THE UNIVERSITY OF THE WITWATERSRAND

MASTERS DISSERTATION

An investigation of *Plasmodium falciparum* SORTILIN in trafficking of invasion proteins of the human malaria parasite

---

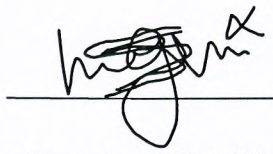
Serena SHUNMUGAM

A dissertation submitted to the Faculty of Health Sciences, The University of the Witwatersrand, in fulfilment of the requirements for the degree of Master of Science in Molecular Medicine and Haematology.

Johannesburg, 2018

Declaration of Authorship

I, Serena Shunmugam declare that this dissertation is my own work. It is being submitted for the degree of Master of Science in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at this or any other university.



(Signature of Candidate)

6th day of June 2018 in Grenoble, France

Ethics Clearance

Ethics clearance was obtained for this project from The University of the Witwatersrand, Committee for Research on Human Subjects (medical).

Ethics number M171088 (Sub-Study of M130569).

The ethics certificate is supplied in Appendix F.

Research Outputs

Manuscript in preparation

Title: Trafficking of invasion proteins to the apical organelles of *Plasmodium falciparum* (subject to change)

Journal: Parasites and Vectors (impact factor: 3.234) or PLOS one (Impact factor 4.411)

Authors: Alisje Churchyard, Serena Shunmugam, Melanie Wepener and Thérèsa L. Coetzer

Presentations arising from this study

This work was presented at three different conferences and entitled 'The trafficking of MAEBL in the human malaria parasite, *P. falciparum*' by Serena Shunmugam and Thérèsa L. Coetzer:

Poster presentation: International Keystone Symposium: Vectors, Pathogens and Diseases: Current Trends and Emerging Challenges (T1), 10-14 September 2017, Durban, South Africa

Oral presentation: 3rd Southern Africa Malaria Research Conference, 7 – 9 November, 2017, Johannesburg, South Africa

Oral and poster presentation: Joint International Tropical Medicine Meeting, 6-8 December 2017, Bangkok, Thailand

Abstract

Malaria is arguably one of the most overwhelming infectious diseases throughout the world's existence. The most virulent parasite, *Plasmodium falciparum*, has a redundancy of invasion proteins, allowing it to switch between different receptors on the host red blood cell. These invasion proteins are stored in the apical organelles, the rhoptries and micronemes, but very little is known about how newly synthesized proteins are transported to these organelles. The hypothesis in this study was that a common protein is involved in trafficking invasion proteins from the trans-Golgi network and *Pf*SORTILIN was investigated as a potential escorter protein. The CCys domain of *Pf*MAEBL, a rhoptry protein, and the prodomain of *Pf*AMA-1, a microneme protein, have been implicated in trafficking to the apical organelles. These domains and the VPS10 domain of *Pf*SORTILIN were cloned into expression vectors encoding a GST- or His-tag. Recombinant proteins were expressed in *E. coli* and purified by affinity chromatography on glutathione- or Ni-particles. *In vitro* binding assays were performed, which showed that *Pf*SORTILIN VPS10 bound to *Pf*MAEBL CCys but not to the *Pf*AMA-1 prodomain, suggesting that *Pf*SORTILIN is a rhoptry protein escorter and is not involved in microneme trafficking. To identify novel binding partners of *Pf*SORTILIN VPS10, the protein was biopanned against a *P. falciparum* phage display library. No binding partners were identified, most likely because the library is not schizont-stage specific, which is when *Pf*SORTILIN and invasion proteins are predominantly expressed. The results from this study were integrated with other studies and a trafficking model for *Pf*MAEBL was proposed. This study enhances our knowledge of trafficking pathways and suggests that *Pf*SORTILIN may serve as a common rhoptry protein escorter.

Acknowledgements

I would like to thank the National Research Foundation for a scarce skills bursary and the University of the Witwatersrand's Medical Faculty Research Endowment Fund for providing project funding. A very special thank you to my supervisor, Prof T Coetzer whose patience, dedication and investment in me has made me a better scientist. Without the support of the entire PMRU, past and present, this would never have been possible. I will be forever grateful to Dr K Naidoo, Dr S Lauterbach, Dr B Bezuidenhout, Dr A Churchyard and Dr D Liebenberg. To Mom, Dad, Aunty Viloshnee and Mamma; thank you for always listening. Last but not least, I would like to thank Dr B Brocco for his genuine interest, support and insightful discussions, which have made me excited about the research environment and science in general.

Contents

Declaration of Authorship i

Ethics Clearance ii

Research Outputs iii

Abstract iv

Acknowledgements v

Contents vi

List of Figures xii

List of Tables xvi

Nomenclature xviii

1 Introduction 1

1.1 The Malaria epidemic 1

1.1.1 The life cycle of *P. falciparum* parasites 2

1.2 Erythrocyte invasion 3

1.3 Invasion proteins 6

1.3.1 *Pf*MAEBL 7

1.3.2 *Pf*AMA-1 9

1.4 Eukaryotic protein trafficking 10

1.4.1 Clathrin and adaptor protein complexes in eukaryotes 11
1.4.2 Protein sorting ability of sortilin in eukaryotes 13
1.4.3 Apicomplexan protein trafficking and sorting 13
1.4.4 Trafficking signals in *T. gondii* 14
1.4.5 Trafficking escorters in *T. gondii* 14
1.4.6 Protein trafficking in *P. falciparum* parasites 15
1.4.6.1 Rhoptry targeting 16
1.4.6.2 Microneme targeting 16
1.4.6.3 Trafficking escorters 18
1.4.6.4 *Pf*SORTILIN 18
1.5 Purpose of the study 19
1.5.1 Study aim 19
1.5.2 Study objectives 19

2 Methods and Materials **21**

2.1 Culturing techniques 21
2.1.1 Parasite culture composition and maintenance 21
2.1.2 Preparation of erythrocytes 22
2.1.3 Thawing frozen parasite cultures 22
2.1.4 Freezing parasite cultures 23
2.1.5 Smear Preparation 23
2.1.6 Synchronisation of cultures 24
2.2 *Pf*DNA preparation and analysis 24
2.2.1 Genomic DNA extraction from *P. falciparum* 24
2.2.2 Determination of DNA concentration, purity and integrity 25
2.3 Polymerase chain reactions (PCR) 26
2.3.1 Primer specifications and parameters 26
2.3.2 PCR conditions 27

2.3.3 Restriction enzyme digestion of PCR products 28
2.4 Cloning 29
2.4.1 Extraction of Plasmid DNA from DH5α cells 29
2.4.2 Restriction enzyme digestion of vector plasmids 30
2.4.3 Ligation of DNA inserts and plasmids 31
2.4.4 Transformations 31
2.4.4.1 DH5α cells 31
2.4.4.2 Rosetta 2™ (DE3) cells 33
2.5 Recombinant protein expression 34
2.5.1 IPTG induction 34
2.5.2 Induction using Overnight Express™ Instant TB Medium 35
2.5.3 Recombinant protein extraction 36
2.5.4 SDS-PAGE analysis and Western blot 37
2.5.5 Immunoblot analysis 38
2.6 Recombinant protein purification 39
2.6.1 Purification of His-tagged proteins 40
2.6.2 Purification of GST-tagged proteins 41
2.7 *In vitro* binding assays 41
2.7.1 Dialysis of protein samples 41
2.7.2 Binding assay of immobilized *Pf*MAEBL CCys or *Pf*AMA-1 prodomain to *Pf*SORTILIN VPS10 42
2.7.3 Binding assay of immobilized *Pf*SORTILIN VPS10 to *Pf*MAEBL CCys or *Pf*AMA-1 prodomain 43
2.8 Biopanning using phage display libraries 44
2.8.1 Bacterial cultures for biopanning 44
2.8.2 Phage amplification and plaque assays 44
2.8.3 Preparation of *Pf*SORTILIN VPS10 for biopanning 45
2.8.4 Biopanning 45

2.8.5 BLT5403 and bacteriophage glycerol stocks 46
2.8.6 PCR amplification and sequencing of selected phage 47

3 Results **48**

3.1 Identification of protein domains 48
3.1.1 *Pf*MAEBL CCys domain (PF3D7_1147800) 48
3.1.2 *Pf*SORTILIN VPS10 domain (PF3D7_1451800) 49
3.2 Genomic DNA extraction from *P. falciparum* 54
3.3 Cloning 55
3.3.1 PCR amplification 55
3.3.2 Ligation reactions 56
3.3.3 Recombinant plasmid construct verification 57
3.4 Recombinant protein expression 58
3.4.1 *Pf*MAEBL CCys domain 58
3.4.2 *Pf*SORTILIN VPS10 domain 60
3.4.3 Final recombinant protein expression conditions 63
3.5 Recombinant protein purification 64
3.5.1 *Pf*MAEBL CCys domain 64
3.5.2 *Pf*AMA-1 prodomain 65
3.5.3 *Pf*SORTILIN VPS10 domain 66
3.6 *In vitro* binding assays 68
3.6.1 Quantitation by densitometry 69
3.6.2 Immobilized *Pf*MAEBL CCys vs *Pf*SORTILIN VPS10 70
3.6.3 Immobilized *Pf*SORTILIN VPS10 vs *Pf*MAEBL CCys 72
3.6.4 Immobilized *Pf*AMA-1 prodomain vs *Pf*SORTILIN VPS10 74
3.6.5 Immobilized *Pf*SORTILIN VPS10 vs *Pf*AMA-1 prodomain 76
3.7 Biopanning *Pf*SORTILIN VPS10 using phage display library 77

4 Discussion **80**

4.1 Protein domain selection 81
4.2 Challenges with recombinant protein expression 81
4.2.1 Expressing eukaryotic proteins in a prokaryotic system 82
4.2.2 Expression levels and solubility 82
4.3 Unexpected migration of recombinant proteins on SDS-PAGE 85
4.4 *In vitro* binding assays determine interactions of *Pf*SORTILIN VPS10 86
4.4.1 *Pf*SORTILIN VPS10 interacts with *Pf*MAEBL CCys 87
4.4.2 *Pf*SORTILIN VPS10 does not interact with *Pf*AMA-1 prodomain . . . 87
4.5 Comparison between the *Pf*MAEBL CCys domain and *Pf*AMA-1 prodomain . 88
4.6 Utility of the phage display library 91
4.7 *Pf*SORTILIN as a rhoptry protein escorter 92
4.7.1 Escorter Model 92
4.7.1.1 Rhoptry escorter model 92
4.7.1.2 Microneme escorter model 93
4.8 Future work 95

A Primers, DNA and protein Sequences **97**
A.1 PCR primers 97
A.2 Primers used for construct sequencing 98
A.3 Sizes of vectors and *P. falciparum* inserts 98
A.4 Protein sequences 99
A.4.1 *Pf*MAEBL CCys 99
A.4.2 *Pf*AMA-1 prodomain 99
A.4.3 *Pf*SORTILIN VPS10 100

B Commercial chemical reagents **106**

C Equipment and consumables **110**

D *P. falciparum* phage display library and PCR amplification of phage isolates **112**

E Vector maps **114**

F Ethics clearance certificate **116**

G Turn-it-in Plagiarism report **117**

References **118**

List of Figures

1.1 The life cycle of *P. falciparum* parasite 3

1.2 Diagram of a *P. falciparum* merozoite 4

1.3 Merozoite invasion of red blood cells 5

1.4 Schematic representation of the DBL-EBP family of proteins in *P. falciparum* . 7

1.5 Schematic representation of *Pf*MAEBL 8

1.6 Schematic representation of *Pf*AMA-1 9

1.7 Vesicular transport from the ER to the Golgi in eukaryotic organisms 11

1.8 Schematic representation of the interaction between clathrin and AP1 12

1.9 A model for *Tg*SORTLR functions in protein sorting and the biogenesis of apical secretory organelles 15

1.10 Schematic representation of the regions needed for rhoptry or microneme trafficking 17

1.11 Schematic representation of *Pf*SORTILIN 18

3.1 Schematic representation of *Pf*MAEBL 49

3.2 MUSCLE alignment of VPS10 domain of *Pf*SORTILIN and VPS10 domain of *Tg*SORTLR 51

3.3 Schematic representation of *Pf*SORTILIN 52

3.4 3-D models of the tertiary structure of *Tg*SORTLR and *Pf*SORTILIN 53

3.5 Agarose gel (1%) electrophoresis of 3D7 *P. falciparum* DNA 54

3.6 Agarose gel (1%) electrophoresis of PCR amplicons 55

3.7 Agarose gel electrophoresis of digested plasmids, *Pf*MAEBL CCys and *Pf*SORTILIN VPS10 inserts 56
3.8 Colony PCR verification of transformed DH5α cells 57
3.9 Restriction enzyme digestion of plasmids isolated from clones positive for colony PCR 58
3.10 SDS-PAGE (17% acrylamide) and anti-His immunoblot analyses of the expression and solubility of *Pf*MAEBL CCys using IPTG induction at varying temperatures 59
3.11 SDS-PAGE (17% acrylamide) and immunoblot analyses of the expression and solubility of *Pf*MAEBL CCys using IPTG induction at varying incubation periods 60
3.12 SDS-PAGE (15% acrylamide) and anti-GST immunoblot analyses of *Pf*SORTILIN VPS10 using Overnight Express™ or IPTG induction 61
3.13 SDS-PAGE (15% acrylamide) and anti-GST immunoblot analyses of the expression and solubility of *Pf*SORTILIN VPS10 using IPTG induction at varying temperatures 62
3.14 SDS-PAGE (15% acrylamide) and anti-GST immunoblot analyses of *Pf*SORTILIN VPS10 using IPTG induction at varying incubation periods 62
3.15 Low complexity regions of the *Pf*SORTILIN VPS10 protein 63
3.16 Purification of His-tagged *Pf*MAEBL CCys protein 65
3.17 Purificatiion of His-tagged *Pf*AMA-1 prodomain protein 66
3.18 Purification of *Pf*SORTILIN VPS10 protein 68
3.19 SDS-PAGE (15% acrylamide) of proteins used for *in vitro* binding assays of immoblized *Pf*AMA-1 prodomain and *Pf*SORTILIN VPS10 69
3.20 Standard curve of BSA (ng) versus peak volume 70
3.21 SDS-PAGE (15% acrylamide) analysis of *in vitro* binding assays of His-tagged *Pf*MAEBL CCys and GST-tagged *Pf*SORTILIN VPS10 71
3.22 Graph showing the dose-dependent binding of immobilised His-tagged *Pf*MAEBL CCys to free GST-tagged *Pf*SORTILIN VPS10 72

