by inoculating 240 ml Overnight Express™ Instant TB Medium with the 10 ml *P. falciparum* construct-containing culture and growing the cells at room temperature for 24 hours using a Red Rotor PR70-230V shaker (Hoefer Scientific Instruments, San Francisco, USA). Whole cell extracts were analysed by separating the proteins on a 10 % Laemmli SDS-polyacrylamide gel, followed by western transfer with Towbin buffer (25 mM Tris, 192 mM glycine, 20 % methanol (v/v), pH 8.3) (Towbin *et al.*, 1979) onto Hybond™-C Extra Nitrocellulose membrane (Amersham Biosciences, Ltd., Buckinghamshire, UK). The fusion-tags were detected with 1:1,500 Penta•His HRP Conjugate antibody (Qiagen GmbH, Hilden, Germany) dilution or 1:120,000 Anti-GST HRP Conjugate antibody (Amersham Biosciences, Ltd., Buckinghamshire, UK) dilution and chemiluminescence with the SuperSignal® West Pico

Chemiluminescent Substrate (Pierce Biotechnology Incorporated, Rockford, USA).

3.2.4 Recombinant *P. falciparum* aminopeptidase purification

An *E. coli* culture expressing the *P. falciparum* aminopeptidase (PFI1570c) was lysed by sonication or chemical lysis with BugBuster™ HT (Novagen, Inc., Madison, USA) and the histidine-tagged protein isolated from the bacterial proteins by affinity selection with nickel coated HIS-Select™ Magnetic Agarose Beads (Sigma-Aldrich Corporation, St. Louis, USA).

A 250 ml BL21-Codon Plus® (DE3)-RIL *E. coli* culture grown in Overnight Express™ Instant TB Medium was centrifuged at 10,000 *g* for 10 minutes at 4 ºC using a Beckman® J2-21 centrifuge. The cell pellet was frozen at -70 ºC for 5 minutes and resuspended in 2 ml His-lysis/binding buffer (50 mM sodium phosphate buffer, pH 8, 150 mM NaCl) containing 1 µl Protein Inhibitor Cocktail Set III (100 mM 4-(2-Aminoethyl) benzenesulphonyl fluoride hydrochloride, 80 µM Aprotinin, 5 mM Bestatin; 1.5 mM E-64; 2 mM Leupeptin, 1 mM Pepstatin A; Calbiochem®, San Diego, USA). The suspension was sonicated 6 times on ice water with a Bandelin Sonoplus UW 2070 sonicator (Bandelin Electronics, Berlin, Germany) set to 15 seconds with 6 x 10 % cycles at 75 % power. The suspension was allowed to cool for 45 seconds between each sonication. The suspension was centrifuged at 16,000 *g* for 10 minutes at 4 ºC, using an Eppendorf 5415 R Centrifuge and the supernatant containing the soluble protein collected. Imidazole (Sigma-Aldrich Corporation, St. Louis, USA) was added to the soluble fraction to a final concentration of 1 mM and the solution added to 30 µl packed HIS-Select™ Magnetic Agarose Beads and allowed to bind for 1 hour at room temperature using the GFL® 3025 rotator (Gesellschaft für Labortechnik m.b.H. & Co., Burgwedel, Germany). The beads were placed on a magnetic separator and the supernatant removed. The beads were washed 5 times for 1 minute with 300 µl His-wash buffer (50 mM sodium phosphate buffer, pH 8, 150 mM NaCl, 20mM imidazole, 1 mM PMSF) and the protein eluted from the beads by

incubating them in 150µl His-elution buffer (50 mM sodium phosphate buffer, pH 8, 150 mM NaCl, 200mM imidazole, 1 mM PMSF) for 15 min with occasional shaking by hand. A second elution was performed with 100 µl His-elution buffer. The supernatant containing the purified protein was collected and stored at 4 ºC until further use. Purification was monitored by separating the protein fractions on a 12 % Laemmli SDS-polyacrylamide gel, western transfer with Towbin buffer onto Hybond™-C Extra Nitrocellulose membrane and detection of the histidine-tag with 1:1,500 Penta•His HRP Conjugate antibody dilution and chemiluminescence with the SuperSignal® West Pico Chemiluminescent Substrate.

3.3 Results

3.3.1 Cloning and expression of spectrin-binding *P. falciparum* proteins

Segments of the genes coding for the five spectrin-binding peptides were used for the cloning and expression procedure. Figure 34 shows the linearised vector DNA and the PCR products prepared by double digestion with *Nde*I/*Bam*HI (pET-15b) or *Xho*I/*Bam*HI (pGEX-4T-2). PFI1570c was cloned into both vectors because the gene did not contain *Xho*I or *Nde*I digestion sites and there was a one base pair difference between the two PFI1570c amplicons because different primers were used for amplification.

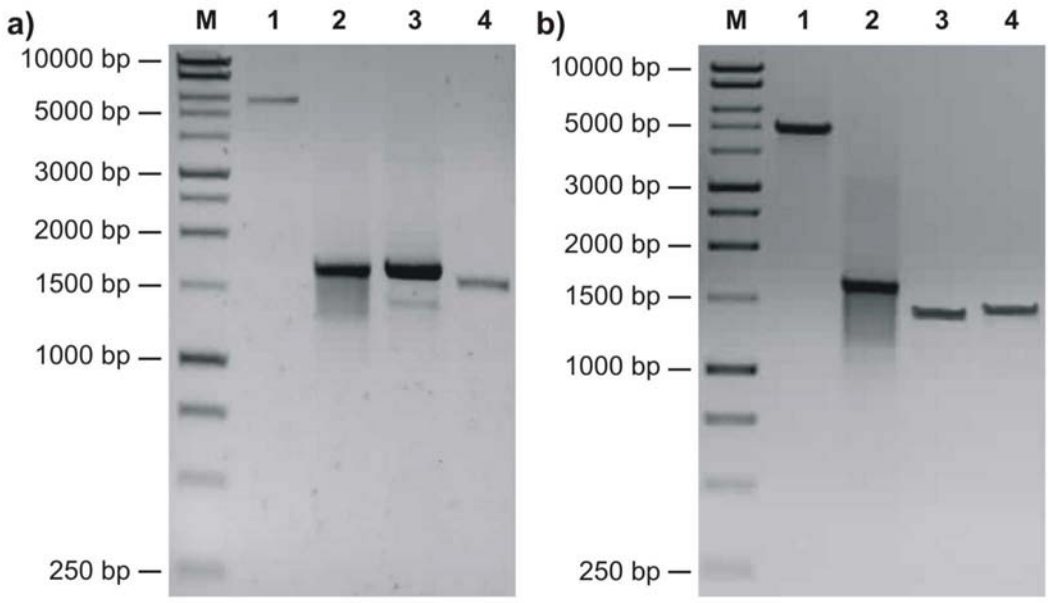


Figure 34: Vectors and *P. falciparum* amplicons before ligation.
Agarose gels showing (a) the *Nde*I/*Bam*HI digestion of the pET-15b constructs and (b) the *Xho*I/*Bam*HI digestion of the pGEX-4T-2 constructs. Lane M in both gels contains 10 µl Promega 1 kb Ladder. Gel (a) shows: 0.5 µl pET-15b plasmid (a1, 5696 bp); 1 µl of the *P. falciparum* genes in pET: PFI1570c (a2, 1691 bp), MAL8P1.104 (a3, 1715 bp), and PFA0125c (a4, 1577 bp). Gel (b) shows: 0.5 µl pGEX-4T-2 plasmid (b1, 4945 bp); 1 µl of the *P. falciparum* genes in pGEX: PFI1570c (b2, 1690 bp), MAL8P1.151 (b3, 1485 bp), and PFE0980c (b4, 1524 bp).

The presence of the correct inserts in the vector plasmids was confirmed by a double digestion of the plasmid constructs. The vector maps and the unique DNA patterns after the double digestion can be seen in Figure 35 and Figure 36. All the pET-15b clones (Figure 35b) digested with *Eco*RI and *Sty*I, have three DNA bands in common, namely 246 bp (very faint on the gel), 2325 bp, and 2992 bp. The different *P. falciparum* DNA inserts give rise to the additional two DNA bands because they all contain an internal *Eco*RI site. The 145 bp fragment generated from pET-15b digestion is too small to be visible. The pGEX-4T-2 clones (Figure 36b) digested with *Eco*RI and *Eco*RV do not have DNA fragments in common because the *Eco*RI restriction site located in the pGEX-4T-2 cloning cassette was removed from the vector when the *P. falciparum* amplicons were inserted into the plasmid.

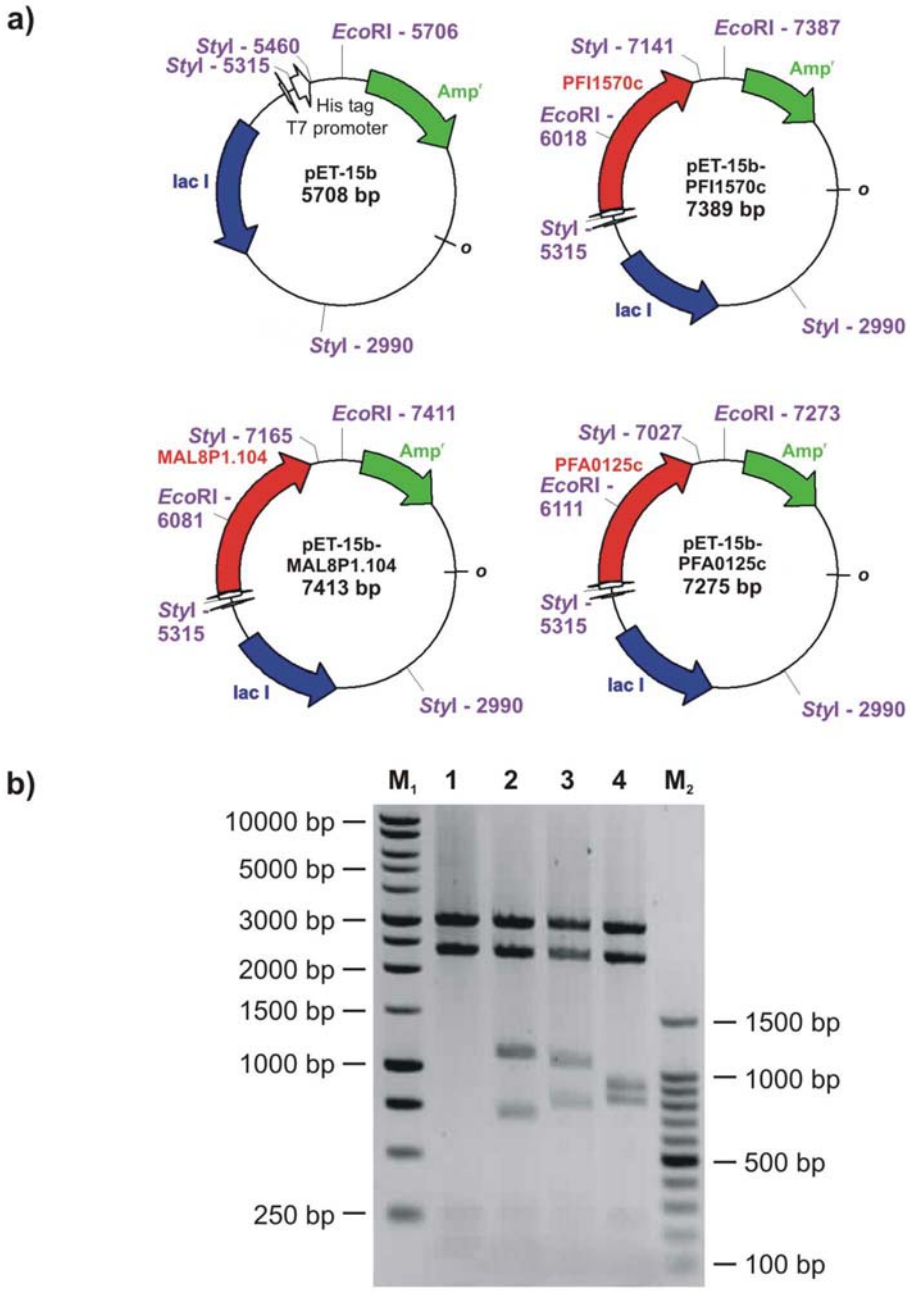


Figure 35: pET-15b vector containing *P. falciparum* inserts.
Vector maps (a) and an agarose gel showing (b) the *Eco*RI/*Sty*I double digestion of pET-15b plasmid without and with the *P. falciparum* inserts. All the constructs share the 246, 2325 and 2992 bp fragments from pET-15b. The *P. falciparum* inserts have two DNA fragments that result from an internal *Eco*RI restriction site. Vector maps: red = *P. falciparum* gene; white = His-tag; green = ampicillin resistance gene; blue = lac repressor gene; purple = restriction enzyme sites. Agarose gel: lanes M_1 and M_2 = 10 µl Promega 1 kb and 100 bp Ladder; lanes 1-4 = 10 µl of the double digests of each construct, namely: 1 = pET-15b; 2 = pET-15b-PFI1570c (703 & 1123 bp); 3 = pET-15b-MAL8P1.104 (766 & 1084 bp); 4 = pET-15b-PFA0125c (796 & 916 bp).

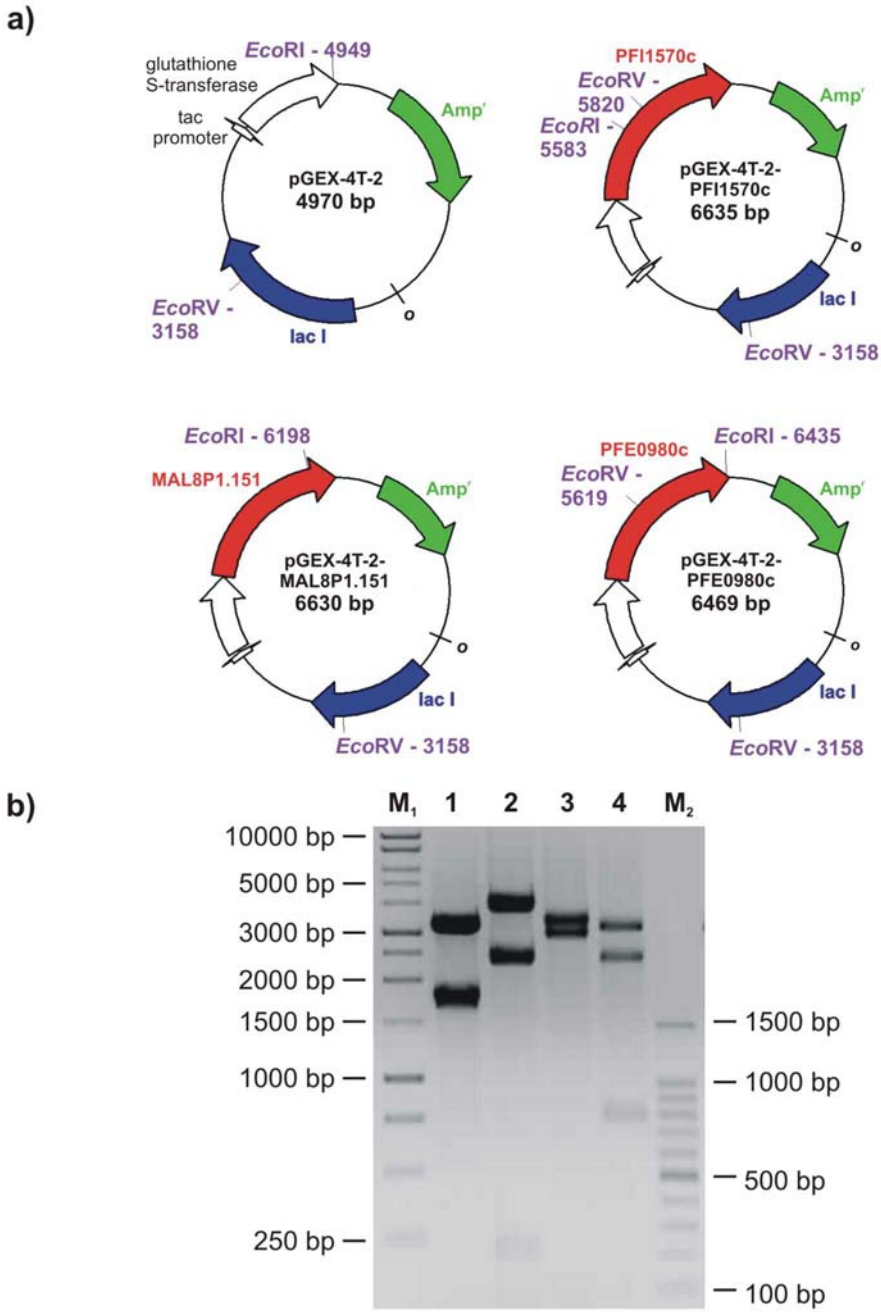


Figure 36: pGEX-4T-2 vector containing *P. falciparum* inserts.
Vector maps (a) and an agarose gel showing (b) the *Eco*RI/*Eco*RV double digestion of pGEX-4T-2 plasmid without and with the *P. falciparum* inserts. pGEX-4T-2 *Eco*RI-4949 is part of the cloning cassette and is therefore not present in the *P. falciparum* constructs. The *P. falciparum* constructs contain internal *Eco*RI and/or *Eco*RV sites. Vector maps: red = *P. falciparum* gene; white = glutathione-S-transferase; green = ampicillin resistance gene; blue = lac repressor gene; purple = restriction enzyme sites. Agarose gels: lanes M_1 and M_2 = 10 µl Promega 1 kb and 100 bp Ladder; lanes 1-4 = 10 µl of the double digests of each construct, namely: 1 = pGEX-4T-2 (1791 & 3179 bp); 2 = pGEX-4T-2-PFI1570c (2425 & 3973 bp); 3 = pGEX-4T-2-MAL8P1.151 (3040 & 3590 bp); 4 = pGEX-4T-2-PFE0980c (816, 2461 & 3192 bp).

One colony of each construct was selected for automated sequencing and transformation into *E. coli* BL-21 (DE3) cells. The pGEX-4T-2-PFI1570c clone contained no *Taq* DNA polymerase errors. The pET-15b-PFI1570c clone contained one silent error (T1542C), which still coded for lysine (Appendix A4). pET-15b-PFA0125c had a 27 nucleotide insertion at base 595, resulting in the translation of 9 extra amino acids (QEQEEEKQE). This construct also had an A503G substitution that caused the translation of an arginine instead of a glutamine. pGEX-4T-2-MAL8P1.151 and pGEX-4T-2-PFE0980c contained a deletion in a stretch of adenosines (1205-1219 and 777-788 respectively). The deletions caused a frame shift in the sequences, resulting in the introduction of an early stop codon. The deletions could however also have been a sequencing artefact because the errors occurred in a stretch of adenosines. Sequencing of pET-15b-MAL8P1.104 was not successful.

All the proteins were expressed in BL21 (DE3) *E. coli* cells by induction with IPTG. The tagged proteins could not be distinguished from the *E. coli* proteins when stained with Coomassie dye and could only be detected by probing with tag-specific antibodies (Figure 37), indicating that the proteins were not expressed at high levels within *E. coli*. Only the histidine-tagged PFI1570c was expressed as one band at the correct molecular weight of 66.919 kDa (3a). The GST-tagged PFI1570c was expressed at the correct molecular weight of 91.463 kDa (b3), but it also contained two additional bands similar in size to GST (26 kDa). The remaining constructs expressed multiple recombinant proteins (except a5) that were smaller than expected (a4-5, b4-5). The *Taq* DNA polymerase errors should have caused PFA0125c (a5), MAL8P1.151 (b4) and PFE0980c (b5) to be expressed at 61.2, 53.6 and 35 kDa (due to premature stops in the sequence), but the immunoblot showed protein bands at 43, 43 and 28 kDa respectively. This indicates that expression of these *P. falciparum* proteins in *E. coli* causes ribosomal stalling and thus early termination of translation. There is however also a possibility that not all the tagged proteins were visualised on the immunoblots because of the low protein expression levels.

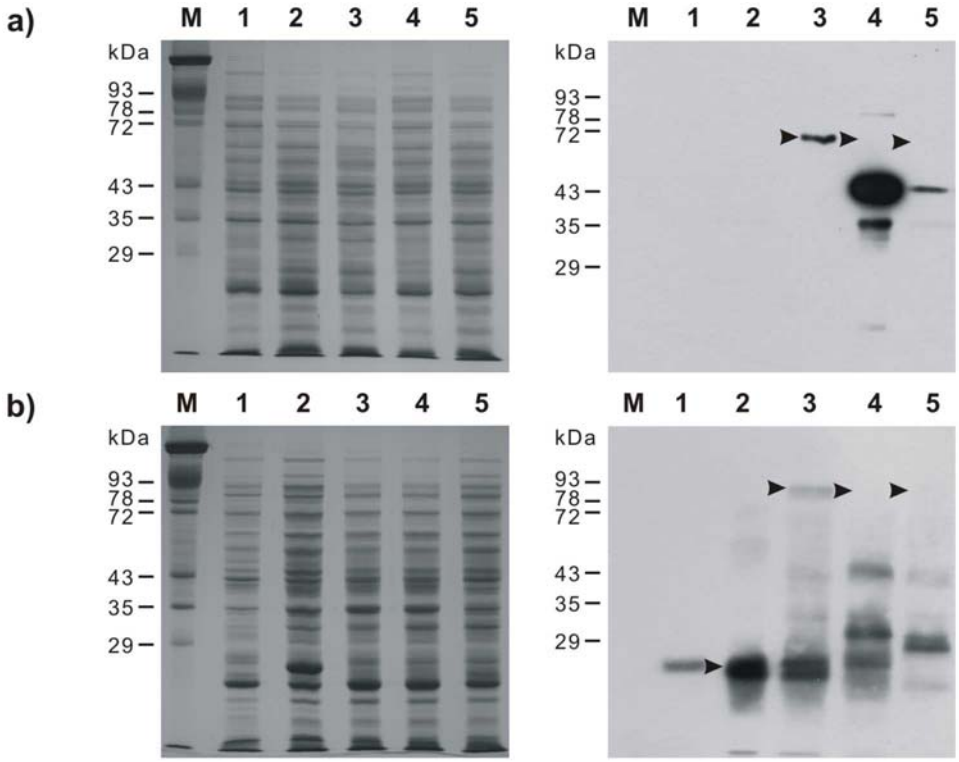


Figure 37: Expression of spectrin-binding *P. falciparum* proteins.
Laemmli gels (left) and immunoblots with the Penta•His™ and anti-GST HRP Conjugate antibodies (right) showing the whole cell extracts from *E. coli* containing the pET-15b (a) and pGEX-4T-2 (b) plasmid constructs. The black arrows indicate where the tagged proteins should occur in the immunoblot. Lane M contains 3 µg erythrocyte membrane proteins. The remaining lanes contain 10 µl BL21 (DE3) *E. coli* whole cell extracts from each construct. Lanes 1 = uninduced vector; 2 = induced vector (GST = 26 kDa); 3 = induced PFI1570c (66.9 kDa & 91.5 kDa); a4 = induced MAL8P1.104 (65.7 kDa); a5 = induced PFA0125c (60.1 kDa); b4 = induced MAL8P1.151 (86.6 kDa); b5 = PFE0980c (85.8 kDa).

A small amount of GST was present in the uninduced pGEX-4T-2 protein sample (Figure 37b1) indicating that low level transcription of the tag occurs when the cells have not been induced. Leaky expression from the pGEX-4T-2 plasmids occurs because transcription of the gene of interest is performed by an inherent *E. coli* RNA polymerase and the transcription levels are controlled by only one promoter, namely the tac promoter. Expression from the pET plasmid is controlled by two T7 promoters. The first promoter controls the expression of T7 RNA polymerase which has been cloned into the *E. coli* genome. This T7 RNA polymerase is required for the transcription of the gene from the pET-15b plasmid, which is controlled by the second T7 promoter (Novagen, 2003). The regulation

of expression in pET-15b is therefore much more controlled than that of pGEX-4T-2. The histidine-tag of the pET-15b plasmid can not be seen on the polyacrylamide gel or the immunoblot, because it is very small and therefore runs off the gel. Leaky expression of this tag could therefore also occur in *E. coli*.

3.3.2 Purification of the *P. falciparum* aminopeptidase

Only the *P. falciparum* aminopeptidase (pET-15b-PFI1570c) was expressed as a single band at the correct molecular weight in the expression study and was therefore chosen for purification. Aminopeptidase expression was initially induced with 0.4 and 1 mM IPTG at 37 ºC or room temperature, but most of the protein formed inclusion bodies (Figure 38 lane IPTG 37 ºC IS and result not shown for room temperature experiment). Therefore Overnight Express™ Instant TB Medium was used instead of IPTG. The cultures were also grown at room temperature, which improved the solubility of the aminopeptidase as seen in Figure 38 (lane ON Exp RT S).

Due to low expression levels and the large proportion of insoluble protein, large culture volumes had to be used to obtain sufficient amounts of soluble protein. In the expression, induction and solubility studies the *E. coli* cells were lysed by chemical lysis with BugBuster™ HT. This type of lysis only works effectively when the correct ratio of wet cell paste to chemical lysis reagent is used. This was not feasible for large scale expression of the aminopeptidase because the protein expressed at very low levels within the host cells and a large proportion of the protein was insoluble. Sonication, which lyses cells by shear force and cavitation, was thus applied to lyse the cells that had been resuspended in a small volume of His-lysis/binding buffer.

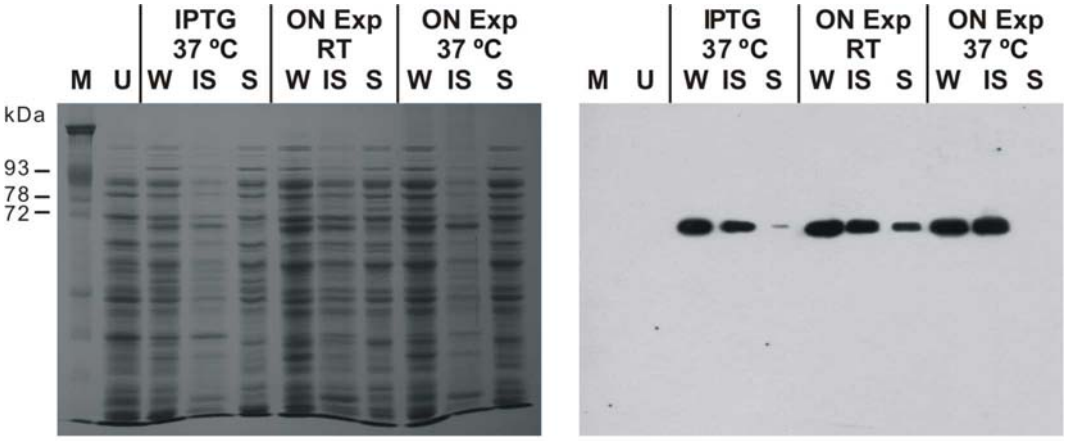


Figure 38: Induction and solubility study of recombinant *P. falciparum* aminopeptidase.
Laemmli gel (left) and immunoblot with the Penta•His™ HRP Conjugate antibody (right) showing the whole cell, insoluble and soluble protein fractions obtained from *E. coli* pET-15b-PFI1570c cultures grown and induced under different culture conditions (listed above lanes). Lane M contains 3 µg erythrocyte membrane proteins and lane U contains 10 µl of the uninduced BL21 (DE3) *E. coli* whole cell extract. The remaining lanes each contain 10 µl *E. coli* whole cell extracts (W), 10 µl of the insoluble protein fraction (IS) and 10 µl of the soluble protein fraction (S) when the cells were either grown in LB and induced with IPTG at 37 ºC, or grown in Overnight Express™ Instant TB Medium at room temperature or 37 ºC.

Figure 39 shows the progressive lysis of the BL21 (DE3) *E. coli* cells after each sonication step. BugBuster™ HT was used to completely lyse the whole cell extract and the insoluble pellets before electrophoresis on the polyacrylamide gel. The amount of aminopeptidase present in the soluble fractions increased with each sonication step. After the sixth sonication step a second band appeared below the aminopeptidase protein band (6 IS) in the insoluble fraction. Sonication generates heat, which in conjunction with the shear forces can damage the *P. falciparum* aminopeptidase as well as the protease inhibitors that were added to the His-lysis/binding buffer prior to sonication. Inactivation of the protease inhibitors could therefore have caused the reactivation of a protease that cleaves the *P. falciparum* aminopeptidase, resulting in the additional protein band in the immunoblot. The fact that this band only occurs in the insoluble fractions can however not be explained.

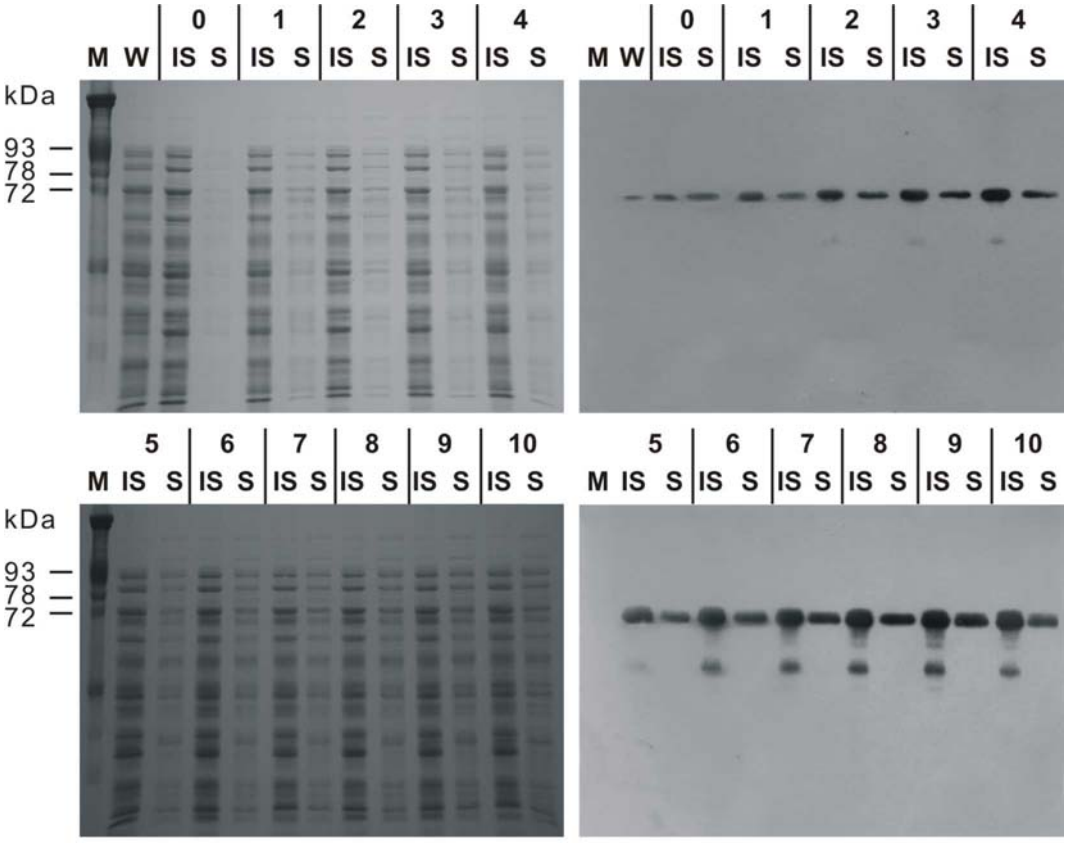


Figure 39: Sonication study of recombinant *P. falciparum* aminopeptidase. Laemmli gels (left) and immunoblots with the Penta•His™ HRP Conjugate antibody (right) showing insoluble and soluble *E. coli* pET-15b-PFI1570c cell fractions obtained after sonication. Lane M = 3 µg erythrocyte membrane proteins; lane W = 10 µl *E. coli* whole cell extracts; the remaining lanes contain 10 µl of soluble (S) and insoluble (IS) protein after each sonication step (listed above lanes).

The aminopeptidase was isolated from the *E. coli* proteins by affinity selection with the HIS-Select™ Magnetic Agarose Beads (Figure 40). To prevent degradation of the aminopeptidase a protease inhibitor mix was added to the *E. coli* cells before lysis. Twenty mM imidazole was used to remove non-binding proteins from the beads and 200 mM imidazole eluted the aminopeptidase from the magnetic beads.

Purification from a 250 ml *E. coli* culture yielded ~120 µg soluble aminopeptidase from the first elution and ~70 µg from the second elution. Only a small amount of protein was obtained at the end of the purification procedure because of the low protein expression and insolubility of the *P. falciparum* aminopeptidase. A

substantial amount of protein was also lost during the washing steps (Figure 40 1-5), especially in the last two steps, which only removed a small amount of contaminants. These two washes were not omitted because pure *P. falciparum* aminopeptidase was required for enzymatic the activity studies (see Chapter 4).

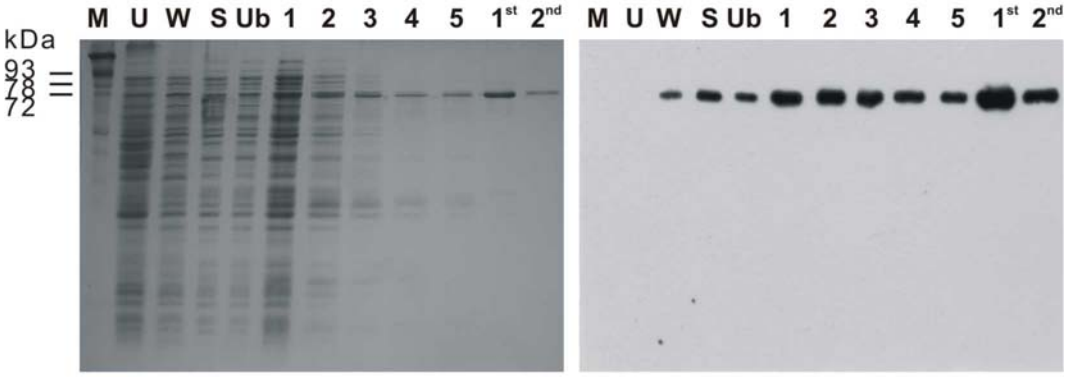


Figure 40: Affinity purification of recombinant *P. falciparum* aminopeptidase.
Laemmli gel (left) and immunoblot with the Penta•His™ HRP Conjugate antibody (right) showing the purification of *P. falciparum* aminopeptidase. Lane M = 3 µg erythrocyte membrane proteins; U & W = 10 µl uninduced and induced whole cell extract of the 10 ml *E. coli* culture; S = 10 µl of the 2 ml soluble protein fraction after six sonication steps; Ub = 10 µl of the 2 ml protein fraction that did not bind to the HIS-Select™ Magnetic Agarose Beads; 1-5 = 20 µl of each 300 µl 20 mM imidazole wash; 1^{st} = 1 µl of the first 150 µl 200 mM imidazole elution; 2^{nd} = 1 µl of the 100 µl second 200 mM imidazole elution.

3.3.3 *P. falciparum* spectrin-binding protein features

Table 12 gives some of the information available on the PlasmoDB database for the five cloned *P. falciparum* proteins. The putative Ebl-1 like protein is involved in the invasion pathway and is the only one of the five proteins that contains transmembrane domains and a hydrophobic N-terminal signal sequence. This signal sequence ensures that the protein is processed via the endoplasmic reticulum. None of the proteins contain a VTS or PEXEL motif that codes for a vacuolar transport signal, which is responsible for the translocation of the protein across the parasitophorous vacuole into the erythrocyte. According to the PlasmoDB database the putative aminopeptidase is involved in haemoglobin

digestion and one hypothetical protein, coded for by MAL8P1.104, forms part of the structure of the nuclear pore. The spectrin binding region of the putative aminopeptidase does not fall within a low complexity region in contrast to the four other proteins.

Table 12: PlasmoDB data for the spectrin-binding *P. falciparum* proteins.

Gene name	PFA0125c	MAL8P1.104	MAL8P1.151	PFE0980c	PFI1570c
Protein name	putative Ebl-1 like protein	hypothetical protein	hypothetical protein	hypothetical protein	putative amino-peptidase
PlasmoDB putative pathways	merozoite invasion	structure of the nuclear pore	none	none	haemo-globin digestion
Signal peptide	yes	no	no	no	no
Transmembrane domains	yes	no	no	no	no
Spectrin binding region in a low complexity region	yes	yes	yes	yes	no

Yeast two-hybrid data (Table 13) (LaCount *et al.*, 2005) are only available for two hypothetical proteins (MAL8P1.104 & PFE0980c) and the putative Ebl-1 like protein (PFA0125c). The putative histone acetyltransferase Gcn5 (PF08_0034) interacted with two of the spectrin-binding proteins (PFA0125c & MAL8P1.104) and is involved in transcription or chromatin metabolism. The authors also categorised PFA0125c and MAL8P1.104 in protein networks involved in host cell invasion and transcription/chromatin metabolism respectively.

Table 13: Yeast two-hybrid data for the spectrin-binding *P. falciparum* proteins.

Gene	Interacting gene	Interacting protein
PFA0125c	PF08_0034*	putative histone acetyltransferase Gcn5
	PF10_0081*	putative 26S proteasome regulatory subunit 4
	PF11_0278*	hypothetical protein
	PF13_0228#	putative 40S ribosomal subunit protein S6
	PF14_0100#	cytidine triphosphate synthetase
	PF14_0241*	putative basic transcription factor 3b
	PFC0465c*	hypothetical protein
	PFI0975c#	hypothetical protein
	PFI1090w*	putative s-adenosylmethionine synthetase
	PFI1445w#	RhopH2
	PFL0830w*	hypothetical protein
	PFL2335w#	hypothetical protein
MAL8P1.104	MAL7P1.171*	hypothetical protein
	MAL8P1.104*#	hypothetical protein
	MAL8P1.153#	hypothetical protein
	PF08_0034*	putative histone acetyltransferase Gcn5
	PF08_0060#	asparagine-rich antigen
	PF11_0241#	hypothetical protein
	PF14_0334*	putative NAD(P)H-dependent glutamate synthase
	PFA0430c#	hypothetical protein
	PFB0130w#	putative polyprenyl synthetase
	PFE0130c#	hypothetical protein
	PFE1590w#	early transcribed membrane protein
	PFF0590c*	homologue of human HSPC025
	PFI0225w*#	hypothetical protein
	PFL0275w#	hypothetical protein
	PFL0350c*	hypothetical protein
	PFL0815w#	putative DNA-binding chaperone
PFE0980c	PF11_0402#	hypothetical protein

* gene was prey and # gene was bait
(Data were obtained from LaCount *et al.*, 2005)

Table 14 and Figure 41 give the mass spectroscopy (Florens *et al.*, 2002, Florens *et al.*, 2004) and microarray data (Le Roch *et al.*, 2003) for the spectrin-binding *P. falciparum* proteins. The DeRisi Lab microarray results (Bozdech *et al.*, 2003) were very similar and are not shown. Mass spectrometry analysis of the putative aminopeptidase detected the protein at the infected red blood cell membrane, which supports the fact that this protein interacts with spectrin. The proteomics data are not as extensive as the microarray data and therefore the mRNA and protein expression profiles do not always correlate. For example, the highest

mRNA levels for MAL8P1.104 occurred in early rings and merozoites (Figure 41), but the protein was only detected in gametocytes and sporozoites. The differences in the mRNA and protein expression profiles could also be due to post-transcriptional regulation in the malaria parasite by unknown signals that delay the translation of mRNA into the protein (Le Roch *et al.*, 2004).

Table 14: Protein and mRNA expression data for the spectrin-binding *P. falciparum* proteins.

Gene name	Protein name	Life cycle stage in which protein is expressed	Life cycle stage in which mRNA is expressed at high levels
PFA0125c	putative Ebl-1 like protein	merozoite	merozoite
		sporozoite	-
		-	late schizont
MAL8P1.104	hypothetical protein	gametocyte	-
		sporozoite	-
		-	merozoite
MAL8P1.151	hypothetical protein	sporozoite	sporozoite
		-	early and late schizont
PFE0980c	hypothetical protein	gametocyte	-
		merozoite	-
		sporozoite	sporozoite
		-	early and late trophozoite
		-	early schizont
PFI1570c	putative amino-peptidase	trophozoite & schizont	early and late trophozoite
		gametocyte	gametocyte
		iRBC membrane	-
		merozoite	-
		sporozoite	-

iRBC = infected red blood cell.
Data sets are from *P. falciparum* strain 3D7 (Florens *et al.*, 2002, Florens *et al.*, 2004, Le Roch *et al.*, 2003, Bozdech *et al.*, 2003).