3.23 Graph showing the binding of immobilised GST-tagged *Pf*SORTILIN VPS10 to free His-tagged *Pf*MAEBL CCys 74
3.24 SDS-PAGE (15% acrylamide) analysis of *in vitro* binding assay of *Pf*AMA-1 prodomain vs *Pf*SORTILIN VPS10 75
3.25 Graph showing the dose-dependent binding of immobilised His-tagged *Pf*AMA-1 prodomain to free GST-tagged *Pf*SORTILIN VPS10 76
3.26 SDS-PAGE (15% acrylamide) of *in vitro* binding assays of immobilized *Pf*SORTILIN VPS10 vs *Pf*AMA-1 prodomain 77
3.27 A schematic depicting biopanning of *Pf*SORTILIN VPS10 against a *P. falciparum* phage display library 78
3.28 A segment of a titering plate after four rounds of biopanning 78
3.29 PCR of the cDNA inserts present in the phage isolated after four rounds of biopanning against *Pf*SORTILIN VPS10 79
3.30 DNA sequence of one phage insert which bound to the *Pf*SORTILIN VPS10 domain 79

4.1 Predicted solubility profiles of recombinant His-tagged *Pf*MAEBL CCys and GST-tagged *Pf*SORTILIN VPS10 83
4.2 Lys residues in GST-tagged *Pf*SORTILIN VPS10 86
4.3 ClustalW alignment of *Pf*MAEBL CCys and *Pf*AMA-1 prodomain 88
4.4 3D surface model of *Pf*MAEBL and crystal structure of *Pf*AMA-1 90
4.5 3D surface model of *Pf*SORTILIN VPS10 showing positive and negative residues. 91
4.6 Model of *Pf*SORTILIN VPS10 as a rhoptry protein escorter 94

A.1 Protein sequence of recombinant *Pf*MAEBL CCys 99
A.2 Protein sequence of recombinant *Pf*AMA-1 prodomain 99
A.3 Protein sequence of recombinant *Pf*SORTILIN VPS10 100
A.4 The chromatogram of the pET-15-b-*Pf*MAEBL CCys construct from colony 1 . 101
A.5 Alignment of *Pf*MAEBL CCys construct sequence from 5' sequencing 102

A.6 Forward sequence of *Pf*SORTILIN VPS10 vector construct 103

A.7 Reverse sequence of *Pf*SORTILIN VPS10 vector construct chromatogram . . . 104

A.8 Alignment of *Pf*SORTILIN VPS10 sequence from 5'to 3' sequencing 105

D.1 Schematic representation of the T7Select10-3b vector which was used to generate the *P. falciparum* phage display library from mRNA. 112

E.1 The pGEX-4T-2 plasmid 114

E.2 The pET-15b plasmid 115

List of Tables

2.1 Reaction mixture for PCRs of *Pf*MAEBL CCys and *Pf*SORTILIN VPS10 . . . 27

2.2 Optimized PCR conditions for each insert 28

2.3 Reaction mixtures for PCR product restriction digests using FastDigest™ enzymes 29

2.4 Reaction mixture for plasmid restriction digests using FastDigest™ enzymes . . 30

2.5 Components of SDS-polyacrylamide gels 38

3.1 Absorbance values and concentration of genomic DNA of 3D7 parasites 54

3.2 Ligation reaction insert:plasmid (3:1) molar ratios 57

3.3 Final growth conditions for expression of soluble recombinant proteins 64

3.4 Final yield and concentration of proteins for binding assay of *Pf*AMA-1 prodomain and *Pf*SORTILIN VPS10 based on densitometry of SDS polyacrylamide gel (Figure 3.19). 70

3.5 Densitometry analysis of the *in vitro* binding assay between immobilized His-tagged *Pf*MAEBL CCys and free GST-tagged *Pf*SORTILIN VPS10 72

3.6 Densitometry analysis of the *in vitro* binding assay between immobilized GST-tagged *Pf*SORTILIN VPS10 and free His-tagged *Pf*MAEBL CCys 73

3.7 Densitometry analysis of the *in vitro* binding assay between immobilized His-tagged *Pf*AMA-1 prodomain and free GST-tagged *Pf*SORTILIN VPS10 . . . 75

4.1 Summary of properties of recombinant proteins generated in this study 83

4.2 Summary of molecular weights of recombinant proteins generated in this study 86

A.1 Sizes of vectors and *P. falciparum* inserts 98

D.1 Primers used for sequencing of bacteriophage that bound to *Pf*SORTILIN VPS10 113

D.2 PCR conditions for amplifying bacteriophage after biopanning 113

Nomenclature

2YT	2X Yeast extract and tryptone
aa	amino acid
AMA-1	apical membrane antigen-1
AP	adaptor protein
ARM	Armidillo Repeats-Only
BAEBL	erythrocyte binding antigen-140
BSA	bovine serum albumin
CCV	clathrin-coated vesicle
CCys	C-terminal cysteine rich region
CD	cytoplasmic domain
COP	coat protein
CT	cytoplamic tail
DB	Duffy binding
DBL	Duffy binding-like protein family
DBP	Duffy binding protein
EBA-175	erythrocyte binding antigen-175
EBL	erythrocyte-binding ligand
EDTA	ethylenediaminetetraacetic acid
ER	endoplasmic reticulum
ESCRT	endosomal sorting complexes required for transport
GGA	Golgi localized, γ-ear-containing, ADP-ribosylation factor-binding proteins
GPI	glycosyl-phosphatidyl inositol

GSH	reduced glutathione
GST	glutathione S-transferase
HRP	horse radish peroxidase
IPTG	isopropyl-β-D-1-thiogalactopyranoside
JESEBL	erythrocyte binding antigne-181
LB	Luria Bertani
MAEBL	merozoite adhesive erythrocyte binding-like protein
MIC	microneme protein
MSP	merozoite surface protein
MTRAP	merozoite thrombospondin-related anonymous protein
MW	molecular weight
ORF	open reading frame
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PEBL	erythrocyte binding antigen-165
pfu	plaque forming unit
pI	isoelectric point
pRBCs	parasitized red blood cells
PTRAMP	thrombospondin-related apical membrane protein
PV	parasitophorous vacuole
OD	optical density
RAMA	rhoptry-associated membrane antigen
RAP1	rhoptry-associated protein 1
RBCs	red blood cells
RBL	reticulocyte-binding like protein
RhopH2	high molecular weight rhoptry protein 2
RO	reverse osmosis

RON	rhoptry neck protein
ROP2	rhoptry protein 2
SORTLR	sortilin-like receptor
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
SP	signal peptide
SUB1/2	subtilisin-like serine protease 1/2
T_a	annealing temperature
T_m	melting temperature
TBS	Tris buffered saline
TE	Tris EDTA
TM	transmembrane domain
TRAP	thrombospondin-related anonymous protein
Tris	Trisaminomethane
UV	ultra violet
VPS	vacoular protein sorting

Chapter 1

Introduction

1.1 The Malaria epidemic

Malaria is arguably one of the most overwhelming infectious diseases throughout the world's existence. It has been documented as early as 2700 BC (Tuteja, 2007). This disease is highly prevalent in poverty-ridden regions of the world such as sub-Saharan Africa, South Asia and parts of South America. According to the World Malaria Report (2017), nearly half of the world's population is at risk of being infected with malaria (WHO, 2017). Malaria is caused by five species of *Plasmodium* parasites namely *P. falciparum*, *P. vivax*, *P. ovale*, *P. malariae* and *P. knowlesi*. The female *Anopheles* mosquito is responsible for the transmission of the parasite between human hosts (Tuteja, 2007).

P. falciparum causes the most severe form of the disease and is the dominant species found in Africa. In rural communities in Africa where proper healthcare is limited the effects of malaria are most noticeable. More than 90% of malaria deaths occur in Africa and the number of global cases increased by 5 million compared to the previous year (WHO, 2017). The spread of drug resistant parasites, the lack of an effective vaccine, and the inconsistency of malaria control programmes have increased the need for new innovations in *Plasmodium* research (Phillips *et al.*, 2017).

1.1.1 The life cycle of *P. falciparum* parasites

The multistage life cycle of *P. falciparum* parasites can be separated into three distinct cycles: The sporogonic cycle (in the mosquito), the exo-erythrocytic cycle and the intra-erythrocytic developmental cycle (both in the human host), shown in Figure 1.1. A human plasmodial infection is initiated when a feeding Anopheles mosquito injects mature sporozoites of the parasite into the bloodstream of the host dermis. This is the beginning of the exo-erythrocytic stage. The sporozoites, transmitted via the saliva of the mosquito, make their way in the bloodstream to the liver of the new host (Silvie *et al.*, 2008).

The initial movement from the dermis to a blood vessel is random and not every sporozoite will reach the hepatocytes in the liver. Those that reach the liver will penetrate the hepatocytes where they remain for 9–16 days and undergo asexual replication known as exo-erythrocytic schizogony (Tuteja, 2007). During this stage a schizont forms within this hepatocyte that is made up of multiple haploid merozoites. The infected hepatocyte ruptures to release thousands of merozoites into the bloodstream to infect erythrocytes. Asexual division begins inside the erythrocyte, which is also the site for the formation of the other morphological stages of the parasites in the human host. After invasion merozoites develop into ring-stage parasites that will mature to form trophozoites (Phillips *et al.*, 2017).

At the end of the trophic phase, multiple rounds of nuclear division occur through asexual reproduction (erythrocytic schizogony). Schizonts are produced in the red blood cells (RBCs) and consist of up to 32 individual merozoites (Leete and Rubin, 1996). The infected RBCs eventually rupture, releasing the merozoites into the bloodstream where they invade new RBCs. The erythrocytic schizogony process is repeated every 48 h, (known as the intra-erythrocytic developmental cycle- Figure 1.1) and coincides with the clinical symptoms of malaria (Tuteja, 2007; Phillips *et al.*, 2017).

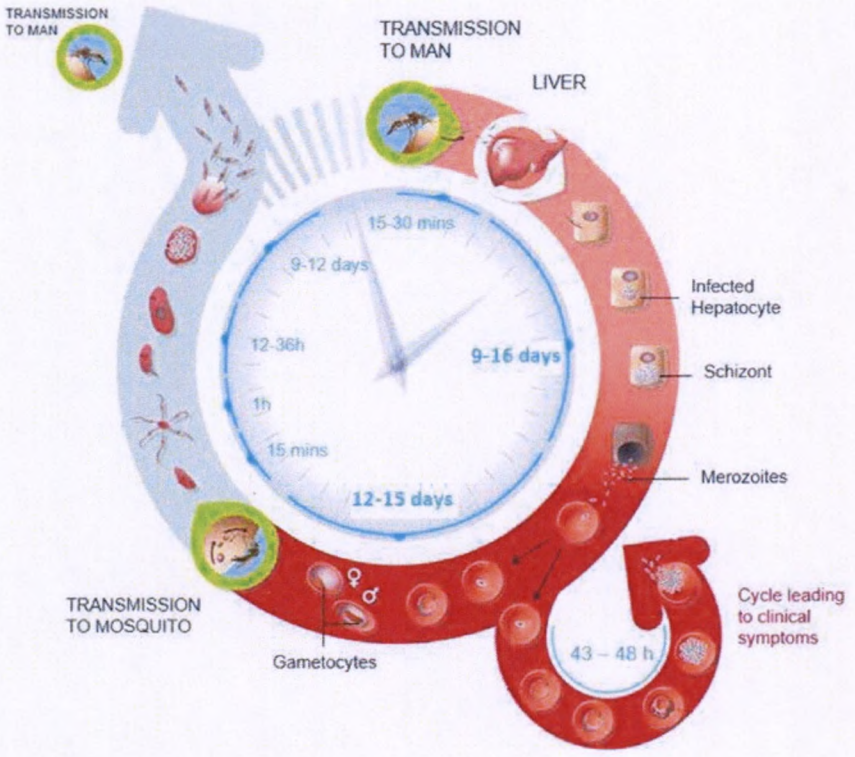


Figure 1.1: **The life cycle of *P. falciparum* parasite. The life cycle includes three separate stages namely the exo-erythrocytic stage inside the hepatocyte, the intra-erythrocytic developmental cycle and the sporogonic stage inside the mosquito vector. This image is from Medicines for Malaria Venture (2016).**

A few of the intra-erythrocytic parasites form male and female gametocytes. These are ingested by the mosquito where the sexual stage of the life cycle takes plce. The gametocytes develop into gametes which fuse to form a zygote within the insect's gut. The zygote develops into an oocyst, where sporozoites are formed and released. These enter the salivary glands, allowing the parasite to be spread among new human hosts when the mosquito feeds (Miller *et al.*, 2013).

1.2 Erythrocyte invasion

The malaria parasite amongst other protozoans, such as *Toxoplasma*, belongs to the phylum Apicomplexa, defined by the presence of specialised secretory organelles called the apical complex (Cowman *et al.*, 2012). These organelles (Figure 1.2) include the rhoptries, micronemes

and dense granules, and secrete crucial invasion proteins, which use sialic acid-dependent and -independent pathways during red blood cell invasion (Cowman and Crabb, 2006).

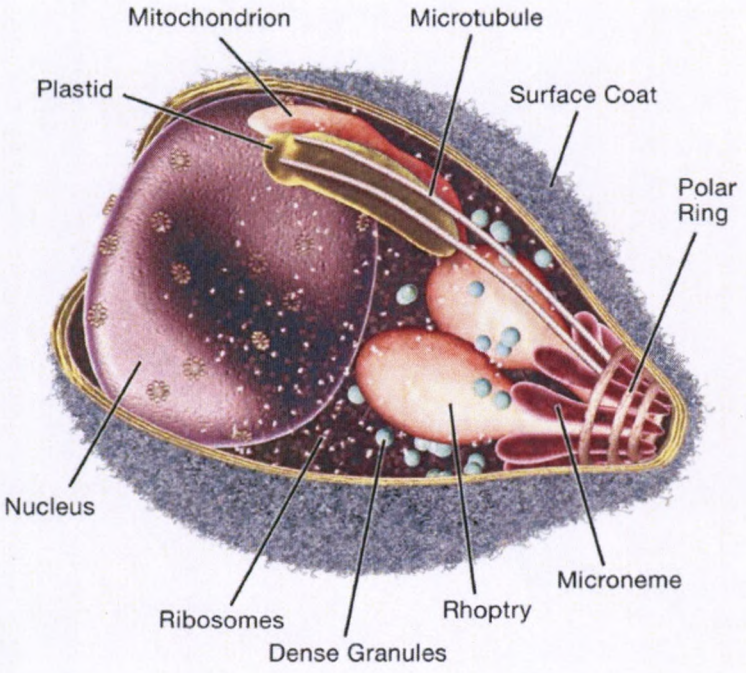


Figure 1.2: **Diagram of a *P. falciparum* merozoite. Merozoites are asexual malaria parasites, which invade host erythrocytes. The secretory organelles – rhoptries, micronemes and dense granules – are visible at the apical end of the parasite. The nucleus is visible at the posterior end. This image was taken from Cowman and Crabb (2006).**

Erythrocyte invasion is a complex process. The interactions and events that take place during invasion are predominantly based on *in vitro* studies (Weiss *et al.*, 2015; Boyle *et al.*, 2013). When the merozoite first comes into contact with the host erythrocyte (Figure 1.3-(1)), a cascade of events takes place (Weiss *et al.*, 2015). Any part of the surface of the parasite can form a reversible initial attachment with the RBC which is mediated by the interaction of the transmembrane erythrocyte band 3 receptor with glycosyl-phosphatidyl inositol (GPI)-anchored merozoite surface proteins (MSPs) (Li *et al.*, 2004). After attachment, specialized invasion proteins, such as apical membrane antigen-1 (AMA-1) which are found in micronemes, are secreted and enable re-orientation as seen in Figure 1.3-(2) although the precise role of AMA-1 is still under debate (Cowman and Crabb, 2006; Riglar *et al.*, 2011). Once the parasite orientation is such that the apical pole points directly to the RBC membrane a second interaction between the

merozoite and host cell results in permanent attachment and the formation of a tight junction (Bannister and Dluzewski, 1990). The tight junction links the host cell and parasite membrane to the actin-myosin motor of the merozoite. The thrombospondin-related apical membrane protein (PTRAMP) and merozoite thrombospondin-related anonymous protein (MTRAP) have been suggested to be involved in this event (Beeson *et al.*, 2016). The rhoptry neck protein (RON) complex, comprised of RON2-5, also appears to play key roles at the junction (Cowman *et al.*, 2012).

The merozoite is then driven by its unique actin-myosin motor (Figure 1.3-(3)) into the host cell through movement of the tight junction from the apical to the posterior end (Keeley and Soldati, 2004). The surface ligands of the merozoite are shed (Figure 1.3-(4)), by subtilisin-like serine protease 1 (SUB1) and SUB2, after which the parasite is enveloped in a parasitophorous vacuole (PV). This creates a compartment ideal for parasite development (Cowman and Crabb, 2006; Beeson *et al.*, 2016).

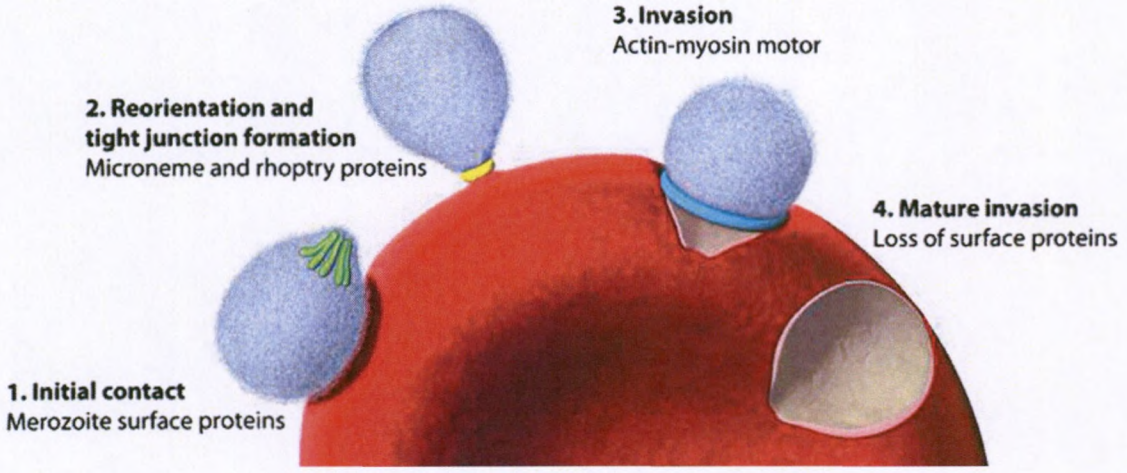


Figure 1.3: **Merozoite invasion of red blood cells. Merozoite invasion appears to occur in a step-wise process. (1) Initial contact; the apical organelles (micronemes and rhoptries) are coloured green. (2) Merozoite re-orientation leading to tight junction formation (yellow) between the parasite and the human red blood cell. (3) Actin–myosin motor movement; the moving junction is coloured blue. (4) Shedding and start of the intra-erythrocytic developmental cycle. Figure taken from Boyle *et al.* (2013).**

1.3 Invasion proteins

Invasion proteins can be divided into two classes; adhesins which bind directly to RBC receptors and invasins which are defined as a group of invasion proteins that function without making direct contact with the RBC (Cowman *et al.*, 2012). *P. falciparum* parasites can use alternative invasion pathways, therefore several parasite ligands exist that interact with and bind to various erythrocyte receptors. Most of these proteins have been shown to be dispensable (Paul *et al.*, 2015). This confirms that the parasite can facilitate invasion through any of these ligand-receptor interactions (Beeson *et al.*, 2016).

Plasmodium merozoite adhesins recognise and bind to receptors on the surface of RBC, which include Duffy binding (DB) group antigens and sialic acids on surface glycophorins. These adhesins are comprised of two super families, namely reticulocyte binding-like (RBL) proteins and Duffy binding-like (DBL) proteins. Duffy binding proteins (DBPs) which belong to the DBL family, bind to Duffy antigens through an N-terminal cysteine rich domain, region II (Adams *et al.*, 2001). A DBL subfamily of proteins is the erythrocyte-binding ligands (EBLs) which bind to other RBC receptors, facilitated by an N-terminal domain in the EBL proteins (Rodriguez *et al.*, 2000).

The most well-known of the EBL proteins is Erythrocyte Binding Antigen-175 (EBA-175), which utilises sialic acid-dependent and -independent pathways to mediate invasion (Duraisingh *et al.*, 2003). The region of the EBA-175 protein responsible for erythrocyte binding is a cysteine-rich domain, homologous to region II of the DBPs (Sim *et al.*, 1994). The DBPs and sialic acid-binding EBA-175 belong to a common gene family (*ebl* gene family) due to similarity in exon/intron gene structure and amino acid homology (Adams *et al.*, 1992; Figure 1.4). This protein family was named the erythrocyte binding protein (EBP) family and includes proteins like EBA-181 (JESEBL), EBA-140 (BAEBL), EBA-165 (PEBL) and merozoite adhesive erythrocyte binding-like protein (MAEBL) (Mitchell *et al.*, 2004; Treeck *et al.*, 2006). Members of the DBL-EBP family have multi-exon structures; conserved boundaries

between exons and introns; a conserved N-terminal cysteine-rich domain (DBL domain-F1/F2) and a C-terminal cysteine-rich (CCys) domain (Adams *et al.*, 1992; Tham *et al.*, 2011).

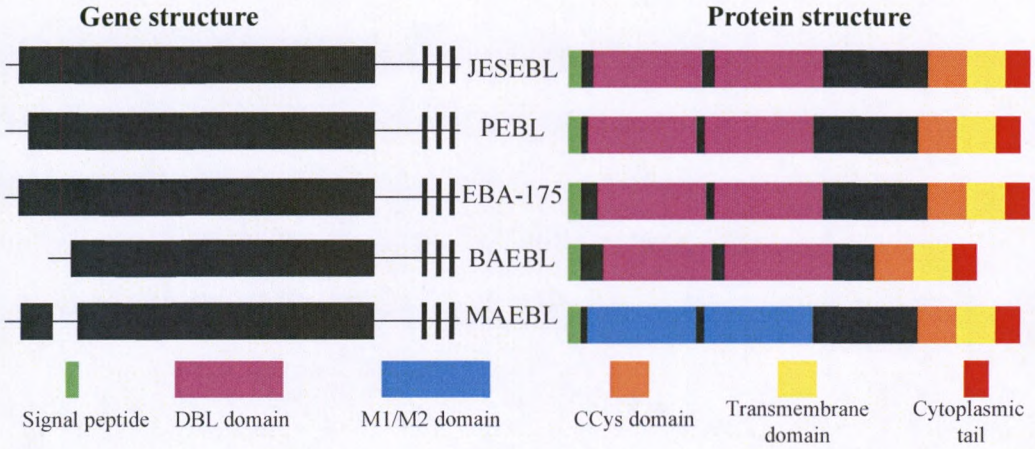


Figure 1.4: **Schematic representation of the DBL-EBP family of proteins in *P. falciparum*. The gene and protein structures are shown. All members of the DBL-EBP family contain a signal peptide, transmembrane domain, cytoplasmic tail and a DBL domain or M1/M2 domain in the case of MAEBL and a CCys domain. Exons are shown as black boxes and introns as solid horizontal lines. This diagram is adapted from Adams *et al.* (2001) and Tham *et al.* (2012).**

1.3.1 *Pf*MAEBL

One of the *P. falciparum* invasion proteins of interest in this study is *Pf*MAEBL. The *maebl* gene was first identified in the rodent parasites *P. yoelii* and *P. berghei* (Kappe *et al.*, 1992). The *Pfmaebl* gene (PF3D7_1147800) which is located on chromosome 11 belongs to the *ebl* family and has a multi-exon structure (Blair *et al.*, 2002). Alternative splicing results in two different open reading frames (ORFs) and transcripts which encode two isoforms of the protein (Singh *et al.*, 2004; Figure 1.5). The alternative splice site exists between exon 2 and 3 in ORF 2. This results in a stop codon in exon 4 resulting in early termination of the coding sequence. This forms a shorter transmembrane domain and cytoplasmic tail in the second isoform (Singh *et al.*, 2004).