The levels of mRNA expression of the five spectrin-binding *P. falciparum* proteins varied considerably when compared to each other (Figure 41). For example, the putative aminopeptidase had very high mRNA expression levels, suggesting that the protein is required in large quantities and could therefore have many enzymatic functions to perform during the erythrocytic life cycle. In contrast, the PFE0980c gene expressed low levels of mRNA, suggesting that the hypothetical protein product may only have a single specific function.

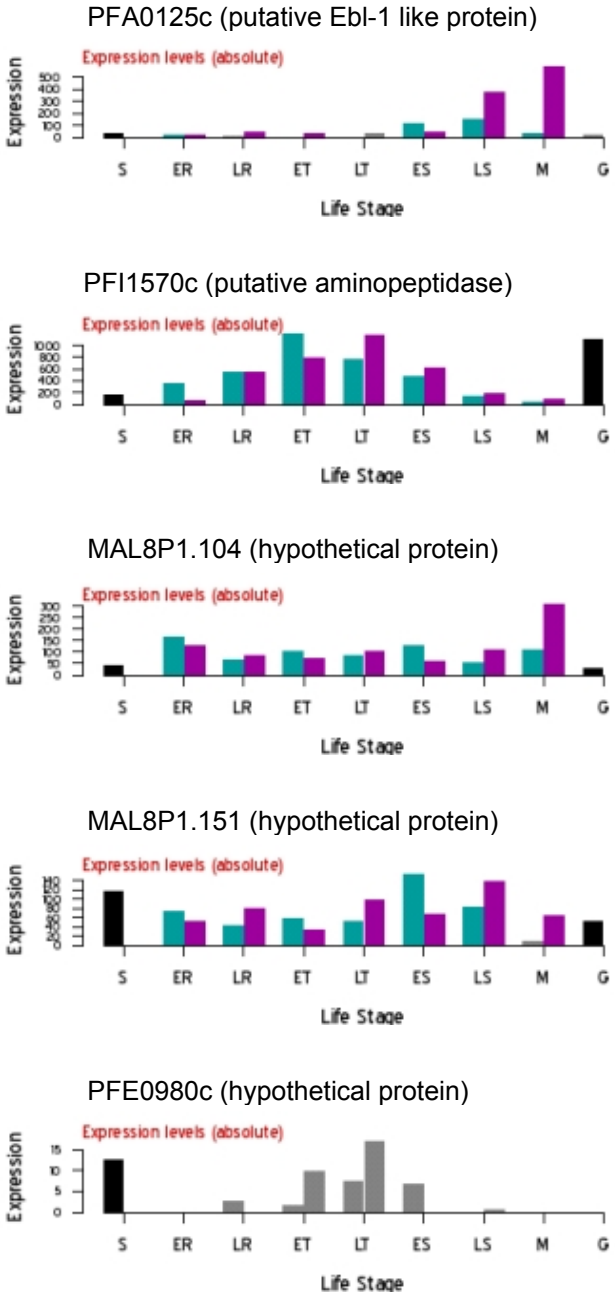


Figure 41: Absolute mRNA expression levels for the spectrin-binding *P. falciparum* proteins.

Graphs showing the Winzeler Lab (Le Roch *et al.*, 2003) microarray absolute expression profiles for the mRNA transcripts of the Ebl-1 like protein, putative aminopeptidase and the three hypothetical proteins. Colours and sampling points: blue = sorbitol synchronised cultures; pink = temperature synchronised cultures; grey = expression level less than 10 (too close to background); S = sporozoite; ER = early ring, LR = late ring; ET = early trophozoite; LT = late trophozoite; ES = early schizont; LS = late schizont; M = merozoite; and G = gametocyte.

3.4 Discussion

3.4.1 Cloning and *P. falciparum* protein expression

Automated sequencing revealed one silent mutation in the putative aminopeptidase sequence and insertions and deletions in the other *P. falciparum* constructs. *Taq* DNA polymerase has a total error rate of 1 x 10^{-4} to 2 x 10^{-5} errors per base pair (Tindall and Kunkel, 1988) and therefore a high fidelity mix, which also contained a DNA polymerase with a 3'-5' exonuclease proofreading activity, was used for the amplification procedure. This *Tgo* DNA polymerase, however, only increases the fidelity of DNA synthesis ~3-fold and does not guarantee error-free amplified DNA sequences. The AT-richness of the *P. falciparum* genome and the presence of adenosine and thymidine stretches within gene sequences could also have caused the DNA polymerases to slip, thereby creating deletions. In addition, the silent mutation and the 27 nucleotide (9 amino acids) insertion may not be true errors, but may reflect a discrepancy between the *P. falciparum* FCR-3 and 3D7 genomes. The PlasmoDB database is based on the 3D7 genome and genes were amplified from FCR-3 genomic DNA. Finally, errors could have been introduced during the sequencing reaction or when the sequencing template was prepared by PCR from the bacterial colony.

Low expression levels were exhibited by all the proteins and tag-specific antibodies had to be used to detect them in the BL21 (DE3) *E. coli* protein extracts. Large *P. falciparum* proteins (>60 kDa) are thought to be less stable in *E. coli* (Mehlin *et al.*, 2006) and all the plasmid constructs were created from gene fragments that were ~1.5 kb in size, thus translating into proteins that would be larger than 60 kDa. Some *P. falciparum* proteins are also toxic to *E. coli* cells (Cinquin *et al.*, 2001), which would slow the growth rate of the cells and less protein would be produced. To improve expression of recombinant proteins a different heterologous host may be used, for example, the baculovirus/insect cell system, yeast or mammalian cells.

Two *P. falciparum* constructs contained deletions responsible for frameshifts and the introduction of premature stop codons in the DNA sequences. Induction of *E. coli* cells containing these constructs, however, resulted in the expression of multiple smaller protein bands that did not match the expected truncated protein sizes. *E. coli* sometimes does not faithfully maintain the coding sequence of the inserted gene (Mehlin *et al.*, 2006) and introduces mutations and frameshifts into the sequences during replication. The translated proteins are subsequently expressed as several differently sized proteins or as proteins with an incorrect amino acid sequence. The instability of the transcribed mRNA and the presence of secondary structures within the mRNA could also result in the translation of fewer and smaller proteins (Wu *et al.*, 2004).

Only the PFI1570c gene that was inserted in the pET-15b plasmid expressed a single protein of the correct size and was therefore chosen for protein purification. Induction with IPTG, however, produced only small amounts of the aminopeptidase, mainly as inclusion bodies. *E. coli* cultures were therefore grown in Overnight Express™ Instant TB Medium at room temperature. The medium and growth conditions ensured that the bacteria grew slowly to a post-log phase with glucose as their main nutrient source. Once the glucose was depleted, lactose was used as a substitute nutrient source and natural inducing agent when converted to allolactose by β-galactosidase in the bacterial cell (Lehninger *et al.*, 1993). These changed conditions decelerated the action of the translation machinery inside the host cell, which facilitated the smooth passage of the ribosomal unit along the mRNA template to allow it to produce a full-length polypeptide chain. Slower translation also allowed the recombinant protein to transfer from the unstable intermediate phase to the correctly folded protein (Flick *et al.*, 2004) while the protein was still being translated, thereby decreasing the formation of inclusion bodies.

3.4.2 Characteristics of the *P. falciparum* spectrin-binding proteins

One of the five analysed *P. falciparum* proteins, namely the Ebl-1 like protein, which is also known as EBA-181/JESEBL, had a hydrophobic signal peptide and transmembrane domains. This protein is a microneme protein and the signal sequence is responsible for transport of the protein to the endoplasmic reticulum, from where it will eventually be redirected, via other signals, to the microneme. EBA-181/JESEBL does not require a vacuolar localisation signal (Marti *et al.*, 2004, Hiller *et al.*, 2004) for transport into the parasitophorous vacuole or the erythrocyte cytoplasm, because it is involved in erythrocyte invasion. None of the other proteins contained signal sequences or vacuolar localisation signals, but a lack of known signals does not exclude the relocation of these four proteins into the erythrocyte cytosol, because transport could be facilitated by signals and mechanisms that have as yet not been identified. The *P. falciparum* aminopeptidase was detected at the infected red blood cell membrane by mass spectroscopy (Florens *et al.*, 2004) which correlates with findings from this study.

The spectrin binding regions in four of the five proteins were located within low complexity regions. The exception was the binding region of the putative aminopeptidase. *P. falciparum* low complexity regions are composed of soluble unstructured amino acid loops (particularly lysine and asparagine) and are located between secondary structural elements of the protein. They are oriented toward the external surface of the protein and may not interfere with the function of the rest of the protein (Aravind *et al.*, 2003). Low complexity regions are common in several *Plasmodium* species and therefore they have been thought to facilitate immune evasion by eliciting a non-productive antibody response against the repetitive amino acid stretches (Anders, 1986). These regions do however play a functional role in protein-protein interactions in other organisms (Liu *et al.*, 2002), and the presence of spectrin-binding regions within the *P. falciparum* low complexity areas provides proof that the parasite also uses these sequences to interact with host proteins.

3.4.3 Possible functions of the spectrin-binding proteins in the erythrocytic phase of the *P. falciparum* life cycle

The Ebl-1 like protein (EBA-181/JESEBL) has been found to interact with the trypsin-resistant receptor E on the erythrocyte surface (Gilberger *et al.*, 2003) during invasion. The yeast two-hybrid data showed that EBA-181/JESEBL also interacted with several other *P. falciparum* proteins, one of them being RhopH2 (PFI1445w) (LaCount *et al.*, 2005), which is a component of the rhoptry high molecular weight complex (RhopH). In ring stage parasites, components of the RhopH complex are distributed throughout the erythrocyte and parasitophorous vacuolar membranes and are presumed to play a role in the establishment of the parasitophorous vacuole (Kaneko *et al.*, 2005). EBA-181/JESEBL also interacts with the 10 kDa domain of protein 4.1. This domain stabilises the interaction between spectrin and actin in the junctional complex, and it is thought that binding of EBA-181/JESEBL to protein 4.1 blocks the repair of the erythrocyte membrane during invasion (Lanzillotti and Coetzer, 2006). Additionally, simultaneous binding of EBA-181/JESEBL to protein 4.1 and spectrin could disrupt the structure of the junctional complex, thereby destabilising the erythrocyte membrane skeleton and allowing the parasite to pass into the erythrocyte.

According to the PlasmoDB database, the hypothetical protein coded for by MAL8P1.104, is a structural component of the nuclear pore ring and the yeast two-hybrid data showed that this protein was involved in transcription and chromatin metabolism (LaCount *et al.*, 2005). The nuclear pore is used to shuttle mRNA between the nucleus and the cytoplasm and therefore the MAL8P1.104 mRNA should be transcribed throughout the parasite erythrocytic life cycle, which is confirmed by microarray data (Le Roch *et al.*, 2003, Bozdech *et al.*, 2003). MAL8P1.104 could also be a structural component of the new transporters and channels of the NPP that are introduced into the erythrocyte membrane and the underlying skeleton as the parasite grows. The interaction between

MAL8P1.104 and spectrin could form an anchoring point for the transporters and channels.

MAL8P1.151 and PFE0980c code for hypothetical proteins and have no assigned function in the PlasmoDB database. This study provides evidence that they interact with spectrin and could thus play a role in any of the processes requiring contact with the erythrocyte membrane.

PFI1570c codes for an aminopeptidase which plays a role in haemoglobin digestion (Dalal and Klemba, 2007). However, the enzyme also interacts with the erythrocyte membrane as evidenced by mass spectrometry data (Florens *et al.*, 2004) and results from this study, indicating that it could therefore perform additional functions in the parasite and infected erythrocyte. These functions will be discussed further in Chapter 4.

Chapter 4: Characterisation of *P. falciparum* M18 aspartyl aminopeptidase

4.1 Introduction

4.1.1 Enzymes and proteases

Enzymes are the reaction catalysts of biological systems and are thus the functional units that degrade or synthesise macromolecules, transform chemical energy, or regulate other enzymes by activating or deactivating them. Proteases, or peptidases are enzymes that hydrolyse the amide bonds of the polypeptide backbone of proteins (Lehninger *et al.*, 1993).

Enzymes work by lowering the activation energy of a reaction, thereby allowing the reaction to occur a lot faster when compared to an uncatalysed reaction. Enzyme activity can be affected by inhibitors, or activators, as well as temperature, pH, and substrate concentration. Because enzymes are proteins, their activity depends upon the integrity of the native protein conformation, and denaturation of the protein, or the dissociation of the enzyme into subunits, therefore destroys the catalytic activity (Lehninger *et al.*, 1993).

4.1.2 Metalloproteases

Proteases that bind metal ions in their active site are known as metalloproteases. In contrast to proteases that are only activated by metal ions, the metal ion in metalloproteases is normally tightly bound in the active site and is therefore retained in the enzyme during purification (Palmer, 1995). Metalloproteases can be distinguished from other enzymes by treatment with metal chelating agents such as EDTA or 1,10-phenanthroline, which remove the metal ion from the active site, thereby inactivating the enzyme (Bugg, 2004).

Metal cofactors play either a structural or catalytic role in metalloenzymes. Structural cofactors are normally bound by four amino acids, which are generally

cysteines (Figure 42a). In contrast, catalytic metal ions are coordinated by three amino acids and an activated water molecule (Figure 42b) (Auld, 2001). Magnesium is probably the most commonly used catalytic cofactor and functions as a nucleophile or a Lewis acid in the active site. Zinc is often used to maintain the tertiary structure of the enzyme, but proteases do exist, that use zinc as a Lewis acid to either activate a water molecule and/or a carbonyl group of a substrate for subsequent base-catalysis or nucleophilic attack by an amino acid group in the active site. Some metalloproteases bind two cofactors in a co-catalytic active site (Figure 42c). In this instance, the metal ions are held in place by five amino acids and one water molecule. One amino acid, usually an aspartate, coordinates both metal ions and the water molecule bridges them. Finally, enzymes also employ metal ions as redox reagents when none of the amino acid side groups in the active site can act in this capacity.

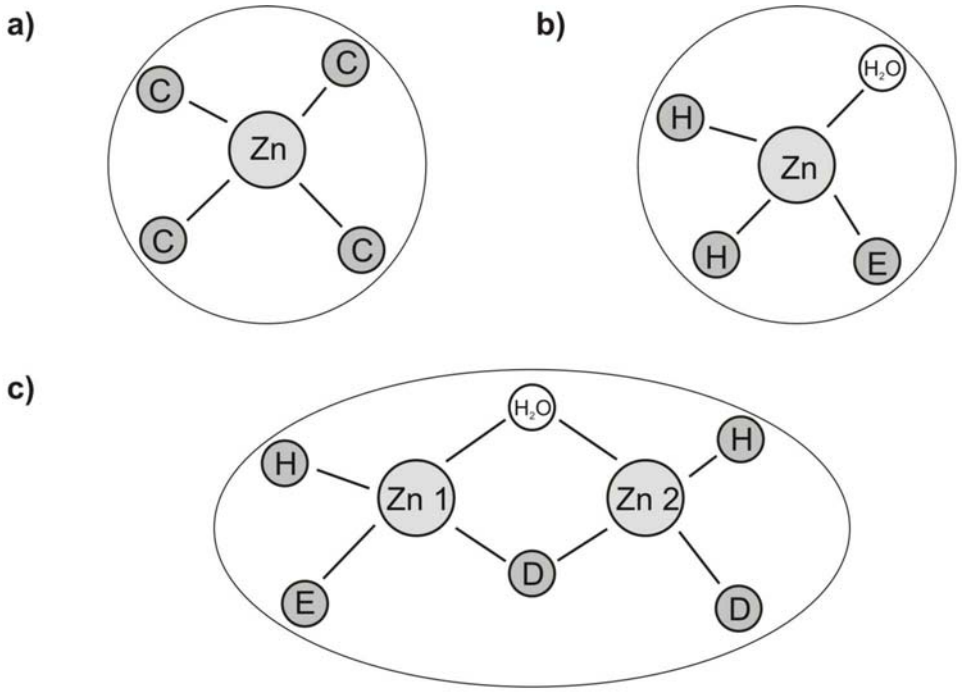


Figure 42: Zinc binding in the active site of a metalloprotease.
Diagram showing how zinc is coordinated by amino acids in a structural (a), catalytic (b) and co-catalytic (c) active site. The zinc atom in the structural site is bound by four amino acids, whereas the zinc ion in the catalytic site is only bound by three amino acids and a water molecule. The zinc ions in the co-catalytic site are bound by five amino acids and water. One of the amino acids coordinates both metal ions. Zn = zinc; H_2O = water; C = cysteine; D = aspartate; E = glutamate; H = histidine (adapted from Auld, 2001).

4.1.3 *P. falciparum* metalloproteases

Wu *et al*. (2003) identified 92 putative *P. falciparum* proteases, which is 59 more than those originally recognised by Gardner in the first genome publication (Gardner *et al.*, 2002). Some of the new proteases include a calpain, a metacaspase and a signal peptidase that are essential for parasitic activity.

Only a small number of *Plasmodium* proteases have been characterised *in vitro* and *in vivo*. As mentioned in the first chapter, *Plasmodium* proteases are involved in erythrocyte entry, the digestion of haemoglobin and exit from the erythrocyte. *P. falciparum* entry proteases include PfSUB-1 (Blackman *et al.*, 1998), PfSUB-2 (Hackett *et al.*, 1999), gp76 (Roggwiller *et al.*, 1996), and the rhomboids (Baker *et al.*, 2006). Digestion proteases include the plasmepsins (Banerjee *et al.*, 2002), falcipains (Shenai *et al.*, 2000, Sijwali and Rosenthal, 2004, Sijwali *et al.*, 2001) and falcilysin (Eggleson *et al.*, 1999). Some of the digestion proteases as well as a 37 kDa acidic protease (Deguercy *et al.*, 1990) and the SERAs (Li *et al.*, 2002) are responsible for parasite escape.

When this study was initiated, only two *Plasmodium* metalloproteases had been characterised, namely PfA-M1, which belongs to the M1 family of zinc metallopeptidases (Florent *et al.*, 1998) and falcilysin, a M16 metalloprotease (Eggleson *et al.*, 1999). In January 2006, the M17 leucyl aminopeptidase (*Pf*M17LAP) was characterised by Gardiner *et al.* (2006) and in August 2007, when this study had just been completed, Teuscher *et al.* (2007) released a publication describing the *P. falciparum* aspartyl aminopeptidase (*Pf*M18AAP) encoded by PFI1570c, which is the subject of this study. All these metalloproteases are involved in haemoglobin digestion, but additional enzymatic roles have been described for PfA-M1 and falcilysin during erythrocyte reinvasion (Allary *et al.*, 2002) and apicoplast transit peptide degradation (Ponpuak *et al.*, 2007), respectively. This study provides evidence for additional roles of *Pf*M18AAP.

4.2 Methods

4.2.1 Aminopeptidase sequence and structure analysis

The *Pf*M18AAP amino acid sequence was used to search the Protein Data Bank (PDB) (http://www.rcsb.org/pdb/home/home.do) (Berman *et al.*, 2000) to identify crystal structures that could be used to predict the tertiary structure of *Pf*M18AAP. The protein names of the identified crystal structures, *Pf*M18AAP and the *Homo sapiens* aspartyl aminopeptidase were used to search the MEROPS database (http://merops.sanger.ac.uk/) (Rawlings *et al.*, 2006) to identify the family and clan that each protease belongs to and the protease alignment. This, as well as alignments created with ClustalW (http://www.ebi.ac.uk/clustalw) (Chenna *et al.*, 2003), were used to compare the proteins with each other and to locate putative motifs in the homologues. Motifs and transmembrane domains were detected with ScanProsite (http://au.expasy.org/tools/scanprosite/) (Gattiker *et al.*, 2002) and SOSUI (http://bp.nuap.nagoya-u.ac.jp/sosui/) (Hirokawa *et al.*, 1998) respectively, and secondary structures were predicted with Jpred (http://www.compbio.dundee.ac.uk/~www-jpred/) (Cuff *et al.*, 1998). Jpred was used, because the program gives the results of six secondary structure prediction algorithms and also integrates these results into one user-friendly output sequence. Crystal structures were viewed and manipulated with the DeepView Swiss-PdbViewer, version 3.7 (SP5) (http://www.expasy.org/spdbv/) (Guex and Peitsch, 1997).

4.2.2 Characterisation of recombinant *Pf*M18AAP

His-tag removal and protein precipitation

A one litre *E. coli* culture containing pET-15b-PFI1570c was grown and recombinant (r) *Pf*M18AAP purified as described in section 3.2.4. Five hundred microlitres from the first elution and 250 µl from the second elution (both elutions performed without PMSF) were combined and dialysed for 15 minutes in seven 100 µl Slide-A-Lyzer® Mini Dialysis Units (Pierce Biotechnology Incorporated, Rockford, USA) against Cleavage Buffer (50 mM Tris-HCl, 10 mM $CaCl_2$, pH

8.0). Two hundred microlitres Thrombin-Agarose (Sigma-Aldrich Corporation, St. Louis, USA) (50% slurry) was centrifuged at 500 *g* for 5 minutes at 4 ºC using the Eppendorf 5415 R Centrifuge, the supernatant removed, the Thrombin-Agarose resuspended in 1 ml Cleavage Buffer and centrifuged for 5 minutes. The wash with 1 ml Cleavage Buffer was repeated and the Thrombin-Agarose finally resuspended in 75 µl 10 x Cleavage Buffer. The dialysed r*Pf*M18AAP sample was added to the Thrombin-Agarose and placed on the Mini Labroller™ (Labnet International, Inc., Woodridge, USA) for 4 hours at room temperature to remove the 6His-tag. The sample was centrifuged for 5 minutes at 500 *g* and the supernatant collected. Residual protein was removed from the Thrombin-Agarose with 50 µl Cleavage Buffer and added to the first supernatant. PMSF was added to a final concentration of 1 mM to inhibit any secondary proteases and the thrombin-cleaved r*Pf*M18AAP added to 30 µl HIS-Select™ Magnetic Agarose Beads for 15 minutes at room temperature. The beads were removed and 2 µl thrombin-cleaved r*Pf*M18AAP kept aside to determine the amount of protein present by electrophoresing the sample through a Laemmli SDS-polyacrylamide gel against 1, 0.5 and 0.25 µg BSA and staining the gel with Coomassie Blue. Nine volumes cold 100 % ethanol were added to the remaining thrombin-cleaved r*Pf*M18AAP sample and the mixture incubated overnight at -20 ºC, followed by centrifugation at 16,000 *g* for 20 minutes at 4 ºC. The ethanol was removed and 1 ml cold 90 % ethanol (v/v) added to wash the pellet. The sample was centrifuged at 16,000 *g* for 5 minutes, the protein pellet air-dried for 30 minutes at room temperature and resuspended in isoelectric focusing (IEF) buffer (O'Farrell, 1975) (9.5 M urea, 2 % Nonidet P40 (NP-40) (w/v) (Roche Diagnostics Gmbh, Mannheim, Germany), 1.6 % Bio-Lyte® pH 5/8 Ampholyte (v/v) (Bio-Rad Laboratories, Hercules, USA), 0.4 % Bio-Lyte® pH 3/10 Ampholyte (v/v) (Bio-Rad Laboratories, Hercules, USA), 5 % β-mercaptoethanol (v/v)) to give a final thrombin-cleaved r*Pf*M18AAP concentration of ~ 1µg/µl. The volume of buffer required for resuspension was calculated from the amount of protein present in the thrombin-cleaved r*Pf*M18AAP sample before precipitation.

Isoelectric focusing

The pI of r*Pf*M18AAP without the His-tag was determined by IEF using the O'Farrell method (O'Farrell, 1975). Glass tubes (15 cm x 3 mm) were soaked overnight in chromic acid (33 ml saturated potassium dichromate, 1 litre sulphuric acid) and rinsed thoroughly with distilled water and 70 % ethanol (v/v). The bottoms of the tubes were sealed with Parafilm® (Pechiney Plastic Packaging, Menasha, USA). A 3.5 % polyacrylamide mixture (2.75 g urea, 665 µl 28.4 % acrylamide/1.6 % bisacrylamide (w/v) mix, 250 µl Bio-Lyte® 3/10 Ampholyte, 1 ml 10 % NP-40 (v/v) in a final volume of 5 ml) was prepared and degassed for 1 minute. Twenty-three microlitres freshly prepared 10 % ammonium persulphate and 3.5 µl TEMED were added and the tubes filled up to 10 cm with the polyacrylamide mixture. The gel solution was overlaid with water and allowed to set for 2 hours. The water was removed and replaced with 20 µl IEF buffer and the gels allowed to set for another 2 hours. The IEF buffer was replaced with fresh 20 µl IEF buffer. The lower chamber of the Shandon Vertical Electrophoresis Chamber (Shandon Scientific Company Ltd., London, UK) was filled with 0.01 M phosphoric acid (Merck KGaA, Darmstadt, Germany). The IEF tube gel was placed in the system and the upper chamber filled with 0.02 M NaOH (degassed for 1 hour). The gel was prefocused at 200 V for 15 minutes, followed by 300 V for 30 minutes and finally 400 V for 30 minutes. The upper buffer was removed and 30 µl 2-D SDS-PAGE Standards (Bio-Rad Laboratories, Hercules, USA) mixed with 3-5 µg thrombin-cleaved r*Pf*M18AAP (1 µg/µl) loaded onto the gel. The sample was overlaid with 10 µl sample overlay solution (9 M urea, 0.8 % Bio-Lyte® 5/8 Ampholyte (v/v), 0.2 % Bio-Lyte® 3/10 Ampholyte (v/v)) and the gel electrophoresed at 400 V for 16 hours followed by 500V for one hour. One microlitre of concentrated bromophenol blue (50 % glycerol (v/v), 0.01 mg/ml bromophenol blue (BDH Laboratory Supplies, Poole, UK) was added to the top of the gel and allowed to electrophorese into the gel for 3-5 minutes to be able to identify the top end of the gel after it has been extruded from the glass tube. The glass tube was removed from the apparatus and the gel carefully extruded with water into 10 ml equilibration buffer (124 mM Tris-HCl, 10 % glycerol (v/v), 2 %

SDS (w/v), 8.6 mM DTT, pH 6.8). The gel was soaked for 10 minutes and carefully loaded onto a 14 cm x 16 cm x 3 mm Laemmli SDS-polyacrylamide gel consisting of a resolving gel containing 13 % polyacrylamide and a stacking gel without wells containing 4 % polyacrylamide. The IEF gel was secured onto the second dimension gel and two wells formed with a single comb tooth at both ends of the IEF tube gel using 0.5 % agarose solution (0.5 % SeaKem LE agarose (w/v) in Laemmli running buffer). Fifteen microlitres denatured whole erythrocyte membrane marker and 15 µl Page Ruler™ Prestained Protein Ladder (Fermentas International Inc., Burlington, Canada) were loaded in the wells and the gel electrophoresed with Laemmli running buffer at 75 V for 18 hours. The gel was fixed overnight in one litre fixing solution (50 % isopropanol (v/v), 7 % glacial acetic acid (v/v)) on the Red Rotor PR70-230V shaker (set to 2) followed by several rinses in milli-Q water. The gel was incubated overnight in milli-Q water and silver stained (Hoefer Scientific Instruments, 1994) by incubating the gel in one litre 0.032 mM DTT for 1 hour with shaking using the Red Rotor PR70-230V shaker (set to 2). The solution was removed and replaced with 400 ml 0.1 % silver nitrate (w/v) (Sigma-Aldrich Corporation, St. Louis, USA) and the gel incubated for another hour. The silver nitrate was removed and the gel briefly rinsed with milli-Q water. The gel was rinsed twice with 200 ml developing solution (800 ml 3 % Na_2CO_3 (w/v) with 400 µl 37 % formaldehyde (v/v) (Sigma-Aldrich Corporation, St. Louis, USA) added just before use). Finally, 400 ml developing solution was added to the gel, which was placed on a shaker for 5-10 minutes. As soon as protein spots appeared the developing solution was removed and the gel rinsed several times with milli-Q water and left overnight in milli-Q water. A dark background was destained (Gharahdaghi *et al.*, 1999) by incubating the gel for 30 seconds to 2 minutes in 400 ml destaining solution (100 mM sodium thiosulphate (Sigma-Aldrich Corporation, St. Louis, USA), 30 mM potassium ferricyanide (Sigma-Aldrich Corporation, St. Louis, USA), prepared separately and mixed in equal volumes prior to use), followed by several rinses in milli-Q water. The pI of thrombin-cleaved r*Pf*M18AAP was estimated by comparing the r*Pf*M18AAP protein spots in relation to those of the 2-D SDS-PAGE Standards.

Molecular weight determination by denaturing polyacrylamide electrophoresis

The molecular weight of denatured r*Pf*M18AAP was calculated from the relative mobility (R_m) values of the protein when 500 ng r*Pf*M18AAP, 500 ng BSA and 20 µg erythrocyte membrane proteins were electrophoresed through a Laemmli SDS-polyacrylamide gel containing a 12 % resolving gel and a 4 % stacking gel. A standard curve consisting of the relative mobility (Y-axis) and molecular weight (X-axis) of each standard was plotted on semi-log paper and used to determine the molecular weight of r*Pf*M18AAP.

Native oligomeric state determination using Ferguson plots

The number of subunits present in the native form of r*Pf*M18AAP was determined by Ferguson plots (Ferguson, 1964, Hedrick and Smith, 1968, Sigma-Aldrich, 2005). The plots are constructed from the R_m values of standards and the protein of interest after the proteins have been electrophoresed through non-denaturing gels (Liu *et al.*, 1982). Tube gels with varying polyacrylamide percentages were prepared the following way. Glass tubes (12 cm x 6 mm) were cleaned and prepared as in the isoelectric focusing section. A 1 % agarose solution was placed in a 43 ºC waterbath. The non-denaturing gel mixture containing the desired percentage polyacrylamide (from a 40 % acrylamide/1.5 % bisacrylamide stock (w/v)) and 1 x Tris acetate buffer (pH 7.4) in a final volume of 2.4 ml was placed in the 43 ºC waterbath. Two hundred microlitres 1.5 % ammonium persulphate, 200 µl 0.5 % TEMED and 1.2 ml 1 % agarose were added and the glass tubes filled up to 11.5 cm with the gel mixture. The gel was overlaid with water and placed in ice water for 3 minutes. Gels were left to polymerise overnight at room temperature and placed in the Shandon Vertical Electrophoresis Chamber. Ten microlitres bromophenol tracking dye (50 % sucrose (w/v), 0.05 % bromophenol blue (w/v)) was added to 20 µg r*Pf*M18AAP (containing 0.01 mM DTT), 10 µg BSA or 10 µg spectrin and the samples loaded onto the different percentage tube gels according to Table 15. The gels were electrophoresed in the cold room with 1

x Tris acetate buffer containing 0.008 % β-mercaptoethanol (v/v) at 4 mA per gel until the bromophenol blue tracking dye reached the end of the gels. The gels were carefully extruded from the glass tubes with water and stained with Coomassie Blue stain.

Table 15: Polyacrylamide percentages for the native gels.

Protein	Percentage polyacrylamide
Spectrin	2; 2.5; 3; 3.5
BSA	6; 7; 8; 9
r*Pf*M18AAP	2; 3; 4; 5; 6; 7; 8; 9; 10

The R_m values of each standard and r*Pf*M18AAP were calculated by dividing the distance the protein travelled by the distance the bromophenol tracking dye travelled. The 100 [Log (R_m x 100)] values for each protein were plotted against the gel concentration and the retardation coefficient (K_R) of each protein standard determined from the slope of each standard (Hedrick and Smith, 1968, Sigma-Aldrich, 2005). The negative values of K_R were plotted against the molecular weight of each standard on log-log paper and the approximate molecular weight of r*Pf*M18AAP read from the second graph (Hedrick and Smith, 1968, Sigma-Aldrich, 2005).

Immunoblotting of non-denaturing gels

Fifteen micrograms r*Pf*M18AAP was electrophoresed on non-denaturing polyacrylamide gels as described in the previous section. The gels were cut in half length wise and one half stained with Coomassie Blue stain. The other half was incubated for 10 minutes in 48 mM Tris, 39 mM glycine, 0.0375 % SDS (w/v), 20 % methanol (v/v), pH 9.2 (Bjerrum and Schafer-Nielsen, 1986) and the proteins transferred onto a Hybond™-C Extra Nitrocellulose membrane with the same buffer for 36 hours at 4 ºC. The histidine-tag of r*Pf*M18AAP was detected with 1:1,500 dilution of the Penta•His™ HRP Conjugate antibody and chemiluminescence with the SuperSignal® West Pico Chemiluminescent Substrate.

4.2.3 Protein overlays

Binding studies were performed by blot overlays of r*Pf*M18AAP with spectrin, trypsin digested spectrin (Lawler *et al.*, 1984) and erythrocyte membrane proteins. The proteins of interest were transferred or spotted onto nitrocellulose membranes in the following ways:

Dot blot preparation: One hundred nanograms r*Pf*M18AAP (Section 3.2.4), 100 ng BSA, 2 µg spectrin tetramers and 2 µg spectrin dimers (Section 2.2.2) were spotted onto two strips of Hybond™-C Extra Nitrocellulose membrane and air dried.

Slot blot preparation: One hundred nanograms r*Pf*M18AAP, 100 ng BSA, 1 µg spectrin tetramers and 1 µg spectrin dimers were applied to two Hybond™-C Extra Nitrocellulose membranes in a Bio-Rad Bio-Dot® SF chamber (Bio-Rad Laboratories, Hercules, USA) connected to Millipore Vacuum Pump XF 54 230 50 (Millipore Corporation, Bedford, USA).

Erythrocyte membrane protein overlay: One hundred nanograms r*Pf*M18AAP (positive control), 100 ng BSA (negative control) and 20 µg erythrocyte membrane proteins were electrophoresed on a 14 cm x 16 cm x 1.5 mm 9 % Laemmli and a 14 cm x 16 cm x 1.5 mm 3.5-17.5 % Fairbanks SDS-polyacrylamide gel (Section 2.2.2) and the protein bands transferred onto a Hybond™-C Extra Nitrocellulose membrane (Section 4.2.2). Five hundred nanograms r*Pf*M18AAP, 500 ng BSA and 20 µg erythrocyte membrane proteins were electrophoresed on a second Laemmli and Fairbanks SDS-polyacrylamide gel and stained with Coomassie Blue stain.

Spectrin tryptic digest overlay: Spectrin was isolated from human erythrocytes as described in section 2.2.2. Four hundred micrograms spectrin was prepared for the digestion reaction by adding 10 x digestion buffer (10 x: 100 mM sodium phosphate buffer, pH 7.5, 1.5 M NaCl, 50 mM Na_2EDTA, 50 mM β-

mercaptoethanol) followed by 1 µg trypsin (0.2 µg/µl) (Sigma-Aldrich Corporation, St. Louis, USA) per 100 µg spectrin. The sample was incubated at 4 ºC for 18 hours and the digest stopped by adding 1 % (v/v) 0.1 M Pefabloc SC. One hundred nanograms r*Pf*M18AAP (positive control), 100 ng BSA (negative control) and three lanes containing 40 µg trypsin digested spectrin were electrophoresed on a 14 cm x 16 cm x 1.5 mm Laemmli SDS-polyacrylamide gel at 65 V for 18 hours. The Laemmli gel contained a 10 % separating gel (40 % acrylamide/1.5 % bisacrylamide stock (w/v) in lower gel buffer (375 mM Tris-HCl, pH 8.3), 0.1 % SDS (w/v) and 2 mM EDTA (pH 7.4). The 2.66 % stacking gel contained upper gel buffer (125 mM Tris-HCl, pH 6.8), 0.1 % SDS (w/v) and 2 mM EDTA (pH 7.4). The proteins were transferred onto a Hybond™-C Extra Nitrocellulose membrane (Section 4.2.2). Two of the lanes containing the trypsin digested spectrin were cut off from the nitrocellulose membrane and the spectrin fragments analysed with separate 1:500 α- and β-spectrin polyclonal antibodies, followed by a 1:20,000 goat anti-rabbit IgG peroxidase (Section 2.2.2) and detection by chemiluminescence with the SuperSignal® West Pico Chemiluminescent Substrate. Five hundred nanograms r*Pf*M18AAP, 500 ng BSA and 40 µg trypsin digested spectrin were electrophoresed on a second Laemmli SDS-polyacrylamide gel and stained with Coomassie Blue stain.

The overlays were performed as follows: All the membranes were blocked for 1 hour at room temperature in 5 % BSA Fraction V (w/v) in TBS. One slot or dot blot membrane was placed in 2 ml 50 mM Tris-HCl buffer (pH 7.5) containing 2.5 ng/µl r*Pf*M18AAP, while the other slot or dot blot membrane was placed into 2 ml 50 mM Tris-HCl buffer (pH 7.5) and incubated for 1 hour at room temperature on the Mini Labroller™. The membranes were then washed two times for 10 minutes in wash buffer A (20 mM Tris-HCl, 500 mM NaCl, 0.12 % Tween®-20 (v/v), 0.2 % Triton® X-100 (v/v), pH 7.5) and once for 10 minutes in HRP-TBS buffer (10 mM Tris-HCl, 150 mM NaCl, pH 7.5). The membranes containing the erythrocyte membrane proteins and the spectrin tryptic digest were incubated for 1 hour in 30 ml 50 mM Tris-HCl buffer (pH 7.5) containing 1.25 ng/µl r*Pf*M18AAP and subsequently washed 4 times for 10 minutes in wash

buffer A and once for 10 minutes in HRP-TBS buffer. All membranes were fixed for 20 minutes at room temperature with HRP-TBS buffer containing 0.5 % (v/v) formaldehyde, followed by incubation for 20 minutes in HRP-TBS buffer containing 2 % glycine. Glycine was used to inactivate remaining reactive aldehyde groups (Tisdale, 2002). The membranes were washed twice for 10 minutes in Tween/Triton buffer (20 mM Tris-HCl, 500 mM NaCl, 0.05 % Tween®-20 (v/v), 0.2 % Triton® X-100 (v/v), pH 7.5) and once for 10 minutes with HRP-TBS buffer and subsequently incubated for 1 hour at room temperature with a 1:1,500 dilution of the Penta•His™ HRP Conjugate antibody. The membranes were washed twice for 10 minutes with Tween/Triton buffer and once for 10 minutes with HRP-TBS buffer and the antibody detected by chemiluminescence with the SuperSignal® West Pico Chemiluminescent Substrate.

4.2.4 Enzyme activity assay

The enzyme activity assay was based on a coupled enzyme reaction (Wilk *et al.*, 1998). If r*Pf*M18AAP is active, it will cleave the aspartate from the Asp-Ala-Pro-β-Napthylamide substrate, thereby liberating Ala-Pro-β-Napthylamide (Figure 43a). This product becomes the substrate for dipeptidyl peptidase IV, which cleaves alanine-proline from the β-Napthylamine chromogen (Figure 43b). The chromogen is chemically converted to an azo dye, which can be quantitated with a spectrophotometer at 580 nm. Azo dye creation entails diazotising β-Napthylamine with sodium nitrite under acidic conditions (Figure 43c). Excess nitrite is removed with ammonium sulphamate before *N*-1-Naphthylethylenediamine is coupled to the diazotised β-Napthylamine to form the azo dye (Figure 43d) (Goldbarg and Rutenburg, 1958, Alur *et al.*, 2001).