The chimaeric *Pf*MAEBL protein consists of six domains: N-terminal signal peptide; two tandem cysteine-rich regions termed M1 and M2; repeat region; C-terminal cysteine-rich domain;

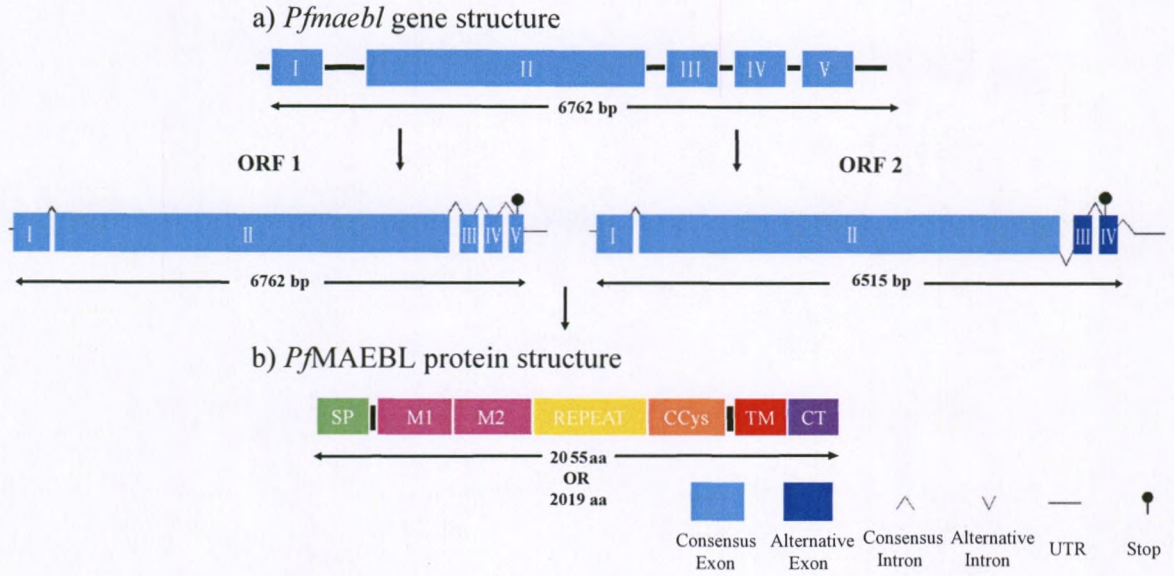


Figure 1.5: **Schematic representation of a) *Pfmaebl* gene (PF3D7_1147800) structure and b) *Pf*MAEBL protein structure. Schematic structures of the canonical (ORF 1) and alternatively spliced (ORF 2) forms of *maebl*. SP: signal peptide; M1/M2: *Pf*MAEBL-specific binding domain; CCys: C-terminal cysteine-rich region; TM: transmembrane domain; CT: cytoplasmic tail. This image was adopted from Singh *et al.* (2004).**

transmembrane domain and cytoplasmic domain (Ghai *et al.*, 2002). Instead of the DBL functional domain, which is needed in EBLs for recognition and binding to the correct erythrocyte receptor, the N-terminal of *Pf*MAEBL has tandem duplicates known as M1 and M2 domains (Blair *et al.*, 2002; Kappe *et al.*, 1992).

*Pf*MAEBL is expressed in both the sporozoite and the schizont stages of the parasite (Ghai *et al.*, 2002). It plays a role in invasion of erythrocytes when conditions are not optimal for other EBL proteins and it uses a unique invasion mechanism which is sialic acid-dependent and trypsin-insensitive (Fu *et al.*, 2005). The erythrocyte receptor of *Pf*MAEBL has not been identified yet (Culleton and Kaneko, 2010). *Pf*MAEBL is also critical in the sporozoite stage of *P. falciparum* parasites and therefore has important roles during infection at distinct points in the life cycle, including the blood, mosquito, and liver stages (Yang *et al.*, 2017). A unique feature of *Pf*MAEBL, unlike *Pf*AMA-1 and the other DBL-members, is that it localizes within the rhoptries of the merozoite and not the micronemes (Blair *et al.*, 2002).

1.3.2 *Pf*AMA-1

The second invasion protein of interest in this study is *Pf*AMA-1. The *Pfama-1* gene, located on chromosome 11, (PF3D7_1133400) only consists of a single exon and is 1869 bp (Figure 1.6). *Pf*AMA-1 is a classical type 1 integral membrane protein and comprises of several domains, of which the most important for invasion is the ectodomain and the cytoplasmic tail (Remarque *et al.*, 2008; Tyler *et al.*, 2011). The bulk of *Pf*AMA-1 is the ectodomain, made up of three sub-domains (Howell *et al.*, 2001) containing 16 conserved cysteines, which contribute to disulphide bonds(Hodder *et al.*, 1996). The ectodomain of *Pf*AMA-1 resembles the binding domains (M1 and M2) found in *Pf*MAEBL (Urquiza *et al.*, 2000; Mitchell *et al.*, 2004; Culleton and Kaneko, 2010). The protein is predominantly expressed in the sporozoite and schizont stages of the parasite, similar to *Pf*MAEBL (Bannister *et al.*, 2003; Yang *et al.*, 2017). *Pf*AMA-1 does not directly interact with the RBC, but rather forms a complex with RON2 (which is in direct contact with the RBC) and is suggested to be the step that commits the merozoite to RBC invasion (Srinivasan *et al.*, 2011). *Pf*AMA-1 localizes to the micronemes in *P. falciparum* parasites (Bannister *et al.*, 2003).

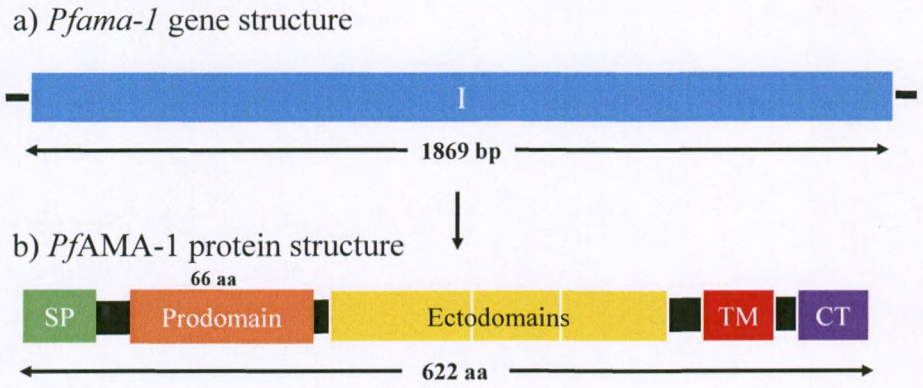


Figure 1.6: **Schematic representation of a) *Pfama-1* gene (PF3D7_1133400) structure and b) *Pf*AMA-1 protein structure. SP: signal peptide; TM: transmembrane domain; CT: cytoplasmic tail. This image was adapted from Treeck *et al.* (2009).**

1.4 Eukaryotic protein trafficking

The endomembrane system in eukaryotic cells is made up of the plasma membrane, nuclear membrane, endoplasmic reticulum (ER), Golgi apparatus, lysosomes, vacuoles, vesicles and endosomes. The ER is the centre of protein synthesis. The Golgi apparatus sorts proteins and other molecules for delivery to various cellular compartments. In eukaryotic organisms, translated proteins cross over the ER membrane from the cytosol by means of a signal peptide (SP) sequence (Fewell and Brodsky, 2009). Trafficking of these proteins from the ER to their final destination requires many sorting and packaging steps by different vesicles within the cell.

Cargo proteins from the ER are packaged into coat protein (COP)II complex transport vesicles, which bud out of ER exit sites. These proteins have an ER exit signal that binds to COPII and allows the proteins to be concentrated in these vesicles and transported out of the ER and to the Golgi apparatus, which is the sorting and dispatching centre of the cell (Rothman and Wieland, 1996). Once the vesicles are released from the ER, the COP-II coat is shed and the individual vesicles begin to fuse together into vesicular tubular clusters (Alberts *et al.*, 2002). This is a short-lived phase and the clusters quickly move along microtubules and fuse with the cis-Golgi network, delivering the cargo proteins (Figure 1.7).

As the vesicular tubular clusters form they begin to be coated with COPI, which is responsible for retrograde transport of proteins back to the ER (Figure 1.7). These ER resident proteins have a C-terminal retrograde signal ('KKXX' for membrane proteins or 'KDEL' for other retrograde proteins) which binds to COPI facilitating this kind of transport (Stornaiuolo *et al.*, 2003). Vacuolar protein sorting (VPS) proteins are also involved in retrograde and endosomal protein sorting (Raymond *et al.*, 1992).

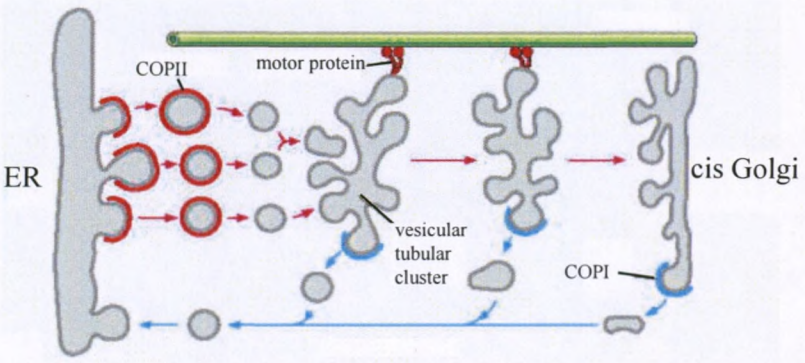


Figure 1.7: **Vesicular transport from the ER to the Golgi in eukaryotic organisms. Vesicular tubular clusters move along microtubules to carry proteins from the ER to the Golgi apparatus. COPII coats mediate the exit of the vesicles out of the ER and COPI coats mediate the budding of vesicles that return to the ER from these clusters. As indicated, the coats quickly disassemble after the vesicles have formed. COP: coat protein, ER: endoplasmic reticulum. Image taken from from Alberts *et al.* (2002).**

1.4.1 Clathrin and adaptor protein complexes in eukaryotes

Proteins not destined for retrograde transport, exit the trans-Golgi and either continue through the exocytic pathway (to the plasma membrane) or enter the endocytic pathway (to internal organelles) (Rothman and Wieland, 1996). The clathrin-depedent trafficking pathway which enables exocytosis is one of the most-characterized in eukaryotes. It involves forming a clathrin-coated vesicle (CCV) by a number of intricate interactions between the cargo proteins, membranes, clathrin and an adaptor protein (AP) complex. The CCVs associate with the microtubules and migrate to the desired subcellular compartments, delivering their contents (Wolfe and Trejo, 2007; Pearse, 1976; Keen *et al.*, 1979).

CCVs are made up of three layers, the outer being composed of a clathrin lattice. Clathrin is composed of three heavy chains and three light chains. It exists as a triskelion within the cell (Bonifacino and Lippincott-Schwartz, 2003). The intermediate layer of the CCVs comprise the AP complex. Currently five known AP complexes are involved in vesicular transport of proteins between different membrane compartments within the cell (Huang, 2011).

The sorting of proteins into CCVs is controlled by the AP1 and AP2 complexes, while the

others seem to have a clathrin-independent nature (Huang, 2011). AP1 and AP2 are heterotetrameic complexes consisting of two large subunits (β and γ), a medium subunit (μ) and small subunit (σ). The μ subunit binds specifically to YxxΦ (x = any amino acid, Φ= bulky hydrophobic amino acid) motifs on the C-terminus of transmembrane cargo proteins (Robinson, 2004; Edeling *et al.*, 2006). The large subunits are responsible for binding between the clathrin triskelions and the Golgi membrane lipids and proteins (Bonifacino and Lippincott-Schwartz, 2003; Figure 1.8).

Finally the internal layer of the CCV is made up of membrane lipids and proteins, and more importantly the cargo proteins. Association between all three layers and clathrin enables the AP1 complex to initiate the formation of the CCV at the trans Golgi membrane (Sebastian *et al.*, 2006).