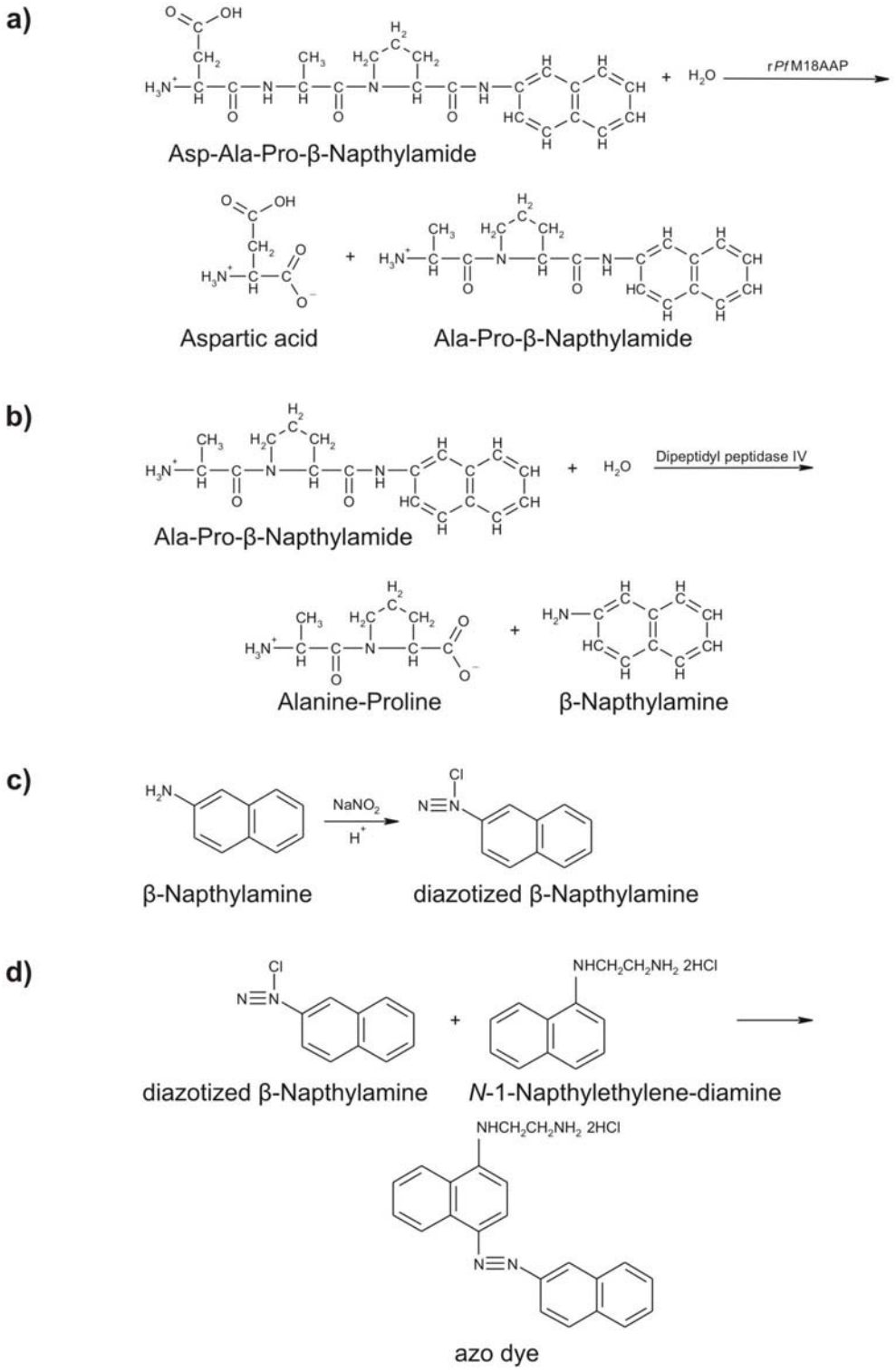


Figure 43: Substrate cleavage by r*Pf*M18AAP, diazotisation and azo dye coupling reaction.

Diagram showing the cleavage of aspartate from Asp-Ala-Pro-β-Napthylamide by r*Pf*M18AAP (a) and the subsequent removal of alanine-proline from the chromogen by dipeptidyl peptidase IV (b). β-Napthylamine is diazotised by sodium nitrite under acidic conditions (c) and *N*-1-Naphthylethylenediamine is added to form the azo dye (b), which can be analysed with a spectrophotometer.

A 500 ml *E. coli* culture containing pET-15b-PFI1570c was grown, lysed and r*Pf*M18AAP purified as described in section 3.2.4. The 250 µl 1st His elution (no PMSF added) was dialysed for 15 minutes against 50 mM Tris-HCl buffer (pH 7.5) using three 100 µl Slide-A-Lyzer® Mini Dialysis Units and the concentration of a 10 µl sample determined with the Coomassie Plus™ Protein Assay Reagent. The desired concentration of H-Asp-Ala-Pro-β-Napthylamide (Peptides International, Louisville, USA) (10 or 100 mM stock dissolved in dimethyl sulphoxide (DMSO; BDH Laboratory Supplies, Poole, UK) in 50 mM Tris-HCl buffer, pH 7.5, containing 0.0025 U dipeptidyl peptidase IV from Porcine kidney (Sigma-Aldrich Corporation, St. Louis, USA), was equilibrated in a 37 ºC waterbath for 5 minutes. r*Pf*M18AAP (2.5 µg) was added to give a final volume of 50 µl and the sample incubated at 37 ºC for 15 minutes in the waterbath. The enzymatic reaction was terminated with 50 µl 10 % trichloroacetic acid (w/v) (Sigma-Aldrich Corporation, St. Louis, USA). Hundred microlitres 0.1 % $NaNO_2$ (w/v) (Riedel de Haën A.G., Hannover, Germany) was added and the sample incubated at room temperature for 3 minutes, followed by the addition of 100 µl 0.5 % ammonium sulphamate (w/v) (Sigma-Aldrich Corporation, St. Louis, USA) and incubation for 3 minutes at room temperature. The azo dye was created by adding 200 µl 0.05 % *N*-(1-napthyl)ethylenediamine dihydrochloride (w/v) (Sigma-Aldrich Corporation, St. Louis, USA) and incubating the solution at room temperature for 45 minutes.

A standard curve containing 0.05, 0.1, 0.3, 0.5, 0.7, 0.8, 1, and 1.2 mM β-Napthylamine was set up using a 2 mM β-Napthylamine solution (prepared from a 100 mM β-Napthylamine stock solution dissolved in DMSO) (Sigma-Aldrich Corporation, St. Louis, USA). The 50 µl samples were treated with trichloroacetic acid, sodium nitrite, ammonium sulphamate and *N*-(1-Naphthyl)ethylenediamine dihydrochloride in the same way as the enzymatic reaction.

All the samples were analysed at 580 nm using the Beckman DU®-65 spectrophotometer (Beckman Coulter, Inc., Fullerton, USA) and the amount of product formed during the enzymatic reaction determined from the standard curve.

The velocity of the enzymatic reaction and the specific activity of the enzyme was calculated with the following formulae:

velocity (nmol/min) = amount product / reaction time

specific activity (nmol/min/mg) = velocity / amount of enzyme

Enzymatic reactions were compared with each other by calculating the relative activity.

pH and ionic strength study

The pH study entailed setting up the enzyme assay containing 0.3 mM H-Asp-Ala-Pro-β-Napthylamide in 50 mM Tris-HCl buffer at pH 6.8, 7.5, 8.0, 8.5 and 9.0, or 0.1 M sodium citrate buffer at pH 5.3, 6.0 and 6.5. The relative activity of the enzymatic reactions was plotted against the pH.

The ionic strength study was performed by setting up the enzyme assay with 0.3 mM H-Asp-Ala-Pro-β-Napthylamide in 50 mM Tris-HCl buffer (pH 7.5) containing 50 or 150 mM NaCl.

Temperature study

The enzyme assay containing 0.3 mM H-Asp-Ala-Pro-β-Napthylamide in 50 mM Tris-HCl buffer (pH 7.5) (pH adjusted to each temperature) was performed at 25, 30, 33, 37 and 39 ºC in the Eppendorf Mastercycler® Gradient machine. The relative activity of the enzymatic reactions was plotted against the temperature.

Metal ion study

The cofactor study was performed by setting up the enzyme assay with 0.3 mM H-Asp-Ala-Pro-β-Napthylamide in 50 mM Tris-HCl buffer (pH 7.5) containing 1 mM $ZnSO_4$, $MgCl_2$, $MnCl_2$, or $CaCl_2$.

Inhibitor study

The inhibitor study entailed setting up the enzyme assay with 0.3 mM H-Asp-Ala-Pro-β-Napthylamide in 50 mM Tris-HCl buffer (pH 7.5) containing 1 mM Pefabloc SC, 10 mM EDTA, 1 mM PMSF, or 10 μM Bestatin (Sigma-Aldrich Corporation, St. Louis, USA). Samples were preincubated at 37 ºC for 10 minutes before the substrate was added.

Maximum velocity and Michaelis constant determination

The maximum velocity (V_{max}) and Michaelis constant (K_m) were determined from enzyme assays containing a range from 0.1-3 mM H-Asp-Ala-Pro-β-Napthylamide in 50 mM Tris-HCl buffer (pH 7.5). The velocity of each reaction was calculated and plotted against the substrate amount used in each reaction and the graph used to read the values of V_{max} and K_m according to the Michaelis-Menten equation:

$$V_0 = (V_{max} \times [S_0]) / ([S_0] + K_m)$$

where V_0 = initial velocity and S_0 = initial substrate concentration

A Lineweaver-Burk plot was constructed by plotting the 1/V readings against the 1/[S] values and K_m and V_{max} read from the x- and y-intercepts respectively.

4.3 Results

4.3.1 *Pf*M18AAP

*Pf*M18AAP is an exopeptidase classified with the enzyme commission (E.C.) number 3.4.11.21: where '3' represents a hydrolase; '4' indicates that the enzyme acts on a peptide bond; '11' indicates that the enzyme cleaves at the amino terminus of the peptide and '21' that the enzyme is an aspartyl aminopeptidase (International Union of Biochemistry, 1992). This enzyme is also classified in a separate system, available in the MEROPS database, that describes only peptidases (Rawlings and Barrett, 1993, Rawlings *et al.*, 2006). The system applies a hierarchical, structure-based approach to assign each peptidase to a family on the basis of statistically significant similarities in amino acid sequence. Furthermore, families that are thought to be homologous are grouped into a clan. In this system *Pf*M18AAP belongs to the MH clan of metallopeptidases and is part of the M18 family which currently contains 253 aminopeptidases (Rawlings *et al.*, 2006). Yokoyama *et al*. (2006) suggested that the M18 family be further divided into two groups based on their substrate specificity. The first group, for example, contains aminopeptidase I from *Saccharomyces cerevisiae* (E.C. 3.4.11.22) (Chang and Smith, 1989), which cleaves N-terminal hydrophobic amino acids from a peptide substrate. An example of an enzyme in the second group is the human aspartyl aminopeptidase (DAP) (E.C. 3.4.11.21) (Wilk *et al.*, 2002), which cleaves N-terminal aspartates or glutamates from a peptide substrate. *Pf*M18AAP was initially categorised in the second group of M18 aminopeptidases, but recently this enzyme and the other *Plasmodium* M18 aspartyl aminopeptidase homologues, as well as the M18 aminopeptidases from *Theileria* and *Cryptosporidium*, were regrouped into a third group based on the amino acid sequence of *Pf*M18AAP.

4.3.2 M18 aminopeptidase family primary sequence features

A search of the crystal structure Protein Data Bank with the *Pf*M18AAP amino acid sequence revealed a similarity to four aminopeptidase protein structures,

namely *Clostridium acetobutylicum*, *Thermotoga maritima*, *Borrelia burgdorferi* and *Pseudomonas aeruginosa*. These four proteins, as well as the *H. sapiens* aspartyl aminopeptidase and *Pf*M18AAP are found in the M18 peptidase family in the MEROPS database (Rawlings *et al.*, 2006). *C. acetobutylicum*, *T. maritima*, and *B. burgdorferi* are listed in the first group of M18 aminopeptidases, which also contains aminopeptidase I from *Saccharomyces cerevisiae*, and *P. aeruginosa* is categorised in the second group of M18 aminopeptidases, which also contains the *H. sapiens* aspartyl aminopeptidase and another aspartyl aminopeptidase from *S. cerevisiae* (Yokoyama *et al.*, 2006). As mentioned above, *Pf*M18AAP and the other *Plasmodium* aspartyl aminopeptidase homologues are grouped into the third group. In addition, a search of the PlasmoDB database showed that *Pf*M18AAP is the only M18 protease in *P. falciparum*.

According to the PlasmoDB database, *Pf*M18AAP has a 30.7 % sequence identity to the *H. sapiens* aspartyl aminopeptidase. This is lower than that obtained from a ClustalW alignment of *Pf*M18AAP and the *H. sapiens* aspartyl aminopeptidase (Table 16). The highest identities are found amongst the *Plasmodium* species (rows 8-11), while other Apicomplexa (rows 12-16) showed only low identities.

Table 16: Sequence comparison of M18 aminopeptidases to *Pf*M18AAP.

Organism	Percentage sequence identity
Saccharomyces cerevisiae (1)	22
Clostridium acetobutylicum	18
Borrelia burgdorferi	18
Thermotoga maritima	16
Homo sapiens	33
Saccharomyces cerevisiae (2)	30
Pseudomonas aeruginosa	30
Plasmodium chabaudi chabaudi	74
Plasmodium yoelii yoelii	73
Plasmodium knowlesi	64
Plasmodium vivax	64
Cryptosporidium parvum	33
Cryptosporidium hominis	33
Theileria annulata	33
Theileria parva	33
Babesia bovis	34

The MEROPS sequence alignment of the two *S. cerevisiae*, *C. acetobutylicum*, *B. burgdorferi*, *T. maritima*, *H. sapiens*, *P. aeruginosa* and *Pf*M18AAP is given in Figure 44. According to the MEROPS database, seven conserved amino acids are involved in the active site, namely two histidines, three aspartates and two glutamates (for clarification purposes the conserved amino acids are numbered according to the human aspartyl aminopeptidase (Wilk *et al.*, 2002) and the amino acid numbers of *Pf*M18AAP are given in brackets). Amongst the conserved amino acids, the two histidines, *Hs*H94 and *Hs*H440 (*Pf*H86 and *Pf*H534), the two aspartates, *Hs*D264 and *Hs*D346 (*Pf*D324 and *Pf*D434), and one glutamate, *Hs*E302 (*Pf*E380), are involved with cofactor or metal ligand binding (all labelled in red in Figure 44). The two remaining amino acids, one glutamate, *Hs*E301 (*Pf*E379), and one aspartate, *Hs*D96 (*Pf*D88), are directly and indirectly involved in substrate cleavage (labelled in pink in Figure 44). A third histidine, *Hs*H170 (*Pf*H160) (labelled in light blue in Figure 44), is conserved amongst all the sequences, and enzyme activity is abolished when it is replaced with phenylalanine in the *H. sapiens* aspartyl aminopeptidase (Wilk *et al.*, 2002). In addition, Wilk *et al*. (2002) showed that substitution of another histidine, *Hs*H352 (*Pf*H440) (labelled in dark blue in Figure 44), causes the subunits of the native *H. sapiens* aspartyl aminopeptidase to dissociate. This histidine is not found in the first group of M18 aminopeptidases and is therefore only present in *H. sapiens*, *P. aeruginosa*, *Pf*M18AAP and the other *Plasmodium* homologues (Figure 44 and Figure 45).

A motif scan revealed a putative protein kinase C (*Pf*535-537; light blue box) and a casein kinase II (*Pf*535-538; dark blue box) phosphorylation site located near the C-terminus of all the proteins (Figure 44 and Figure 45). The SOSUI transmembrane domain scan gave two transmembrane domains in the *Pf*M18AAP sequence (*Pf*504-526 and *Pf*538-560) (labelled in green in Figure 44 and Figure 45), suggesting that this protein is able to associate with membranes. None of the other aminopeptidases contained a transmembrane domain and were thus classified as soluble proteins.

```
Sc(1) EHNYEDIAQEFIDFIYKNPTTYHVVSFFAELLDKHNFKYLSEKSNWQDSIGEDGGK 101
Ca    -KEVFALGDRFKNFISNCKTERECVTELIKTAEKSGYRNIEDILAK-GETLKEGDK  75
Bb    KNQILNFSESYKKFISKFKTEREVTAYALDKAKKLGFINAEEK-----KNLMPGDK  65
Tm    ----EAFSKEYMEFMSKAKTERMTVKEIKRILDESGFVPLEDF-----AGDPMNMT  64
Hs    KEAVQTAAKELLKFVNRSPSPFHAVAECRNRLLQAGFSELKETEKW---NIKPESK  63
Sc(2) KTCKSDYPKEFVSFLNSSHSPYHTVHNIKKHLVSNGFKELSERDSW-AGHVAQKGK  66
Pa    MRAEL--NQGLIDFLKASPTPFHATASLARRLEAAGYRRLDERDAW---HTETGGR  51
Pf    DKKAREYAQDALKFIQRSGSNFLACKNLKERLENNGFINLSEGETW---NLNKNEG  54

Sc(1) FYTIRNGTNLSAFILGKNWRAEK-GVGVIGSHVDALTVKLKPVSFKDTAEGYGRIA 156
Ca    VYANNRGKGLIMFLIGKEPL-YT-GFKILGAHIDSPRLDLKQNPLYED-TDLAMLE 128
Bb    IFYTCREKSVAFAIIGKNPI-ED-GMNFIVSHTDSPRLDAKPSPISEE-NELTFIK 118
Tm    VYAVNRGKAIAAFRVVDDL--KR-GLNLVVAHIDSPRLDFKPNPLIED-EQIALFK 116
Hs    YFMTRNSSTIIAFAVGGQYVPGN-GFSLIGAHTDSPCLRVKRRSRRSQ-VGFQQVG 117
Sc(2) YFVTRNGSSIIAFAVGGKWEPGN-PIAITGAHTDSPALRIKPISKRVS-EKYLQVG 120
Pa    YYVTRNDSSLIAIRLGRRSPLES-GFRLVGAHTDSPCLRVKPNPEIAR-NGFLQLG 105
Pf    YVLCKENRNICGFFVGKNFNIDTGSILISIGHIDSCALKISPNNNVIK-KKIHQIN 109

Sc(1) VAPYGGTLNELWLDRDLGIGGRLLYKKKGTNEIKSALVDSTPLPVCRIPSLAPHF- 211
Ca    THYYGGIKKYQWVTLPLAIHGVIVKKDGT--IVNVCVGEDDNDPVFGVSDILVHLA 182
Bb    TNYYGGIKKYQWLSTPLSIRGVVFLKNGE--KVEINIGDNENDPVFVIPDILPHLD 172
Tm    THYYGGIKKYHWLSIPLEIHGVLFKNDGT--EIEIHIGDKPEDPVFTIPDLLPHLD 170
Hs    VETYGGGIWSTWFDRDLTLAGRVIVKCPTSGRLEQQLVH-VERPILRIPHLAIHL- 171
Sc(2) VETYGGAIWHSWFDKDLGVAGRVFVKDAKTGKSIARLVD-LNRPLLKIPTLAIHL- 174
Pa    VEVYGGALFAPWFDRDLSLAGRVTFRA-N-GKLESRLVD-FRKAIAVIPNLAIHL- 157
Pf    VECYGSGLWHTWFDRSLGLSGQVLYKKGN--KLVEKLIQ-INKSVLFLPSLAIHL- 161

Sc(1) -GKPAEG---PFDKEDQTIPVIGFPTPDEEGN----------------------- 239
Ca    SE-QLEKKASKVIEGEDLNILIGSIPL---------------------------- 208
Bb    RKIQRNKKSDEIVEGENLKILIGSLPI---------------------------- 199
Tm    KE---DAKISEKFKGENLMLIAGTIPL---------------------------- 194
Hs    -QRNINE-NFGPNTEMHLVPILATAIQEELEK----------------------- 201
Sc(2) -DRDVNQ-KFEFNRETQLLPIGGLQEDKTEAK----------------------- 204
Pa    -NRAANE-GWPINAQNELPPIIAQLAPGEAAD----------------------- 187
Pf    QNRTRYDFSVKINYENHIKPIISTTLFNQLNKCKRNNVHHDTILTTDTKFSHKENS 217

Sc(1) ------------------------------------------EPPTDDEKKSPLF 252
Ca    ----------------------------------------------KDGEEK 214
Bb    ----------------------------------------------ETKEK 204
Tm    ----------------------------------------------SGEEK 199
Hs    ------------------------------------------GTPEPGPLNAVD 213
Sc(2) ---------------------------------------TEKEINNGEFTSIKTIV 221
Pa    ------------------------------------------------------ 187
Pf    QNKRDDQMCHSFNDKDVSNHNLDKNTIEHLTNQQNEEKNKHTKDNPNSKDIVEHIN 273

Sc(1) GKHCIHLLRYVAKLAGVE-VSELIQMDLDLFDVQKGTIGGIGKHFLFAPRLDDRLC 307
Ca    QKVKHNIMKILNEKYDIS-EEDFVSAELEIVPAGKARDYGFDRSMVMGYGQDDRIC 269
Bb    NKVKLATLQLIKEKYKIE-EEDFVSSEIEIVPAGTAKDVGFDKALIGAYGQDDKIC 259
Tm    EAVKTNVLKILNEMYGIT-EEDFVSGEIEVVPAFSPREVGMDRSLIGAYGQDDRIC 254
Hs    ERHHSVLMSLLCAHLGLS-PKDIVEMELCLADTQPAVLGGAYDEFIFAPRLDNLHS 268
Sc(2) QRHHAELLGLIAKELAIDTIEDIEDFELILYDHNASTLGGFNDEFVFSGRLDNLTS 277
Pa    --FRLLLDEQLLREHGIT-ADVVLDYELSFYDTQSAAVVGLNDEFIAGARLDNLLS 240
Pf    TDNSYPLLYLLSKELNCK-EEDILDFELCLMDTQEPCFTGVYEEFIEGARFDNLLG 328
```

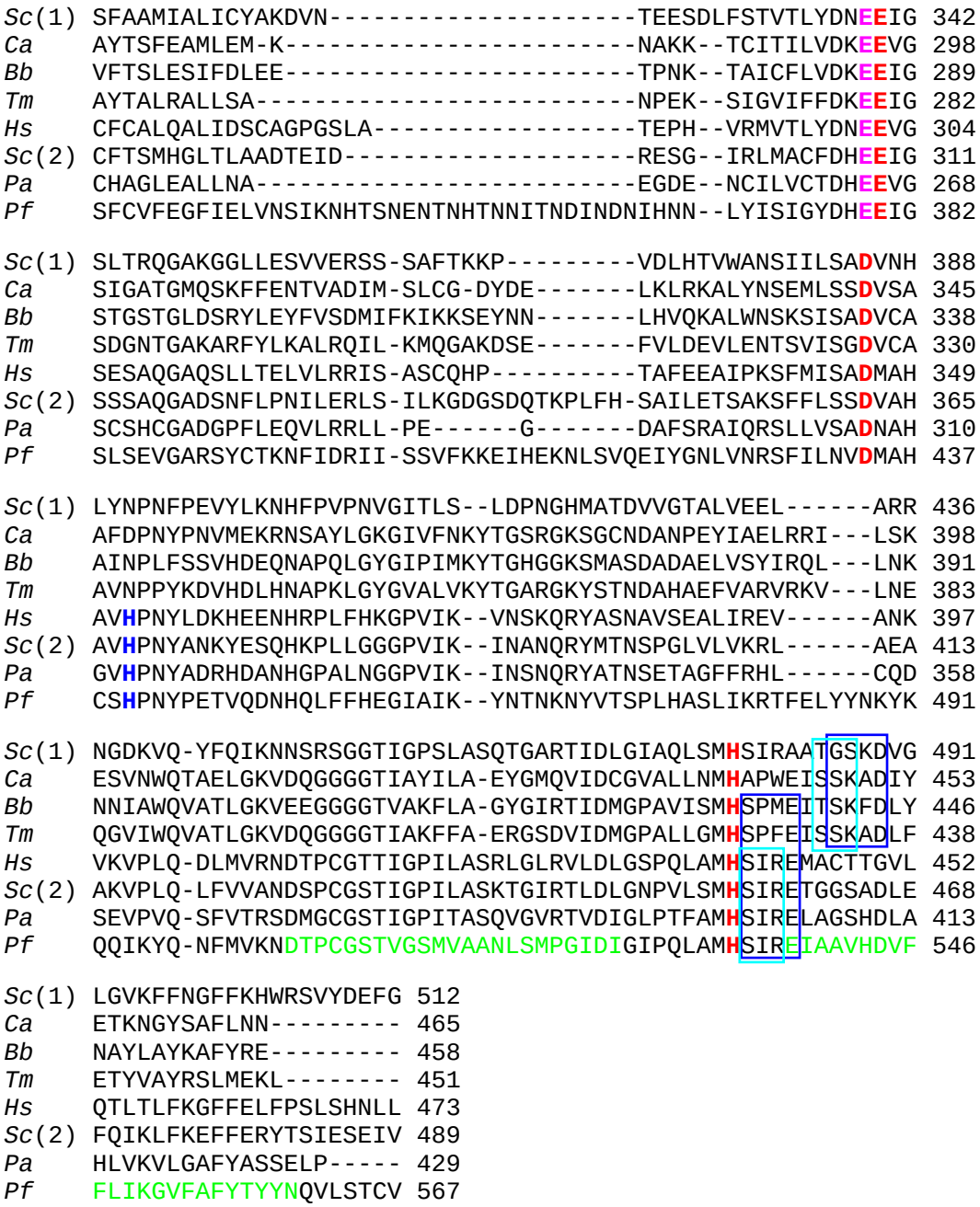


Figure 44: MEROPS database sequence alignment of M18 aminopeptidases. Amino acid sequence alignments of *S. cerevisiae* (*Sc* 1 & 2), *C. acetobutylicum* (*Ca*), *B. burgdorferi* (*Bb*), *T. maritima* (*Tm*), *H. sapiens* (*Hs*), *P. aeruginosa* (*Pa*), and *Pf*M18AAP (*Pf*). The N- and C-terminal amino acids of most of the sequences have been omitted by the creators of this alignment and therefore the first and last three amino acids are missing from the *Pf*M18AAP sequence. Amino acid colours: red = metal binding sites; pink and light blue = active site residues; dark blue = quaternary structure residue; green = putative transmembrane domains; and underlined yellow = spectrin-binding region. Box colours: light blue = putative protein kinase C phosphorylation site; and dark blue = putative casein kinase II phosphorylation site.

```
Pc    ----------MDKKAREYAQEAIKFIQRSGSSFMACKNLREKLESHGLIHIKEGDQ  46
Py    FVSVVILFKQMDKKAREYAQEALKFIQRSGSNFMACKNLREKLESHGLIHIKEGDQ  57
Pk    ---MQPTDDTMEKKVRDYAQGAVKFIKKSGSNFLACKNLREKLEEKGFKRIQEGEK  53
Pv    ----------MEKKAREYAQGAVRFIQKSGSNFLACKNLREKLEERGFKRIHEGEK  46
Pf    ----------MDKKAREYAQDALKFIQRSGSNFLACKNLKERLENNGFINLSEGET  46
                 *:**.*:*** *::**::***.*:*****:*:**..*: .: **:

Pc    WKLQKNQGYVLCKENRNICSFFVGKNFNINNGSILISIGHIDSCALKISPNNKVTK 102
Py    WKLQKNQGYVLCKENRNICSFFIGKNFNINNGSILISIGHIDSCTLKISPNNKVTK 113
Pk    WDLRKNEGYVFSKQNRNICGFFIGKDFNMEKGSILISIGHVDTCCLKISPNNNTVK 109
Pv    WELRKNEGYVLSKQSRNICGFFIGKDFTIEKGSILISIGHIDSCCLKVSPNNNVVK 102
Pf    WNLNKNEGYVLCKENRNICGFFVGKNFNIDTGSILISIGHIDSCALKISPNNNVIK 102
      *.*.**:***:.*:.****.**:**:*.::.*********:*:* **:****:. *

Pc    DQISQLNVECYGSGLWHTWFDRSLGLSGQVVYKKDNKLVEKIIQINKSVLFLPSLA 158
Py    DQISQLNVECYGSGLWHTWFDRGLGLSGQVVYKKDNKLIEKIIQINKSVLFLPSLA 169
Pk    SKVNQLNVECYGSGLWHTWFDRSLGLSGQVVYKKDDKLVEKLIQINRSIIFLPSLA 165
Pv    SKLHQLNVECYGSGLWHTWFDRSLGLSGQVLYKKEGKLVERLIQINKSLLFLPSLA 158
Pf    KKIHQINVECYGSGLWHTWFDRSLGLSGQVLYKKGNKLVEKLIQINKSVLFLPSLA 158
      .:: *:****************.*******:*** .**:*::****:*::******

Pc    IHLQNRTRYDFSVKVNYENHLKPILSTVLYEKLIKG-------------------- 194
Py    IHLQNRTRYDFSVKVNYENHLKPIISTLLYEKLIKG-------------------- 205
Pk    IHLQNRTRFEFSVKVNFENHLKPIISTVLYDQLIKGK------------------- 202
Pv    IHLQNRTRFEFSVKINYEAHLKPILSTLLYEHLVKG-------------------- 194
Pf    IHLQNRTRYDFSVKINYENHIKPIISTTLFNQLNKCKRNNVHHDTILTTDTKFSHK 214
      ********::****:*:* *:***:** *:::*  *

Pc    --------------------NENISEKNNS-----STDDEDKNSKN--------- 215
Py    --------------------NENILEKNISNIDDNNNDDDDMNSKN--------- 231
Pk    --------------------EKQNTDAFTEDTLHAEKIQDKCLNGDD--ASPSCLS 236
Pv    --------------------GKPGAASPTEDATDADNAQEKRLDAED--HSPSCHS 228
Pf    ENSQNKRDDQMCHSFNDKDVSNHNLDKNTIEHLTNQQNEEKNKHTKDNPNSKDIVE 270
                          : .    . .     . ::.  . .:

Pc    ----INSSPLLYLLANELKCKEDDILDFELCLMDTNQPCFTGVYEEFIEGARFDNL 267
Py    ----LNSSPLLYLLANELKCKEEDILDFELCLMDTNKPCFTGVYEEFIEGARFDNL 283
Pk    HQENPNSSPLLYTLAKELQCEEKDILDFELCLMDVNEPCFTGAYEEFIEGARFDNL 292
Pv    HQENPNSSPLLYTLAKELQCQEKDILDFELCLMDVNQPCFTGAYEEFIEGARFDNL 284
Pf    HINTDNSYPLLYLLSKELNCKEEDILDFELCLMDTQEPCFTGVYEEFIEGARFDNL 326
           ** **** *::**:*:*.**********.::*****.*************

Pc    LGTFSVFEAYIELIKMIKSE------------------------NNKNEPLENN 297
Py    LGTFGVFEAYVELIKNLKNE------------------------DNEN--LGNN 311
Pk    LGSYCVFEAFAEMIDMLKG-------------------KTPPSDGAVLPPEAHAN 328
Pv    LGSFCVFEAFAEMVDMLRGGAEAAAGAAAAAEGEASAAGAASAGAAAAPPPGAHAN 340
Pf    LGSFCVFEGFIELVNSIKNHT------------SNENTNHTNNITNDINDNIHNN 369
      **:: ***.: *::. ::.                                   *

Pc    LYICIGYDHEEIGSLSEIGAQSYFTKSFIERIIGNIFKNELKN------------ 340
Py    LYICIGYDHEEIGSLSEIGAQSYFTKNFIERILGNIFKNELKN------------ 354
Pk    LYICIGYDHEEIGSLSEVGAQSYFTQNFIKRILTAISSSQVGQNTHP--------S 376
Pv    LYICIGYDHEEIGSLSEVGAQSYFTQNFIKRILAAVCSSHACDAASATTSSAATAS 396
Pf    LYISIGYDHEEIGSLSEVGARSYCTKNFIDRIISSVFKKEIHE------------ 412
      ***.*************:**:** *:.**.**:  : ...  :
```

```
Pc    NETTVDEIYGSLSSRSLILNVDMAHCGHPNYPETIQQSHHLRFHDGIAIKYNTNKN 396
Py    NDITIDEIYGSLSNRSLILNVDMAHCGHPNYPETIQQNHHLRFHEGIAIKYNTNKN 410
Pk    SPTSIDELYGSLMSRSLILNVDMAHCSHPNYPETVQANHQLFFHEGIAIKYNTNRN 432
Pv    APPSIDELYGSLMSRSLLLNVDMAHCSHPNYPETVQASHQLFFHEGIAIKYNTNKN 452
Pf    KNLSVQEIYGNLVNRSFILNVDMAHCSHPNYPETVQDNHQLFFHEGIAIKYNTNKN 468
         :::*:**.* .**::********.*******:* .*:* **:**********:*

Pc    YVTSPYYACLLKRTFELYQNQYNQKIKYQNFMIKNDTPCGSTVGSMVASNLSMPGM 452
Py    YVTSPYYACLLKRTFELYQNQNNQKIKYQNFMIKNDTPCGSTVGSMVAANLSMPGM 466
Pk    YATSPYYTCLLKRTFELFTSKFNEKIKYQNFMVKNDTPCGSTVGSMVASNLSMPGV 488
Pv    YVTSPYYTCLLKRTFELFASNFNEKIKYQNFMVKNDTPCGSTVGSMVAANLSMPGI 508
Pf    YVTSPLHASLIKRTFELYYNKYKQQIKYQNFMVKNDTPCGSTVGSMVAANLSMPGI 524
      *.*** ::.*:******: .: :::*******:***************:******:

Pc    DIGISQLAMHSIREIAAIHDIYYLIKGVFAFYAYYNQVLSTCVHDA 498
Py    DIGIPQLAMHSIRELAAIHDIYYLVKGIFAFYAYYNQVLSSCVHDS 512
Pk    DIGIPQLAMHSIREIAAVRDIYYLVKGILAFYTYYSHVHASCVPDE 534
Pv    DIGIPQLAMHSIREIAAVRDVYYLVKGVLAFYAYYSHVLASCVPDA 554
Pf    DIGIPQLAMHSIREIAAVHDVFFLIKGVFAFYTYYNQVLSTCVHDK 570
      ****.*********:**::*:::*:**::***:**.:* ::** *
```

Figure 45: Sequence alignment of the *Plasmodium Pf*M18AAP homologues. ClustalW alignment of the *P. chabaudi chabaudi* (*Pc*), *P. yoelii yoelii* (*Py*), *P. knowlesi* (*Pk*), *P. vivax* (*Pv*) and *Pf*M18AAP amino acid sequences. Amino acid colours: red = metal binding sites; pink and light blue = active site residues; dark blue = quaternary structure residue; green = putative transmembrane domains; and underlined yellow = spectrin-binding region. Box colours: light blue = putative protein kinase C phosphorylation site; and dark blue = putative casein kinase II phosphorylation site. Alignment annotation: '*****' = residues that are identical in all sequences; '**:**' = conserved substitutions; and '**.**' = semi-conserved substitutions.

The spectrin-binding region (*Pf*216-257; labelled yellow in Figure 44 and Figure 45) is located within an amino acid stretch that is not found in any of the other aminopeptidases and represents a unique sequence insertion in *Pf*M18AAP. Despite this *P. falciparum*-specific insert (*Pf*195-234) and the larger size of the *P. falciparum* protein (16-72 amino acids longer), there is still a 64-74 % sequence identity (Table 16) between *Pf*M18AAP and the *Plasmodium* homologues.

4.3.3 *Pf*M18AAP secondary structure

The secondary structure of *Pf*M18AAP was analysed to determine what structural elements are found in the unique spectrin-binding region. These data were obtained from the PlasmoDB database or by sequence analysis with the SOSUI or Jpred programme. The PlasmoDB secondary structure prediction for the *Pf*M18AAP-specific spectrin-binding region (*Pf*216-257; orange box in Figure

46) showed that this region contains two alpha helices. This contrasts with the Jpred secondary structure prediction (labelled orange in Figure 47), which only revealed one alpha helix corresponding to the second peak in the PlasmoDB prediction. The SOSUI program predicted two transmembrane domains (*Pf*504-526 and *Pf*538-560; labelled in green in Figure 47), but this was not substantiated by the Kyte-Doolittle hydropathy plot provided by PlasmoDB (green boxes in Figure 46).

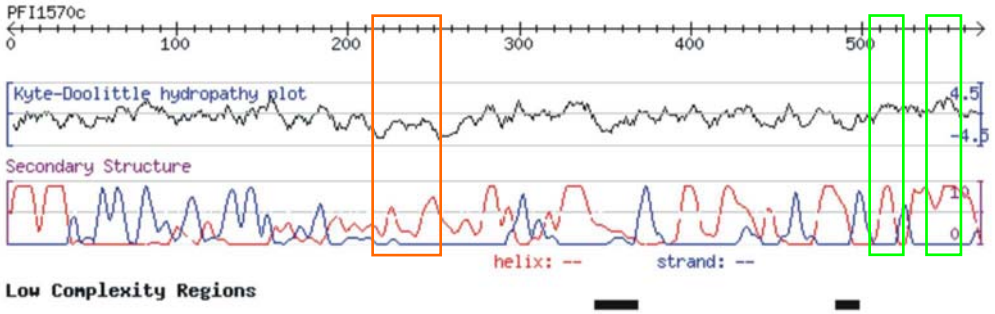


Figure 46: PlasmoDB hydropathy plot, secondary structure and low complexity regions for PFI1570c (*Pf*M18AAP).
Diagram showing the Kyte-Doolittle hydropathy plot, secondary structure and low complexity region prediction for the *Pf*M18AAP amino acid sequence. Secondary structures: red = alpha helix; and blue = beta strand. Box colours: orange = spectrin-binding region (*Pf*216-257); and green = putative transmembrane domains (*Pf*504-526 and *Pf*538-560).

Secondary structure predictions for the whole enzyme were performed with Jpred to ascertain if the different M18 aminopeptidases had a similar alpha helix, beta sheet and random coil pattern. Figure 47 shows the secondary structure alignment of the M18-family aminopeptidase 1 from *T. maritima*, the M18-family aminopeptidase 2 from *P. aeruginosa* and *Pf*M18AAP. A similar helix and sheet pattern was predicted for all the M18 aminopeptidases, indicating that *Pf*M18AAP could fold in a similar way to the other M18 aminopeptidases.

```
Tm    ----EAFSKEYMEFMSKAKTERMTVKEIKRILDESGFVPLEDF---AGDPMNMTVY  66
Pa    MRAEL--NQGLIDFLKASPTPFHATASLARRLEAAGYRRLDERDAW-HTETGGRYY  53
Pf    DKKAREYAQDALKFIQRSGSNFLACKNLKERLENNGFINLSEGETW-NLNKNEGYV  56

Tm    ----HHHHHHHHHH------HHHHHHHHHHHH-------------------EE  66
Pa    --HHH--HHHHHHHH------HHHHHHHHHHHH----E-------------EEE  53
Pf    ---HHHHHHHHHH------HHHHHHHHHHHH----HHH-----------EEE  56

Tm    AVNRGKAIAAFRVVDDL--KR-GLNLVVAHIDSPRLDFKPNPLIEDEQIALFKTHY 119
Pa    VTRNDSSLIAIRLGRRSPLES-GFRLVGAHTDSPCLRVKPNPEIARNGFLQLGVEV 108
Pf    LCKENRNICGFFVGKNFNIDTGSILISIGHIDSCALKISPNNNVIKKKIHQINVEC 112

Tm    EEE---EEEEEEE----------EEEEEE------------------EEEEEEEE 119
Pa    EEE---EEEEEEE----------EEEEEE-------E----------EEEEEEEE 108
Pf    EEE---EEEEEEE----------EEEEEE-------E----------EEEEEEEE 112

Tm    YGGIKKYHWLSIPLEIHGVLFKNDGT-EIEIHIGDKPEDPVFTIPDLLPHLDKE-- 172
Pa    YGGALFAPWFDRDLSLAGRVTFRA-NGKLESRLVD-FRKAIAVIPNLAIHL--NRA 160
Pf    YGSGLWHTWFDRSLGLSGQVLYKKGN-KLVEKLIQ-INKSVLFLPSLAIHL-QNRT 165

Tm    ------------------EEEEE-----EEEEEEE------EE---HHHH------ 172
Pa    ------------------EEEEE-----EEEEEE-------EE------------- 160
Pf    -----EEEE---------EEEEE-----EEEEEEE--------------------- 165

Tm    -DAKISEKFKGENLMLIAGTIPL-------------------------------- 194
Pa    ANE-GWPINAQNELPPIIAQLAPGEAAD--------------------------- 187
Pf    RYDFSVKINYENHIKPIISTTLFNQLNKCKRNNVHHDTILTTDTKFSHKENSQNKR 221

Tm    ----------------EEE------------------------------------ 194
Pa    --------------EEEE------------------------------------- 187
Pf    -----EEEEEEE----EEEEE---------EEEEE------EEEE----------- 221

Tm    -------------------------------------------------SGEEKEAVK 203
Pa    ----------------------------------------------------------FR 189
Pf    DDQMCHSFNDKDVSNHNLDKNTIEHLTNQQNEEKNKHTKDNPNSKDIVEHINTDNS 277

Tm    ----------------------------------------------------------H 203
Pa    ----------------------------------------------------------H 189
Pf    ----------------------HHHHHHHHHHH---------------------- 277

Tm    TNVLKILNEMYGITEEDFVSGEIEVVPAFSPREVGMDRSLIGAYGQDDRICAYTAL 259
Pa    LLLDEQLLREHGITADVVLDYELSFYDTQSAAVVGLNDEFIAGARLDNLLSCHAGL 245
Pf    YPLLYLLSKELNCKEEDILDFELCLMDTQEPCFTGVYEEFIEGARFDNLLGSFCVF 333

Tm    HHHHHHHHHH------HHHHHEEEEE----------HHHHH------HHHHHHHH 259
Pa    HHHHHHHHHHH----------EEE-------------HHH------HHHHHHHH 245
Pf    HHHHHHHHHH---------EEEEEEE-----------HHHH------HHHHHHHH 333

Tm    RALLSA------------------------NPEKSIGVIFFDKEEIGSDGNTGA 289
Pa    EALLNA------------------------EGDENCILVCTDHEEVGSCSHCGA 275
Pf    EGFIELVNSIKNHTSNENTNHTNNITNDINDNIHNNLYISIGYDHEEIGSLSEVGA 389

Tm    HHHH--------------------------------EEEEEE------------ 289
Pa    HHHHH--------------------------------EEEEEE--EEE-------- 275
Pf    HHHHH--------------------------------EEEEEE------------- 389
```

```
Tm    KARFYLKALRQILKMQGAKDSE-------FVLDEVLENTSVISGDVCAAVNPPYKD 338
Pa    DGPFLEQVLRRLLPE------G-------DAFSRAIQRSLLVSADNAHGVHPNYAD 318
Pf    RSYCTKNFIDRIISSVFKKEIHEKNLSVQEIYGNLVNRSFILNVDMAHCSHPNYPE 445

Tm    ---HHHHHHHHHHH------H-------HHHHHHHHH-EEE-------------- 338
Pa    ----HHHHHHHHHH-------H-------HHHHHHHH--HHHH------------- 318
Pf    ---HHHHHHHHHHH--------------HHHHHHHHH---EEE------------- 445

Tm    VHDLHNAPKLGYGVALVKYTGARGKYSTNDAHAEFVARVRKV---LNEQGVIWQVA 391
Pa    RHDANHGPALNGGPVIK--INSNQRYATNSETAGFFRHL------CQDSEVPVQ-S 365
Pf    TVQDNHQLFFHEGIAIK--YNTNKNYVTSPLHASLIKRTFELYYNKYKQQIKYQ-N 498

Tm    -------------EEEEEE-----------HHHHHHHHHHHH---HH-----EEEE 391
Pa    -------------EEEE--E---------HHHHHHHHHH------HHH----EE-E 365
Pf    -------------EEEE--E----------HHHHHHHHHHHHHH--------EE-E 498

Tm    TLGKVDQGGGGTIAKFFA-ERGSDVIDMGPALLGMHSPFEISSKADLFETYVAYRS 446
Pa    FVTRSDMGCGSTIGPITASQVGVRTVDIGLPTFAMHSIRELAGSHDLAHLVKVLGA 421
Pf    FMVKNDTPCGSTVGSMVAANLSMPGIDIGIPQLAMHSIREIAAVHDVFFLIKGVFA 554

Tm    EEEE--------E-HHHH------EEE----------HHHH----HHHHHHHHHHH 446
Pa    EEE---------HHHHHHHH----EEEE----E-----EEE---HHHHHHHHHHH 421
Pf    EEEE-----------HHH------EEE----------HHHH-----HHHHHHHHHH 554

Tm    LMEKL-------- 451
Pa    FYASSELP----- 429
Pf    FYTYYNQVLSTCV 567

Tm    HH----------- 451
Pa    HHH---------- 429
Pf    HHH-------EEE 567
```

Figure 47: Jpred secondary structure prediction of *Pf*M18AAP and its aminopeptidase homologues.
Jpred secondary structure predictions and MEROPS sequence alignments of the *T. maritima* (*Tm*), *P. aeruginosa* (*Pa*), and *Pf*M18AAP (*Pf*) aminopeptidases. Amino acid colours: red = metal binding sites; pink and light blue = active site residues; dark blue = quaternary structure residue; green = putative transmembrane domains; and underlined yellow = spectrin-binding region. Secondary structure predictions: H = alpha helix; E = beta-sheet or strand; '-' = random coil; and orange letters = predicted alpha-helix in spectrin-binding region.

4.3.4 Molecular weight and isoelectric point of the r*Pf*M18AAP monomer

To determine some of the chemical characteristics of the *Pf*M18AAP monomer, the molecular weight and the pI of the denatured enzyme was analysed by SDS-polyacrylamide gel and two-dimensional polyacrylamide gel electrophoresis.

The r*Pf*M18AAP monomer separated on a 12 % Laemmli SDS-polyacrylamide gel with a molecular weight of approximately 67 kDa. This value correlates well with the calculated molecular weight for r*Pf*M18AAP of 66.919 kDa. Removal of the histidine-tag resulted in a molecular weight of ~65 kDa for the r*Pf*M18AAP monomer (Figure 48).

The histidine-tag was removed from r*Pf*M18AAP by thrombin cleavage, so that it would not alter the pI of the aminopeptidase. Figure 48 shows that the His-tag was successfully removed from r*Pf*M18AAP after 4 hours of thrombin cleavage at room temperature (Lane T) and that there was no significant degradation of the protein throughout the procedure. A loss of protein was however observed during the cleavage procedure, because double the volume of protein (Lane T and Tu) had to be electrophoresed through the gel to be visible after Coomassie staining. This was probably due to some of the protein-containing supernatant remaining on the agarose beads at the end of the cleavage procedure, because it was very difficult to remove all the supernatant from the beads.

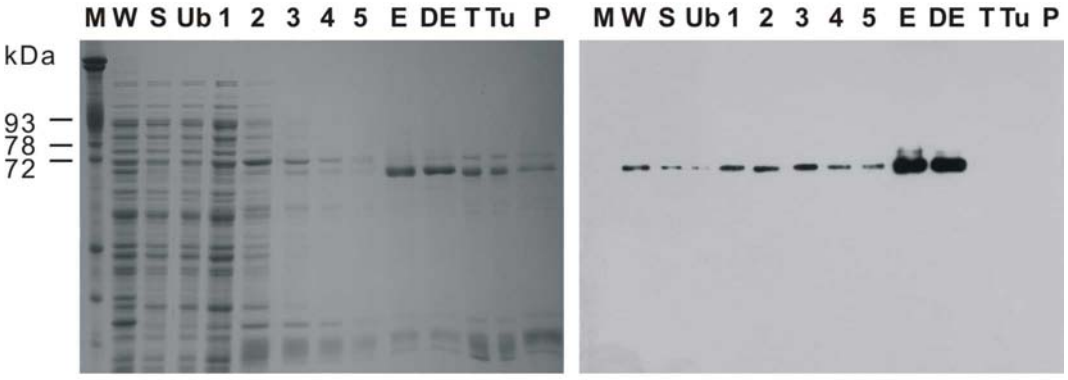


Figure 48: Thrombin cleavage of r*Pf*M18AAP.
Laemmli gel (left) and immunoblot with the Penta•His™ HRP Conjugate antibody (right) showing the purification and thrombin-cleavage of r*Pf*M18AAP. Lane M = 3 µg erythrocyte membrane proteins; W = 10 µl of the 8 ml induced *E. coli* whole cell extract; S = 10 µl of the 8 ml soluble protein fraction after sonication; Ub = 10 µl of the 8 ml protein fraction that did not bind to the HIS-Select™ Magnetic Agarose Beads; 1-5 = 20 µl of each 1.2 ml 20 mM imidazole wash; E = 1 µl of the 750 µl 200 mM imidazole elution; DE = 1 µl of 750 µl dialysed r*Pf*M18AAP; T = 2 µl of 875 µl thrombin-cleaved r*Pf*M18AAP; Tu = 2 µl of 875 µl unbound thrombin-cleaved r*Pf*M18AAP; P = 1 µl of precipitated r*Pf*M18AAP resuspended in 100 µl IEF buffer after thrombin cleavage.

The isoelectric point of r*Pf*M18AAP was determined by two-dimensional gel electrophoresis (Figure 49), where carrier ampholytes were used to set up the pH gradient. Rabbit muscle GAPDH (36 kDa), which has a pI of 8.3 and 8.5, was right at the edge of the gel, indicating that the pH gradient did not go up to pH 10. r*Pf*M18AAP had been resuspended in IEF buffer that contained pH 5/8 ampholytes, causing this segment of the pH gradient to spread over a greater distance of the gel and possibly causing the lower and higher pH ranges to be compacted at the gel ends. The spread of the pH gradient could also have been caused by a phenomenon called cathodic drift, in which the ampholytes responsible for the higher end of the pH gradient, are drawn out of the gel matrix into the upper chamber buffer (Westermeier, 1997).

The approximate pI values of the three charge isomers of r*Pf*M18AAP were estimated from the position of the standards and were found to be ~6.6, ~6.7, and ~6.9. Apart from the standards, additional protein spots can be seen (not circled in Figure 49). These contaminants were also seen during purification (Figure 48) and when performing the purification procedure from *E. coli* cells that only contained the pET-15b plasmid (results not shown). A large amount of r*Pf*M18AAP can also be seen at the edge of the pH gradient (at pH 8.5), indicating that some of the protein did not electrophorese into the IEF gel. This was surprising because r*Pf*M18AAP was resuspended in IEF buffer that contained 9.5 M urea. Protein aggregation could have been due to the r*Pf*M18AAP high protein concentration required for the IEF procedure.

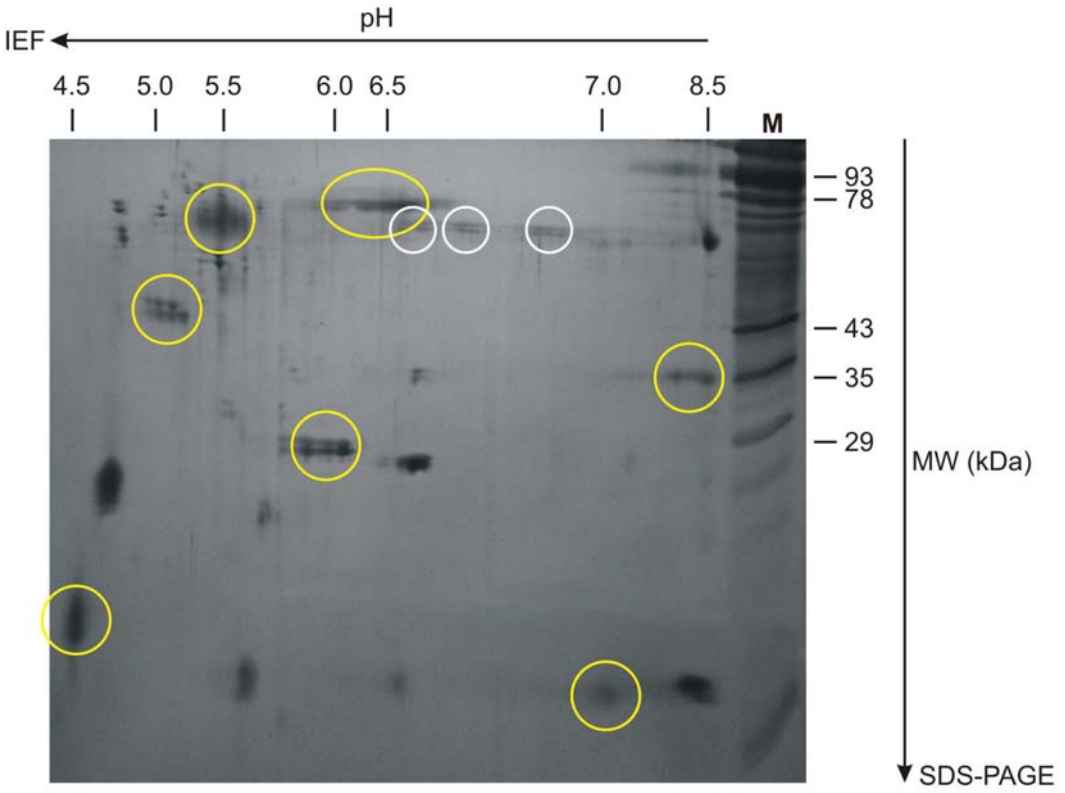


Figure 49: Two-dimensional gel electrophoresis of r*Pf*M18AAP.
Second dimension silver stained Laemmli gel showing the Bio-Rad 2-D SDS-PAGE Standards and r*Pf*M18AAP after they had been separated by IEF. The three charge isomers (circled in white) of r*Pf*M18AAP have a pI of ~6.6, ~6.7 and ~6.9 respectively. Bio-Rad 2-D SDS-PAGE Standards (circled in yellow): Hen egg white conalbumin type 1 (MW = 76 kDa; pI = 6.0, 6.3, 6.6); BSA (MW = 66.2 kDa; pI = 5.4 & 5.6); Bovine muscle actin (MW = 43 kDa; pI = 5.0 & 5.1); Rabbit muscle GAPDH (MW = 36 kDa; pI = 8.3 & 8.5); Bovine carbonic anhydrase (MW = 31 kDa; pI = 5.9 & 6.0); Soybean trypsin Inhibitor (MW = 21.5 kDa; pI = 4.5); Equine myoglobin (MW = 17.5 kDa, pI = 7.0). Laemmli gel marker (M) = 15 µl erythrocyte membrane proteins.

4.3.5 The oligomeric state of native r*Pf*M18AAP

The four M18 aminopeptidase crystal structures show that these enzymes occur as multimers in their native state. r*Pf*M18AAP was therefore electrophoresed through non-denaturing gels to determine the approximate molecular size and hence the number of subunits of the native protein.

Figure 50 shows that several oligomeric forms of r*Pf*M18AAP were separated on the native gels. As the percentage acrylamide increases, protein bands are

separated, which comigrate on the 2 and 3 % polyacrylamide native gels. The first and third band can be seen clearly on the gels and the second band is very faint. A large amount of protein can also be seen at the top of the gels.

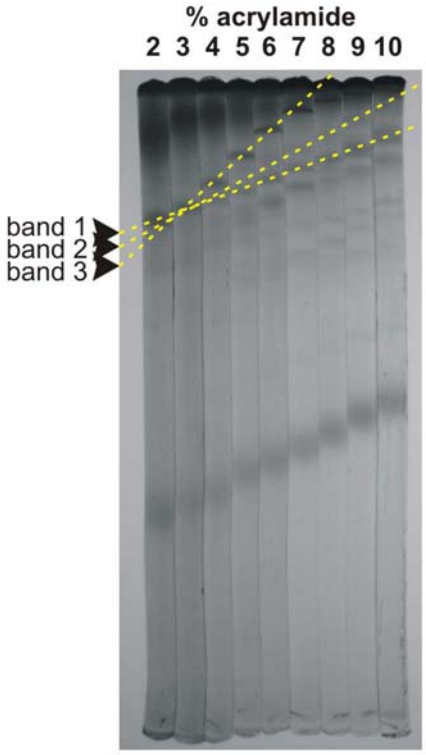


Figure 50: Non-denaturing gel electrophoresis of r*Pf*M18AAP.
Photograph of 2-10 % native tube gels showing the protein bands of r*Pf*M18AAP. Twenty-five micrograms of recombinant r*Pf*M18AAP was electrophoresed through each native gel and stained with Coomassie Blue stain.

The Ferguson plot linear regression lines (Figure 51), which were calculated from the R_m values of each r*Pf*M18AAP band, intersect between 3 and 4 % gel concentration, indicating that these bands are oligomers of r*Pf*M18AAP (Hedrick and Smith, 1968, Rodbard *et al.*, 1974).

The approximate molecular weights of the r*Pf*M18AAP subunits that were estimated from the standard curve (Figure 52) are listed in Table 17. The second band is approximately double the size of the r*Pf*M18AAP monomer (band 1) and is therefore a dimer, and the third band is four times the size of the monomer and is therefore a tetramer of r*Pf*M18AAP (Table 17).

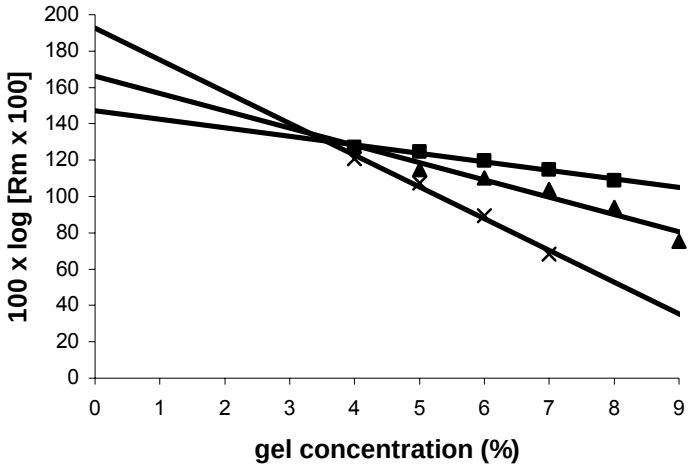


Figure 51: Ferguson plot of the r*Pf*M18AAP protein bands.
Graph showing the log R_m of the r*Pf*M18AAP monomer, dimer and tetramer at different polyacrylamide percentages. These proteins are oligomers of each other because the lines intersect above 3 % gel concentration (Hedrick and Smith, 1968, Rodbard *et al.*, 1974). Squares = first r*Pf*M18AAP band; triangles = second r*Pf*M18AAP band; and crosses = third r*Pf*M18AAP band 3.

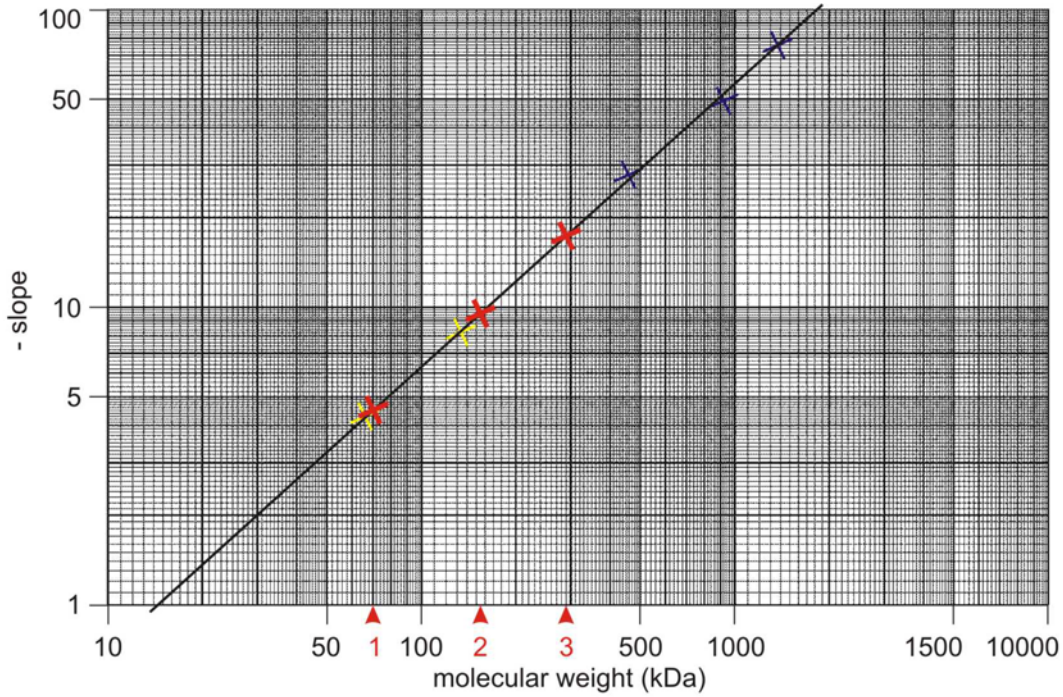


Figure 52: Non-denaturing gel electrophoresis standard curve.
Double-log graph of the negative slopes (obtained from Ferguson plots) versus the molecular weight of each standard. The molecular weight of the three r*Pf*M18AAP bands was read from this graph and found to be ~70, ~155, and ~290 kDa respectively. Yellow crosses = BSA (66 and 132 kDa); blue crosses = spectrin (460, 920, and 1380 kDa); red crosses and numbers = r*Pf*M18AAP.

Table 17: Approximate molecular weights of the r*Pf*M18AAP subunits.

Band	Molecular weight from standard curve (kDa)	Subunit
1	~70	monomer
2	~155	dimer
3	~290	tetramer

The native gel protein bands were transferred onto nitrocellulose membrane and the presence of the histidine-tag in r*Pf*M18AAP detected by chemiluminescence. The r*Pf*M18AAP monomer was seen as a distinct band and the tetramer is part of the smear above the monomer (Figure 53). Several other protein bands were not visualised on the immunoblot indicating that these proteins are contaminants isolated during the purification procedure.

The smear extends to the top of the gels indicating that even higher oligomeric forms of r*Pf*M18AAP are present. These higher oligomers could include an octamer, which is the r*Pf*M18AAP subunit composition that Teuscher *et al.* (2007) determined by assaying HPLC fractions for enzymatic activity. Octamers were also observed for the *H. sapiens* and *S. cerevisiae* aspartyl aminopeptidases (Wilk *et al.*, 2002, Yokoyama *et al.*, 2006). However, the smear could also include a dodecamer, which is the native subunit composition of the crystal structures of the *T. maritima*, *C. acetobutylicum* and *P. aeruginosa* M18 aminopeptidases (Min and Shapiro, 2006).

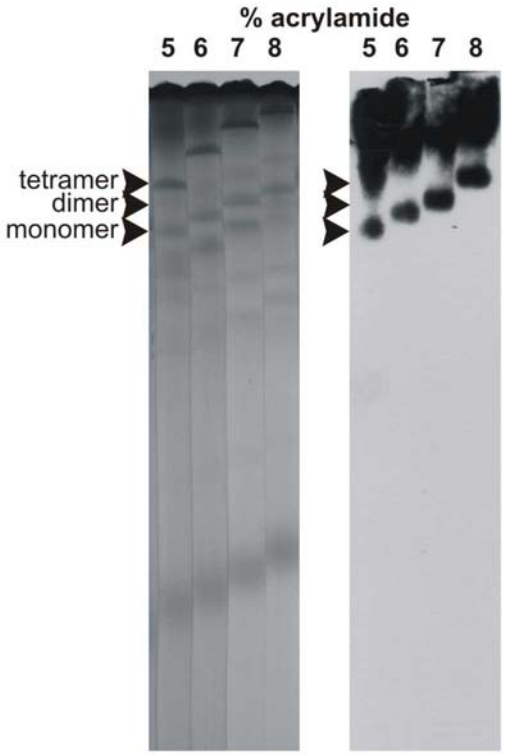


Figure 53: Detection of r*Pf*M18AAP in non-denaturing gels.
Photograph of 5-8 % native tube gels (left) and immunoblot with the Penta•His™ HRP Conjugate antibody (right) showing the monomer and oligomers of r*Pf*M18AAP. Fifteen to twenty-five micrograms of recombinant r*Pf*M18AAP was electrophoresed through each native gel. The tetramer is seen as an intense protein band in the Coomassie Blue stained gels, but a large amount of r*Pf*M18AAP remained at the top of each gel, indicating that higher oligomeric forms of r*Pf*M18AAP are present. This is confirmed by the immunoblot, where the higher oligomeric forms are part of the smears.

4.3.6 *Pf*M18AAP tertiary structure

The tertiary structures of four M18 aminopeptidases were studied to gain insight on the structure of *Pf*M18AAP and to study the location and conformation of the active site and the spectrin-binding region.

The monomers of the *C. acetobutylicum*, *T. maritima*, *B. burgdorferi* and *P. aeruginosa* M18 aminopeptidase crystal structures (Min and Shapiro, 2006) were compared with each other using the Swiss-PdbViewer and showed similar tertiary structures. The monomer of *T. maritima* was the most complete structure (i.e. no

gaps were present in the amino acid sequence) and was therefore used for the structure analysis and comparisons with *Pf*M18AAP.

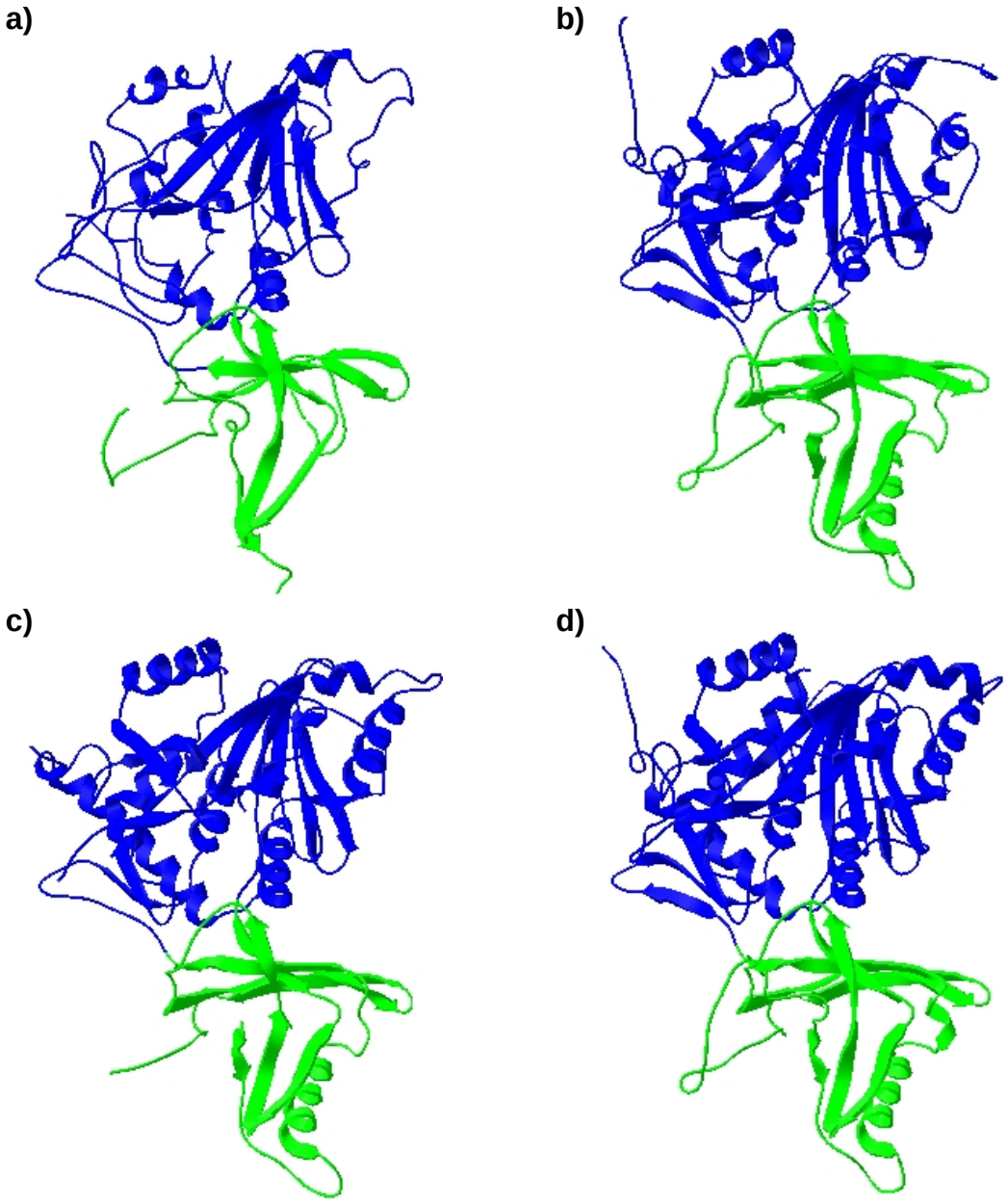


Figure 54: Crystal structures of the monomers of the M18 aminopeptidases. Diagram showing similar ribbon structures for the *Pseudomonas aeruginosa* (2ijz) (a), *Clostridium acetobutylicum* (2glj) (b), *Borrelia burgdorferi* (1y7e) (c) and *Thermotoga maritima* (2glf) (d) M18 aminopeptidases. Ribbon structures were created with DeepView Swiss-PdbViewer, version 3.7 (SP5). The PDB codes of the different proteins are listed in brackets. Blue = domain I and green = domain II.

The *T. maritima* aminopeptidase monomer contains two distinct structural regions that will be termed domain I and II (Figure 55a). Domain I is not contiguous and is formed by the N- and C-terminal amino acids of the enzyme. The active site lies at the edge of domain I, in close proximity to domain II, and forms a pocket in which the cofactor-binding amino acids, *Tm*H93, *Tm*D250, *Tm*E280, *Tm*D327 and *Tm*H426 (*Pf*H86, *Pf*D324, *Pf*E380, *Pf*D434 and *Pf*H534) and the two cofactors are located (Figure 55b inset). The two amino acids, *Tm*E279 (*Pf*E379) and *Tm*D95 (*Pf*D88), which are directly and indirectly involved in substrate cleavage are positioned at the edge of the active site. Domain II is formed by *T. maritima* amino acids *Tm*97-232 and is thus an internal section of the protein. The histidine, *Tm*H168 (*Pf*H160), that Wilk *et al*. (2002) found to be essential for enzyme activity in the *H. sapiens* aspartyl aminopeptidase, is located in an unstructured loop of domain II and faces in the opposite direction to the active site in domain I. When two monomers (A and B in Figure 55b) associate according to the parameters listed in the crystal structure file (Min and Shapiro, 2006), *Tm*H168 of monomer B completes the active site in monomer A (Figure 55b inset) and *vice versa*. The histidine, *Hs*H352 (*Pf*H440), speculated to be involved in quaternary structure stabilisation in the *H. sapiens* aspartyl aminopeptidase (Wilk *et al.*, 2002), is not present in *T. maritima* and is therefore not shown in the crystal structure. However, if this residue were present, it would be located on domain I of one monomer and could interact with domain II of the other monomer, thereby stabilising the quaternary structure. The putative protein kinase C and casein kinase II phosphorylation sites are located next to a histidine, *Tm*H426 (*Pf*H534), which forms part of the active site in domain I and therefore these two phosphorylation sites could be used to regulate enzymatic activity. The active sites are however located inside the native enzyme (Figure 56) and access to the phosphorylation sites by kinases could be limited. However, phosphorylation could also occur before the monomers are assembled into the quaternary structure.

The *T. maritima* M18 aminopeptidase occurs as dodecamer in its native state and has the symmetry of a tetrahedron. Because the dimers lie along the edge of the triangles (Figure 56a), three domain Is, each containing an active site, form a

single vertex and three domain IIs form the flat surface of each triangle in the tetrahedron. Franzetti *et al.* (2002) described the crystal structure of a tetrahedral M42 aminopeptidase, which is also classified in the MH clan in the MEROPS database. In this enzyme, substrate access to the active sites located within the tetrahedron occurs via pores on the faces (21 Å) and in the vertices (17 Å) of the active enzyme.

Secondary structure prediction showed that the *T. maritima* aminopeptidase is similar to r*Pf*M18AAP (Figure 47), and therefore the quaternary structure of these two enzymes could be similar. If this is the case, then the amino acids, which form the spectrin-binding region in the *Pf*M18AAP structure, would be located in domain II on the opposite side to domain I (asterisk in Figure 55a). When the transformation parameters of the *T. maritima* crystal structure file (Min and Shapiro, 2006) are applied to build a dimer and a dodecamer, the spectrin-binding region remains on the surface of these two structures (asterisks in Figure 55b and black amino acids on the surface of the ribbon structure in Figure 56a). If *Pf*M18AAP is a dodecamer, then each triangle face of the tetrahedron would have three spectrin-binding regions extending as loops from its surface (white loops in Figure 56b) and the three active sites would be located on the inside of each vertex of the tetrahedron (red circles in the inset of Figure 56a).

Finally, the putative transmembrane domains of *Pf*M18AAP are found near the C-terminus of the protein. Since the C-terminus of the *T. maritima* aminopeptidase is embedded in domain I (Figure 55) and contributes towards the tertiary structure of the domain (Min and Shapiro, 2006), it is unlikely that the putative transmembrane domains of *Pf*M18AAP function as membrane anchors.

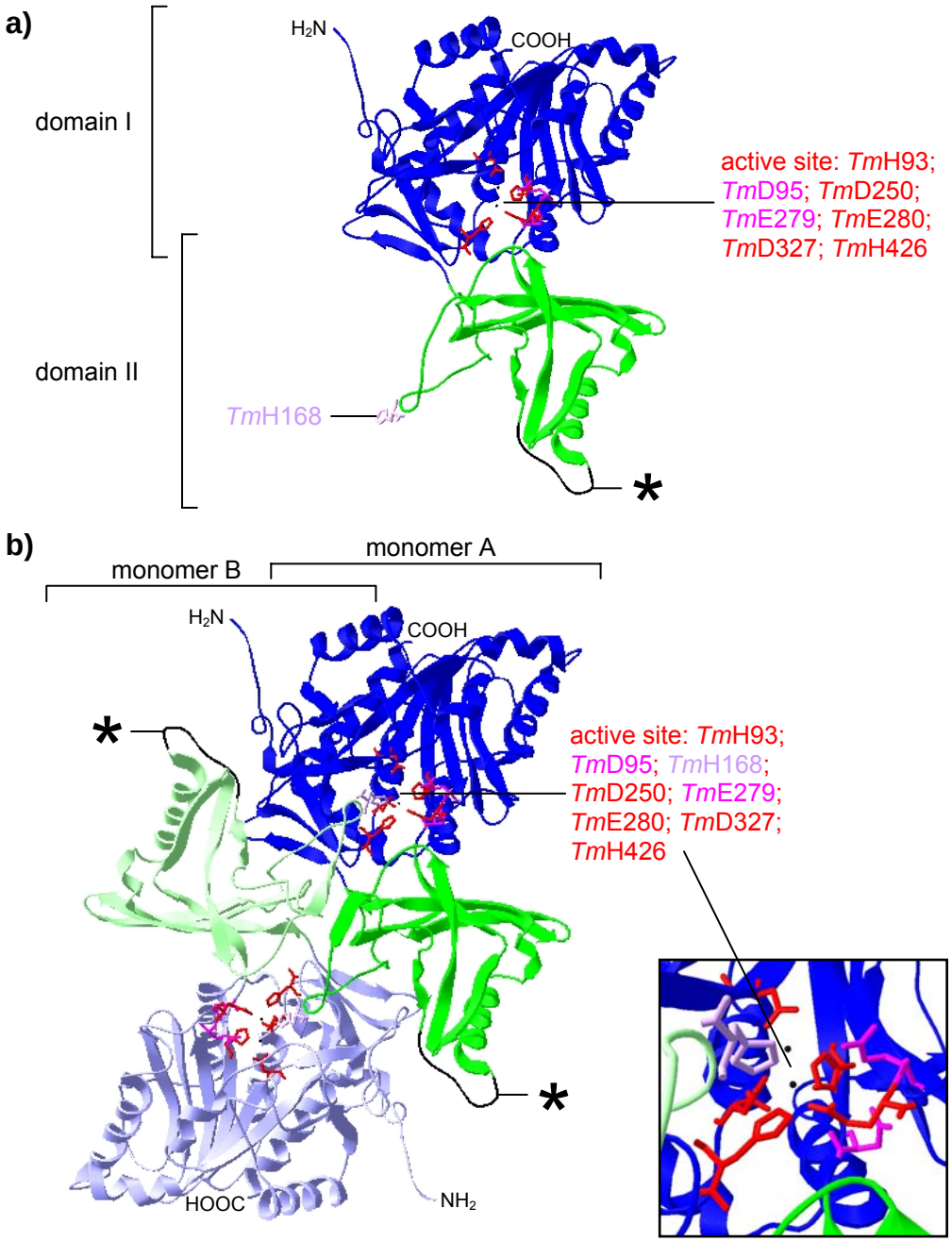


Figure 55: The *T. maritima* M18 aminopeptidase monomer and dimer.
Diagram showing the ribbon structure of a monomer (a) and dimer (b) of *T. maritima* M18-family aminopeptidase 1 (PDB code: 2glf) (Min and Shapiro, 2006). Domain I of monomer A and B are shown in blue and light blue respectively. Domain II of both monomers are depicted in green and light green respectively. The asterisk marks the predicted location of the *Pf*M18AAP spectrin-binding region in the structure. The inset shows the active site with the two manganese atoms (black dots) surrounded by the metal binding amino acids (red residues: *Tm*H93, *Tm*D250, *Tm*E280, *Tm*D327 and *Tm*H426) and the additional histidine involved in enzyme activity (lavender residue: *Tm*H168). The two amino acids that are directly and indirectly involved in substrate cleavage are located at the edge of the active site (pink residues: *Tm*E279 and *Tm*D95). Structures were created with DeepView Swiss-PdbViewer, version 3.7 (SP5).

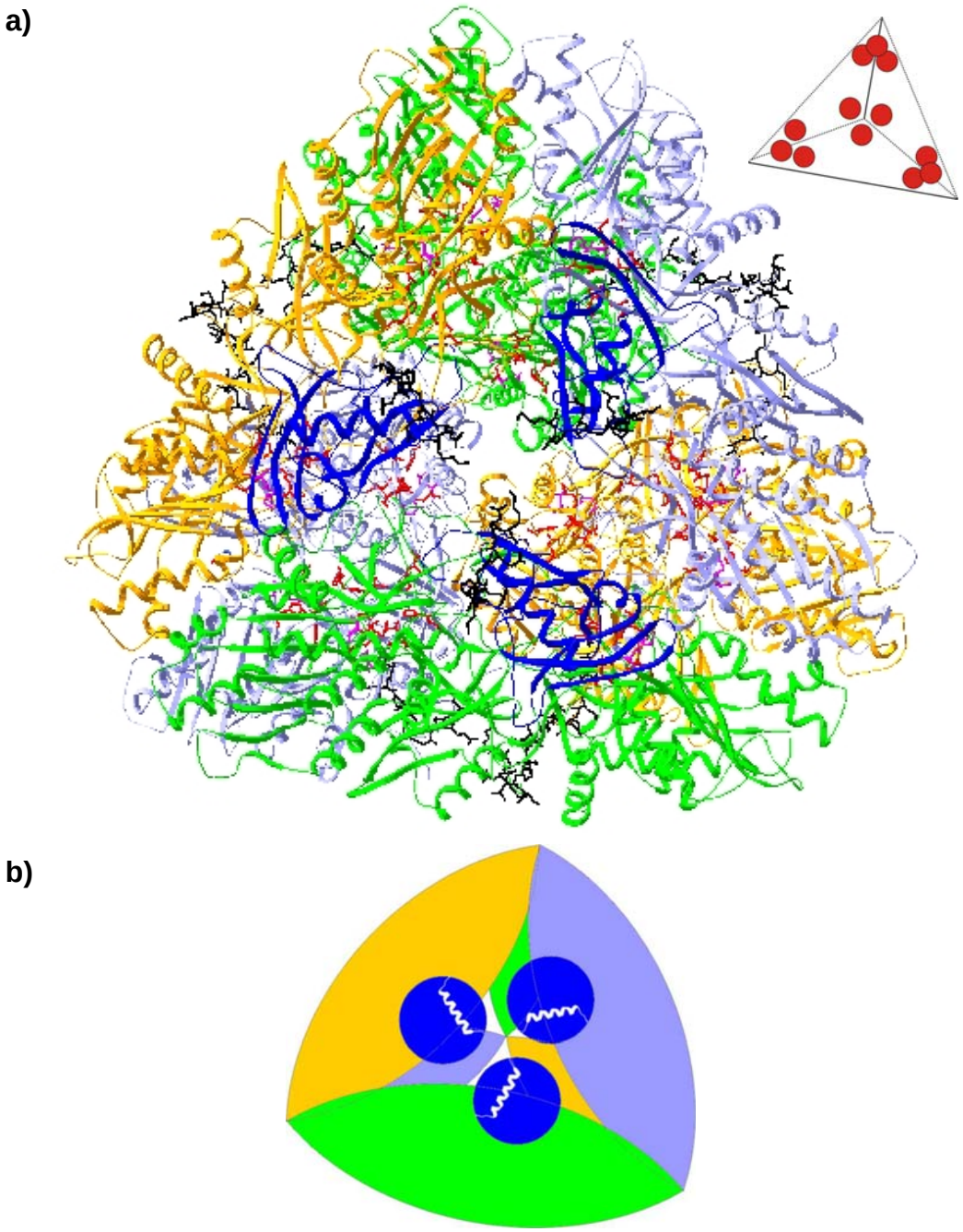


Figure 56: The M18 aminopeptidase dodecamer.
Ribbon structure and diagram showing the quaternary structure of *T. maritima* M18-family aminopeptidase 1 (PDB code: 2glf) (a) (Min and Shapiro, 2006) and diagram of a possible structure outline of *Pf*M18AAP (b). Dimers (yellow, green and blue) form the edges of the tetrahedron and each triangle face contains three domain IIs (dark blue ribbons in (a) and dark blue circles in (b)), on which the spectrin-binding region of *Pf*M18AAP would be located (black amino acids in (a) or white loops in (b)). Three active sites (red amino acids in (a) and red circles in inset) lie within the vertices of the tetrahedron and the substrate can access these sites via pores located in the faces and vertices of the enzyme. The ribbon structure was created with DeepView Swiss-PdbViewer, version 3.7 (SP5).