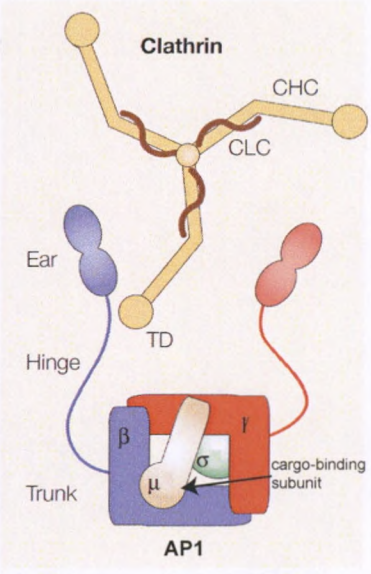


Figure 1.8: **Schematic representation of the interaction between clathrin and AP1. A clathrin triskelion (CHC, clathrin heavy chain; CLC, clathrin light chain) is shown interacting through its terminal domain (TD) with the appendage domains of the β subunit of AP1. The cargo binding subunit is μ1. Image taken from Bonifacino and Lippincott-Schwartz (2003).**

Transport of vesicles from the trans-Golgi to the endosomes or other secretory machinery can occur in a clathrin-independent manner (McMahon and Mills, 2004). Mammalian AP3 and AP4 are homologous to AP1 and AP2, respectively. AP3 and AP4 interact directly with COPII

and play a role in sorting proteins away from the trans-Golgi. Studies have also identified a family of monomeric adaptors that are known as the Golgi localized, γ-ear-containing, ADP-ribosylation factor-binding proteins (GGAs), which are components of clathrin coats and fulfil many of the same functions as the AP complexes (Bonifacino and Lippincott-Schwartz, 2003).

1.4.2 Protein sorting ability of sortilin in eukaryotes

Sortilin is a type I membrane glycoprotein belonging to the VPS10 family of sorting receptors and is most abundantly expressed in the central nervous system of higher level organisms (Andersen *et al.*, 2014). Initially considered a rather peculiar receptor resembling an intracellular sorting protein in yeast, sortilin has emerged as a key player in mammalian protein sorting. Studies have shown sortilin to be the first example of a mammalian receptor targeted by the GGA family of cytosolic sorting proteins (Nielsen *et al.*, 2001).

1.4.3 Apicomplexan protein trafficking and sorting

The phylum Apicomplexa is comprised of parasitic protozoans and includes both *T. gondii* and *P. falciparum*. Apicomplexan parasites have a complex endomembrane system that is built amongst all three of their distinct secretory organelles; the rhoptries, micronemes and dense granules. Proteins that are released from these organelles are needed for invasion. Due to the complexity of the internal organization of these parasites, a range of trafficking factors is required for the correct sorting of proteins towards the various organelles (Tomavo *et al.*, 2013; Figure 1.9). Secretory proteins, such as invasion proteins, co-translationally enter the ER via the SP pathway. Proteins are then specifically trafficked to their final destinations. Some signals and escorter proteins which govern this sorting process have been identified for trafficking to the micronemes and rhoptries (Jimenez-Ruiz *et al.*, 2016).

Compared to the endomembrane system found in higher eukaryotes, the system in Apicomplexans is more primitive and not as well characterized (Tomavo *et al.*, 2013). This phylum has lost

many endosomal sorting complexes required for transport e.g. members of the ESCRT family, which are comprised of VPS proteins. This suggests the importance of ESCRT-III and the VPS proteins that still exist in this phylum (Woo *et al.*, 2015; Jimenez-Ruiz *et al.*, 2016).

1.4.4 Trafficking signals in *T. gondii*

T. gondii is an attractive model in this phylum and is one of the best-characterized in terms of its trafficking pathways. Rhoptry protein 2 (ROP2) is transported from the cis-Golgi to premature rhoptry compartments (Besteiro *et al.*, 2009) by means of a tyrosine motif (YXXΦ) in the cytoplasmic tail (Hoppe *et al.*, 2000). Two tyrosine-based motifs, SYHYY and EIEYE, in the cytoplasmic tail of *T. gondii* microneme proteins (MICs) were shown to be critical for correct sorting to the micronemes (Di Cristina *et al.*, 2000).

1.4.5 Trafficking escorters in *T. gondii*

Evidence is mounting in *T. gondii* that trafficking to the rhoptries and micronemes involves not only the Golgi but also a post-Golgi system that is endosome-like. It has been suggested that the μ1 subunit of *Tg*AP1 is implicated in the transport of invasion proteins from the ER to secretory organelles (Ngô *et al.*, 2003). Microneme proteins appear to be transported in complexes with transmembrane proteins and MIC-6 has been identified as the escorter responsible for the microneme trafficking of MIC-1 and MIC-4 (Reiss *et al.*, 2001).

An escorter responsible for both rhoptry and microneme trafficking has been identified in *T. gondii* (Figure 1.9). The lumenal domain of this protein, SORTILIN-like receptor (*Tg*SORTLR), is needed for correct apical organelle trafficking of several microneme and rhoptry proteins (Sloves *et al.*, 2012). They predicted that the cytoplasmic tail of *Tg*SORTLR associated with cytosolic trafficking proteins such as clathrin, μ1 subunit of AP1, retromer complex (VPS26 and VPS35) or VPS9 (Tomavo *et al.*, 2013).

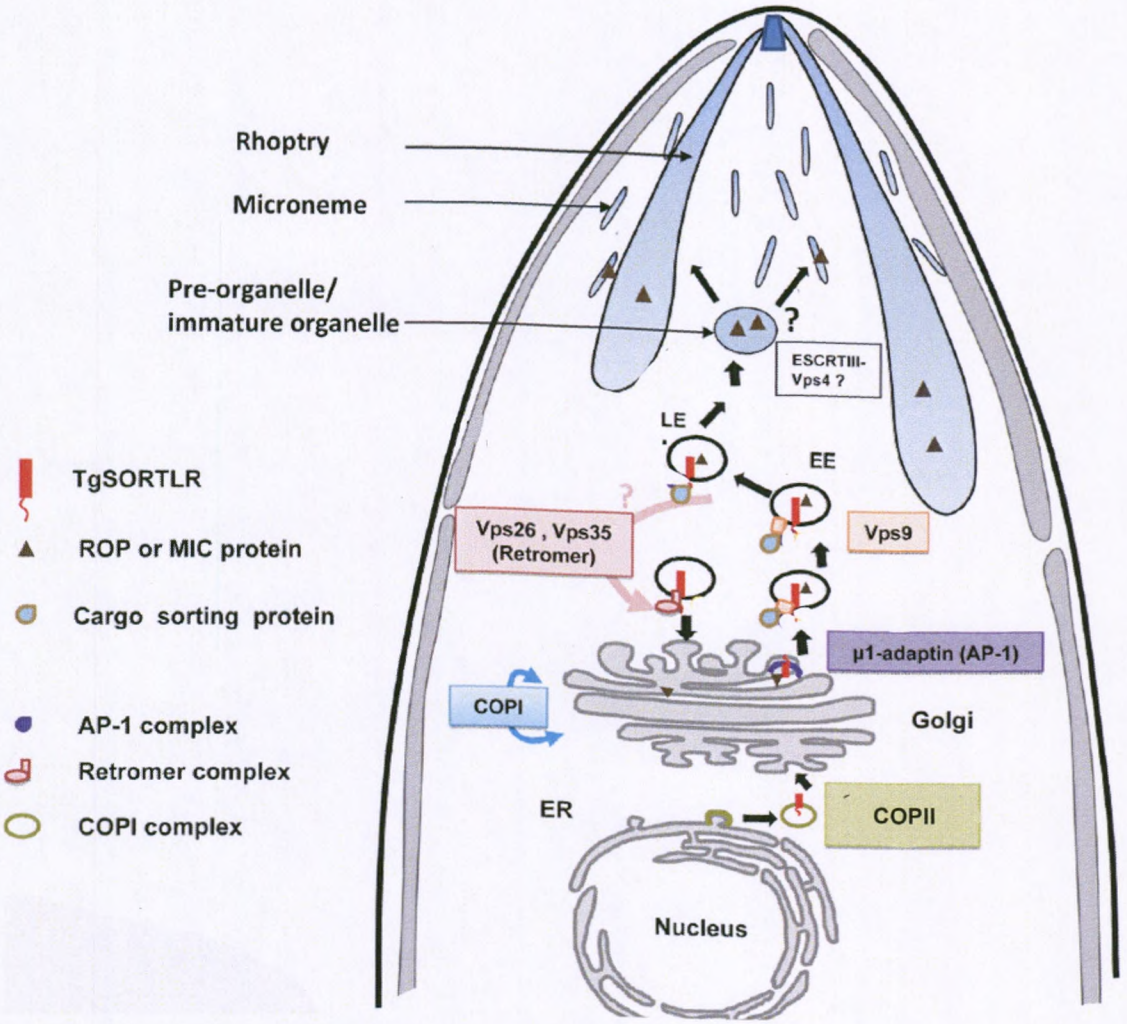


Figure 1.9: **A model for *Tg*SORTLR functions in protein sorting and the biogenesis of apical secretory organelles. It was proposed that *Tg*SORTLR has a distinct role as a type I transmembrane cargo-protein receptor for ROPs and MICs of apicomplexan parasites. The model further proposes that the cytoplasmic tail of *Tg*SORTLR binds to AP1, clathrin, clathrin-associated adaptor protein, and VPS9, and this defines it as a key receptor involved in the anterograde transport of cargo ROP and MIC proteins. The binding of *Tg*SORTLR to the retromer VPS26/VPS35 indicates that this receptor is also involved in the retrograde transport of components. Endosome sorting complex required for transport (ESCRT-III) is suggested to be required for protein transport from the LE to secretory organelles. EE: early endosome, LE: late endosome. Image from Tomavo *et al.* (2013).**

1.4.6 Protein trafficking in *P. falciparum* parasites

Targeting secretory proteins to the apical organelles remains poorly characterized in *P. falciparum*. It is estimated that more than 50 proteins are involved in erythrocyte invasion with

some proteins being expressed on the parasite surface, but the majority are localized within the apical organelles (Kats *et al.*, 2008).

The trafficking of secretory proteins in *P. falciparum* to the ER is the same as in eukaryotes and relies on a signal peptide (Tomavo *et al.*, 2013). This includes proteins belonging to the DBL-EBP family, which then require further targeting signals to direct translocation of these proteins to their destined secretory organelles.

1.4.6.1 Rhoptry targeting

The exact trafficking mechanism to the rhoptries has not been completely elucidated (Kats *et al.*, 2006). In a study using transgenic parasites expressing GFP-fusion proteins, it was shown that the first 20 aa of Armidillo Repeats-Only (ARM) protein and the first 24 aa of high molecular weight rhoptry protein 2 (RhopH2) were required for rhoptry localization (Cabrera *et al.*, 2012; Kats *et al.*, 2008; Ghoneim *et al.*, 2007; Figure 1.10). Residues 22-55 of the N-terminal of rhoptry-associated protein 1 (RAP1) were needed for correct targeting but when this targeting signal was compared to other *Plasmodium* rhoptry proteins (including the ones mentioned above), it was seen that the motif was not conserved (Richard *et al.*, 2009). Melanie Wepener from the Coetzer lab, showed that the CCys domain is necessary for trafficking of *Pf*MAEBL to the apical end of the parasite (Wepener, 2014).

1.4.6.2 Microneme targeting

No common targeting motif has been identified for microneme proteins. In the cytoplasmic tail of the *Pf*TRAP a tyrosine motif exists which is important for sporozoite microneme localization (Bhanot *et al.*, 2003; Withers-Martinez *et al.*, 2004; Figure 1.10). Many other rhoptry and microneme proteins have this motif so it cannot be confirmed if it truly contributes to targeting of proteins to the apical organelles (Kats *et al.*, 2008). The transmembrane domain of *Pf*SUB2 is necessary for trafficking to the microneme (Child *et al.*, 2013), whereas EBA-175, BAEBL,

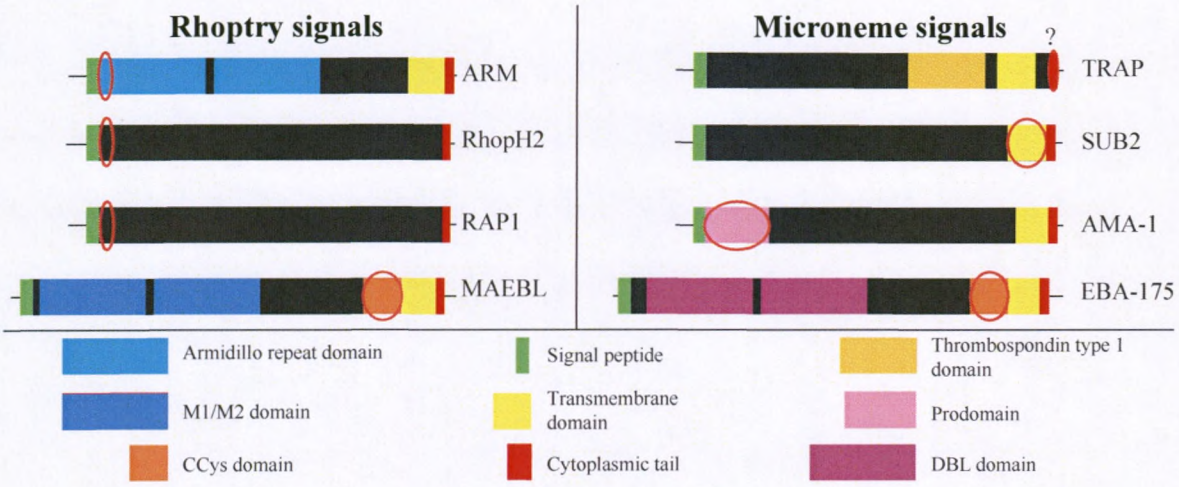


Figure 1.10: **Schematic representation of the regions needed for rhoptry or microneme trafficking. The circled areas indicate which region is needed for correct trafficking of each protein. The domain for thrombospondin-related anonymous protein (TRAP) has not been confirmed. ARM: armidillo repeat-only protein, RhopH2: high molecular weight rhoptry protein 2, RAP1: rhoptry-associated protein 1, MAEBL: merozoite adhesive erythrocyte binding-like protein, TRAP: thrombospondin-related anonymous protein, SUB2: subtilisin-like serine protease 2, AMA-1: apical membrane antigen 1. Image based on Topolska *et al.* (2003), Withers-Martinez *et al.* (2004), Kats *et al.* (2006), Treeck (2006), Kats *et al.* (2008), Richard *et al.* (2009), DeSimone *et al.* (2009), Cabrera *et al.* (2012), Sakura *et al.* (2013), Wepener (2014), Churchyard (2017) and PlasmoDB (2018).**

JESEBL and AMA-1 do not require the transmembrane domain for correct localization (Treeck *et al.*, 2006). The Cys domain in the ectodomain of EBA-175 (Sakura *et al.*, 2013) and the prodomain of AMA-1 are responsible for their localization to the micronemes (Treeck *et al.*, 2009; Churchyard, 2017).

It is of interest to note most rhoptry proteins contain the targeting signal at the N-terminal end, except for *Pf*MAEBL and most microneme proteins contain the targeting signal at the C-terminal end, except for *Pf*AMA-1 (Figure 1.10). Collectively, these studies indicate the variability in the essential domains for correct protein trafficking and that no globally conserved motif for invasion proteins exists (Gilberger *et al.*, 2003; Treeck *et al.*, 2006; Treeck *et al.*, 2009).

1.4.6.3 Trafficking escorters

Little is known about the escorters involved in the trafficking of invasion proteins destined for the apical organelles. The GPI-anchored protein, rhoptry-associated membrane antigen (RAMA) has been shown to interact with and facilitate the transport of RAP1 to the rhoptries (Richard *et al.*, 2009; Topolska *et al.*, 2003). The AP1μsubunit resides near the Golgi compartment and co-localizes with RAP1, which suggests thats this subunit is involved in post-Golgi trafficking, as is the case in most eukaryotes (Kibria *et al.*, 2015). Rhoptry proteins cluster in lipid rafts for trafficking in contrast to microneme proteins (Topolska *et al.*, 2003). A microneme escorter has not been identified in *P. falciparum*.

1.4.6.4 *Pf*SORTILIN

A *P. falciparum* homologue of the *Tg*SORTLR (*Pf*SORTILIN) was identified by Krai *et al.* (2014). The *Pfsortilin* gene (PF3D7_1451800) is 3139 bp, consisting of four exons which encode a 895 aa predicted product. The protein predominantly consists of the 60 kDa VPS10 domain. A cytosolic domain and a very small transmembrane domain exist at the C-terminal end as indicated in Figure 1.11.

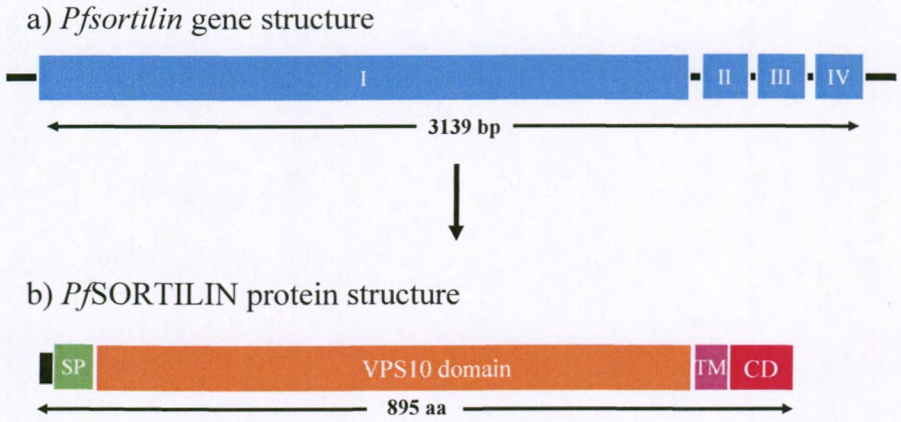


Figure 1.11: **Schematic representation of a) *Pfsortilin* (PF3D7_1451800) gene structure and b) *Pf*SORTILIN protein structure. SP: signal peptide, TM: Transmembrane domain and CD: cytoplasmic domain. Image adapted from Krai *et al.* (2014).**

Florescence microscopy of transgenic parasites expressing the protein with an HA-tag revealed that it localized to the Golgi apparatus (Krai *et al.*, 2014). *Pf*SORTILIN is expressed in late trophozoite and schizont stages (PlasmoDB, 2018), which may suggest a role in targeting proteins to the apical organelles. Recent findings in 2018 that were published after the completion of the current study, revealed that *Pf*SORTILIN interacts with regions of RAMA, a proposed rhoptry escorter (Topolska *et al.*, 2003; Richard *et al.*, 2009; Hallée *et al.*, 2018). This interaction was sufficient for localization to the rhoptries providing support for *Pf*SORTILIN's potential role as a protein escorter of the RAMA-RAP complex to this organelle (Hallée *et al.*, 2018). Its role in microneme trafficking has not been investigated.

1.5 Purpose of the study

Protein trafficking pathways to the rhoptries and micronemes are starting to be unravelled. *P. falciparum* invasion proteins are redundant and are thus not suitable drug targets, but the research proposed in this study may identify a trafficking mechanism/escorter, which could provide clues for novel therapeutic strategies. It is hypothesized that *Pf*SORTILIN may function as an escorter for the transport of invasion proteins.

1.5.1 Study aim

The aim of this study is to investigate the potential role of *Pf*SORTILIN as an escorter protein for *Pf*MAEBL and *Pf*AMA-1.

1.5.2 Study objectives

1. Express selected domains of*Pf*SORTILIN, *Pf*MAEBL and *Pf*AMA-1 as recombinant proteins

2. Perform binding assays to investigate the interaction of *Pf*SORTILIN with *Pf*MAEBL and *Pf*AMA-1
3. Investigate the interaction of *Pf*SORTILIN with unknown cargo proteins by biopanning against a *P. falciparum* phage display library

Chapter 2

Methods and Materials

2.1 Culturing techniques

All solutions were made using ultrapure laboratory grade water generated by a MilliQ^{TM} Water System. Sterile methods and consumables were employed for all culturing techniques, and solutions were sterilised by autoclaving or filtration through a 0.2 µm membrane. The commercial chemical reagents and the equipment used in this study appear in Appendix B and Appendix C.

2.1.1 Parasite culture composition and maintenance

P. falciparum 3D7 parasites were cultured in a sterile P2 hood by the method originally defined by Trager and Jensen (1976), with adjustments. Cultures consisted of parasitized red blood cells (pRBCs) and washed RBCs, kept at 5% haematocrit. Cultures were maintained at 37 °C in 25 cm^3 (small) or 75 cm^3 (medium) sterile culture flasks (Nunc, Denmark), in 5 ml (small) or 20 ml (medium) incomplete medium (RPMI 1640 with L-Glutamine, 25 mM HEPES, 50 mg/l Hypoxanthine, 50 mg/l Gentamycin sulphate, pH 6.72) which was made into complete medium by supplementation with 0.5% Albumax and 0.2% sodium bicarbonate. Cultures were gassed with a mixture of 5.5% CO_2, 2.75% O_2 and 91.25% N_2 for 30 s and the lids were tightly closed. Complete medium was changed daily and cultures were diluted with fresh washed erythrocytes

to maintain a parasitaemia of less than 3% trophozoites and less than 6% rings. Smears were prepared daily to visualise the life stage, parasitaemia and general health of the parasite cultures.

2.1.2 Preparation of erythrocytes

Blood was collected (Ethics number M171088) from volunteers in 6 ml acid citrate dextrose tubes. These tubes were centrifuged at 1000 x *g* for 10 min at 4 °C using a 5702R centrifuge, allowing the blood to separate into three layers. In a sterile hood, the top plasma and buffy layers were aspirated, leaving only a layer of erythrocytes in the tube. Two volumes of incomplete medium were added and the contents of the tube inverted to wash the erythrocytes. Centrifugation was repeated as stated above and the top layer was aspirated. This wash was performed twice. The erythrocytes were then diluted in one volume of incomplete medium and stored at 4 °C for up to two weeks.

2.1.3 Thawing frozen parasite cultures

Frozen stocks (1 ml), stored in cryotubes, were removed from liquid nitrogen and defrosted at 37 °C. Under sterile conditions, the thawed cells were transferred to a 15 ml centrifuge tube (Greiner Bio-one, Germany). For every millilitre of cells, 0.1 ml of 12% sodium chloride was slowly added while swirling. The tube was incubated at room temperature for 5 min before the addition of 9 volumes of 1.6% sodium chloride (Blomqvist, 2008). The suspension was gently mixed and then centrifuged at 400 x *g* for 5 min at 37 °C (Eppendorf Centrifuge 5702R). Removing the supernatant, the pellet was resuspended in 9 volumes of incomplete medium. Centrifugation was repeated and the pellet was made up to a volume of 0.75 ml using freshly washed and diluted erythrocytes (approximately 500 µl of 1:1 diluted erythrocytes). The cells were then transferred to a small culture flask and complete medium was added to make a total volume of 5 ml. The medium included 1% Albumax, instead of the usual 0.5% Albumax; the increased protein concentration of the medium aids in parasite growth. The flask was then

gassed and incubated at 37 °C for 48 h. Smears were prepared and medium was changed daily (using the 1% Albumax medium) for 7 days. After the first week culturing was continued using medium containing 0.5% Albumax.

2.1.4 Freezing parasite cultures

High parasitaemia cultures (> 4% rings) were frozen to make stocks as described by Moll (2008) with adjustments. A 20 ml culture was transferred to two 15 ml centrifuge tubes and centrifuged at 400 x *g* for 5 min at room temperature. After removal of the supernatant, the pellets were resuspended in freezing solution (28% glycerol in Phosphate buffered saline (PBS)- 120 mM NaCl, 83 mM Na_2HPO_4, 32 mM KH_2PO_4, pH 7.4) at a ratio of 1:1. The suspension was then mixed well and transferred to sterile cryotubes (0.5 ml to 1.0 ml per tube). Tubes were incubated at room temperature for 5 min before storing in liquid nitrogen.

2.1.5 Smear Preparation

Smears were prepared daily. Flasks were tilted upward, allowing the culture medium to flow back and exposing the thin layer of erythrocytes at the bottom of the flask. A glass Pasteur pipette was then used to take up ±10 µl of erythrocytes by capillary action. These erythrocytes were transferred to a glass microscope slide and smeared thinly using the back of a second glass slide. The smear air dried inside the flow hood and was then stained using the Rapi-Diff Stain Set. Slides were viewed at 1000X magnification under oil immersion (Axiostar, Zeiss). The parasitaemia of a culture was estimated by the following equation:

$$\%\text{Parasitaemia} = \frac{\text{number of infected RBCs}}{\text{number of total RBCs}} \times 100$$

The mean of 5 fields (≈ 100 cells in each) was used to determine the final parasitaemia.

2.1.6 Synchronisation of cultures

Sorbitol synchronisation is based on the increased permeability of trophozoite-infected RBCs to sorbitol (due to the development of new permeation pathways) that leads to lysis of infected trophozoite stage RBCs (Lambros and Vanderberg, 1979). A culture containing predominantly ring-stages was pelleted by centrifugation for 5 min at 400 x *g* at 37 °C. The spent culture medium was aspirated, the cells resuspended in 5% (w/v) D-sorbitol and incubated at 37 °C for 15 min. Cells were pelleted again by centrifugation at 400 x *g* for 5 min at 37 °C, the supernatant removed and additional culture medium or RBCs added (as described in subsection 2.1.1) to prepare a culture with a 5% haematocrit and 1-3% parasitaemia.