4.3.7 r*Pf*M18AAP enzyme activity

*Pf*M18AAP is described in the PlasmoDB database as an aminopeptidase that has homology to the human aspartyl aminopeptidase. The coupled enzyme assay designed to assay the human aminopeptidase was thus used to test if *Pf*M18AAP is an active aminopeptidase and to establish the optimal cleavage conditions.

r*Pf*M18AAP was active as shown in Figure 57. The middle test tube shows the colour development after r*Pf*M18AAP was allowed to cleave the substrate, in contrast to the blank reaction (left test tube) and one of the standard curve samples (right test tube).



Figure 57: Colour reaction of the r*Pf*M18AAP enzyme activity assay.
Photograph of test tubes after completion of the substrate cleavage reaction and azo dye development. Left = blank; middle = r*Pf*M18AAP incubated with 0.3 mM Asp-Ala-Pro-β-Napthylamide and dipeptidyl peptidase IV; right = 0.8 mM β-Napthylamide (3rd highest standard).

After it was established that r*Pf*M18AAP cleaved Asp-Ala-Pro-β-Napthylamide, enzyme assay conditions were optimised (results not shown) so that product formation was linear with time and the substrate concentration exceeded enzyme concentration. These experimental conditions were used to determine the optimal pH, temperature and ionic strength for the enzyme reaction. Figure 58 shows that r*Pf*M18AAP was most active at pH 7.5 and at 37 ºC and that the enzyme retained most of its enzymatic activity between pH 7-8.5 and 33-39 ºC. Varying the ionic strength of the 50 mM Tris buffer (50 or 150 mM NaCl) had no effect on enzymatic activity (results not shown). Even though overnight storage with 20 % glycerol at 4 ºC only decreased the enzymatic activity of r*Pf*M18AAP by 2 %, all studies were performed directly after purification.

The enzymatic activity of r*Pf*M18AAP was studied in the presence of metal ions (Table 18) and showed that none of the tested metal ions increased the activity of r*Pf*M18AAP. Magnesium and manganese had no effect on substrate cleavage, indicating that these metal ions could be the actual or substitute cofactors of the enzyme. In contrast, zinc and calcium had a strong (>90 %) and moderate (~20 %) inhibitory effect on enzymatic activity, suggesting that these metal ions are not the cofactors of the enzyme. Cobalt was not tested, but Teuscher *et al*. (2007) showed that it increased the activity.

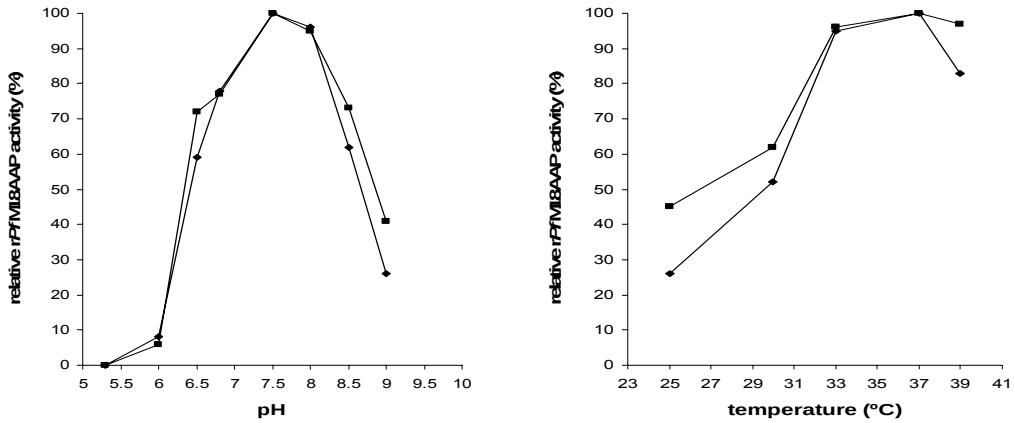


Figure 58: pH and temperature optima of r*Pf*M18AAP.
Graphs showing the relative activity (%) of r*Pf*M18AAP over a pH and temperature range. Left graph = 2.5 µg r*Pf*M18AAP was incubated with 0.3 mM Asp-Ala-Pro-β-Napthylamide and 0.125 U dipeptidyl peptidase IV for 15 minutes at 37 ºC in different buffers. For pH 6.8, 7.5, 8, 8.5, and 9, 50 mM Tris-HCl buffer was used. For pH 5.3 and 6, 0.1 M sodium citrate buffer was used. Right graph = 2.5 µg r*Pf*M18AAP was incubated with 0.3 mM Asp-Ala-Pro-β-Napthylamide and 0.125 U dipeptidyl peptidase IV in 50 mM Tris-HCl buffer (pH 7.5) for 15 minutes at 25, 30, 33, 37, and 39 ºC. The curves represent two independent experiments.

Table 18: r*Pf*M18AAP relative activity in the presence of metal ions.

Cofactor	Relative activity (%) ± SEM*
no cofactor	100
1 mM Zn^{2+}	7 ± 3
1 mM Mg^{2+}	89 ± 6
1 mM Mn^{2+}	89 ± 9
1 mM Ca^{2+}	78 ± 8

* sample size = 4
SEM = standard error of the mean

Enzymatic assays in the presence of enzyme inhibitors (Table 19) showed that EDTA, which is a metal chelator and therefore affects metalloproteases, had the strongest effect on r*Pf*M18AAP (46 % loss of activity). This effect was increased to 57 % when the enzyme was preincubated with EDTA for 10 minutes at 37 ºC. Bestatin, which specifically inhibits aminopeptidases, only inhibited r*Pf*M18AAP by 20 %, but this effect was enhanced to 39 % when r*Pf*M18AAP was preincubated with this competitive inhibitor for 10 minutes at 37 ºC. Presumably this allowed the dipeptide analogue, which cannot be processed by aminopeptidases, to bind in the r*Pf*M18AAP active site.

The serine protease inhibitor, Pefabloc SC, inhibited enzymatic activity, but another serine protease inhibitor, PMSF, showed only a slight inhibition. The difference could be attributed to the instability of PMSF at elevated temperatures and the insolubility of PMSF in aqueous solutions. PMSF stock solutions are prepared in DMSO and addition of this stock solution to the reaction buffer may cause precipitation of the inhibitor, resulting in a decreased amount of available inhibitor. The active site of r*Pf*M18AAP has serine residues adjacent to two of the amino acids that are involved in cofactor binding and substrate cleavage (*Pf*H534 and *Pf*D88). The serine protease inhibitors could have bound to these serine molecules and prevented the substrate from entering the active site, thereby reducing enzymatic activity.

Table 19: r*Pf*M18AAP relative activity in the presence of protease inhibitors.

Inhibitor	Relative activity (%) ± SEM* with no preincubation at 37 ºC
no inhibitor	100
10 mM EDTA	54 ± 13
1mM Pefabloc SC	61 ± 8
1 mM PMSF	90 ± 7
10 µM Bestatin	79 ± 5

* sample size = 4
SEM = standard error of the mean

It was problematic to determine a reliable value for K_m and V_{max} from all the experiments because a fresh enzyme preparation was used each time, with variable activity. In addition, slight substrate inhibition was observed at high

concentrations (>1.5 mM). Figure 59 shows velocity *versus* substrate curves for three experiments with a high, medium and low enzyme activity of r*Pf*M18AAP and due to the inconsistent behaviour of the enzyme, K_m and V_{max} values were only determined from experiments showing medium and high enzymatic activity.

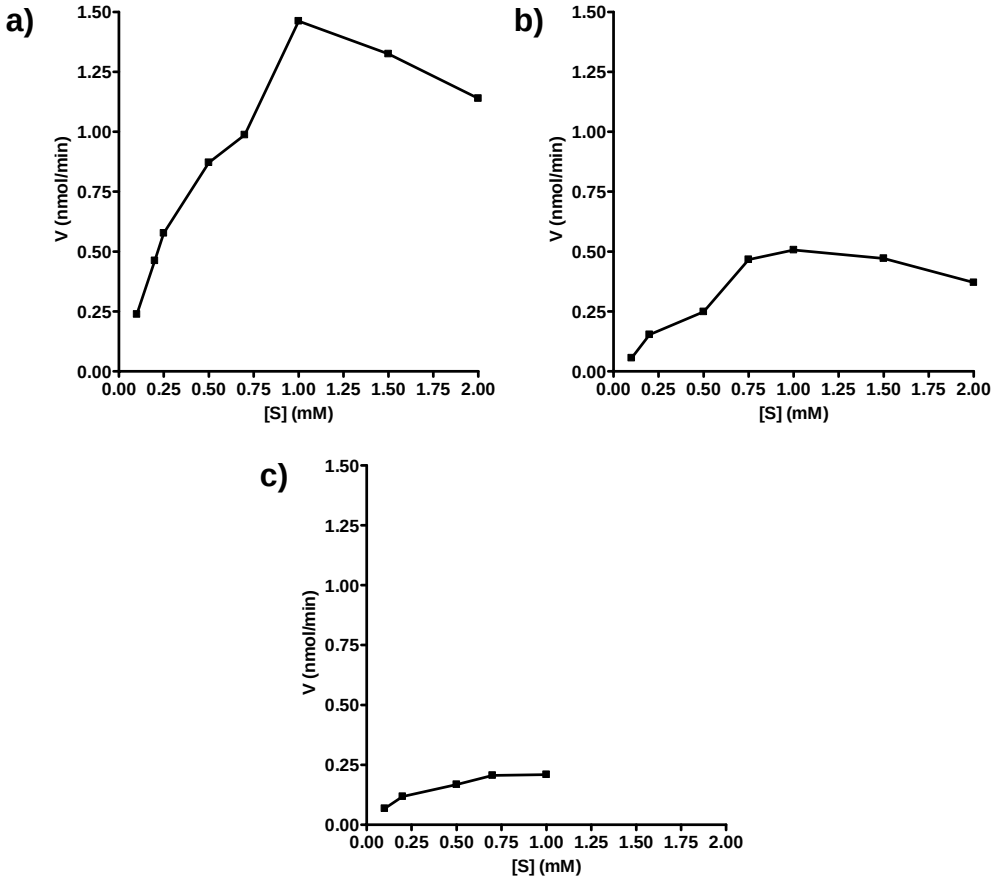


Figure 59: Variation in r*Pf*M18AAP enzymatic activity.
Plots of velocity (V) against substrate (S) according to the Michaelis-Menten equation showing a high (a), medium (b) and low (c) enzymatic activity for different preparations of r*Pf*M18AAP.

The Michaelis-Menten calculation does not take substrate inhibition into consideration (Figure 60a) and therefore a Lineweaver-Burk plot (Figure 60b) was used to determine the following approximate K_m and V_{max} values for r*Pf*M18AAP:

K_m = ~0.8 mM

V_{max} = ~2 nmol/min

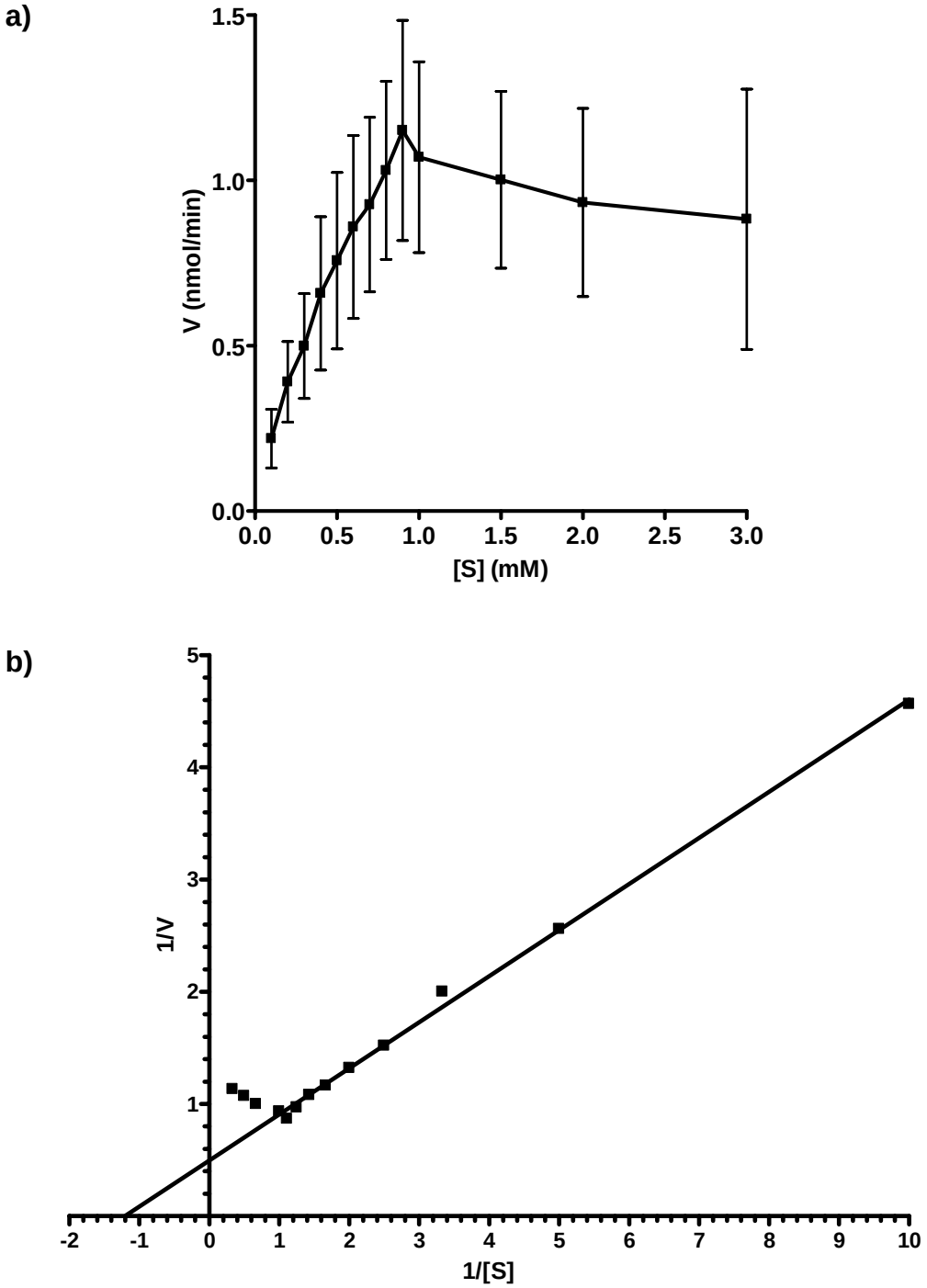


Figure 60: Michaelis-Menten and Lineweaver-Burk plots for r*Pf*M18AAP. Plot of velocity (V) against substrate (S) according to the Michaelis-Menten equation (a) and double-reciprocal plot of 1/V against 1/S (b) for r*Pf*M18AAP. Substrate inhibition occurs at ~1.5 mM Asp-Ala-Pro-β-Napthylamide in the Michaelis-Menten graph and therefore the mean velocity of 1.5, 2, and 3 mM Asp-Ala-Pro-β-Napthylamide was not used for the Lineweaver-Burk plot. Michaelis-Menten plot: N = 3; error bars represent standard error of the mean. Lineweaver-Burk plot: K_m = 0.83 mM; V_{max} = 2.02 nmol/min. Graphs and transformation data were generated with GraphPad Prism® (Version 4.0.3).

4.3.8 Binding studies

The interaction of the *Pf*M18AAP fragment with spectrin, which was identified by phage display, was confirmed by overlaying full length r*Pf*M18AAP on a dot and slot blot containing spectrin dimers and tetramers (Figure 61). r*Pf*M18AAP binds to both spectrin dimers and spectrin tetramers.

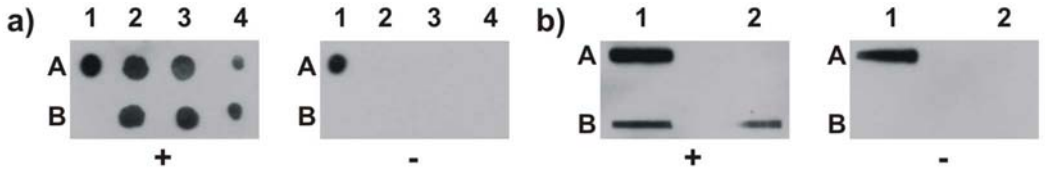


Figure 61: Spectrin dimers and tetramers overlaid with r*Pf*M18AAP.
Photographs of dot (a) and slot (b) blots containing spectrin dimers and tetramers overlaid with 5 µg r*Pf*M18AAP (+) and 50 mM Tris buffer (pH 7.5) (-). r*Pf*M18AAP was detected with the Penta•His™ HRP Conjugate antibody. Dot blots: A1 = 100 ng r*Pf*M18AAP; A2-4 = 2, 1.4 and 0.5 µg spectrin dimers; B1 = 100 ng BSA; B2-4 = 2, 1.4 and 0.5 µg spectrin tetramers. Slot blots: A1 = 100 ng r*Pf*M18AAP; A2 = 100 ng BSA; B1 = 1 µg spectrin dimers; B2 = 1 µg spectrin tetramers.

The blot overlays were also used to determine if r*Pf*M18AAP bound to other erythrocyte membrane proteins. Membrane proteins separated on Laemmli or Fairbanks SDS-polyacrylamide gels, transferred to nitrocellulose and overlaid with r*Pf*M18AAP, showed that the enzyme interacted strongly with β-spectrin, protein 4.1, protein 4.2, actin and G3PD (Figure 62a and b). A weak interaction with band 3 (Laemmli blot) and α-spectrin (Fairbanks blot) was also observed. The transfer of large molecules such as α- and β-spectrin onto the nitrocellulose membrane from the low percentage Fairbanks gel was better than from the high percentage Laemmli gel, accounting for the signals on the Fairbanks blot overlay.

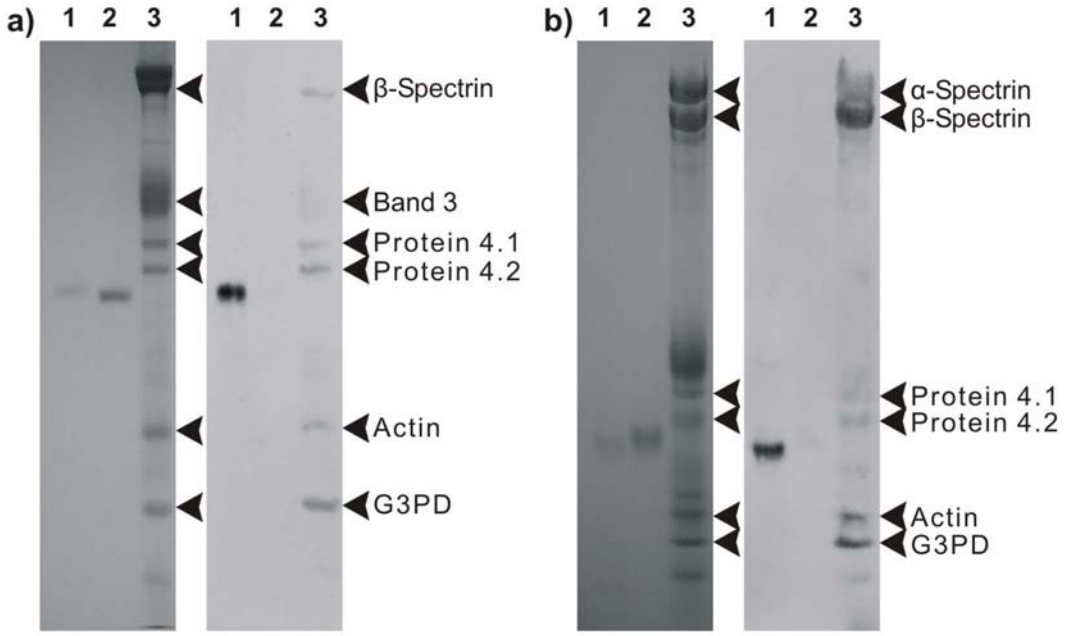


Figure 62: Erythrocyte membrane proteins overlaid with r*Pf*M18AAP.
Laemmli (a) and Fairbanks gel (b) (left) and the corresponding immunoblots with the Penta•His™ HRP Conjugate antibody (right) showing the erythrocyte membrane proteins bound by r*Pf*M18AAP. The enzyme binds strongly to β-spectrin, protein 4.1, protein 4.2, actin, and glyceraldehyde-3-phosphate dehydrogenase and weakly to α-spectrin and band 3. Lane 1 = 500 ng (gel) and 100 ng (immunoblot) r*Pf*M18AAP; Lane 2 = 500 ng (gel) and 100 ng (immunoblot) BSA; Lane 3 = 20 µg erythrocyte membrane proteins.

A blot overlay was subsequently performed on trypsin-digested spectrin to determine which spectrin domains r*Pf*M18AAP binds and to see if the enzyme interacts with the N-terminal α- and β-spectrin domains. The experiment (Figure 63) did however not give clear results because of several factors. Firstly, the trypsin-digested α- and β-spectrin domains often comigrate (Figure 63 α- and β-antibody immunoblot) making it difficult to delineate the domains. Secondly, α- and β-spectrin are composed of homologous repeats that could all bind r*Pf*M18AAP. Thirdly, the tryptic digest could expose glutamates and aspartates at the N-termini of the spectrin protein fragments, because trypsin cleaves proteins on the C-terminal side of arginine and lysine residues. This results, for example, in the exposure of a glutamate and an aspartate on the 80 and 52 kDa α-spectrin domains respectively (Speicher and Marchesi, 1984), which could subsequently serve as substrates for r*Pf*M18AAP. Fourthly, only a limited tryptic digest was performed with a low trypsin to spectrin ratio and digestion at 4 ºC, to prevent the

spectrin domains from being digested into smaller peptides. This however resulted in the presence of some incompletely digested spectrin molecules containing several domains, which are larger than the 80 kDa N-terminal α-spectrin domain (Figure 63) that still bound r*Pf*M18AAP, thereby obscuring the results. Finally, the 74 kDa N-terminal domain of β-spectrin is often broken down into smaller 52, 41 and 30 kDa protein fragments during the tryptic digest (Figure 63 α- and β-antibody immunoblot), which could comigrate with the α-spectrin fragments and thereby making it impossible to distinguish between the α- and β-spectrin binding profile. The generation of the different spectrin domains as separate peptides with recombinant protein expression systems and subsequently allowing these peptides to interact with r*Pf*M18AAP, could resolve the issues with mentioned above.

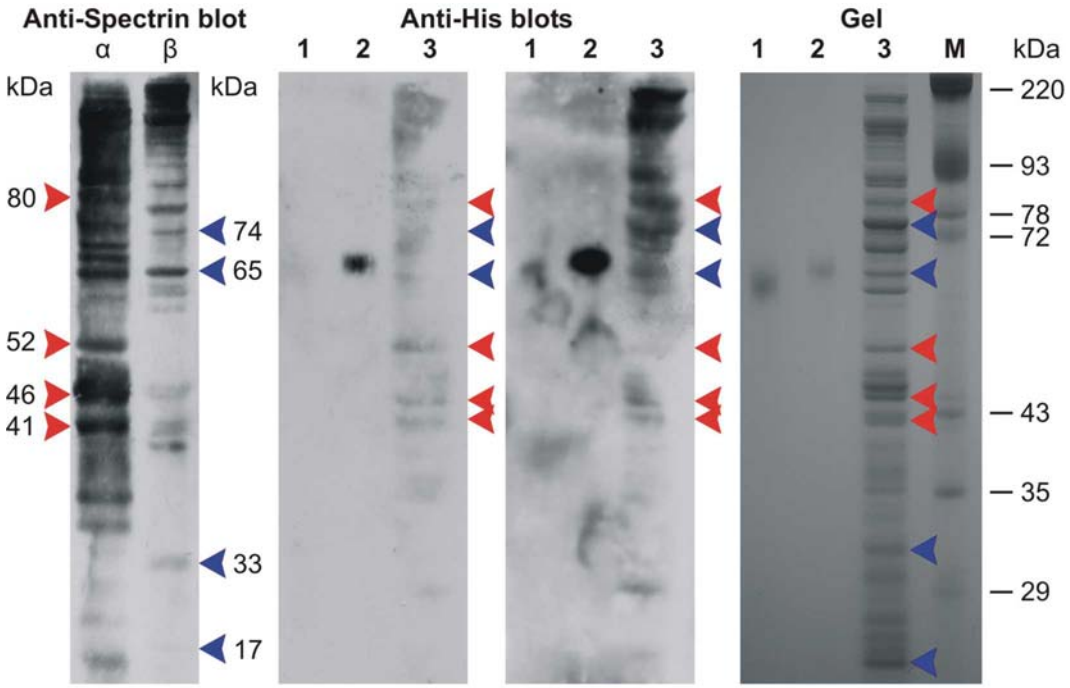


Figure 63: Spectrin tryptic digest overlaid with r*Pf*M18AAP.
Immunoblot with the anti-α- and anti-β-spectrin antibodies (left) showing the α- and β-spectrin tryptic digest peptides, two exposures of the same immunoblot with the Penta•His™ HRP Conjugate antibody (middle) showing the spectrin tryptic digest peptides bound by r*Pf*M18AAP, and Laemmli gel (right) showing the spectrin tryptic digest peptides. Alpha and beta-spectrin immunoblot: 40 µg trypsin digested with 0.4 µg trypsin for 18 hours at 4 ºC. Blot overlay immunoblots and gel: Lane 1 = 100 ng (immunoblot) and 500 ng (gel) BSA; Lane 2 = 100 ng (immunoblot) and 500 ng (gel) r*Pf*M18AAP; Lane 3 = 40 µg spectrin tryptic digest, Lane M = 20 µg erythrocyte membrane proteins. Red arrowheads = α-spectrin peptide domains; blue arrowheads = β-spectrin peptide domains.

4.4 Discussion

4.4.1 The *Pf*M18AAP active site

Teuscher *et al.* (2007) showed that r*Pf*M18AAP was activated by cobalt, which suggests that the active site contains cobalt as cofactors. It was therefore possible to draw a diagram of the inferred *Pf*M18AAP active site based on the crystal structures of the other aspartyl aminopeptidases (Stamper *et al.*, 2001, Schürer *et al.*, 2004). The two Co^{2+} ions are each coordinated with one histidine (*Pf*H86 and *Pf*H534) and an aspartate or a glutamate (*Pf*D434 or *Pf*E380), while another aspartate (*Pf*D324) and a water molecule bridge the two metal ions (Figure 64). A second glutamate residue (*Pf*E379) is located near the active site and is hydrogen bonded to the bridging water molecule and another aspartate (*Pf*D88) decreases the Lewis acidity of the second metal ion by binding to *Pf*H86.

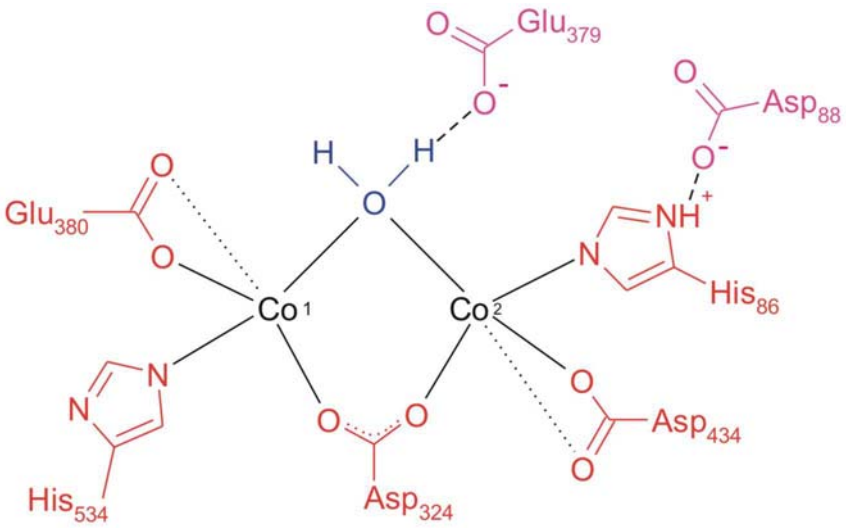


Figure 64: The putative metal binding site of *Pf*M18AAP.
Diagram showing how the two metal ions of *Pf*M18AAP could be bound in the active site of the enzyme. The metal ions, possibly cobalt, are bound by two histidines (*Pf*H86 and *Pf*H534), two aspartates (*Pf*D434 and *Pf*D324) and a glutamate (*Pf*E380). A water molecule, or hydroxide ion, is hydrogen bonded to a second glutamate (*Pf*E379) and bridges both cobalt ions. A third aspartate (*Pf*D88) decreases the Lewis acidity of the second metal ion by binding to *Pf*H86. The final histidine (*Pf*H170), which is involved in enzymatic activity, is located between *Pf*H534 and *Pf*D324, but is too far away from these residues and the metal ions to form bonds with them (not shown).

Aminopeptidases that contain a co-catalytic active site, use a common reaction mechanism to cleave a substrate (Schürer *et al.*, 2004), regardless of the type of metal ions present in the enzyme. The *Aeromonas proteolytica* aminopeptidase active site contains the same metal-binding amino acids as *Pf*M18AAP to bind two zinc ions and therefore the reaction mechanism that is employed by this enzyme may be used to explain the substrate cleavage mechanism of *Pf*M18AAP (Figure 65) (Stamper *et al.*, 2001, Schürer *et al.*, 2004).

The enzymatic reaction is initiated when the substrate (a dipeptide with an N-terminal aspartate) binds with its N-terminal amino group to the second Co^{2+} ion and with the carbonyl oxygen of the scissile peptide bond to the first Co^{2+} ion. The N-terminus (NH_2) of the substrate also interacts with Asp434, which is bound to the second Co^{2+} ion. After substrate-binding, the bridging water molecule loses its coordination to the second Co^{2+} ion, leaving it bound to only the first Co^{2+} ion. This loss of coordination is enhanced by the hydrogen bond between His86 and Asp88, which reduces the Lewis acidity of the second Co^{2+} ion and therefore reduces the metal ion's affinity for the water molecule. Substrate cleavage occurs after the water molecule is deprotonated by Glu379 (base catalysis), resulting in a nucleophilic hydroxide ion that acts on the substrate's peptide bond. Finally, Glu379 may donate two protons to the peptide nitrogen to complete the separation of the substrate into two molecules (acid catalysis).

Wilk *et al*. (2002) showed that an additional histidine (*Pf*H170), which is not part of the conserved metal-binding amino acids, is also required for substrate cleavage. Histidines are generally used by metalloenzymes to coordinate the catalytic metal ion (Auld, 2001), but by analogy to the human enzyme, histidine *Pf*H170 might be acting directly on the substrate as a nucleophile or base (Bugg, 2004), or it may be involved in substrate-binding because it is located on the side of the active site where the substrate enters the pocket.

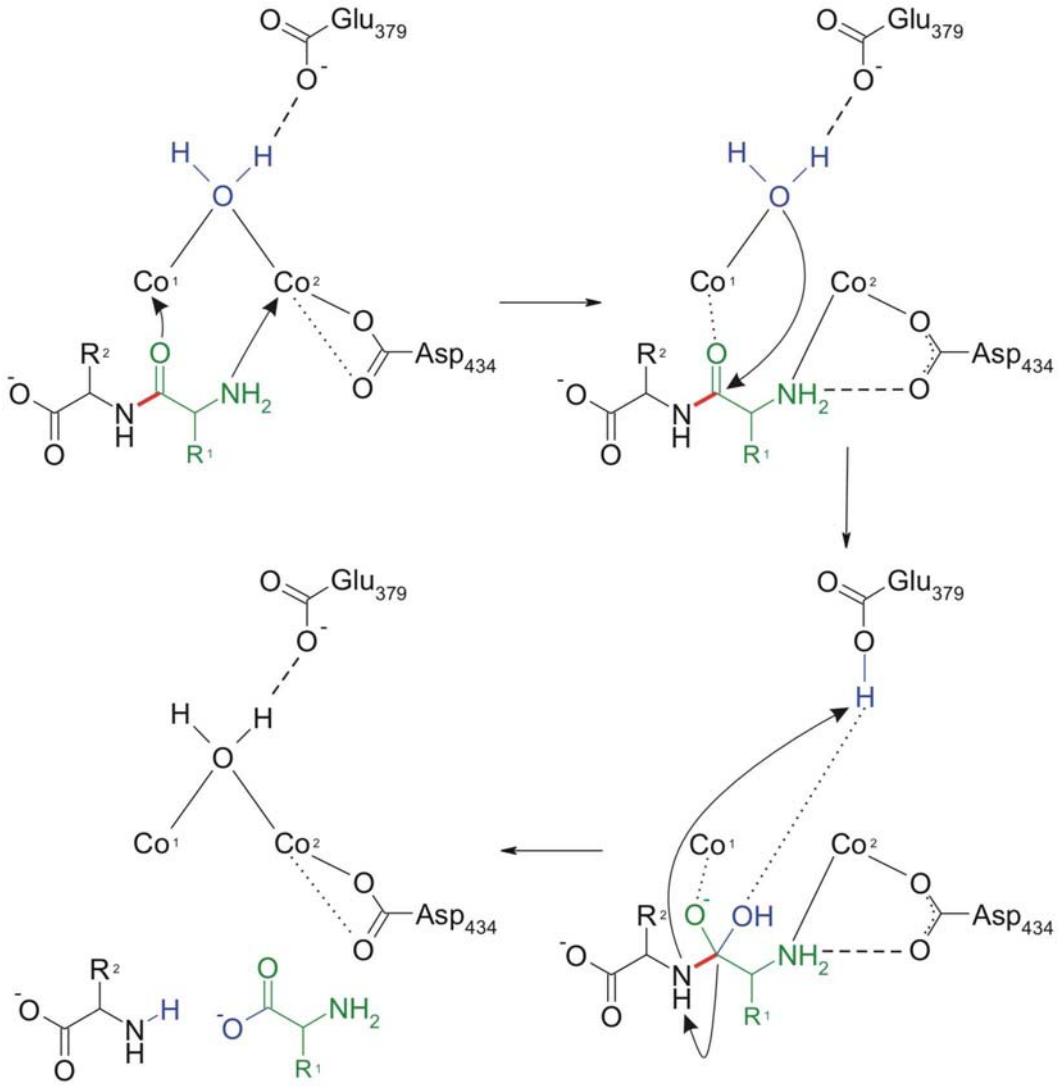


Figure 65: The putative substrate cleavage mechanism of *Pf*M18AAP.
Diagram showing the cleavage of a dipeptide substrate by *Pf*M18AAP. The dipeptide, which has an N-terminal aspartate, binds to both cobalt ions, followed by deprotonation of the water molecule by the glutamate (*Pf*E379). The nucleophilic hydroxide ion attacks the peptide bond in the substrate resulting in the release of two amino acids. The exact function of *Pf*H170 in the cleavage mechanism is not known. *Pf*H86, *Pf*D88, *Pf*H170, *Pf*E380, and *Pf*H534 have been omitted to simplify the diagram (adapted from Schürer *et al.*, 2004, and Stamper *et al.*, 2001).

4.4.2 The *Pf*M18AAP spectrin-binding site

The spectrin-binding region (*Pf*216-257) is only found in *Pf*M18AAP and could therefore represent a special evolutionary adaptation of *P. falciparum*. Secondary structure prediction and a comparison to the quaternary structures of the other

aspartyl aminopeptidases showed that the spectrin-binding region contains an alpha helix and occurs on the outer surface of the native enzyme. This alpha-helical region (*Pf*243-254) lies within an unstructured region of amino acids on the outer surface of the enzyme and can therefore provide an anchoring point for *Pf*M18AAP to the erythrocyte membrane skeleton, without interfering with the enzymatic function. The enzyme could, for example, bind to one of the proteins it interacts with and cleave any of the proteins of the erythrocyte membrane skeleton, since all these proteins are interlinked. The binding region could also guide spectrin, which has a diameter of 20 Å (Yan *et al.*, 1993), or any of the other erythrocyte membrane proteins that *Pf*M18AAP interacts with, into the pores towards the internal active sites of the enzyme.

4.4.3 r*Pf*M18AAP protein and enzyme characteristics

Non-denaturing gel electrophoresis of r*Pf*M18AAP indicated that the native enzyme consisted mainly of a tetramer and higher order oligomers, possibly octamers and dodecamers. Teuscher *et al*. (2007) showed that *Pf*M18AAP occurred as an octamer in its active native state by assaying HPLC fractions for enzymatic activity. This is the same number of subunits as predicted for the *H. sapiens* and *S. cerevisiae* aspartyl aminopeptidases (Wilk *et al.*, 2002, Yokoyama *et al.*, 2006). However, the unpublished crystal structures of the *P. aeruginosa*, *C. acetobutylicum*, *T. maritima* and *B. burgdorferi* M18 aminopeptidases showed that these enzymes occurred as dodecamers. Crystallisation or velocity sedimentation of *Pf*M18AAP could be performed to determine the exact number of monomers that are necessary to form the native enzyme.

Enzyme activity assays showed that r*Pf*M18AAP cleaved the N-terminal aspartate from the tripeptide Asp-Ala-Pro-β-Napthylamide. Teuscher *et al.* (2007) showed that r*Pf*M18AAP also cleaved the N-terminal amino acids from the fluorogenic dipeptide substrates H-Asp-4-methyl-7-coumarinylamide (NHMec) and H-Glu-NHMec and that r*Pf*M18AAP was more active on N-terminal glutamates in comparison to N-terminal aspartates. Both these amino acids are acidic, indicating

that *Pf*M18AAP only cleaves peptides and proteins that have an acidic N-terminus. This was confirmed by Teuscher *et al*. (2007), who also tested substrates containing hydrophobic, basic or neutral N-terminal amino acids and found no enzymatic activity.

A pH study showed that r*Pf*M18AAP was active from pH 7-8.5, with the highest activity at the physiological pH of 7.5. The enzyme had no activity at pH 5.3, which is close to the pH of the food vacuole (5.0-5.2 (Yayon *et al.*, 1984)), indicating that it does not participate in the initial breakdown of host haemoglobin into 2-10 amino acid peptides (Klemba *et al.*, 2004). These small peptides are subsequently exported from the food vacuole into the parasite cytosol where they are converted into amino acids by *Pf*M18AAP and other aminopeptidases (Teuscher *et al.*, 2007, Dalal and Klemba, 2007). This correlates well with the fluorescence microscopy findings of Dalal *et al*. (Dalal and Klemba, 2007) that *Pf*M18AAP (PfDAP in their nomenclature) was excluded from the food vacuole. The parasite cytosol has a pH of 7.3-7.4, allowing *Pf*M18AAP to be active in this parasite compartment. Even though *Pf*M18AAP has no export signals, it has also been localised to the parasitophorous vacuole by immunoblot analysis (Teuscher *et al.*, 2007) and the erythrocyte membrane by mass spectroscopy (Florens *et al.*, 2004), as well as by phage display interactions in this study. The enzyme could therefore also be active in the erythrocyte cytosol which has a neutral pH of 7.2-7.3 (Funder and Wieth, 1966).

The temperature study revealed that r*Pf*M18AAP was active from 33-39 ºC with maximum activity at 37 ºC, which indicates that the enzyme is active when the parasite resides in the human host. r*Pf*M18AAP had 90 % activity at 39 ºC, implying that the enzyme is still functional when the human host experiences fever (maximum temperature 42 ºC), which occurs after the erythrocytes rupture (Sinden and Gilles, 2002). *Pf*M18AAP could therefore be active during the initial growth phase in new erythrocytes. The enzyme only had ~35 % activity at 25 ºC, which is the ambient body temperature of the mosquito (Oaks *et al.*, 1991), suggesting that the enzyme is not fully functional during this stage of its lifecycle.