2.2 *Pf*DNA preparation and analysis

2.2.1 Genomic DNA extraction from *P. falciparum*

DNA extraction was carried out on late-stage parasites from cultures with 3-5% parasitaemia according to the methods of Cowman *et al.* (2008). A 5 ml culture from a small flask was transferred to a sterile 15 ml centrifuge tube and centrifuged at 400 x *g* for 5 min at 4 °C. The supernatant was removed and the pellet was washed in 14 ml PBS followed by centrifugation at 400 x *g* for 5 min at 4 °C. The supernatant was aspirated and then four pellet volumes of hypotonic solution (5 mM KH_2PO_4, pH 7.4 with K_2HPO_4) were added to facilitate RBC lysis. One volume of 18% sodium dodecyl sulphate (SDS) was added to denature the released proteins (Birnboim and Doly, 1979). The suspension was incubated for 3 min at room temperature after which 8 volumes of 1:1 phenol:chloroform mixture were added to remove any soluble proteins that were still present. The tube was inverted several times, centrifuged in a 5702R Eppendorf centrifuge at 3000 x *g* for 10 min at 4 °C and the top aqueous layer, containing the chromosomal DNA, was carefully collected and transferred equally as 400 µl aliquots into 1.5 ml Eppendorf

tubes, on ice. To each tube one tenth of a volume of 3 M sodium acetate (CH_3COONa, pH 5) and 2.5 volumes of ice cold 100% molecular grade ethanol (Merck, South Africa) were added to promote chromosomal DNA aggregation (Moore and Dowhan, 2002). The tubes were incubated for at least 15 min at −70 °C, centrifuged in a 5415R Eppendorf centrifuge at 16100 x *g* for 30 min at 4 °C and the supernatant decanted. One millilitre of 70% ethanol was added to the tube which was centrifuged at 16100 *g* for 5 min at 4 °C to remove salt from the sample, which may hinder downstream applications (Moore and Dowhan, 2002). The supernatant was decanted and the pellets were air dried for 15 min and then resuspended and pooled into a final volume of 600 µl TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8) containing 100 mg/l RNase A to remove contaminating RNA.

The sample was mixed by swirling and incubated for 30 min at 37 °C in a water bath. A single volume of 1:1 phenol:chloroform mixture was added and the tube inverted several times. The tube was then centrifuged as before for 30 min at 4 °C. The top aqueous layer was carefully collected and transferred into a 1.5 ml Eppendorf tube to which one volume (600 µl) of chloroform was added; the tube was inverted and centrifuged as before. The top aqueous layer was collected and transferred into a 2 ml Eppendorf tube, in which the DNA was precipitated with ethanol as described above. The DNA pellet was air dried and suspended in 20 µl to 50 µl TE buffer. The extracted genomic DNA was kept at 4 °C or −20 °C for short or long term storage respectively.

2.2.2 Determination of DNA concentration, purity and integrity

A Nanodrop ND-1000 spectrophotometer was used to estimate the DNA concentration with 1 A_{260} unit equal to 50 µg/ml of dsDNA.

This system allowed the DNA purity to be determined using the A_{260}/A_{280} value (NanoDrop Technologies Inc, 2006). This is facilitated by DNA absorbing electromagnetic radiation strongly at a wavelength of about 260 nm due to the presence of purine and pyrimidine bases (Blackburn

et al., 2006). A value of 1.8 indicates a pure sample. If the A_{260}/A_{280} reading is less than 1.8, protein contamination is present. A_{260}/A_{280} reading greater than 2 indicates the presence of RNA in the sample.

The integrity of the DNA was determined using agarose gel electrophoresis. A 1% agarose gel was made by dissolving 0.5 g agarose in 50 ml 1x TAE buffer (which was made from a 50X TAE stock- 2 M Tris, 0.1 M acetic acid, 50 mM EDTA, pH 8) by heating. To the cooled solution 2.5 µl of 10 µg/µl ethidium bromide was added. The solution was poured and set in a gel casting chamber. Samples and a mixed range base pair standard (80 – 10000 bp) were loaded into the wells and electrophoresed in 400 ml 1x TAE buffer with 2.5 µl of 10 µg/µl ethidium bromide at the anode of the chamber, for 1 h at 100 V. The gel was then visualized under ultra violet (UV) light using the GeneSnap GeneGenius Geldoc scanning system and version 6.05 image acquisition software.

2.3 Polymerase chain reactions (PCR)

2.3.1 Primer specifications and parameters

Sequences for *Pf*MAEBL (PF3D7_1147800) and *Pf*SORTILIN (PF3D7_1451800), were obtained from the Plasmodium genome database, PlasmoDB version 31 (http://plasmodb.org/plasmo/). Primers were designed with the aid of Integrated DNA Technologies SciTools Oligo Analyzer 3.1 (Integrated DNA Technologies, 2016) for the amplification of various *P. falciparum* genes or domains. The primers were assessed to ensure that the formation of any hairpins or primer dimers during the PCR process would be non-consequential. Criteria for the primers also included melting temperatures that were above 45 °C for the *P. falciparum*-specific template. Recognition sites for restriction enzymes *Nde*1 and *Bam*H1 were included in the primers of the C-terminal cysteine-rich domain (CCys) in *Pf*MAEBL for cloning into the pET-15b vector. Restriction enzyme sites for *Bam*H1 and *Xho*1 were included in the primers of the VPS10 domain

in *Pf*SORTILIN for cloning into the pGEX-4-T2 vector. The reverse primers for both vectors included a stop codon. Primer sequences can be found in Appendix A. Primer synthesis was performed by Inqaba Biotech (South Africa). Primers were received in lyophilised form and dissolved in the recommended amounts of nuclease free water, to make a final concentration of 100 µM and stored at −20 °C. Primer working stock solutions of 10 µM were made for PCR reactions and were stored at −20 °C.

2.3.2 PCR conditions

The *Pf*MAEBL and *Pf*SORTILIN domains were amplified using the Phusion® Flash High-Fidelity PCR Master Mix. The mix contained a thermostable Taq DNA polymerase with a high proof reading ability and replication accuracy (ThermoScientific, 2014). The reaction mixtures were prepared as in Table 2.1. The Phusion® Flash High-Fidelity PCR Master Mix has the necessary reaction components for optimal PCR except the template DNA and primers. Reactions were set up on ice in 0.6 ml nuclease free microcentrifuge tubes and were then placed in a Mastercycler gradient thermocycler preheated to 98 °C.

Table 2.1: **Reaction mixture for PCRs of *Pf*MAEBL CCys and *Pf*SORTILIN VPS10**

Master Mix (2X)	**Forward primer** (0.5 µM)	**Reverse primer** (0.5 µM)	**DNA** (250 ng)	**Nuclease free water**
25 µl	2.5 µl (10 µM)	2.5 µl (10 µM)	2.5 µl (100 ng/µl)	Up to 50 µl final volume

Each amplicon required two successive annealing temperatures. Annealing temperatures (T_a) were determined using the online Integrated DNA Technologies SciTools Oligo Analyzer 3.1 (Integrated DNA Technologies, 2016).

The first T_a was calculated for the *P. falciparum* sequence without the added bases making up the restriction sites. Five cycles (denaturation, annealing and elongation) at the *P. falciparum* specific T_a were carried out. The second T_a took the full-length primer sequences into account

and was used to carry out another 30 cycles of amplification. Below are the final reaction programs used:

Table 2.2: **Optimized PCR conditions for each insert**

	***Pf*MAEBL CCys**	***Pf*SORTILIN VPS10**
Initial denaturation	98 °C for 1 s	98 °C for 1 s
Denaturation	98 °C for 5 s	98 °C for 5 s
Primer annealing	46 °C for 5 s, 5 cycles	49 °C for 5 s, 5 cycles
	53 °C for 5 s, 30 cycles	69 °C for 5 s, 30 cycles
Extension	72 °C for 25 s	72 °C for 25 s
Final Extension	72 °C for 1 min	72 °C for 1 min

Samples (1 µl with 1 µl FastDigest™ Green buffer) were electrophoresed on a 1% agarose gel for 1 h at 100 V to assess amplification of the correct size product. The amplified PCR products were purified using Macherey-Nagel NucleoSpin® Gel and PCR Clean-up Kit, according to the manufacturer's instructions, and eluted into 50 µl nuclease free water (Macherey-Nagel, 2017). Amplicon purity and concentration were assessed using the Nanodrop ND-1000 spectrophotometer.

2.3.3 Restriction enzyme digestion of PCR products

Purified PCR products (≈1 µg) were double-digested using FastDigest™ enzymes. Reaction mixtures (final volume 30 µl) were prepared in 0.6 ml nuclease free microcentrifuge tubes and are outlined in Table 2.3. Depending on the PCR product, double digestion with *Bam*H1 and *Xho*1 or *Nde*1 was carried out in 10X FastDigest™ Green Buffer for 1 h at 37 °C. One microlitre of each of these restriction enzymes is able to digest up to 1 µg of DNA in 15 min. According to the supplier, under the prescribed conditions the enzymes have 100% activity with the absence of star activity. Star activity is the non-specific cleavage of DNA by restriction enzymes when reaction conditions are not optimal.

Total samples were electrophoresed on a 1% agarose gel (as previously described in subsection 2.2.2) to verify the size of the amplicons (in case of star activity). Purified samples were cut

Table 2.3: **Reaction mixtures for PCR product restriction digests using FastDigest™ enzymes**

PCR product	***Bam*H1**	***Xho*1**	***Nde*1**	**Digest Buffer (10X)**	**Nuclease free water**
*Pf*MAEBL CCys	2 µl	-	2 µl	3 µl	Up to 30 µl final volume
*Pf*SORTILIN VPS10 13 µl	2 µl	2 µl	-	3 µl	Up to 30 µl final volume

from the gel and cleaned using Macherey-Nagel NucleoSpin® Gel and PCR Clean-up Kit and eluted in 50 µl nuclease free water (Macherey-Nagel, 2017). The concentrations of the products were determined using a Nano-drop ND-1000 spectrophotometer (NanoDrop Technologies Inc, 2006).

2.4 Cloning

2.4.1 Extraction of Plasmid DNA from DH5α cells

Plasmids were extracted from DH5α *E. coli* cells using the Macherey-Nagel NucleoSpin® Plasmid DNA purification Kit according to the manufacturer's protocol (Macherey-Nagel, 2017). The method relies on an SDS-based lysis of *E. coli* cells, followed by the adsorption of plasmid DNA onto a silica membrane in the presence of high salt concentrations. A spin-wash step removed any contaminants such as salts, metabolites, and soluble macromolecular cellular components. The bound DNA was eluted into 50 µl nuclease free water.

DH5α cells were grown in sterile 100 ml Erlenmeyer flasks containing 5 ml Luria Bertani (LB) medium (1% Bacto tryptone, 1% yeast extract, 0.5% NaCl, 10 mM Tris−HCl-pH 7.5) which has the nutrients and a buffer providing optimum growth conditions for the cells. Cells were grown overnight at 37 °C in an orbital shaker (250 rpm). Cells were harvested using the Eppendorf Centrifuge 5415R at 11000 x *g* for 30 s at 4 °C. The pelleted bacteria were resuspended in

250 µl Buffer A1 of the kit. Cell were lysed with the addition of 250 µl of SDS/alkaline Buffer A2. The sample was inverted and incubated for 5 min at room temperature. The lysate was neutralized with 300 µl of Buffer A3, creating the appropriate conditions for the liberated plasmid to bind to the silica membrane of the NucleoSpin® column. The sample was mixed thoroughly by inversion and precipitated protein, genomic DNA, and cell debris were pelleted by centrifugation at 11000 x *g* for 10 min at 4 °C. The supernatant was transferred to the NucleoSpin® column and centrifuged at 11000 x *g* for 1 min to allow for the binding of the plasmid to the membrane. Contaminants like salts, metabolites, and soluble macromolecular cellular components were removed by washing with 600 µl of the ethanolic Buffer A4. The device was centrifuged at 11000 x *g* for 1 min and the flow-through discarded. The membrane was dried by centrifugation at 11000 x *g* for 2 min. The column was removed and transferred into a 1.5 ml tube. A volume of 50 µl of nuclease free water and centrifugation at 11000 x *g* for 2 min facilitated the elution of the pure plasmid from the membrane with a decrease in the ionic strength conditions and increase in pH (Macherey-Nagel, 2017).

2.4.2 Restriction enzyme digestion of vector plasmids

To prepare the plasmids for ligation with the *Pf*DNA amplicons, 1 µg of each plasmid isolated in subsection 2.4.1 underwent double restriction digests using FastDigest™ enzymes. The reactions were prepared as listed in Table 2.4.