Studies utilising metal ions and protease inhibitors revealed some interesting features of r*Pf*M18AAP. EDTA inhibited enzymatic activity confirming that r*Pf*M18AAP is a metalloprotease and that the metal ions remained bound in the enzyme during the purification procedure. Bestatin, an aminopeptidase inhibitor, also inhibited the enzyme, but the effect was not as pronounced as it was with the metal chelator, EDTA. Teuscher *et al*. (2007) observed an 86 % inhibition of enzymatic activity when r*Pf*M18AAP was preincubated with EDTA for 30 minutes. This is higher than the 39 % inhibitory effect observed in this study, presumably because the enzyme was preincubated for 10 minutes only. The metal ion study revealed that zinc inhibited enzymatic activity and that calcium, magnesium and manganese had only a slight effect. The zinc ions could have displaced the natural cofactor of r*Pf*M18AAP in the active site and therefore obliterated the enzymatic activity. The *in vitro* study performed by Teuscher *et al.* (2007) with pure enzyme and parasite extracts, showed that cobalt was the only metal ion that enhanced *Pf*M18AAP enzymatic activity. They also confirmed the inhibitory effect of zinc.

The substrate concentration versus velocity curve for r*Pf*M18AAP showed that substrate inhibition occurred at ≥1.5 mM Asp-Ala-Pro-β-Napthylamide. Substrate inhibition is a form of uncompetitive inhibition, where two substrate molecules simultaneously bind in the active site, at the positions where only one substrate molecule would normally bind during enzymatic cleavage, thereby forming a dead-end complex. The second substrate molecule therefore becomes an inhibitor. In the case of r*Pf*M18AAP, a second substrate molecule could for example bind to the cobalt ion that normally binds the carbonyl oxygen of the scissile peptide bond of the first substrate molecule, thereby preventing the nucleophilic attack on the peptide bond. The presence of substrate inhibition made it difficult to determine accurate K_m and V_{max} values for r*Pf*M18AAP using the non-linear least squares procedure based on the Michaelis-Menten equation and therefore K_m and V_{max} values were estimated from the Lineweaver-Burk plot.

Enzyme activity was very variable amongst different enzyme preparations. This could be due to numerous factors. Firstly, the proportion of contaminating proteins varied in the different purifications, which would decrease the amount of r*Pf*M18AAP in the assay and yield lower activity. Secondly, folding of the protein could be different each time the protein is expressed in *E. coli* and this could compromise activity. Thirdly, manipulation of the enzyme during purification could inactivate some of the enzyme and the degree of inactivation could vary between preparations. Finally, the variable length of time from purification to assay could have caused a decrease in enzyme activity.

4.4.4 r*Pf*M18AAP-erythrocyte membrane protein interactions

In the phage display studies (Chapter 2) an interaction between spectrin and a *P. falciparum*-specific 42 amino acid stretch of *Pf*M18AAP was identified. To verify this association with the full length protein, blot overlays were performed. These demonstrated strong specific binding to spectrin dimers, spectrin tetramers and in the case of the spectrin monomers, marked preferential binding to β-spectrin compared to α-spectrin. In addition, r*Pf*M18AAP interacted with protein 4.1, protein 4.2 and actin. All these erythrocyte membrane proteins contain one or more aspartates or glutamates within the first 10 N-terminal residues (Figure 66), making them putative substrates for r*Pf*M18AAP. However, the host glycolytic enzyme, G3PD, which associates with the cytoplasmic domain of band 3, also showed a strong binding of r*Pf*M18AAP on the blot overlays. The significance of this interaction between the parasite and the human enzyme is unclear, since G3PD does not have any aspartates or glutamates in the first 10 N-terminal residues (Figure 66).

Actin or α-spectrin could be prime targets for *Pf*M18AAP, because they have three aspartates (actin) and one glutamate (α-spectrin) after the N-terminal methionine (Figure 66), which would be cleaved from newly synthesised proteins by human methionine aminopeptidases (reviewed in Bradshaw *et al.*, 1998). There is however a possibility that *Pf*M18AAP does not cleave actin because it is

acetylated (Figure 66). Yeast aspartyl aminopeptidase, for example, cannot remove an N-acetylated aspartate from a N-terminally blocked substrate (Yokoyama *et al.*, 2006).



Figure 66: The first 10 amino acids of the r*Pf*M18AAP-binding erythrocyte membrane proteins.
Diagram showing the first ten N-terminal amino acids of α-spectrin, β-spectrin, protein 4.1, protein 4.2, actin, and glyceraldehyde-3-phosphate dehydrogenase (G3PD). Aspartates and glutamates are labelled in red. Post-translationally modified amino acids are boxed in blue: actin and β-spectrin = acetylated; and protein 4.2 = myristoylated.

The N-termini of β-spectrin molecules are linked to actin and protein 4.1 in the junctional complex (Ungewickell *et al.*, 1979) and protein 4.2 is part of the band 3 complex of the erythrocyte membrane (Korsgren and Cohen, 1988). Cleavage of any of these proteins could therefore destabilise and disrupt the junctional complexes and the band 3 complexes. The N-termini of α-spectrin are located at the self-association sites of spectrin tetramers (Morrow *et al.*, 1980) and since *Pf*M18AAP shows a marked preferential binding to β-spectrin compared to α-spectrin, the enzyme could bind to β-spectrin near the self-association site and cleave the N-termini of α-spectrin, thereby disrupting the spectrin tetramers.

4.4.5 The function of *Pf*M18AAP in the infected erythrocyte

The primary function of *Pf*M18AAP in *P. falciparum* is to cleave aspartates or glutamates from the oligopeptides that are exported from the food vacuole into the parasite cytosol after haemoglobin digestion (Teuscher *et al.*, 2007, Dalal and Klemba, 2007). These and other free amino acids are utilised by the parasite for protein synthesis. In addition, 84 % of the amino acids derived from haemoglobin

digestion, are released from the cell and this regulates the colloid-osmotic pressure within the infected erythrocyte to prevent premature cell lysis and to establish a concentration gradient that facilitates the entry of rare amino acids into the infected erythrocyte (Lew *et al.*, 2003). *Pf*M18AAP is therefore indirectly involved in regulating the volume of the host cell to accommodate the increasing size of the growing parasite.

Results from this study, as well as *Pf*M18AAP mRNA and protein data from other laboratories, indicate that the enzyme could also perform additional functions in the parasitophorous vacuole, the erythrocyte cytosol and particularly at the infected erythrocyte membrane skeleton. Microarray data showed that *Pf*M18AAP mRNA is expressed throughout the erythrocytic stages, with the highest expression levels in early and late trophozoites and in gametocytes (Le Roch *et al.*, 2003, Bozdech *et al.*, 2003), whilst Northern blot analysis revealed predominant expression in rings (Teuscher *et al.*, 2007). Protein data localised *Pf*M18AAP in merozoites, trophozoites and schizonts (Florens *et al.*, 2002) and at the infected erythrocyte membrane in trophozoite/schizont stage parasites (Florens *et al.*, 2004). Western blot analysis and immunolocalisation experiments, utilising anti-*Pf*M18AAP antiserum, also revealed the enzyme in rings, the parasite cytosol and the parasitophorous vacuole (Teuscher *et al.*, 2007). *Pf*M18AAP mRNA is therefore translated into the enzyme at several stages in the erythrocytic life cycle and the *Pf*M18AAP protein is located in several compartments within the parasite-infected erythrocyte. These data imply that *Pf*M18AAP could play a role in parasite invasion, growth and/or exit from the host cell, since parasite proteins interact with the erythrocyte membrane and the underlying erythrocyte membrane skeleton during all phases of the erythrocytic stage of the parasite. Transfection of parasite-infected erythrocytes with a plasmid containing *Pf*M18AAP fused to green fluorescent protein could shed more light on the location of the enzyme during the different stages of the parasite life cycle.

Merozoites are the invasive form in erythrocytic schizogony and utilise proteases to gain entry into the new host cell. Inhibition studies with 1,10-phenanthroline

have shown that metalloproteases are involved in parasite invasion (Kitjaroentham *et al.*, 2006) and hence *Pf*M18AAP could be involved in this process. During invasion the red cell membrane and membrane skeleton are rearranged in localised areas to facilitate entry of the parasite (Bannister and Mitchell, 2003). *Pf*M18AAP could weaken the spectrin skeleton and thus assist in invasion.

The host experiences febrile paroxysms triggered by the release of toxins during parasitised erythrocyte rupture (Sinden and Gilles, 2002). Since *Pf*M18AAP is functional at elevated temperatures, it could thus be responsible for regulating, processing and/or activating other parasite proteins within the early ring stage by removing their N-termini.

During further growth the parasite introduces NPPs into the erythrocyte membrane to facilitate the import and export of nutrients and waste products from the red cell. *Pf*M18AAP could aid in the insertion of these NPPs in the erythrocyte membrane by disrupting the membrane skeleton. The enzyme could also be responsible for altering the activity of erythrocyte membrane channels and transporters. Maturation of the parasite in the erythrocyte is also concomitant with the formation of knobs on the erythrocyte surface. Knobs are anchored in the erythrocyte membrane skeleton (Cooke *et al.*, 2004) and therefore *Pf*M18AAP could be responsible for weakening and disrupting the erythrocyte membrane skeleton to facilitate the insertion of these structures.

Several parasite proteases, such as plasmepsin-II (Le Bonniec *et al.*, 1999) and falcipain-2 (Dua *et al.*, 2001), which are primarily involved in haemoglobin digestion inside the food vacuole, also facilitate the release of the parasite from its erythrocytic host by cleaving membrane skeleton proteins such as spectrin, protein 4.1 and ankyrin. Studies with inhibitors have proven that parasite proteases first weaken the parasitophorous vacuole membrane and then the erythrocyte membrane prior to parasite release (Wickham *et al.*, 2003), which is caused by an osmotic pressure build-up within the erythrocyte (Glushakova *et al.*, 2005). Weakening of the erythrocyte membrane would involve disruption of the spectrin-

junctional complex network below the plasma membrane and therefore *Pf*M18AAP could also play a role in the release of the parasite from its erythrocyte host.

4.4.6 *Pf*M18AAP as a novel drug target

A knockout experiment utilising a single-crossover strategy to integrate an episome into the PFI1570c open reading frame, was used by Dalal and Klemba (2007) to express *Pf*M18AAP, which lacked 236 amino acids from its C-terminus. Parasites expressing the truncated enzyme showed less than 10 % enzymatic activity when compared to wild-type parasites, indicating that enzymatic activity was nearly completely abolished. Alteration of the PFI1570c gene did however not have deleterious effects on parasite replication, possibly because the small amount of active enzyme was enough to perform all the necessary *Pf*M18AAP enzymatic reactions in the parasite. This interpretation may explain the contrasting findings from Teuscher *et al*. (2007), who performed knockdown experiments utilising a plasmid containing an antisense copy of the PFI1570c gene. They showed an 80-fold decrease in *Pf*M18AAP enzyme activity in parasite extracts that resulted in a lethal phenotype due to significant morphological damage, which included rupture of the food vacuole and the accumulation of haemoglobin and lipid droplets in the parasite. The same group tested phosphorus-containing inhibitors on *Pf*M18AAP and found that these compounds were only moderately effective *in vitro* and had no inhibitory effect on the growth of the parasite in culture (Teuscher *et al.*, 2007).

These experiments show that *Pf*M18AAP is required for parasite survival and therefore further studies on the function of the enzyme and its reaction mechanism could lead to the development of new drugs that can be employed to fight malaria infections.

Chapter 5: Conclusion

The malaria parasite depends on parasite-host protein interactions during invasion of, development in and exit from the human erythrocyte and interactions with human erythrocyte membrane proteins are of particular importance. The aim of this study was therefore to identify and characterise *P. falciparum* proteins that interact with human spectrin, which is the main structural component of the erythrocyte membrane skeleton. To achieve this goal, a novel approach based on the construction of a *P. falciparum* phage display library and subsequent biopanning of this library against immobilised spectrin, was applied to identify five spectrin-binding *P. falciparum* peptides. These results show that phage display is a very powerful tool to probe host-parasite interactions. In addition, this species-specific library can be used to identify other protein-protein interactions, provided a target protein can be purified and immobilised for the panning procedure.

*Pf*M18AAP is an aminopeptidase, which contains a *P. falciparum*-specific insert that interacts with spectrin, as evidenced by phage display studies. The full length protein was expressed as a 6His-recombinant protein in *E. coli* and characterised. The enzyme cleaves an N-terminal aspartate from a peptide substrate and participates in the final stages of haemoglobin digestion. Results from this study suggest that *Pf*M18AAP performs additional functions in the parasite and infected erythrocyte, by acting on spectrin and other erythrocyte membrane proteins. Cleavage of spectrin could for example destabilise the erythrocyte membrane skeleton and facilitate parasite entry or exit, or it could allow for the insertion of NPPs into the erythrocyte membrane during parasite development. Further characterisation of this multifunctional enzyme could identify additional substrates of *Pf*M18AAP in the parasite and the human host and could ultimately lead to the development of new therapeutics for malaria treatment.

APPENDICES

A1: METHODS

0.4 mM PMSF stock preparation and addition to erythrocyte lysis buffer (Section 2.2.2)

A 0.4 mM PMSF stock solution was prepared in DMSO and slowly added to the erythrocyte lysis buffer with rapid stirring. PMSF is a serine and cysteine protease inhibitor that has a half-life of 110 minutes when incubated in phosphate buffer, pH 7 at 25 ºC. The hydrolysis of PMSF is accelerated with an increase in pH and temperature.

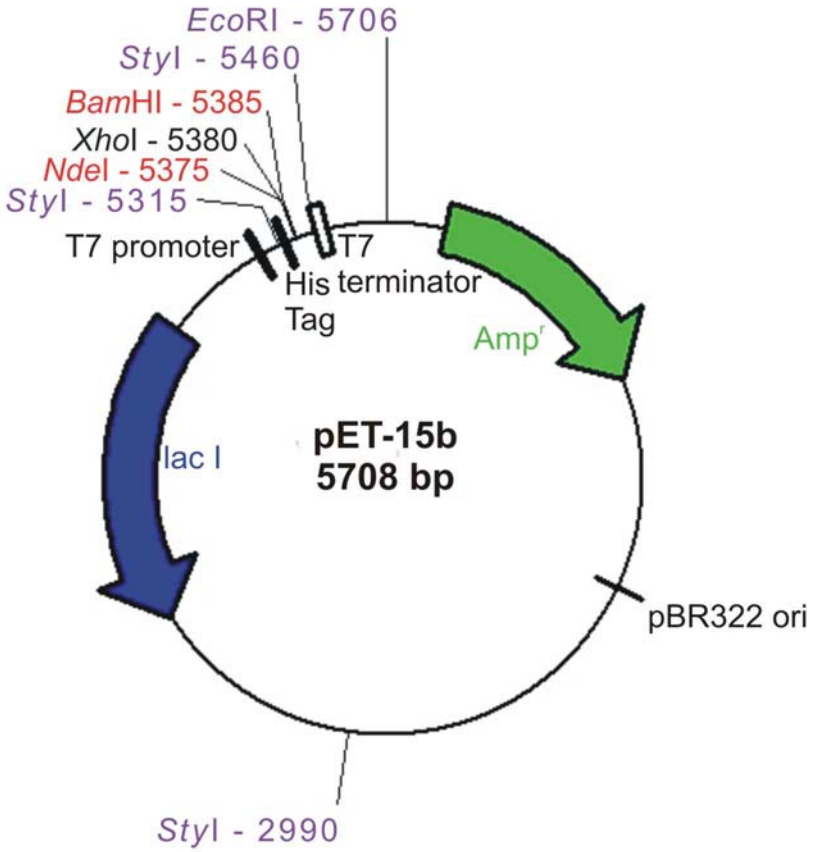
PCR for radioactively labelled size markers

Size markers were amplified using *TaKaRa Taq*™ (Takara Bio Inc., Otsu, Japan). The 347 bp fragment was amplified using the P2 and P3 primers (Inqaba Biotechnical Industries, Pretoria, South Africa), which are specific for human erythrocyte α-spectrin (exon 2). The 722 bp fragment was amplified with the P86 and P87 primers (Inqaba Biotechnical Industries, Pretoria, South Africa) which are specific for human erythrocyte band 3 (exon 18 and 19). Approximately 250 ng human DNA was amplified with 300 nM of each primer, 2.5 µl 10 x PCR Buffer, 2 µl dNTP Mixture (2.5 mM of each nucleotide), 2.5 µCi [α-^{32}P]dATP (specific activity = 3000 Ci/mmol), and 0.625 U enzyme (reaction volume = 25 µl) under the following conditions using the Eppendorf Mastercycler® Gradient machine:

Hot start:	95 ºC for 1 min
30 cycles:	94 ºC for 1 min
	55 ºC (P2/P3) or 63 ºC (P86/P87) for 1 min
	72 ºC for 1 min

A2: VECTOR MAPS

a)



b)

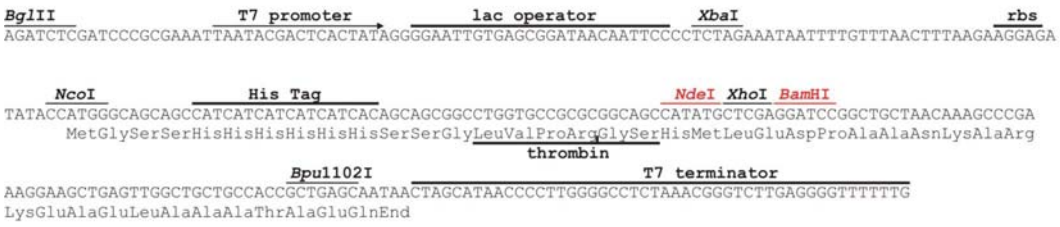
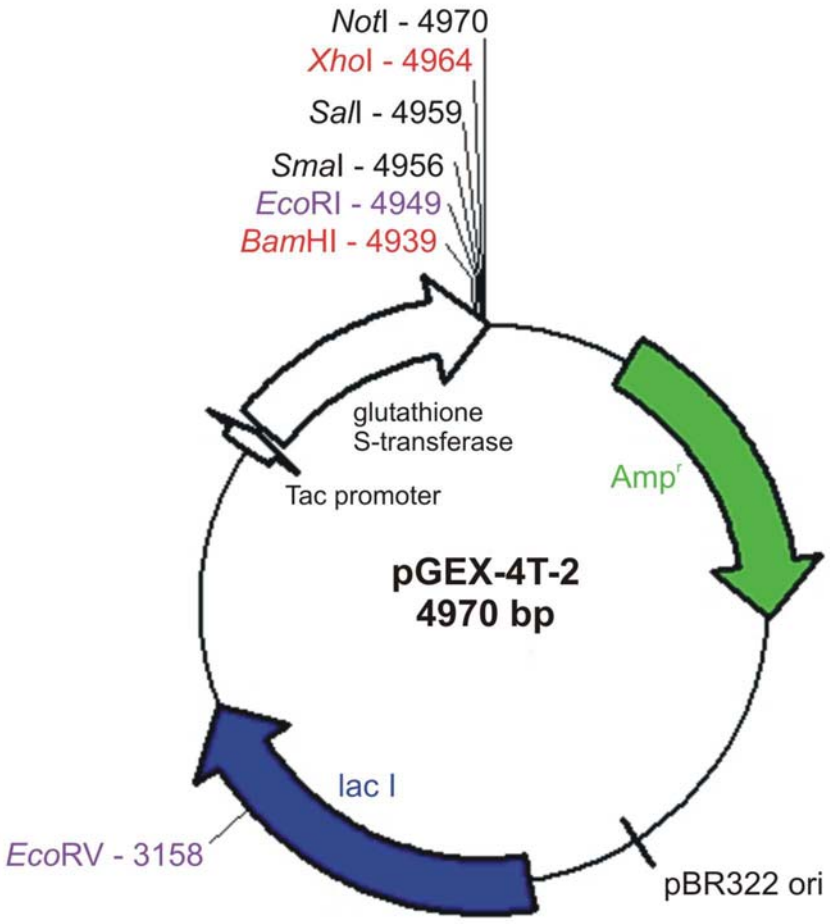


Figure A1: pET-15b vector map (a) and multiple cloning cassette (b) (adapted from Novagen, 1999b).

a)



b)



Figure A2: pGEX-4T-2 vector map (a) and multiple cloning cassette (b) (adapted from GE Healthcare, 2007).

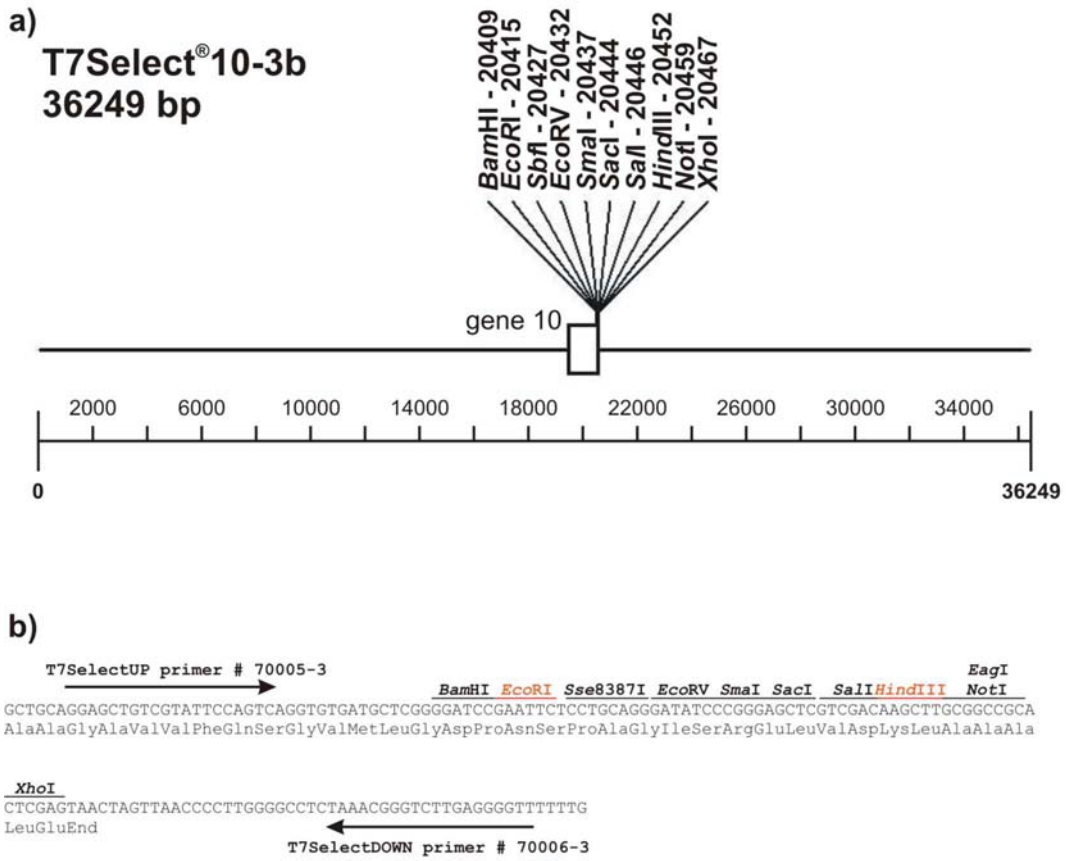


Figure A3: T7 bacteriophage vector map (a) and multiple cloning cassette (b) (adapted from Novagen, 2000a).

A3: CODON USAGE TABLES

```
Escherichia coli 536 [gbbct]: 4629 CDS's (1445921 codons) fields:
[triplet] [frequency: per thousand] ([number])

UUU 22.6( 32725)  UCU  8.8( 12736)  UAU 16.6( 24004)  UGU  5.3(  7700)
UUC 16.3( 23512)  UCC  8.8( 12703)  UAC 12.1( 17425)  UGC  6.4(  9192)
UUA 14.2( 20514)  UCA  7.7( 11067)  UAA  2.0(  2896)  UGA  1.0(  1389)
UUG 13.9( 20074)  UCG  8.8( 12676)  UAG  0.2(   344)  UGG 15.2( 21918)

CUU 11.5( 16623)  CCU  7.4( 10695)  CAU 13.3( 19209)  CGU 20.8( 30107)
CUC 11.0( 15952)  CCC  5.6(  8159)  CAC  9.6( 13895)  CGC 21.1( 30550)
CUA  4.0(  5840)  CCA  8.4( 12206)  CAA 15.5( 22441)  CGA  3.7(  5395)
CUG 52.1( 75338)  CCG 22.5( 32573)  CAG 29.0( 41888)  CGG  5.6(  8146)

AUU 30.3( 43849)  ACU  9.2( 13234)  AAU 18.3( 26426)  AGU  9.3( 13376)
AUC 24.6( 35528)  ACC 22.8( 32999)  AAC 21.1( 30545)  AGC 15.8( 22784)
AUA  5.0(  7181)  ACA  7.8( 11212)  AAA 33.4( 48324)  AGA  2.4(  3417)
AUG 27.5( 39775)  ACG 14.5( 20918)  AAG 10.3( 14831)  AGG  1.5(  2144)

GUU 18.4( 26611)  GCU 15.6( 22541)  GAU 32.3( 46650)  GGU 24.7( 35700)
GUC 15.1( 21798)  GCC 25.2( 36443)  GAC 18.8( 27186)  GGC 28.6( 41296)
GUA 11.1( 16030)  GCA 20.7( 29960)  GAA 39.1( 56532)  GGA  8.6( 12413)
GUG 25.7( 37188)  GCG 32.7( 47214)  GAG 17.8( 25669)  GGG 11.2( 16255)

Coding GC 51.51% 1st letter GC 58.68% 2nd letter GC 40.74% 3rd letter GC 55.11%
```

```
Plasmodium falciparum [gbinv]: 1140 CDS's (693485 codons) fields:
[triplet] [frequency: per thousand] ([number])

UUU 31.9( 22146)  UCU 17.0( 11793)  UAU 39.0( 27038)  UGU 16.2( 11256)
UUC  7.2(  5013)  UCC  6.0(  4170)  UAC  4.7(  3289)  UGC  2.3(  1612)
UUA 50.6( 35098)  UCA 19.7( 13628)  UAA  1.2(   847)  UGA  0.2(   117)
UUG  9.9(  6870)  UCG  2.5(  1713)  UAG  0.3(   193)  UGG  5.3(  3661)

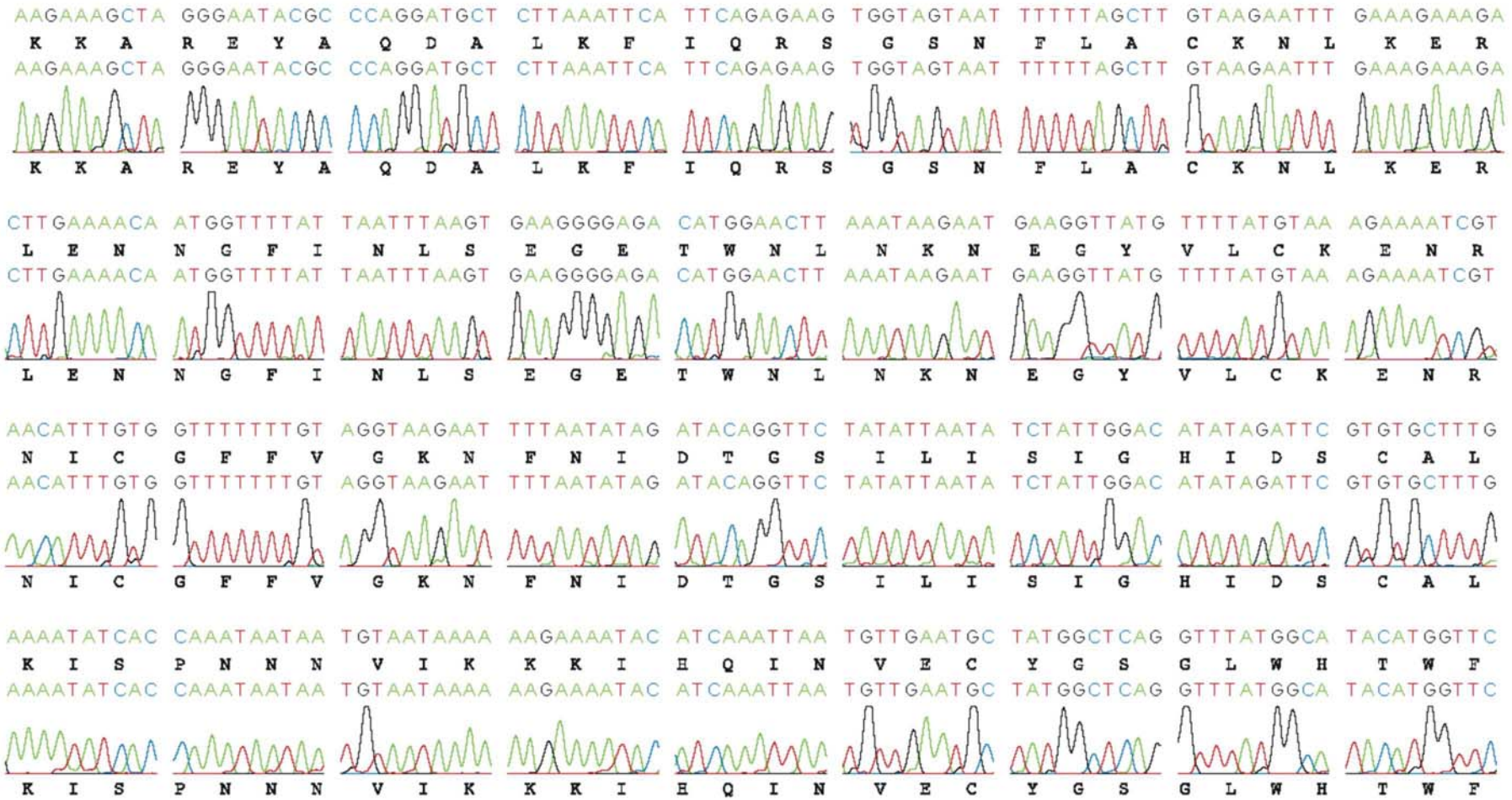
CUU  8.8(  6098)  CCU 10.5(  7271)  CAU 17.6( 12201)  CGU  3.3(  2267)
CUC  1.5(  1048)  CCC  2.9(  2038)  CAC  3.7(  2594)  CGC  0.3(   225)
CUA  5.5(  3830)  CCA 17.8( 12325)  CAA 25.8( 17874)  CGA  2.0(  1360)
CUG  1.1(   735)  CCG  0.9(   619)  CAG  3.2(  2226)  CGG  0.2(   110)

AUU 32.5( 22524)  ACU 14.0(  9742)  AAU 97.6( 67702)  AGU 21.8( 15090)
AUC  5.6(  3893)  ACC  5.4(  3770)  AAC 19.2( 13331)  AGC  3.7(  2560)
AUA 41.4( 28705)  ACA 22.0( 15242)  AAA 84.9( 58908)  AGA 17.0( 11773)
AUG 18.9( 13114)  ACG  3.5(  2417)  AAG 17.7( 12309)  AGG  3.3(  2294)

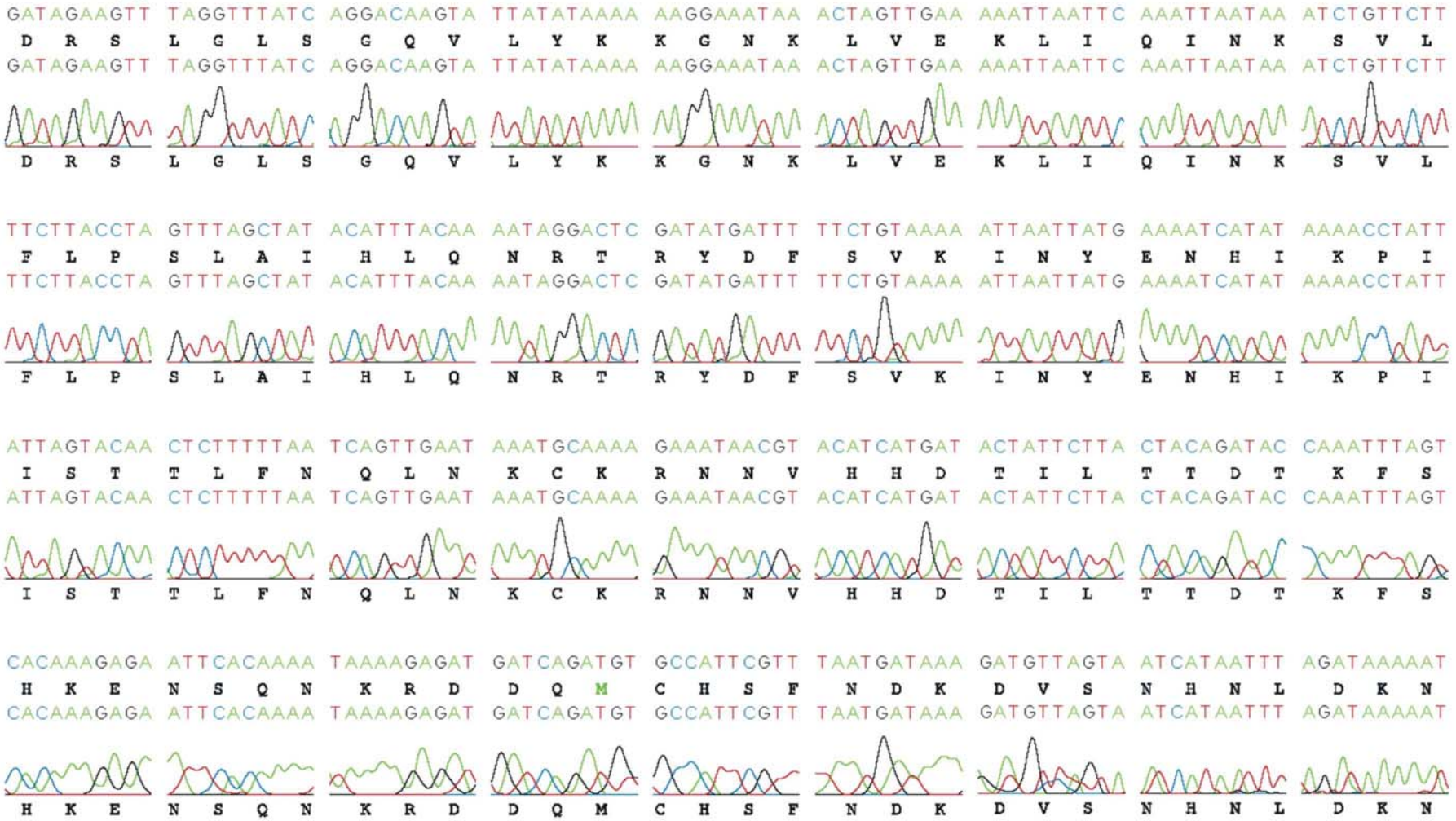
GUU 20.3( 14068)  GCU 14.5( 10036)  GAU 54.9( 38057)  GGU 17.6( 12193)
GUC  2.6(  1821)  GCC  4.0(  2765)  GAC  8.7(  6005)  GGC  2.0(  1364)
GUA 19.4( 13449)  GCA 16.9( 11723)  GAA 73.4( 50933)  GGA 19.8( 13736)
GUG  3.8(  2632)  GCG  1.3(   895)  GAG 10.1(  6978)  GGG  3.2(  2191)

Coding GC 27.31% 1st letter GC 37.73% 2nd letter GC 27.68% 3rd letter GC 16.50%
```

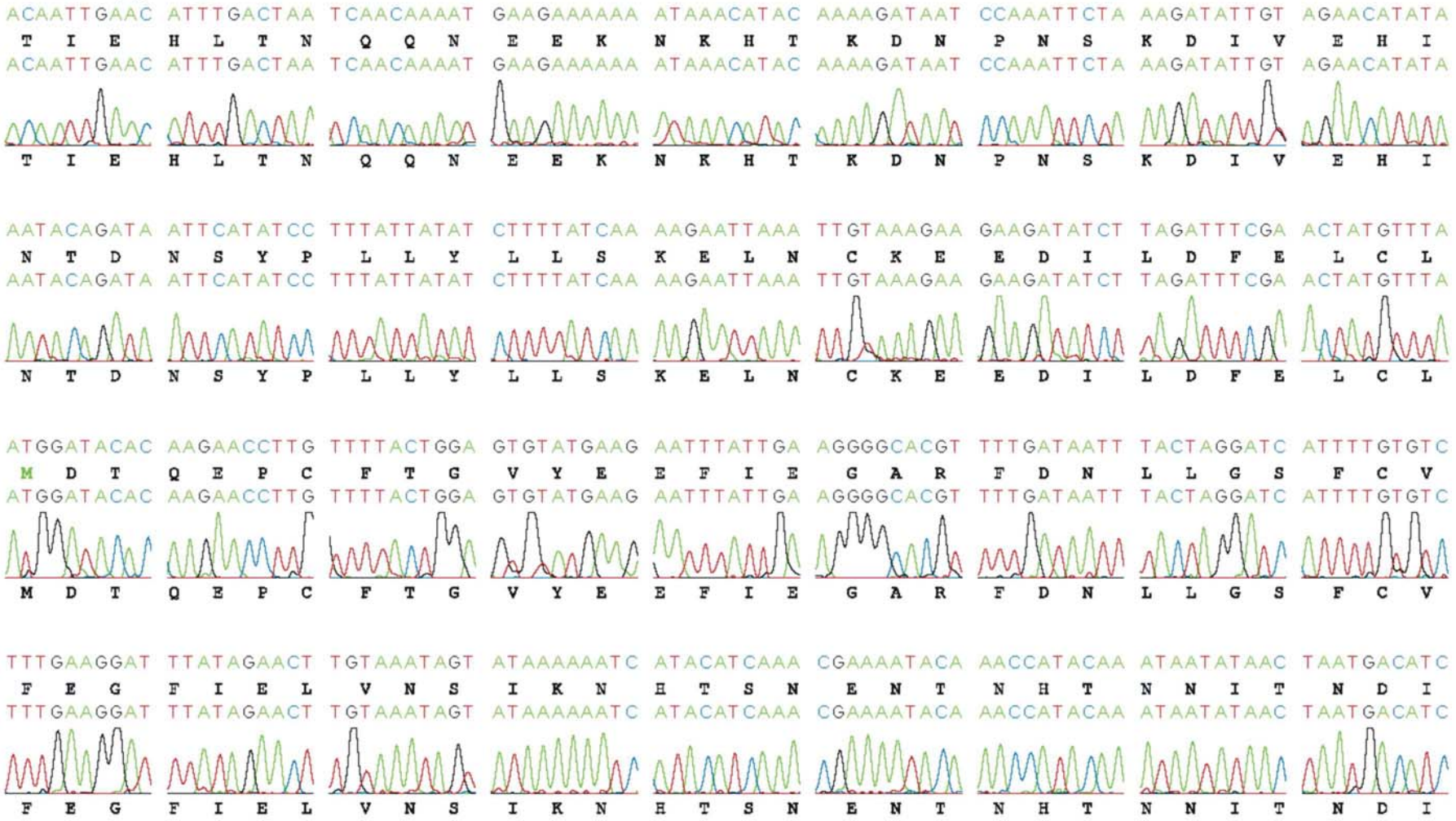
A4: AUTOMATED SEQUENCING RESULTS OF PFI1570c



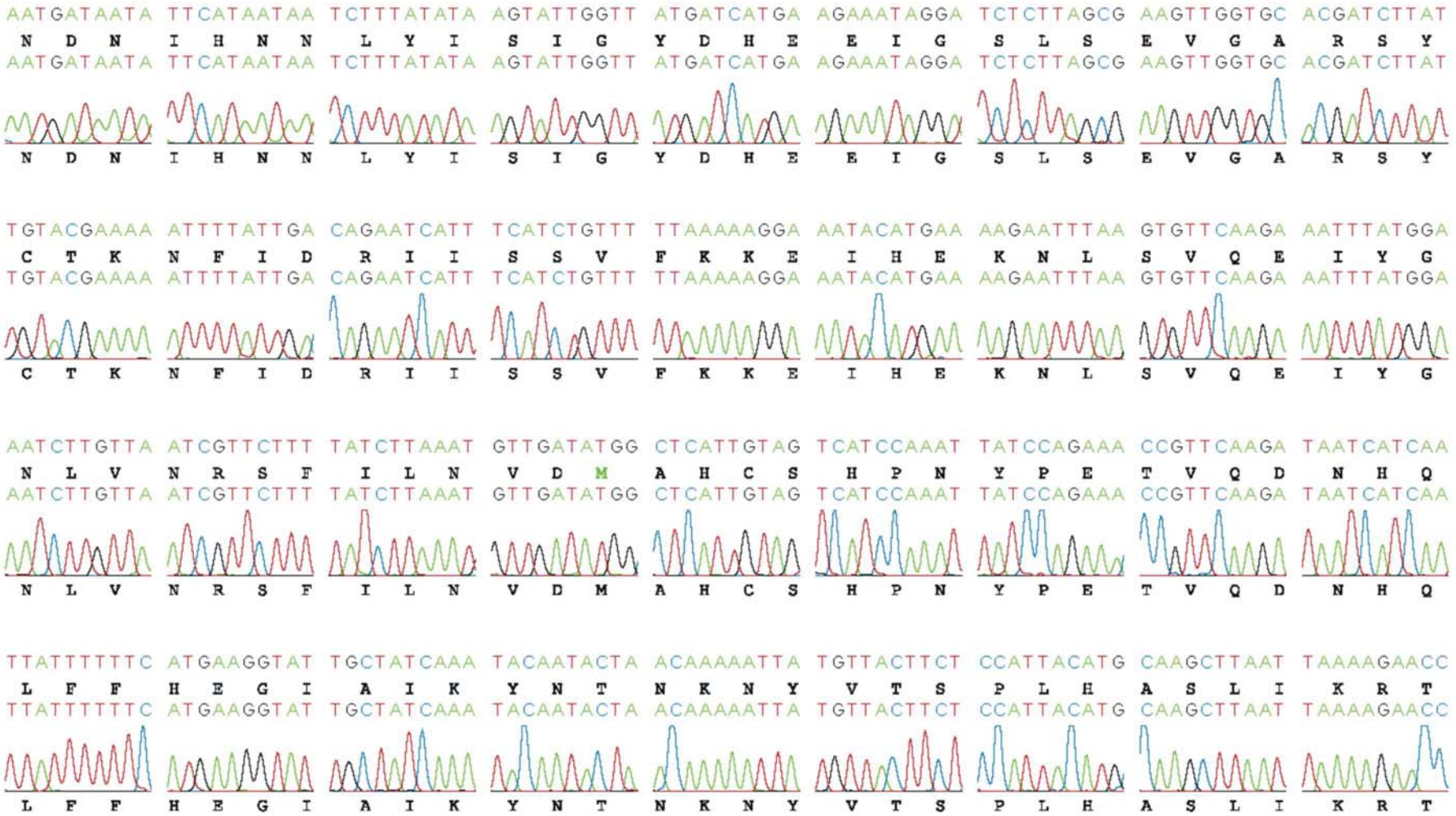
continued from page 179



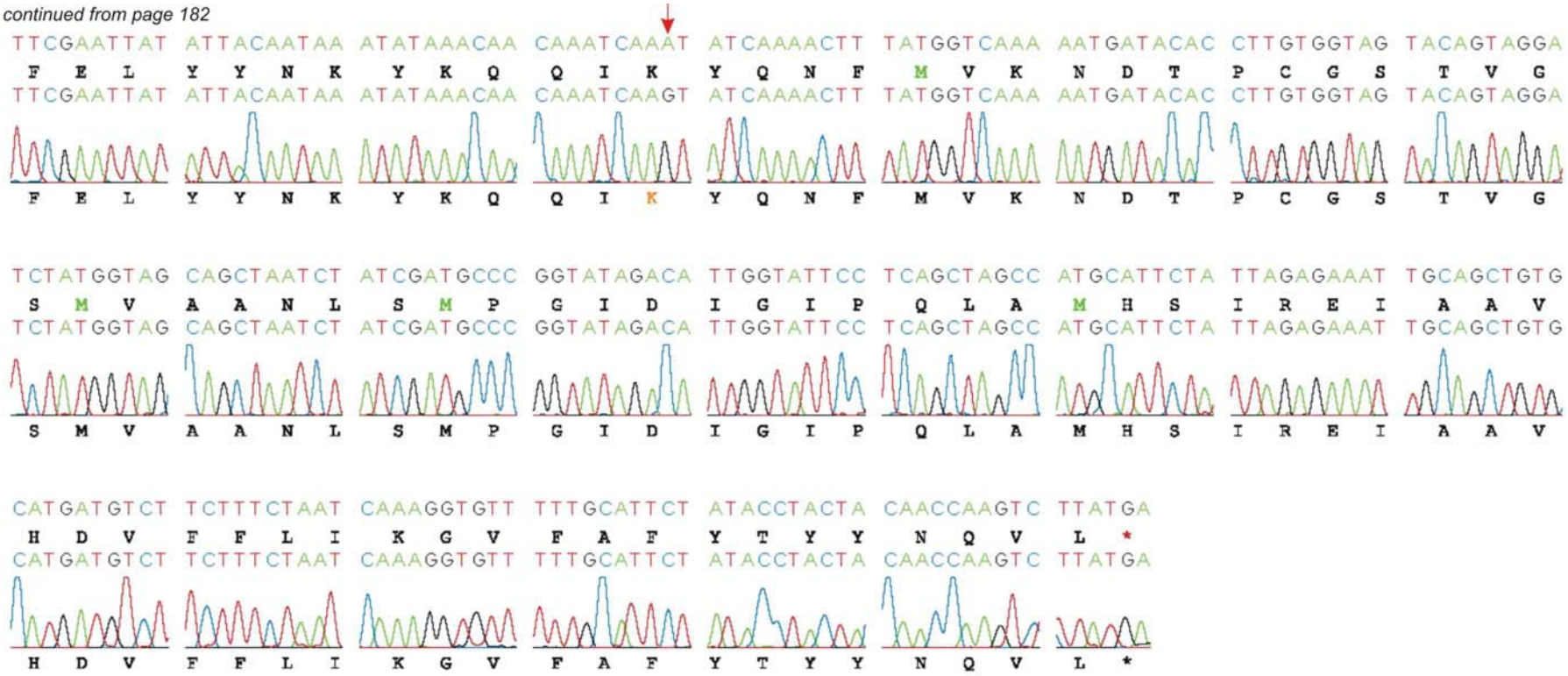
continued from page 180



continued from page 181



continued from page 182



Automated sequencing results of pET-15b-PFI1570c.

The PlasmoDB PFI1570c DNA and protein sequence is given above the automated sequencing DNA and protein sequence. The position of the A to G silent mutation, which translates into a lysine residue, is marked with a red arrow. Blue nucleotides = C; black nucleotides = G; red nucleotides = T; and green nucleotides = A.

REFERENCES

Adda, C.G., Tilley, L., Anders, R.F. and Foley, M. (1999) Isolation of peptides that mimic epitopes on a malarial antigen from random peptide libraries displayed on phage. *Infect Immun*, **67**, 4679-4688.

Aikawa, M. (1971) *Plasmodium*: the fine structure of malarial parasites. *Exp Parasitol*, **30**, 284-320.

Aikawa, M., Miller, L.H., Johnson, J. and Rabbege, J. (1978) Erythrocyte entry by malarial parasites. A moving junction between erythrocyte and parasite. *J Cell Biol*, **77**, 72-82.

Aikawa, M., Uni, Y., Andrutis, A.T. and Howard, R.J. (1986) Membrane-associated electron-dense material of the asexual stages of *Plasmodium falciparum*: evidence for movement from the intracellular parasite to the erythrocyte membrane. *Am J Trop Med Hyg*, **35**, 30-36.

Allary, M., Schrevel, J. and Florent, I. (2002) Properties, stage-dependent expression and localization of *Plasmodium falciparum* M1 family zinc-aminopeptidase. *Parasitology*, **125**, 1-10.

Allen, S.J., O'Donnell, A., Alexander, N.D., Mgone, C.S., Peto, T.E., Clegg, J.B., Alpers, M.P. and Weatherall, D.J. (1999) Prevention of cerebral malaria in children in Papua New Guinea by Southeast Asian Ovalocytosis band 3. *Am J Trop Med Hyg*, **60**, 1056-1060.

Alloisio, N., Dalla Venezia, N., Rana, A., Andrabi, K., Texier, P., Gilsanz, F., Cartron, J.P., Delaunay, J. and Chishti, A.H. (1993) Evidence that red blood cell protein p55 may participate in the skeleton-membrane linkage that involves protein 4.1 and glycophorin C. *Blood*, **82**, 1323-1327.

Alur, H.H., Desai, R.P., Mitra, A.K. and Johnston, T.P. (2001) Inhibition of a model protease-pyroglutamate aminopeptidase by a natural oligosaccharide gum from *Hakea gibbosa*. *Int J Pharm*, **212**, 171-176.

An, X., Guo, X., Sum, H., Morrow, J., Gratzer, W. and Mohandas, N. (2004) Phosphatidylserine binding sites in erythroid spectrin: location and implications for membrane stability. *Biochemistry*, **43**, 310-315.

Anders, R.F. (1986) Multiple cross-reactivities amongst antigens of *Plasmodium falciparum* impair the development of protective immunity against malaria. *Parasite Immunol*, **8**, 529-539.

Anderson, J.P. and Morrow, J.S. (1987) The interaction of calmodulin with human erythrocyte spectrin. Inhibition of protein 4.1-stimulated actin binding. *J Biol Chem*, **262**, 6365-6372.

Anderson, R.A. and Lovrien, R.E. (1984) Glycophorin is linked by band 4.1 protein to the human erythrocyte membrane skeleton. *Nature*, **307**, 655-658.

Ansorge, I., Benting, J., Bhakdi, S. and Lingelbach, K. (1996) Protein sorting in *Plasmodium falciparum*-infected red blood cells permeabilized with the pore-forming protein streptolysin O. *Biochem J*, **315**, 307-314.

Aravind, L., Iyer, L.M., Wellems, T.E. and Miller, L.H. (2003) *Plasmodium* biology: genomic gleanings. *Cell*, **115**, 771-785.

Atkinson, C.T. and Aikawa, M. (1990) Ultrastructure of malaria-infected erythrocytes. *Blood Cells*, **16**, 351-368.

Auld, D.S. (2001) Zinc coordination sphere in biochemical zinc sites. *Biometals*, **14**, 271-313.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (Eds) (1994) *Current Protocols in Molecular Biology*. John Wiley & Sons, Inc., New Jersey.

Ayi, K., Min-Oo, G., Serghides, L., Crockett, M., Kirby-Allen, M., Quirt, I., Gros, P. and Kain, K.C. (2008) Pyruvate kinase deficiency and malaria. *N Engl J Med*, **358**, 1805-1810.

Bahl, A., Brunk, B., Crabtree, J., Fraunholz, M.J., Gajria, B., Grant, G.R., Ginsburg, H., Gupta, D., Kissinger, J.C., Labo, P., Li, L., Mailman, M.D., Milgram, A.J., Pearson, D.S., Roos, D.S., Schug, J., Stoeckert, C.J., Jr. and Whetzel, P. (2003) PlasmoDB: the *Plasmodium* genome resource. A database integrating experimental and computational data. *Nucleic Acids Res*, **31**, 212-215.

Baker, R.P., Wijetilaka, R. and Urban, S. (2006) Two *Plasmodium* rhomboid proteases preferentially cleave different adhesins implicated in all invasive stages of malaria. *PLoS Pathog*, **2**, e113.

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**, 990-995.

Bannister, L. and Mitchell, G. (2003) The ins, outs and roundabouts of malaria. *Trends Parasitol*, **19**, 209-213.

Bannister, L.H., Hopkins, J.M., Fowler, R.E., Krishna, S. and Mitchell, G.H. (2000) A brief illustrated guide to the ultrastructure of *Plasmodium falciparum* asexual blood stages. *Parasitol Today*, **16**, 427-433.

Bennett, V. and Stenbuck, P.J. (1979) The membrane attachment protein for spectrin is associated with band 3 in human erythrocyte membranes. *Nature*, **280**, 468-473.

Bennett, V. and Stenbuck, P.J. (1980) Association between ankyrin and the cytoplasmic domain of band 3 isolated from the human erythrocyte membrane. *J Biol Chem*, **255**, 6424-6432.

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**, 235-242.

Bjerrum, O.J. and Schafer-Nielsen, C. (1986) *Analytical Electrophoresis*. Verlag Chemie, Weinheim.

Blackman, M.J., Fujioka, H., Stafford, W.H., Sajid, M., Clough, B., Fleck, S.L., Aikawa, M., Grainger, M. and Hackett, F. (1998) A subtilisin-like protein in secretory organelles of *Plasmodium falciparum* merozoites. *J Biol Chem*, **273**, 23398-23409.

Blair, P.L., Kappe, S.H., Maciel, J.E., Balu, B. and Adams, J.H. (2002) *Plasmodium falciparum* MAEBL is a unique member of the ebl family. *Mol Biochem Parasitol*, **122**, 35-44.

Blisnick, T., Morales Betoulle, M.E., Barale, J.C., Uzureau, P., Berry, L., Desroses, S., Fujioka, H., Mattei, D. and Braun Breton, C. (2000) *Pfsbp*1, a Maurer's cleft *Plasmodium falciparum* protein, is associated with the erythrocyte skeleton. *Mol Biochem Parasitol*, **111**, 107-121.

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**, e5.

Bradshaw, R.A., Brickey, W.W. and Walker, K.W. (1998) N-terminal processing: the methionine aminopeptidase and N alpha-acetyl transferase families. *Trends Biochem Sci*, **23**, 263-267.

Brenner, S.L. and Korn, E.D. (1979) Spectrin-actin interaction. Phosphorylated and dephosphorylated spectrin tetramer cross-link F-actin. *J Biol Chem*, **254**, 8620-8627.

Brown, D.A. and London, E. (1998) Functions of lipid rafts in biological membranes. *Annu Rev Cell Dev Biol*, **14**, 111-136.

Bruce, L.J., Beckmann, R., Ribeiro, M.L., Peters, L.L., Chasis, J.A., Delaunay, J., Mohandas, N., Anstee, D.J. and Tanner, M.J. (2003) A band 3-based macrocomplex of integral and peripheral proteins in the RBC membrane. *Blood*, **101**, 4180-4188.

Bugg, T.D., H. (2004) *Introduction to Enzyme and Coenzyme Chemistry*. 2nd ed. Blackwell Publishing Ltd., Oxford.

Bull, B.S. and Breton-Gorius, J. (1995) Morphology of the erythron. In *Williams Hematology*. 5th ed. (Eds, Beutler, E., Lichtman, M.A., Coller, B.S. and Kipps, T.J.) McGraw-Hill, Inc., New York.

Buscaglia, C.A., Coppens, I., Hol, W.G. and Nussenzweig, V. (2003) Sites of interaction between aldolase and thrombospondin-related anonymous protein in *Plasmodium*. *Mol Biol Cell*, **14**, 4947-4957.

Byers, T.J. and Branton, D. (1985) Visualization of the protein associations in the erythrocyte membrane skeleton. *Proc Natl Acad Sci U S A*, **82**, 6153-6157.

Calderone, T.L., Stevens, R.D. and Oas, T.G. (1996) High-level misincorporation of lysine for arginine at AGA codons in a fusion protein expressed in *Escherichia coli*. *J Mol Biol*, **262**, 407-412.

Campanella, M.E., Chu, H. and Low, P.S. (2005) Assembly and regulation of a glycolytic enzyme complex on the human erythrocyte membrane. *Proc Natl Acad Sci U S A*, **102**, 2402-2407.

Casey, J.L., Coley, A.M., Anders, R.F., Murphy, V.J., Humberstone, K.S., Thomas, A.W. and Foley, M. (2004) Antibodies to malaria peptide mimics inhibit *Plasmodium falciparum* invasion of erythrocytes. *Infect Immun*, **72**, 1126-1134.

Cerritelli, M.E. and Studier, F.W. (1996) Assembly of T7 capsids from independently expressed and purified head protein and scaffolding protein. *J Mol Biol*, **258**, 286-298.

Cesareni, G. (1992) Peptide display on filamentous phage capsids. A new powerful tool to study protein-ligand interaction. *FEBS Lett*, **307**, 66-70.

Chang, Y.H. and Smith, J.A. (1989) Molecular cloning and sequencing of genomic DNA encoding aminopeptidase I from *Saccharomyces cerevisiae*. *J Biol Chem*, **264**, 6979-6983.

Chasis, J.A. and Mohandas, N. (1992) Red blood cell glycophorins. *Blood*, **80**, 1869-1879.

Chebloune, Y., Pagnier, J., Trabuchet, G., Faure, C., Verdier, G., Labie, D. and Nigon, V. (1988) Structural analysis of the 5' flanking region of the beta-globin gene in African sickle cell anemia patients: further evidence for three origins of the sickle cell mutation in Africa. *Proc Natl Acad Sci U S A*, **85**, 4431-4435.

Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D.G. and Thompson, J.D. (2003) Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res*, **31**, 3497-3500.

Chishti, A.H., Maalouf, G.J., Marfatia, S., Palek, J., Wang, W., Fisher, D. and Liu, S.C. (1994) Phosphorylation of protein 4.1 in *Plasmodium falciparum*-infected human red blood cells. *Blood*, **83**, 3339-3345.

Chishti, A.H., Palek, J., Fisher, D., Maalouf, G.J. and Liu, S.C. (1996) Reduced invasion and growth of *Plasmodium falciparum* into elliptocytic red blood cells with a combined deficiency of protein 4.1, glycophorin C, and p55. *Blood*, **87**, 3462-3469.

Chitnis, C.E. and Blackman, M.J. (2000) Host cell invasion by malaria parasites. *Parasitol Today*, **16**, 411-415.

Chomczynski, P. and Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem*, **162**, 156-159.

Cinquin, O., Christopherson, R.I. and Menz, R.I. (2001) A hybrid plasmid for expression of toxic malarial proteins in *Escherichia coli*. *Mol Biochem Parasitol*, **117**, 245-247.

Coetzer, T., Palek, J., Lawler, J., Liu, S.C., Jarolim, P., Lahav, M., Prchal, J.T., Wang, W., Alter, B.P., Schewitz, G. and et al. (1990) Structural and functional heterogeneity of alpha spectrin mutations involving the spectrin heterodimer self-association site: relationships to hematologic expression of homozygous hereditary elliptocytosis and hereditary pyropoikilocytosis. *Blood*, **75**, 2235-2244.

Coetzer, T.L., Lawler, J., Liu, S.C., Prchal, J.T., Gualtieri, R.J., Brain, M.C., Dacie, J.V. and Palek, J. (1988) Partial ankyrin and spectrin deficiency in severe, atypical hereditary spherocytosis. *N Engl J Med*, **318**, 230-234.

Coetzer, T.L. and Palek, J. (1986) Partial spectrin deficiency in hereditary pyropoikilocytosis. *Blood*, **67**, 919-924.

Cohen, A.M., Liu, S.C., Lawler, J., Derick, L. and Palek, J. (1988) Identification of the protein 4.1 binding site to phosphatidylserine vesicles. *Biochemistry*, **27**, 614-619.

Coley, A.M., Campanale, N.V., Casey, J.L., Hodder, A.N., Crewther, P.E., Anders, R.F., Tilley, L.M. and Foley, M. (2001) Rapid and precise epitope mapping of monoclonal antibodies against *Plasmodium falciparum* AMA1 by combined phage display of fragments and random peptides. *Protein Eng*, **14**, 691-698.

Condron, B.G., Atkins, J.F. and Gesteland, R.F. (1991) Frameshifting in gene 10 of bacteriophage T7. *J Bacteriol*, **173**, 6998-7003.

Cooke, B.M., Mohandas, N. and Coppel, R.L. (2004) Malaria and the red blood cell membrane. *Semin Hematol*, **41**, 173-188.

Cortes, A., Benet, A., Cooke, B.M., Barnwell, J.W. and Reeder, J.C. (2004) Ability of *Plasmodium falciparum* to invade Southeast Asian ovalocytes varies between parasite lines. *Blood*, **104**, 2961-2966.

Cowman, A.F. and Crabb, B.S. (2006) Invasion of red blood cells by malaria parasites. *Cell*, **124**, 755-766.

Cuff, J.A., Clamp, M.E., Siddiqui, A.S., Finlay, M. and Barton, G.J. (1998) JPred: a consensus secondary structure prediction server. *Bioinformatics*, **14**, 892-893.

Curley, G.P., O'Donovan, S.M., McNally, J., Mullally, M., O'Hara, H., Troy, A., O'Callaghan, S.A. and Dalton, J.P. (1994) Aminopeptidases from *Plasmodium falciparum*, *Plasmodium chabaudi chabaudi* and *Plasmodium berghei*. *J Eukaryot Microbiol*, **41**, 119-123.

Daily, J.P., Le Roch, K.G., Sarr, O., Fang, X., Zhou, Y., Ndir, O., Mboup, S., Sultan, A., Winzeler, E.A. and Wirth, D.F. (2004) *In vivo* transcriptional profiling of *Plasmodium falciparum*. *Malar J*, **3**, 30.

Dalal, S. and Klemba, M. (2007) Roles for two aminopeptidases in vacuolar hemoglobin catabolism in *Plasmodium falciparum*. *J Biol Chem*, **282**, 35978-35987.

Das, A., Elmendorf, H.G., Li, W.I. and Haldar, K. (1994) Biosynthesis, export and processing of a 45 kDa protein detected in membrane clefts of erythrocytes infected with *Plasmodium falciparum*. *Biochem J*, **302**, 487-496.

de la Cruz, V.F., Lal, A.A. and McCutchan, T.F. (1988) Immunogenicity and epitope mapping of foreign sequences via genetically engineered filamentous phage. *J Biol Chem*, **263**, 4318-43122.

Decherf, G., Egee, S., Staines, H.M., Ellory, J.C. and Thomas, S.L. (2004) Anionic channels in malaria-infected human red blood cells. *Blood Cells Mol Dis*, **32**, 366-371.

Deguercy, A., Hommel, M. and Schrevel, J. (1990) Purification and characterization of 37-kilodalton proteases from *Plasmodium falciparum* and *Plasmodium berghei* which cleave erythrocyte cytoskeletal components. *Mol Biochem Parasitol*, **38**, 233-244.

Desai, S.A. and Rosenberg, R.L. (1997) Pore size of the malaria parasite's nutrient channel. *Proc Natl Acad Sci U S A*, **94**, 2045-2049.

Devlin, J.J., Panganiban, L.C. and Devlin, P.E. (1990) Random peptide libraries: a source of specific protein binding molecules. *Science*, **249**, 404-406.

Dhawan, S., Dua, M., Chishti, A.H. and Hanspal, M. (2003) Ankyrin peptide blocks falcipain-2-mediated malaria parasite release from red blood cells. *J Biol Chem*, **278**, 30180-30186.

Dluzewski, A.R., Nash, G.B., Wilson, R.J., Reardon, D.M. and Gratzer, W.B. (1992) Invasion of hereditary ovalocytes by *Plasmodium falciparum in vitro* and its relation to intracellular ATP concentration. *Mol Biochem Parasitol*, **55**, 1-7.

Dluzewski, A.R., Rangachari, K., Gratzer, W.B. and Wilson, R.J. (1983) Inhibition of malarial invasion of red cells by chemical and immunochemical linking of spectrin molecules. *Br J Haematol*, **55**, 629-637.

Dodge, J.T., Mitchell, C. and Hanahan, D.J. (1963) The preparation and chemical characteristics of hemoglobin-free ghosts of human erythrocytes. *Arch Biochem Biophys*, **100**, 119-130.

Dolan, S.A., Proctor, J.L., Alling, D.W., Okubo, Y., Wellems, T.E. and Miller, L.H. (1994) Glycophorin B as an EBA-175 independent *Plasmodium falciparum* receptor of human erythrocytes. *Mol Biochem Parasitol*, **64**, 55-63.

Dua, M., Raphael, P., Sijwali, P.S., Rosenthal, P.J. and Hanspal, M. (2001) Recombinant falcipain-2 cleaves erythrocyte membrane ankyrin and protein 4.1. *Mol Biochem Parasitol*, **116**, 95-99.

Duraisingh, M.T., Triglia, T., Ralph, S.A., Rayner, J.C., Barnwell, J.W., McFadden, G.I. and Cowman, A.F. (2003) Phenotypic variation of *Plasmodium falciparum* merozoite proteins directs receptor targeting for invasion of human erythrocytes. *Embo J*, **22**, 1047-1057.

Durand, P.M. and Coetzer, T.L. (2008) Pyruvate kinase deficiency protects against malaria in humans. *Haematologica*, **93**, 939-940.

Dvorak, J.A., Miller, L.H., Whitehouse, W.C. and Shiroishi, T. (1975) Invasion of erythrocytes by malaria merozoites. *Science*, **187**, 748-750.

Dynal® (2000) *Dynabeads® mRNA Direct™ Kit*. Dynal A.S., Oslo.

Eda, K., Eda, S. and Sherman, I.W. (2004) Identification of peptides targeting the surface of *Plasmodium falciparum*-infected erythrocytes using a phage display peptide library. *Am J Trop Med Hyg*, **71**, 190-195.

Eggleson, K.K., Duffin, K.L. and Goldberg, D.E. (1999) Identification and characterization of falcilysin, a metallopeptidase involved in hemoglobin catabolism within the malaria parasite *Plasmodium falciparum*. *J Biol Chem*, **274**, 32411-32417.

Elmendorf, H.G. and Haldar, K. (1993) Secretory transport in *Plasmodium*. *Parasitol Today*, **9**, 98-102.

Elmendorf, H.G. and Haldar, K. (1994) *Plasmodium falciparum* exports the Golgi marker sphingomyelin synthase into a tubovesicular network in the cytoplasm of mature erythrocytes. *J Cell Biol*, **124**, 449-462.

Etzion, Z., Murray, M.C. and Perkins, M.E. (1991) Isolation and characterization of rhoptries of *Plasmodium falciparum*. *Mol Biochem Parasitol*, **47**, 51-61.

Facer, C.A. (1989) Malaria, hereditary elliptocytosis, and pyropoikilocytosis. *Lancet*, **1**, 897.

Fairbanks, G., Steck, T.L. and Wallach, D.F. (1971) Electrophoretic analysis of the major polypeptides of the human erythrocyte membrane. *Biochemistry*, **10**, 2606-2617.

Ferguson, K.A. (1964) Starch-gel electrophoresis - application to the classification of pituitary proteins and polypeptides. *Metabolism*, **13**, 985-1002.

Fields, S. and Song, O. (1989) A novel genetic system to detect protein-protein interactions. *Nature*, **340**, 245-246.

Finn, R.D., Mistry, J., Schuster-Bockler, B., Griffiths-Jones, S., Hollich, V., Lassmann, T., Moxon, S., Marshall, M., Khanna, A., Durbin, R., Eddy, S.R., Sonnhammer, E.L. and Bateman, A. (2006) Pfam: clans, web tools and services. *Nucleic Acids Res*, **34**, D247-251.

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**, 50.

Florens, L., Liu, X., Wang, Y., Yang, S., Schwartz, O., Peglar, M., Carucci, D.J., Yates, J.R. and Wub, Y. (2004) Proteomics approach reveals novel proteins on the surface of malaria-infected erythrocytes. *Mol Biochem Parasitol*, **135**, 1-11.

Florens, L., Washburn, M.P., Raine, J.D., Anthony, R.M., Grainger, M., Haynes, J.D., Moch, J.K., Muster, N., Sacci, J.B., Tabb, D.L., Witney, A.A., Wolters, D., Wu, Y., Gardner, M.J., Holder, A.A., Sinden, R.E., Yates, J.R. and Carucci, D.J. (2002) A proteomic view of the *Plasmodium falciparum* life cycle. *Nature*, **419**, 520-526.

Florent, I., Derhy, Z., Allary, M., Monsigny, M., Mayer, R. and Schrevel, J. (1998) A *Plasmodium falciparum* aminopeptidase gene belonging to the M1 family of zinc-metallopeptidases is expressed in erythrocytic stages. *Mol Biochem Parasitol*, **97**, 149-160.

Foley, M. and Tilley, L. (1995) Home improvements: malaria and the red blood cell. *Parasitol Today*, **11**, 436-439.

Foley, M., Tilley, L., Sawyer, W.H. and Anders, R.F. (1991) The ring-infected erythrocyte surface antigen of *Plasmodium falciparum* associates with spectrin in the erythrocyte membrane. *Mol Biochem Parasitol*, **46**, 137-147.

Foo, L.C., Rekhraj, V., Chiang, G.L. and Mak, J.W. (1992) Ovalocytosis protects against severe malaria parasitemia in the Malayan aborigines. *Am J Trop Med Hyg*, **47**, 271-275.

Fowler, V.M. (1996) Regulation of actin filament length in erythrocytes and striated muscle. *Curr Opin Cell Biol*, **8**, 86-96.

Franzetti, B., Schoehn, G., Hernandez, J.F., Jaquinod, M., Ruigrok, R.W. and Zaccai, G. (2002) Tetrahedral aminopeptidase: a novel large protease complex from archaea. *Embo J*, **21**, 2132-2138.

Fujioka, H. and Aikawa, M. (1999) The malaria parasite and its life-cycle. In *Malaria: Molecular and Clinical Aspects*. (Eds, Wahlgren, M. and Perlmann, P.) Harwood Academic Publishers, Singapore.

Funder, J. and Wieth, J.O. (1966) Chloride and hydrogen ion distribution between human red cells and plasma. *Acta Physiol Scand*, **68**, 234-245.

Garcia, C.R., Takeuschi, M., Yoshioka, K. and Miyamoto, H. (1997) Imaging *Plasmodium falciparum*-infected ghost and parasite by atomic force microscopy. *J Struct Biol*, **119**, 92-98.

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**, 1741-1745.

Gardner, M.J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R.W., Carlton, J.M., Pain, A., Nelson, K.E., Bowman, S., Paulsen, I.T., James, K., Eisen, J.A., Rutherford, K., Salzberg, S.L., Craig, A., Kyes, S., Chan, M.S., Nene, V., Shallom, S.J., Suh, B., Peterson, J., Angiuoli, S., Pertea, M., Allen, J., Selengut, J., Haft, D., Mather, M.W., Vaidya, A.B., Martin, D.M., Fairlamb, A.H., Fraunholz, M.J., Roos, D.S., Ralph, S.A., McFadden, G.I., Cummings, L.M., Subramanian, G.M., Mungall, C., Venter, J.C., Carucci, D.J., Hoffman, S.L., Newbold, C., Davis, R.W., Fraser, C.M. and Barrell, B. (2002) Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature*, **419**, 498-511.

Gattiker, A., Gasteiger, E. and Bairoch, A. (2002) ScanProsite: a reference implementation of a PROSITE scanning tool. *Appl Bioinformatics*, **1**, 107-108.

GE Healthcare (2007) *pGEX Vectors, GST Gene Fusion System*. GE Healthcare Bio-Sciences AB, Uppsala.

Gharahdaghi, F., Weinberg, C.R., Meagher, D.A., Imai, B.S. and Mische, S.M. (1999) Mass spectrometric identification of proteins from silver-stained polyacrylamide gel: a method for the removal of silver ions to enhance sensitivity. *Electrophoresis*, **20**, 601-605.

Ghosh, A.K., Ribolla, P.E. and Jacobs-Lorena, M. (2001) Targeting *Plasmodium* ligands on mosquito salivary glands and midgut with a phage display peptide library. *Proc Natl Acad Sci U S A*, **98**, 13278-13281.

Gilberger, T.W., Thompson, J.K., Triglia, T., Good, R.T., Duraisingh, M.T. and Cowman, A.F. (2003) A novel erythrocyte binding antigen-175 paralogue from *Plasmodium falciparum* defines a new trypsin-resistant receptor on human erythrocytes. *J Biol Chem*, **278**, 14480-14486.

Glushakova, S., Yin, D., Li, T. and Zimmerberg, J. (2005) Membrane transformation during malaria parasite release from human red blood cells. *Curr Biol*, **15**, 1645-1650.

Gluzman, I.Y., Francis, S.E., Oksman, A., Smith, C.E., Duffin, K.L. and Goldberg, D.E. (1994) Order and specificity of the *Plasmodium falciparum* hemoglobin degradation pathway. *J Clin Invest*, **93**, 1602-1608.

Goel, V.K., Li, X., Chen, H., Liu, S.C., Chishti, A.H. and Oh, S.S. (2003) Band 3 is a host receptor binding merozoite surface protein 1 during the *Plasmodium falciparum* invasion of erythrocytes. *Proc Natl Acad Sci U S A*, **100**, 5164-5169.

Goldbarg, J.A. and Rutenburg, A.M. (1958) The colorimetric determination of leucine aminopeptidase in urine and serum of normal subjects and patients with cancer and other diseases. *Cancer*, **11**, 283-291.

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**, 2931-2935.

Greenwood, J., Willis, A.E. and Perham, R.N. (1991) Multiple display of foreign peptides on a filamentous bacteriophage. Peptides from *Plasmodium falciparum* circumsporozoite protein as antigens. *J Mol Biol*, **220**, 821-827.

Guex, N. and Peitsch, M.C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, **18**, 2714-2723.

Hackett, F., Sajid, M., Withers-Martinez, C., Grainger, M. and Blackman, M.J. (1999) PfSUB-2: a second subtilisin-like protein in *Plasmodium falciparum* merozoites. *Mol Biochem Parasitol*, **103**, 183-195.

Hadley, T.J., Klotz, F.W., Pasvol, G., Haynes, J.D., McGinniss, M.H., Okubo, Y. and Miller, L.H. (1987) *Falciparum* malaria parasites invade erythrocytes that lack glycophorin A and B (M^kM^k). Strain differences indicate receptor heterogeneity and two pathways for invasion. *J Clin Invest*, **80**, 1190-1193.

Hang, V.T., Be, T.V., Tran, P.N., Thanh, L.T., Hien, L.V., O'Brien, E. and Morris, G.E. (1995) Screening donor blood for malaria by polymerase chain reaction. *Trans R Soc Trop Med Hyg*, **89**, 44-47.

Harris, H.W., Jr. and Lux, S.E. (1980) Structural characterization of the phosphorylation sites of human erythrocyte spectrin. *J Biol Chem*, **255**, 11512-11520.

Heal, K.G., Hill, H.R., Stockley, P.G., Hollingdale, M.R. and Taylor-Robinson, A.W. (1999) Expression and immunogenicity of a liver stage malaria epitope presented as a foreign peptide on the surface of RNA-free MS2 bacteriophage capsids. *Vaccine*, **18**, 251-258.

Hedrick, J.L. and Smith, A.J. (1968) Size and charge isomer separation and estimation of molecular weights of proteins by disc gel electrophoresis. *Arch Biochem Biophys*, **126**, 155-164.

Herrera, S., Rudin, W., Herrera, M., Clavijo, P., Mancilla, L., de Plata, C., Matile, H. and Certa, U. (1993) A conserved region of the MSP-1 surface protein of *Plasmodium falciparum* contains a recognition sequence for erythrocyte spectrin. *Embo J*, **12**, 1607-1614.

Hiller, N.L., Bhattacharjee, S., van Ooij, C., Liolios, K., Harrison, T., Lopez-Estrano, C. and Haldar, K. (2004) A host-targeting signal in virulence proteins reveals a secretome in malarial infection. *Science*, **306**, 1934-1937.

Hillier, C.J., Ware, L.A., Barbosa, A., Angov, E., Lyon, J.A., Heppner, D.G. and Lanar, D.E. (2005) Process development and analysis of liver-stage antigen 1, a preerythrocyte-stage protein-based vaccine for *Plasmodium falciparum*. *Infect Immun*, **73**, 2109-2115.

Hirokawa, T., Boon-Chieng, S. and Mitaku, S. (1998) SOSUI: classification and secondary structure prediction system for membrane proteins. *Bioinformatics*, **14**, 378-379.

Hoefer Scientific Instruments (1994) *Protein Electrophoresis Applications Guide*. Hoefer, Inc., San Francisco.

Hoess, R.H. (1993) Phage display of peptides and protein domains. *Curr Opin Struct Biol*, **3**, 572-579.

Hoess, R.H. (2001) Protein design and phage display. *Chem Rev*, **101**, 3205-3218.

Holder, A.A. and Freeman, R.R. (1982) Biosynthesis and processing of a *Plasmodium falciparum* schizont antigen recognized by immune serum and a monoclonal antibody. *J Exp Med*, **156**, 1528-1538.

Houshmand, H., Froman, G. and Magnusson, G. (1999) Use of bacteriophage T7 displayed peptides for determination of monoclonal antibody specificity and biosensor analysis of the binding reaction. *Anal Biochem*, **268**, 363-370.

Howell, S.A., Well, I., Fleck, S.L., Kettleborough, C., Collins, C.R. and Blackman, M.J. (2003) A single malaria merozoite serine protease mediates shedding of multiple surface proteins by juxtamembrane cleavage. *J Biol Chem*, **278**, 23890-23898.

Huse, W.D., Stinchcombe, T.J., Glaser, S.M., Starr, L., MacLean, M., Hellstrom, K.E., Hellstrom, I. and Yelton, D.E. (1992) Application of a filamentous phage pVIII fusion protein system suitable for efficient production, screening, and mutagenesis of F(ab) antibody fragments. *J Immunol*, **149**, 3914-3920.

International Union of Biochemistry (1992) *Enzyme Nomenclature*. Academic Press, San Diego.

Jacobsson, K. and Frykberg, L. (1998) Gene VIII-based, phage-display vectors for selection against complex mixtures of ligands. *Biotechniques*, **24**, 294-301.

Jarolim, P., Palek, J., Amato, D., Hassan, K., Sapak, P., Nurse, G.T., Rubin, H.L., Zhai, S., Sahr, K.E. and Liu, S.C. (1991) Deletion in erythrocyte band 3 gene in malaria-resistant Southeast Asian ovalocytosis. *Proc Natl Acad Sci U S A*, **88**, 11022-11026.

Jenkins, J.D., Madden, D.P. and Steck, T.L. (1984) Association of phosphofructokinase and aldolase with the membrane of the intact erythrocyte. *J Biol Chem*, **259**, 9374-9378.

Jespers, L.S., Messens, J.H., De Keyser, A., Eeckhout, D., Van den Brande, I., Gansemans, Y.G., Lauwereys, M.J., Vlasuk, G.P. and Stanssens, P.E. (1995) Surface expression and ligand-based selection of cDNAs fused to filamentous phage gene VI. *Biotechnology*, **13**, 378-382.

Jiang, J., Abu-Shilbayeh, L. and Rao, V.B. (1997) Display of a PorA peptide from *Neisseria meningitidis* on the bacteriophage T4 capsid surface. *Infect Immun*, **65**, 4770-4777.

Jones, M.L., Kitson, E.L. and Rayner, J.C. (2006) *Plasmodium falciparum* erythrocyte invasion: a conserved myosin associated complex. *Mol Biochem Parasitol*, **147**, 74-84.

Kanaani, J. and Ginsburg, H. (1991) Transport of lactate in *Plasmodium falciparum*-infected human erythrocytes. *J Cell Physiol*, **149**, 469-476.

Kaneko, O., Tsuboi, T., Ling, I.T., Howell, S., Shirano, M., Tachibana, M., Cao, Y.M., Holder, A.A. and Torii, M. (2001) The high molecular mass rhoptry protein, RhopH1, is encoded by members of the clag multigene family in *Plasmodium falciparum* and *Plasmodium yoelii*. *Mol Biochem Parasitol*, **118**, 223-231.

Kaneko, O., Yim Lim, B.Y., Iriko, H., Ling, I.T., Otsuki, H., Grainger, M., Tsuboi, T., Adams, J.H., Mattei, D., Holder, A.A. and Torii, M. (2005) Apical expression of three RhopH1/Clag proteins as components of the *Plasmodium falciparum* RhopH complex. *Mol Biochem Parasitol*, **143**, 20-28.

Kant, J.A. and Steck, T.L. (1973) Specificity in the association of glyceraldehyde 3-phosphate dehydrogenase with isolated human erythrocyte membranes. *J Biol Chem*, **248**, 8457-8464.

Kapust, R.B. and Waugh, D.S. (1999) *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci*, **8**, 1668-1674.

Kats, L.M., Black, C.G., Proellocks, N.I. and Coppel, R.L. (2006) *Plasmodium* rhoptries: how things went pear-shaped. *Trends Parasitol*, **22**, 269-276.

Keeley, A. and Soldati, D. (2004) The glideosome: a molecular machine powering motility and host-cell invasion by Apicomplexa. *Trends Cell Biol*, **14**, 528-532.

Kennedy, S.P., Warren, S.L., Forget, B.G. and Morrow, J.S. (1991) Ankyrin binds to the 15th repetitive unit of erythroid and nonerythroid beta-spectrin. *J Cell Biol*, **115**, 267-277.

Kidson, C., Lamont, G., Saul, A. and Nurse, G.T. (1981) Ovalocytic erythrocytes from Melanesians are resistant to invasion by malaria parasites in culture. *Proc Natl Acad Sci U S A*, **78**, 5829-5832.

Kiefhaber, T., Rudolph, R., Kohler, H.H. and Buchner, J. (1991) Protein aggregation *in vitro* and *in vivo*: a quantitative model of the kinetic competition between folding and aggregation. *Biotechnology (N Y)*, **9**, 825-829.

Kilejian, A. (1979) Characterization of a protein correlated with the production of knob-like protrusions on membranes of erythrocytes infected with *Plasmodium falciparum*. *Proc Natl Acad Sci U S A*, **76**, 4650-4653.

Kirk, K. (2001) Membrane transport in the malaria-infected erythrocyte. *Physiol Rev*, **81**, 495-537.

Kitjaroentham, A., Suthiphongchai, T. and Wilairat, P. (2006) Effect of metalloprotease inhibitors on invasion of red blood cell by *Plasmodium falciparum*. *Acta Trop*, **97**, 5-9.

Klemba, M., Gluzman, I. and Goldberg, D.E. (2004) A *Plasmodium falciparum* dipeptidyl aminopeptidase I participates in vacuolar hemoglobin degradation. *J Biol Chem*, **279**, 43000-43007.

Konigk, E. and Mirtsch, S. (1977) *Plasmodium chabaudi*-infection of mice: specific activities of erythrocyte membrane-associated enzymes and patterns of proteins and glycoproteins of erythrocyte membrane preparations. *Tropenmed Parasitol*, **28**, 17-22.

Korsgren, C. and Cohen, C.M. (1988) Associations of human erythrocyte band 4.2. Binding to ankyrin and to the cytoplasmic domain of band 3. *J Biol Chem*, **263**, 10212-10218.

Kost, T.A., Condreay, J.P. and Jarvis, D.L. (2005) Baculovirus as versatile vectors for protein expression in insect and mammalian cells. *Nat Biotechnol*, **23**, 567-575.

Krugliak, M., Zhang, J. and Ginsburg, H. (2002) Intraerythrocytic *Plasmodium falciparum* utilizes only a fraction of the amino acids derived from the digestion of host cell cytosol for the biosynthesis of its proteins. *Mol Biochem Parasitol*, **119**, 249-256.

Kun, J.F., Hibbs, A.R., Saul, A., McColl, D.J., Coppel, R.L. and Anders, R.F. (1997) A putative *Plasmodium falciparum* exported serine/threonine protein kinase. *Mol Biochem Parasitol*, **85**, 41-51.

Kushwaha, A., Perween, A., Mukund, S., Majumdar, S., Bhardwaj, D., Chowdhury, N.R. and Chauhan, V.S. (2002) Amino terminus of *Plasmodium falciparum* acidic basic repeat antigen interacts with the erythrocyte membrane through band 3 protein. *Mol Biochem Parasitol*, **122**, 45-54.

LaCount, D.J., Vignali, M., Chettier, R., Phansalkar, A., Bell, R., Hesselberth, J.R., Schoenfeld, L.W., Ota, I., Sahasrabudhe, S., Kurschner, C., Fields, S. and Hughes, R.E. (2005) A protein interaction network of the malaria parasite *Plasmodium falciparum*. *Nature*, **438**, 103-107.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680-685.

Lander, E.S., Linton, L.M., Birren, B.*, et al.* (2001) Initial sequencing and analysis of the human genome. *Nature*, **409**, 860-921.

Langreth, S.G., Jensen, J.B., Reese, R.T. and Trager, W. (1978) Fine structure of human malaria *in vitro*. *J Protozool*, **25**, 443-452.

Lanzillotti, R. and Coetzer, T.L. (2006) The 10 kDa domain of human erythrocyte protein 4.1 binds the *Plasmodium falciparum* EBA-181 protein. *Malar J*, **5**, 100.

Lawler, J., Liu, S.C., Palek, J. and Prchal, J. (1984) A molecular defect of spectrin in a subset of patients with hereditary elliptocytosis. Alterations in the alpha-subunit domain involved in spectrin self-association. *J Clin Invest*, **73**, 1688-1695.

Le Bonniec, S., Deregnaucourt, C., Redeker, V., Banerjee, R., Grellier, P., Goldberg, D.E. and Schrevel, J. (1999) Plasmepsin II, an acidic hemoglobinase from the *Plasmodium falciparum* food vacuole, is active at neutral pH on the host erythrocyte membrane skeleton. *J Biol Chem*, **274**, 14218-14223.

Le Roch, K.G., Johnson, J.R., Florens, L., Zhou, Y., Santrosyan, A., Grainger, M., Yan, S.F., Williamson, K.C., Holder, A.A., Carucci, D.J., Yates, J.R. and Winzeler, E.A. (2004) Global analysis of transcript and protein levels across the *Plasmodium falciparum* life cycle. *Genome Res*, **14**, 2308-2318.

Le Roch, K.G., Zhou, Y., Blair, P.L., Grainger, M., Moch, J.K., Haynes, J.D., De La Vega, P., Holder, A.A., Batalov, S., Carucci, D.J. and Winzeler, E.A. (2003) Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science*, **301**, 1503-1508.

Lehninger, A.L., Nelson, D.L. and Cox, M.M. (1993) *Principles of Biochemistry*. 2nd ed. Worth Publishers, New York.

Lew, V.L. (2005) Malaria: endless fascination with merozoite release. *Curr Biol*, **15**, R760-761.

Lew, V.L., Tiffert, T. and Ginsburg, H. (2003) Excess hemoglobin digestion and the osmotic stability of *Plasmodium falciparum*-infected red blood cells. *Blood*, **101**, 4189-4194.

Li, J., Matsuoka, H., Mitamura, T. and Horii, T. (2002) Characterization of proteases involved in the processing of *Plasmodium falciparum* serine repeat antigen (SERA). *Mol Biochem Parasitol*, **120**, 177-186.

Li, X. and Bennett, V. (1996) Identification of the spectrin subunit and domains required for formation of spectrin/adducin/actin complexes. *J Biol Chem*, **271**, 15695-15702.

Lindqvist, B.H. and Naderi, S. (1995) Peptide presentation by bacteriophage P4. *FEMS Microbiol Rev*, **17**, 33-39.

Ling, E., Danilov, Y.N. and Cohen, C.M. (1988) Modulation of red cell band 4.1 function by cAMP-dependent kinase and protein kinase C phosphorylation. *J Biol Chem*, **263**, 2209-2216.

Lingelbach, K. and Przyborski, J.M. (2006) The long and winding road: protein trafficking mechanisms in the *Plasmodium falciparum* infected erythrocyte. *Mol Biochem Parasitol*, **147**, 1-8.

Liu, J., Tan, H. and Rost, B. (2002) Loopy proteins appear conserved in evolution. *J Mol Biol*, **322**, 53-64.

Liu, S.C., Derick, L.H. and Palek, J. (1987) Visualization of the hexagonal lattice in the erythrocyte membrane skeleton. *J Cell Biol*, **104**, 527-536.

Liu, S.C., Palek, J. and Prchal, J.T. (1982) Defective spectrin dimer-dimer association with hereditary elliptocytosis. *Proc Natl Acad Sci U S A*, **79**, 2072-2076.

Lovett, J.L. and Sibley, L.D. (2003) Intracellular calcium stores in *Toxoplasma gondii* govern invasion of host cells. *J Cell Sci*, **116**, 3009-3016.

Lowman, H.B., Bass, S.H., Simpson, N. and Wells, J.A. (1991) Selecting high-affinity binding proteins by monovalent phage display. *Biochemistry*, **30**, 10832-10838.

Lu, P.W., Soong, C.J. and Tao, M. (1985) Phosphorylation of ankyrin decreases its affinity for spectrin tetramer. *J Biol Chem*, **260**, 14958-14964.

Lundquist, R., Nielsen, L.K., Jafarshad, A., Soesoe, D., Christensen, L.H., Druilhe, P. and Dziegiel, M.H. (2006) Human recombinant antibodies against *Plasmodium falciparum* merozoite surface protein 3 cloned from peripheral blood leukocytes of individuals with immunity to malaria demonstrate antiparasitic properties. *Infect Immun*, **74**, 3222-3231.

Lyon, J.A. and Haynes, J.D. (1986) *Plasmodium falciparum* antigens synthesized by schizonts and stabilized at the merozoite surface when schizonts mature in the presence of protease inhibitors. *J Immunol*, **136**, 2245-2251.

Maenaka, K., Furuta, M., Tsumoto, K., Watanabe, K., Ueda, Y. and Kumagai, I. (1996) A stable phage-display system using a phagemid vector: phage display of hen egg-white lysozyme (HEL), *Escherichia coli* alkaline phosphatase, and anti-HEL monoclonal antibody, HyHEL10. *Biochem Biophys Res Commun*, **218**, 682-687.

Magowan, C., Coppel, R.L., Lau, A.O., Moronne, M.M., Tchernia, G. and Mohandas, N. (1995) Role of the *Plasmodium falciparum* mature-parasite-infected erythrocyte surface antigen (MESA/PfEMP-2) in malarial infection of erythrocytes. *Blood*, **86**, 3196-3204.

Magowan, C., Nunomura, W., Waller, K.L., Yeung, J., Liang, J., Van Dort, H., Low, P.S., Coppel, R.L. and Mohandas, N. (2000) *Plasmodium falciparum* histidine-rich protein 1 associates with the band 3 binding domain of ankyrin in the infected red cell membrane. *Biochim Biophys Acta*, **1502**, 461-470.

Maier, A.G., Duraisingh, M.T., Reeder, J.C., Patel, S.S., Kazura, J.W., Zimmerman, P.A. and Cowman, A.F. (2003) *Plasmodium falciparum* erythrocyte invasion through glycophorin C and selection for Gerbich negativity in human populations. *Nat Med*, **9**, 87-92.

Makowski, L. (1993) Structural constraints on the display of foreign peptides on filamentous bacteriophages. *Gene*, **128**, 5-11.

Manno, S., Takakuwa, Y. and Mohandas, N. (2005) Modulation of erythrocyte membrane mechanical function by protein 4.1 phosphorylation. *J Biol Chem*, **280**, 7581-7587.

Manno, S., Takakuwa, Y., Nagao, K. and Mohandas, N. (1995) Modulation of erythrocyte membrane mechanical function by beta-spectrin phosphorylation and dephosphorylation. *J Biol Chem*, **270**, 5659-5665.

Marchesi, V.T. (1979) Spectrin: present status of a putative cyto-skeletal protein of the red cell membrane. *J Membr Biol*, **51**, 101-131.

Marti, M., Good, R.T., Rug, M., Knuepfer, E. and Cowman, A.F. (2004) Targeting malaria virulence and remodeling proteins to the host erythrocyte. *Science*, **306**, 1930-1933.

Mastico, R.A., Talbot, S.J. and Stockley, P.G. (1993) Multiple presentation of foreign peptides on the surface of an RNA-free spherical bacteriophage capsid. *J Gen Virol*, **74**, 541-548.

Mattei, D., Ward, G.E., Langsley, G. and Lingelbach, K. (1999) Novel secretory pathways in *Plasmodium*? *Parasitol Today*, **15**, 235-237.

Mayer, D.C., Kaneko, O., Hudson-Taylor, D.E., Reid, M.E. and Miller, L.H. (2001) Characterization of a *Plasmodium falciparum* erythrocyte-binding protein paralogous to EBA-175. *Proc Natl Acad Sci U S A*, **98**, 5222-5227.

Mayer, D.C., Mu, J.B., Feng, X., Su, X.Z. and Miller, L.H. (2002) Polymorphism in a *Plasmodium falciparum* erythrocyte-binding ligand changes its receptor specificity. *J Exp Med*, **196**, 1523-1528.

Mehlin, C., Boni, E., Buckner, F.S., Engel, L., Feist, T., Gelb, M.H., Haji, L., Kim, D., Liu, C., Mueller, N., Myler, P.J., Reddy, J.T., Sampson, J.N., Subramanian, E., Van Voorhis, W.C., Worthey, E., Zucker, F. and Hol, W.G. (2006) Heterologous expression of proteins from *Plasmodium falciparum*: results from 1000 genes. *Mol Biochem Parasitol*, **148**, 144-160.

Miller, L.H., Baruch, D.I., Marsh, K. and Doumbo, O.K. (2002) The pathogenic basis of malaria. *Nature*, **415**, 673-679.

Miller, L.H., Mason, S.J., Clyde, D.F. and McGinniss, M.H. (1976) The resistance factor to *Plasmodium vivax* in blacks. The Duffy-blood-group genotype, FyFy. *N Engl J Med*, **295**, 302-304.

Min, T. and Shapiro, L. (2006) to be published. www.rcsb.org/pdb/home/home.do.

Min-Oo, G., Fortin, A., Tam, M.F., Nantel, A., Stevenson, M.M. and Gros, P. (2003) Pyruvate kinase deficiency in mice protects against malaria. *Nat Genet*, **35**, 357-362.

Min-Oo, G. and Gros, P. (2005) Erythrocyte variants and the nature of their malaria protective effect. *Cell Microbiol*, **7**, 753-763.

Mogk, A., Mayer, M.P. and Deuerling, E. (2002) Mechanisms of protein folding: molecular chaperones and their application in biotechnology. *Chembiochem*, **3**, 807-814.

Mohandas, N., Winardi, R., Knowles, D., Leung, A., Parra, M., George, E., Conboy, J. and Chasis, J. (1992) Molecular basis for membrane rigidity of hereditary ovalocytosis. A novel mechanism involving the cytoplasmic domain of band 3. *J Clin Invest*, **89**, 686-692.

Morrissette, N.S. and Sibley, L.D. (2002) Cytoskeleton of Apicomplexan parasites. *Microbiol Mol Biol Rev*, **66**, 21-38.

Morrow, J.S., Speicher, D.W., Knowles, W.J., Hsu, C.J. and Marchesi, V.T. (1980) Identification of functional domains of human erythrocyte spectrin. *Proc Natl Acad Sci U S A*, **77**, 6592-6596.

Nardin, E.H., Nussenzweig, V., Nussenzweig, R.S., Collins, W.E., Harinasuta, K.T., Tapchaisri, P. and Chomcharn, Y. (1982) Circumsporozoite proteins of human malaria parasites *Plasmodium falciparum* and *Plasmodium vivax*. *J Exp Med*, **156**, 20-30.

Novagen (1999a) *OrientExpress™ cDNA Manual*. Novagen, Inc., Madison.

Novagen (1999b) *pET-15b Vector*. Novagen, Inc., Madison.

Novagen (2000a) *T7Select®10-3b Vector*. Novagen, Inc., Madison.

Novagen (2000b) *T7Select® System Manual*. Novagen, Inc., Madison.

Novagen (2003) *pET System Manual*. 10th ed. Novagen, Inc., Madison.

Oaks, S.C., Mitchell, V.S., Pearson, G.W. and Carpenter, C.C.J. (1991) *Malaria: Obstacles and Opportunities*. National Academy Press, Washington.

O'Farrell, P.H. (1975) High resolution two-dimensional electrophoresis of proteins. *J Biol Chem*, **250**, 4007-4021.

Oh, S.S., Voigt, S., Fisher, D., Yi, S.J., LeRoy, P.J., Derick, L.H., Liu, S. and Chishti, A.H. (2000) *Plasmodium falciparum* erythrocyte membrane protein 1 is anchored to the actin-spectrin junction and knob-associated histidine-rich protein in the erythrocyte skeleton. *Mol Biochem Parasitol*, **108**, 237-247.

Pagni, M., Ioannidis, V., Cerutti, L., Zahn-Zabal, M., Jongeneel, C.V. and Falquet, L. (2004) MyHits: a new interactive resource for protein annotation and domain identification. *Nucleic Acids Res*, **32**, W332-335.

Palmer, T. (1995) *Understanding Enzymes*. 4th ed. Ellis Horwood Ltd., Hempstead.

Parker, P.D., Tilley, L. and Klonis, N. (2004) *Plasmodium falciparum* induces reorganization of host membrane proteins during intraerythrocytic growth. *Blood*, **103**, 2404-2406.

Parmley, S.F. and Smith, G.P. (1988) Antibody-selectable filamentous fd phage vectors: affinity purification of target genes. *Gene*, **73**, 305-318.

Pedelacq, J.D., Cabantous, S., Tran, T., Terwilliger, T.C. and Waldo, G.S. (2006) Engineering and characterization of a superfolder green fluorescent protein. *Nat Biotechnol*, **24**, 79-88.

Pei, X., An, X., Guo, X., Tarnawski, M., Coppel, R. and Mohandas, N. (2005) Structural and functional studies of interaction between *Plasmodium falciparum* knob-associated histidine-rich protein (KAHRP) and erythrocyte spectrin. *J Biol Chem*, **280**, 31166-31171.

Pei, X., Guo, X., Coppel, R., Bhattacharjee, S., Haldar, K., Gratzer, W., Mohandas, N. and An, X. (2007) The ring-infected erythrocyte surface antigen (RESA) of *Plasmodium falciparum* stabilizes spectrin tetramers and suppresses further invasion. *Blood*, **110**, 1036-1042.

Perkins, M.E. (1984) Surface proteins of *Plasmodium falciparum* merozoites binding to the erythrocyte receptor, glycophorin. *J Exp Med*, **160**, 788-798.

Perkins, M.E. and Ziefer, A. (1994) Preferential binding of *Plasmodium falciparum* SERA and rhoptry proteins to erythrocyte membrane inner leaflet phospholipids. *Infect Immun*, **62**, 1207-1212.

Peterson, M.G., Marshall, V.M., Smythe, J.A., Crewther, P.E., Lew, A., Silva, A., Anders, R.F. and Kemp, D.J. (1989) Integral membrane protein located in the apical complex of P*lasmodium falciparum*. *Mol Cell Biol*, **9**, 3151-3154.

Piche, C. and Schernthaner, J.P. (2005) Optimization of in vitro transcription and full-length cDNA synthesis using the T4 bacteriophage gene 32 protein. *J Biomol Tech*, **16**, 239-247.

Pierce (2005) *Protein Assay: Technical Handbook*. Pierce Biotechnology, Inc., Rockford.

Pinder, J.C., Fowler, R.E., Dluzewski, A.R., Bannister, L.H., Lavin, F.M., Mitchell, G.H., Wilson, R.J. and Gratzer, W.B. (1998) Actomyosin motor in the merozoite of the malaria parasite, *Plasmodium falciparum*: implications for red cell invasion. *J Cell Sci*, **111**, 1831-1839.

Pizzi, E. and Frontali, C. (2001) Low-complexity regions in *Plasmodium falciparum* proteins. *Genome Res*, **11**, 218-229.

Ponpuak, M., Klemba, M., Park, M., Gluzman, I.Y., Lamppa, G.K. and Goldberg, D.E. (2007) A role for falcilysin in transit peptide degradation in the *Plasmodium falciparum* apicoplast. *Mol Microbiol*, **63**, 314-334.

Przyborski, J.M., Wickert, H., Krohne, G. and Lanzer, M. (2003) Maurer's clefts - a novel secretory organelle? *Mol Biochem Parasitol*, **132**, 17-26.

Rangachari, K., Dluzewski, A., Wilson, R.J. and Gratzer, W.B. (1986) Control of malarial invasion by phosphorylation of the host cell membrane cytoskeleton. *Nature*, **324**, 364-365.

Rawlings, N.D. and Barrett, A.J. (1993) Evolutionary families of peptidases. *Biochem J*, **290**, 205-218.

Rawlings, N.D., Morton, F.R. and Barrett, A.J. (2006) MEROPS: the peptidase database. *Nucleic Acids Res*, **34**, D270-272.

Rayner, J.C., Vargas-Serrato, E., Huber, C.S., Galinski, M.R. and Barnwell, J.W. (2001) A *Plasmodium falciparum* homologue of *Plasmodium vivax* reticulocyte binding protein (PvRBP1) defines a trypsin-resistant erythrocyte invasion pathway. *J Exp Med*, **194**, 1571-1581.

Robson, K.J., Hall, J.R., Jennings, M.W., Harris, T.J., Marsh, K., Newbold, C.I., Tate, V.E. and Weatherall, D.J. (1988) A highly conserved amino-acid sequence in thrombospondin, properdin and in proteins from sporozoites and blood stages of a human malaria parasite. *Nature*, **335**, 79-82.

Roche Applied Science (1999) *D-Biotin-Nhydroxysuccinimide ester*. Roche Diagnostics Gmbh, Mannheim.

Roche Applied Science (2004) *High fidelity PCR master*. Roche Diagnostics GmbH, Penzberg.

Rodbard, D., Chrambach, A. and Weiss, G.H. (1974) Optimization of resolution in analytical and preparative polyacrylamide gel electrophoresis. In *Electrophoresis and Isolelectric Focusing in Polyacrylamide Gel*. (Eds, Allen, R.C. and Maurer, H.R.) Walter de Gruyter & Co., Berlin.

Roeffen, W.F., Raats, J.M., Teelen, K., Hoet, R.M., Eling, W.M., van Venrooij, W.J. and Sauerwein, R.W. (2001) Recombinant human antibodies specific for the Pfs48/45 protein of the malaria parasite *Plasmodium falciparum*. *J Biol Chem*, **276**, 19807-19811.

Roggwiller, E., Betoulle, M.E., Blisnick, T. and Braun Breton, C. (1996) A role for erythrocyte band 3 degradation by the parasite gp76 serine protease in the formation of the parasitophorous vacuole during invasion of erythrocytes by *Plasmodium falciparum*. *Mol Biochem Parasitol*, **82**, 13-24.

Roll Back Malaria, World Health Organization and UNICEF (2005) *World Malaria Report 2005*. World Health Organization, Switzerland.

Rosenberg, A., Griffin, K., Studier, F.W., McCormick, M., Berg, J., Novy, R. and Mierendorf, R. (1996) T7Select® Phage Display System: A powerful new protein display system based on bacteriophage T7. *inNovations*, **6**, 1-6.

Roth, E., Jr. (1990) *Plasmodium falciparum* carbohydrate metabolism: a connection between host cell and parasite. *Blood Cells*, **16**, 453-460.

Rybicki, A.C., Heath, R., Lubin, B. and Schwartz, R.S. (1988) Human erythrocyte protein 4.1 is a phosphatidylserine binding protein. *J Clin Invest*, **81**, 255-260.

Saliba, K.J., Horner, H.A. and Kirk, K. (1998) Transport and metabolism of the essential vitamin pantothenic acid in human erythrocytes infected with the malaria parasite *Plasmodium falciparum*. *J Biol Chem*, **273**, 10190-10195.

Salmon, B.L., Oksman, A. and Goldberg, D.E. (2001) Malaria parasite exit from the host erythrocyte: a two-step process requiring extraerythrocytic proteolysis. *Proc Natl Acad Sci U S A*, **98**, 271-276.

Sanders, P.R., Gilson, P.R., Cantin, G.T., Greenbaum, D.C., Nebl, T., Carucci, D.J., McConville, M.J., Schofield, L., Hodder, A.N., Yates, J.R. and Crabb, B.S. (2005) Distinct protein classes including novel merozoite surface antigens in raft-like membranes of *Plasmodium falciparum*. *J Biol Chem*, **280**, 40169-40176.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*, **74**, 5463-5467.

Schneider, E.L., King, D.S. and Marletta, M.A. (2005) Amino acid substitution and modification resulting from *Escherichia coli* expression of recombinant *Plasmodium falciparum* histidine-rich protein II. *Biochemistry*, **44**, 987-995.

Schulman, S., Roth, E.F., Jr., Cheng, B., Rybicki, A.C., Sussman, I., Wong, M., Wang, W., Ranney, H.M., Nagel, R.L. and Schwartz, R.S. (1990) Growth of *Plasmodium falciparum* in human erythrocytes containing abnormal membrane proteins. *Proc Natl Acad Sci U S A*, **87**, 7339-7343.

Schumann, W. and Ferreira, L.C.S. (2004) Production of recombinant proteins in *Escherichia coli*. *Genet Mol Biol*, **27**, 442-453.

Schürer, G., Lanig, H. and Clark, T. (2004) *Aeromonas proteolytica* aminopeptidase: an investigation of the mode of action using a quantum mechanical/molecular mechanical approach. *Biochemistry*, **43**, 5414-5427.

Shenai, B.R., Sijwali, P.S., Singh, A. and Rosenthal, P.J. (2000) Characterization of native and recombinant falcipain-2, a principal trophozoite cysteine protease and essential hemoglobinase of *Plasmodium falciparum*. *J Biol Chem*, **275**, 29000-29010.

Sherman, I.W. (1977) Amino acid metabolism and protein synthesis in malarial parasites. *Bull World Health Organ*, **55**, 265-276.

Sherman, I.W. and Jones, L.A. (1979) *Plasmodium lophurae*: membrane proteins of erythrocyte-free plasmodia and malaria-infected red cells. *J Protozool*, **26**, 489-501.

Sidhu, S.S., Fairbrother, W.J. and Deshayes, K. (2003) Exploring protein-protein interactions with phage display. *ChemBioChem*, **4**, 14-25.

Siegel, D.L. and Branton, D. (1985) Partial purification and characterization of an actin-bundling protein, band 4.9, from human erythrocytes. *J Cell Biol*, **100**, 775-785.

Sigma-Aldrich (2005) *Nondenaturesd protein molecuar weight kit*. Sigma-Aldrich, Inc., Saint Louis.

Sijwali, P.S. and Rosenthal, P.J. (2004) Gene disruption confirms a critical role for the cysteine protease falcipain-2 in hemoglobin hydrolysis by Plas*modium falciparum*. *Proc Natl Acad Sci U S A*, **101**, 4384-4389.

Sijwali, P.S., Shenai, B.R., Gut, J., Singh, A. and Rosenthal, P.J. (2001) Expression and characterization of the *Plasmodium falciparum* haemoglobinase falcipain-3. *Biochem J*, **360**, 481-489.

Sim, B.K., Chitnis, C.E., Wasniowska, K., Hadley, T.J. and Miller, L.H. (1994) Receptor and ligand domains for invasion of erythrocytes by *Plasmodium falciparum*. *Science*, **264**, 1941-1944.

Sinden, R.E. and Gilles, H.M. (2002) The malaria parasites. In *Essential Malariology*. 4th ed. (Eds, Warrell, D.A. and Gilles, H.M.) Arnold Publishers, London.

Singh, G.P., Chandra, B.R., Bhattacharya, A., Akhouri, R.R., Singh, S.K. and Sharma, A. (2004) Hyper-expansion of asparagines correlates with an abundance of proteins with prion-like domains in *Plasmodium falciparum*. *Mol Biochem Parasitol*, **137**, 307-319.

Smith, G.P. (1985) Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science*, **228**, 1315-1317.

Smith, G.P. and Petrenko, V.A. (1997) Phage display. *Chem Rev*, **97**, 391-410.

Smythe, J.A., Coppel, R.L., Brown, G.V., Ramasamy, R., Kemp, D.J. and Anders, R.F. (1988) Identification of two integral membrane proteins of *Plasmodium falciparum*. *Proc Natl Acad Sci U S A*, **85**, 5195-5199.

Snow, R.W., Guerra, C.A., Noor, A.M., Myint, H.Y. and Hay, S.I. (2005) The global distribution of clinical episodes of *Plasmodium falciparum* malaria. *Nature*, **434**, 214-217.

Sorensen, H.P. and Mortensen, K.K. (2005a) Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *J Biotechnol*, **115**, 113-128.

Sorensen, H.P. and Mortensen, K.K. (2005b) Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*. *Microb Cell Fact*, **4**, 1.

Sowa, K.M., Cavanagh, D.R., Creasey, A.M., Raats, J., McBride, J., Sauerwein, R., Roeffen, W.F. and Arnot, D.E. (2001) Isolation of a monoclonal antibody from a malaria patient-derived phage display library recognising the Block 2 region of *Plasmodium falciparum* merozoite surface protein-1. *Mol Biochem Parasitol*, **112**, 143-147.

Speicher, D.W. and Marchesi, V.T. (1984) Erythrocyte spectrin is comprised of many homologous triple helical segments. *Nature*, **311**, 177-180.

Speicher, D.W., Weglarz, L. and DeSilva, T.M. (1992) Properties of human red cell spectrin heterodimer (side-to-side) assembly and identification of an essential nucleation site. *J Biol Chem*, **267**, 14775-14782.

Stamper, C., Bennett, B., Edwards, T., Holz, R.C., Ringe, D. and Petsko, G. (2001) Inhibition of the aminopeptidase from *Aeromonas proteolytica* by L-leucinephosphonic acid. Spectroscopic and crystallographic characterization of the transition state of peptide hydrolysis. *Biochemistry*, **40**, 7035-7046.

Sternberg, N. and Hoess, R.H. (1995) Display of peptides and proteins on the surface of bacteriophage lambda. *Proc Natl Acad Sci U S A*, **92**, 1609-1613.

Stewart, M.J., Schulman, S. and Vanderberg, J.P. (1986) Rhoptry secretion of membranous whorls by *Plasmodium falciparum* merozoites. *Am J Trop Med Hyg*, **35**, 37-44.

Stratagene (2005) *BL21-CodonPlus® Competent Cells Instruction Manual*. Stratagene, La Jolla.

Swanson, D. (1999) The ties that bind: peptide display technology. *The Scientist*, **13**, 22.

Taraschi, T.F., Trelka, D., Martinez, S., Schneider, T. and O'Donnell, M.E. (2001) Vesicle-mediated trafficking of parasite proteins to the host cell cytosol and erythrocyte surface membrane in *Plasmodium falciparum* infected erythrocytes. *Int J Parasitol*, **31**, 1381-1391.

Tasseron-de Jong, J.G., Aker, J. and Giphart-Gassler, M. (1988) The ability of the restriction endonuclease *Eco*RI to digest hemi-methylated versus fully cytosine-methylated DNA of the herpes tk promoter region. *Gene*, **74**, 147-149.

Teuscher, F., Lowther, J., Skinner-Adams, T.S., Spielmann, T., Dixon, M.W., Stack, C.M., Donnelly, S., Mucha, A., Kafarski, P., Vassiliou, S., Gardiner, D.L., Dalton, J.P. and Trenholme, K.R. (2007) The M18 aspartyl aminopeptidase of the human malaria parasite *Plasmodium falciparum*. *J Biol Chem*, **282**, 30817-30826.

Tindall, K.R. and Kunkel, T.A. (1988) Fidelity of DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Biochemistry*, **27**, 6008-6013.

Tisdale, E.J. (2002) Glyceraldehyde-3-phosphate dehydrogenase is phosphorylated by protein kinase Ciota /lambda and plays a role in microtubule dynamics in the early secretory pathway. *J Biol Chem*, **277**, 3334-3341.

Topolska, A.E., Wang, L., Black, C.G. and Coppel, R.L. (2004) Merozoite biology. In *Malaria Parasites: Genomes and Molecular Biology*. (Eds, Waters, A.P. and Janse, C.J.) Caister Academic Press, Wymondham.

Towbin, H., Staehelin, T. and Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci U S A*, **76**, 4350-4354.

Trager, W. (1956) The intracellular position of malarial parasites. *Trans R Soc Trop Med Hyg*, **50**, 419-420.

Trager, W. and Jensen, J.B. (1976) Human malaria parasites in continuous culture. *Science*, **193**, 673-675.

Trepod, C.M. and Mott, J.E. (2002) A spontaneous runaway vector for production-scale expression of bovine somatotropin from *Escherichia coli*. *Appl Microbiol Biotechnol*, **58**, 84-88.

Trucco, C., Fernandez-Reyes, D., Howell, S., Stafford, W.H., Scott-Finnigan, T.J., Grainger, M., Ogun, S.A., Taylor, W.R. and Holder, A.A. (2001) The merozoite surface protein 6 gene codes for a 36 kDa protein associated with the *Plasmodium falciparum* merozoite surface protein-1 complex. *Mol Biochem Parasitol*, **112**, 91-101.

Tse, W.T., Lecomte, M.C., Costa, F.F., Garbarz, M., Feo, C., Boivin, P., Dhermy, D. and Forget, B.G. (1990) Point mutation in the beta-spectrin gene associated with alpha I/74 hereditary elliptocytosis. Implications for the mechanism of spectrin dimer self-association. *J Clin Invest*, **86**, 909-916.

Turgut-Balik, D., Shoemark, D.K., Moreton, K.M., Sessions, R.B. and Holbrook, J.J. (2001) Over-production of lactate dehydrogenase from *Plasmodium falciparum* opens a route to new antimalarials. *Biotechnol Lett*, **23**, 917–921.

Tuteja, R. (2007) Malaria - an overview. *Febs J*, **274**, 4670-4679.

Ungewickell, E., Bennett, P.M., Calvert, R., Ohanian, V. and Gratzer, W.B. (1979) *In vitro* formation of a complex between cytoskeletal proteins of the human erythrocyte. *Nature*, **280**, 811-814.

Ungewickell, E. and Gratzer, W. (1978) Self-association of human spectrin. A thermodynamic and kinetic study. *Eur J Biochem*, **88**, 379-385.

Ursitti, J.A., Kotula, L., DeSilva, T.M., Curtis, P.J. and Speicher, D.W. (1996) Mapping the human erythrocyte beta-spectrin dimer initiation site using recombinant peptides and correlation of its phasing with the alpha-actinin dimer site. *J Biol Chem*, **271**, 6636-6644.

van Regenmortel, M.H.V., Fauquet, C.M., Bishop, D.H.L., Carstens, E.B., Estes, M.K., Lemon, S.M., Maniloff, J., Mayo, M.A., McGeoch, D.J., Pringle, C.R. and Wickner, R.B. (Eds) (2000) *Virus Taxonomy: Classification and Nomenclature of Viruses*. Academic Press, San Diego.

Van Wye, J., Ghori, N., Webster, P., Mitschler, R.R., Elmendorf, H.G. and Haldar, K. (1996) Identification and localization of rab6, separation of rab6 from ERD2 and implications for an 'unstacked' Golgi, in *Plasmodium falciparum*. *Mol Biochem Parasitol*, **83**, 107-120.

Vedadi, M., Lew, J., Artz, J., Amani, M., Zhao, Y., Dong, A., Wasney, G.A., Gao, M., Hills, T., Brokx, S., Qiu, W., Sharma, S., Diassiti, A., Alam, Z., Melone, M., Mulichak, A., Wernimont, A., Bray, J., Loppnau, P., Plotnikova, O., Newberry, K., Sundararajan, E., Houston, S., Walker, J., Tempel, W., Bochkarev, A., Kozieradzki, I., Edwards, A., Arrowsmith, C., Roos, D., Kain, K. and Hui, R. (2007) Genome-scale protein expression and structural biology of *Plasmodium falciparum* and related Apicomplexan organisms. *Mol Biochem Parasitol*, **151**, 100-110.

Walder, J.A., Chatterjee, R., Steck, T.L., Low, P.S., Musso, G.F., Kaiser, E.T., Rogers, P.H. and Arnone, A. (1984) The interaction of hemoglobin with the cytoplasmic domain of band 3 of the human erythrocyte membrane. *J Biol Chem*, **259**, 10238-10246.

Walensky, L.D., Mohandas, N. and Lux, S.E. (2003) Disorders of the red blood cell membrane. In *Blood: Principles and Practice of Hematology*. 2nd ed. (Eds, Handin, R.I., Lux, S.E. and Stossel, T.P.) Lippincott Williams & Wilkins, Philadelphia.

Waller, K.L., Cooke, B.M., Nunomura, W., Mohandas, N. and Coppel, R.L. (1999) Mapping the binding domains involved in the interaction between the *Plasmodium falciparum* knob-associated histidine-rich protein (KAHRP) and the cytoadherence ligand *P. falciparum* erythrocyte membrane protein 1 (PfEMP1). *J Biol Chem*, **274**, 23808-23813.

Waller, K.L., Nunomura, W., An, X., Cooke, B.M., Mohandas, N. and Coppel, R.L. (2003) Mature parasite-infected erythrocyte surface antigen (MESA) of *Plasmodium falciparum* binds to the 30-kDa domain of protein 4.1 in malaria-infected red blood cells. *Blood*, **102**, 1911-1914.

Wang, L.F. and Yu, M. (2004) Epitope identification and discovery using phage display libraries: applications in vaccine development and diagnostics. *Curr Drug Targets*, **5**, 1-15.

Weatherall, D.J. (2008) Genetic variation and susceptibility to infection: the red cell and malaria. *Br J Haematol*, **141**, 276-286.

Webb, S.E., Fowler, R.E., O'Shaughnessy, C., Pinder, J.C., Dluzewski, A.R., Gratzer, W.B., Bannister, L.H. and Mitchell, G.H. (1996) Contractile protein system in the asexual stages of the malaria parasite *Plasmodium falciparum*. *Parasitology*, **112**, 451-457.

Weickert, M.J., Doherty, D.H., Best, E.A. and Olins, P.O. (1996) Optimization of heterologous protein production in E*scherichia coli*. *Curr Opin Biotechnol*, **7**, 494-499.

Weidekamm, E., Wallach, D.F., Lin, P.S. and Hendricks, J. (1973) Erythrocyte membrane alterations due to infection with *Plasmodium berghei*. *Biochim Biophys Acta*, **323**, 539-546.

Westermeier, R. (1997) *Electrophoresis in Practice: A Guide to Methods and Applications of DNA and Protein Separations*. 2nd ed. VCH Verlagsgesellschaft mbH, Weinheim.

World Health Organisation (2006) Malaria. http://www.who.int/topics/malaria/en/.

Wickert, H., Gottler, W., Krohne, G. and Lanzer, M. (2004) Maurer's cleft organization in the cytoplasm of *Plasmodium falciparum*-infected erythrocytes: new insights from three-dimensional reconstruction of serial ultrathin sections. *Eur J Cell Biol*, **83**, 567-582.

Wickham, M.E., Culvenor, J.G. and Cowman, A.F. (2003) Selective inhibition of a two-step egress of malaria parasites from the host erythrocyte. *J Biol Chem*, **278**, 37658-37663.

Wilk, S., Wilk, E. and Magnusson, R.P. (1998) Purification, characterization, and cloning of a cytosolic aspartyl aminopeptidase. *J Biol Chem*, **273**, 15961-15970.

Wilk, S., Wilk, E. and Magnusson, R.P. (2002) Identification of histidine residues important in the catalysis and structure of aspartyl aminopeptidase. *Arch Biochem Biophys*, **407**, 176-183.

Williams, T.N. (2006) Red blood cell defects and malaria. *Mol Biochem Parasitol*, **149**, 121-127.

Winograd, E. and Sherman, I.W. (2004) Malaria infection induces a conformational change in erythrocyte band 3 protein. *Mol Biochem Parasitol*, **138**, 83-87.

Wu, X., Jornvall, H., Berndt, K.D. and Oppermann, U. (2004) Codon optimization reveals critical factors for high level expression of two rare codon genes in *Escherichia coli*: RNA stability and secondary structure but not tRNA abundance. *Biochem Biophys Res Commun*, **313**, 89-96.

Wu, Y., Wang, X., Liu, X. and Wang, Y. (2003) Data-mining approaches reveal hidden families of proteases in the genome of malaria parasite. *Genome Res*, **13**, 601-616.

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**, 159-162.

Yan, Y., Winograd, E., Viel, A., Cronin, T., Harrison, S.C. and Branton, D. (1993) Crystal structure of the repetitive segments of spectrin. *Science*, **262**, 2027-2030.

Yayon, A., Cabantchik, Z.I. and Ginsburg, H. (1984) Identification of the acidic compartment of *Plasmodium falciparum*-infected human erythrocytes as the target of the antimalarial drug chloroquine. *Embo J*, **3**, 2695-2700.

Yokoyama, R., Kawasaki, H. and Hirano, H. (2006) Identification of yeast aspartyl aminopeptidase gene by purifying and characterizing its product from yeast cells. *Febs J*, **273**, 192-198.

Yokoyama, S. (2003) Protein expression systems for structural genomics and proteomics. *Curr Opin Chem Biol*, **7**, 39-43.

Yu, J., Fischman, D.A. and Steck, T.L. (1973) Selective solubilization of proteins and phospholipids from red blood cell membranes by nonionic detergents. *J Supramol Struct*, **1**, 233-248.

Zail, S.S. and Coetzer, T.L. (1984) Defective binding of spectrin to ankyrin in a kindred with recessively inherited hereditary elliptocytosis. *J Clin Invest*, **74**, 753-762.

Zhang, K. and Rathod, P.K. (2002) Divergent regulation of dihydrofolate reductase between malaria parasite and human host. *Science*, **296**, 545-547.

Zheng, L., Baumann, U. and Reymond, J.L. (2003) Production of a functional catalytic antibody ScFv-NusA fusion protein in bacterial cytoplasm. *J Biochem (Tokyo)*, **133**, 577-581.

Ziemnicka-Kotula, D., Xu, J., Gu, H., Potempska, A., Kim, K.S., Jenkins, E.C., Trenkner, E. and Kotula, L. (1998) Identification of a candidate human spectrin Src homology 3 domain-binding protein suggests a general mechanism of association of tyrosine kinases with the spectrin-based membrane skeleton. *J Biol Chem*, **273**, 13681-13692.