Table 2.4: **Reaction mixture for plasmid restriction digests using FastDigest™ enzymes**

Vector	***Bam*H1**	***Xho*1**	***Nde*1**	**Digest Buffer (10X)**	**Fast AP™**	**Nuclease free water**
pET-15b	2 µl	-	2 µl	6 µl	2 µl	Up to 60 µl final volume
pGEX-4T-2	2 µl	2 µl	-	6 µl	2 µl	Up to 60 µl final volume

The solutions were incubated at 37 °C for 1 h to allow for complete digestion of the plasmids. Fast AP™ (ThermoScientific, USA) was included in the reactions to dephosphorylate ends of the plasmid and prevent re-ligation. Samples were cleaned using NucleoSpin® Gel and PCR Clean-up Kit and eluted in 50 µl nuclease free water (Macherey-Nagel, 2017). An aliquot of each sample was resolved on a 1% agarose gel (as described in subsection 2.2.2) along with 5 µl MassRuler™ DNA Ladder Mix. This ladder, which contains bands with known DNA concentrations, and the GeneTools™ software were used together to determine the concentrations of the digested plasmids and inserts.

2.4.3 Ligation of DNA inserts and plasmids

Ligation reactions were carried out using the Roche Rapid™ Ligation Kit. A plasmid to DNA insert molar ratio of 1:3 was added to 2 µl of 5X DNA dilution buffer and made up to 10 µl with nuclease free water. The total amount of DNA did not exceed 200 ng. Then 10 µl of 2X T4 DNA Ligation Buffer was added to this solution and mixed. Lastly, 1 µl of T4 DNA Ligase was added and the solution was incubated at 16 °C for 30 min. A 10 µl aliquot of the ligation reaction was used for transformation and the remainder was stored at 4 °C.

2.4.4 Transformations

2.4.4.1 DH5α cells

DH5α *E. coli* (Invitrogen Ltd, UK) cells were made competent with PIPES buffer (10 µM 1,4-piperazinediethanesulphonic acid, 100 mM $CaCl_2$, 15% glycerol (v/v), pH 7). Ten millilitres LB medium was inoculated with DH5α *E. coli* cells and grown for 16 h at 37 °C in an Orbital Shaker (225 rpm). The 10 ml culture was used to inoculate 100 ml LB medium and grown for another 90 min at 37 °C while shaking. The cells were centrifuged at 4000 x *g* for 15 min at 4 °C using the Beckman® J2-21 centrifuge and the spent medium removed. The pellet was

resuspended in 10 ml cold sterile PIPES buffer and incubated on ice for 20 min. The cells were centrifuged at 4000 x *g* for 10 min at 4 °C and the supernatant discarded. The pellet was resuspended in 2 ml PIPES buffer and the competent cells stored at −70 °C as 100 µl aliquots.

Aliquots of in-house DH5α chemically competent cells were thawed on ice. To each 100 µl aliquot 10 µl of each ligation reaction was added to separate DH5α cell aliquots, mixed gently by tapping and incubated on ice for 30 min. Control reactions contained the digested plasmid with no insert. All samples were heat-shocked for 90 s at 42 °C in a water bath and placed on ice for 5 min.

Aliquots of 10 µl and 90 µl were spread in a sterile manner onto agar plates. The agar plates were made by autoclaving 200 ml of LB medium with 6 g of bacteriological agar, cooling the solution and adding 200 µl of 100 mg/ml ampicillin. This solution was poured into 85 mm petri dishes. The plates were left to set overnight. Once the transformed DH5α cells had been spread onto the plates, they were inverted and placed in an incubator overnight at 37 °C. The plasmid contains an ampicillin resistance gene; therefore the addition of the antibiotic to the agar plate ensures that only plasmid-containing, ampicillin-resistant bacterial colonies will proliferate. After 16 h incubation, colonies were compared to the control plate to check for any self-ligated vectors or inefficient restriction enzyme digestion.

Colony PCR verification: Ten single colonies were removed from the plate using a pipette tip and placed in 10 µl MilliQ™ water. Of this solution, 5 µl was used in a PCR reaction. The bacteria were first lysed by heating at 95 °C for 5 min in a thermocycler and PCR reagents and primers specific for the *P. falciparum* insert were added as outlined in Table 2.1 (subsection 2.3.2). The PCR products were resolved on a 1% agarose gel (subsection 2.2.2) to check that the inserts were present. The rest of the sample was stored at 4 °C.

Restriction digestion verification of extracted plasmid constructs: The remaining 5 µl of the positive colonies was added to 14 ml Falcon™ round bottom tubes containing 2 ml LB

medium with 2 µl of 100 µg/ml ampicillin and placed in a 37 °C orbital shaker overnight at 250 rpm. The next day, stock solutions were made by adding 500 µl of the transformed DH5α cells to 500 µl of 60% sterile glycerol. The stocks were stored at –70 °C. The plasmid constructs were extracted from the remaining cells using the Macherey-Nagel NucleoSpin® Plasmid DNA purification kit as described in subsection 2.4.1. The plasmid constructs (6 µl, ≈ 1 µg) were digested for 1 h with the relevant restriction endonucleases to excise the insert (subsection 2.3.3). Whole and digested plasmids were resolved on a 1% agarose gel (subsection 2.2.2) to verify the presence of the insert.

Positive clones were sent to Inqaba Biotech for DNA sequencing using primers for the T7 promoter in the pET15-b *Pf*MAEBL CCys construct and both the pGEX-4T-2 sense and antisense set in the pGEX-4T-2 *Pf*SORTILIN VPS10 construct due to the large size of the insert. Primer sequences can be found in Appendix A. All sequencing was performed with BigDye® Terminator v3.1 Cycle Sequencing on a ABI3500XL sequencer.

2.4.4.2 Rosetta 2™ (DE3) cells

Colonies which had error-free inserts were grown in 5 ml of LB medium in a 100 ml flask with 100 µg/ml ampicillin. Plasmids were extracted using Macherey-Nagel NucleoSpin® Plasmid DNA purification kit as described in subsection 2.4.1.

Rosetta 2™ (DE3) cells (Novagen, USA) contain tRNAs for seven rare codons (AGA/AGG/CGG-Arg, AUA-Ile, CUA-Leu, GGA-Gly and CCC-Pro) that are not usually present in *E. coli* cells, which enhances the expression of eukaryotic proteins. Aliquots of Rosetta 2 (DE3) cells (20 µl in a 1.5 ml Eppendorf tube) were thawed on ice and 200 ng of the plasmid constructs were added and mixed gently by tapping. The cells were incubated on ice for 30 min, followed by a heat shock at 42 °C for 30 s. The sample was immediately placed on ice for 2 min and 80 µl of LB medium was added then placed in an orbital shaker at 37 °C for 1 h at 250 rpm.

Aliquots of 10 µl and 90 µl were spread in a sterile manner onto agar plates containing ampicillin (100 µg/ml) and chloramphenicol (50 µg/ml) and placed in an incubator at 37 °C over-night. The addition of the two antibiotics selects for the recombinant ampicillin-resistant plasmids that have been taken up by Rosetta 2 (DE3) cells which contain chloramphenicol-resistant plasmids (Novagen, 2016). Stock solutions were made from overnight cultures by adding 500 µl of the Rosetta 2™ (DE3) cells to 500 µl of 60% sterile glycerol.

2.5 Recombinant protein expression

Two induction methods were used to express *Pf*MAEBL CCys and *Pf*SORTILIN VPS10 isopropyl-β-D-1-thiogalactopyranoside (IPTG) and Overnight Express™ Instant TB Medium (Novagen, USA):

2.5.1 IPTG induction

A scraping of a glycerol stock of the transformed Rosetta 2™ (DE3) cells, containing the appropriate vector construct, was added to 5 ml LB (in a 100 ml flask) containing ampicillin (100 µg/ml) and chloramphenicol (50 µg/ml), then placed in a shaking incubator (250 rpm) at 37 °C overnight. This was used to inoculate a 250 ml flask containing 20 ml LB (with ampicillin and chloramphenicol as above). The cell culture was incubated at 37 °C in a shaking incubator (250 rpm) until an $OD_{600\,nm}$ of 0.6-1 was reached. IPTG was added to the culture to a final concentration of 1 mM and different conditions were tested, such as time and temperature. Uninduced cell culture controls were incubated under the same conditions without IPTG. The $OD_{600\,nm}$ was determined as before and protein extraction was carried out. To increase protein production, 2X Yeast extract and tryptone (2YT- 1.6% tryptone, 10% yeast extract, 0.5% NaCl) medium was chosen for expression of proteins for binding assays (Rosano and Ceccarelli, 2014).

2.5.2 Induction using Overnight Express™ Instant TB Medium

A scraping of a glycerol stock of the transformed Rosetta 2™ (DE3) cells, containing the appropriate vector construct, was added to 5 ml LB (in a 100 ml flask) containing ampicillin (100 µg/ml) and chloramphenicol (50 µg/ml), then placed in a shaking incubator (250 rpm) at 37 °C overnight. A volume of 500 µl of the overnight culture was used to seed 20 ml of Overnight Express™ Instant TB Medium (60 g Overnight Express™ Instant TB Medium Powder, 10 ml glycerol, made up to 1 l and autoclaved) containing ampicillin (100 µg/ml) and chloramphenicol (50 µg/ml) then incubated at 37 °C for 6 h to increase cell numbers, followed by a 10 h incubation at 20 °C to limit the accumulation of inclusion bodies (Novagen, 2016). These conditions were used for the His-tagged *Pf*AMA-1 prodomain and the GST-tag control, as determined by Churchyard (2017).

Overnight Express™ Instant TB Medium promotes high density cell growth, several fold higher than that of conventional protocols such as IPTG inducible bacterial systems, due to its specific carbon source blend and magnesium. The unique blend of nitrogen sources sustains high protein expression, induced through lactose activating the *Lac* promoter (Grabski *et al.*, 2005). For a comparative un-induced control of the *E. coli* protein expression system, 500 µl of the 5 ml LB grow was seeded into a 2% glucose LB mixture. In this growth medium a glucose effect will be induced, which will reduce *Lac* transcription and so prevent the expression of the recombinant proteins (Novagen, 2016). The next day the optical density was checked at a wavelength of 600 nm ($OD_{600\,nm}$) using a Thermo Biomate 5 Spectrophotometer. The $OD_{600\,nm}$ indicates the amount of bacterial cells present in the medium and therefore the approximate phase of growth that has been reached. If the $OD_{600\,nm}$ was >1, the bacterial cells were estimated to have reached the stationary phase and protein extraction was carried out.

To compare all samples, similar number of cells were used for sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) analysis (subsection 2.5.4), so the $OD_{600\,nm}$ of all uninduced and induced cultures were normalized to the culture with the lowest $OD_{600\,nm}$.

2.5.3 Recombinant protein extraction

A bacterial cell culture (20 ml) was transferred into a 50 ml Falcon tube and centrifuged (Centrifuge 5702R) at 4 °C for 20 min at 3000 x *g*. The supernatant was removed and the cell pellet was resuspended in 1.5 ml His-binding buffer (50 mM Na_2HPO_4/NaH_2PO_4 buffer, 150 mM NaCl, pH 8) for His-tagged proteins or GST binding/wash buffer (4.2 mM Na_2HPO_4, 2 mM K_2HPO_4, 140 mM NaCl, 10 mM KCl, pH 7.2) for GST-tagged proteins with 1.5 µl Protease 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) and 1.5 µl DNaseI (ThermoScientific, USA). The solution was frozen at −70 °C for 15 min and then thawed at 37 °C for 5 min. This was repeated twice. Cell lysis was predominately facilitated by the use of a Sonoplus HD 3400 sonicator. The thawed suspension was transferred into a 15 ml Nunc tube and sonicated with the MS73 probe for 4 cycles, while immersed in ice water, at an 80% amplitude; each cycle lasting for 30 s, with pulse on for 1 sec, off for 0.5 sec and 5 s between each cycle. An aliquot of 75 µl was collected after the final round of sonication, representing the total protein aliquot. The sample was centrifuged at 4 °C for 20 min at 3000 x *g*. A 75 µl aliquot of the supernatant was collected representing the soluble protein aliquot. The rest was decanted and kept at 4 °C until needed. The pellet was resuspended in 1.5 ml of the appropriate buffer and 75 µl was kept to represent the insoluble protein aliquot. The total, soluble and insoluble fractions were diluted with 75 µl of the respective buffer to allow for efficient loading of samples onto the SDS-polyacrylamide gel. The unbound fraction from the purification was diluted as above. The remaining samples from the purification (subsection 2.6.1 and subsection 2.6.2) were not diluted.

A sample preparation solution was made to prepare samples for SDS-PAGE. This was composed of 40 µl 5x suspension solution (50 mM Tris-HCl, 5 mM EDTA, 5% SDS, 25% sucrose, pH 8.0), 5 µl dye mix (2.5% sucrose, 0.5% bromophenol blue) and 5 µl β-mercaptoethanol. For

every 4 µl of a sample, 1 µl of the sample preparation solution was added. For example to prepare a 50 µl sample 12.5 µl of the sample preparation solution was added. The samples were then vortexed thoroughly and boiled for 2 min.

2.5.4 SDS-PAGE analysis and Western blot

Protein samples were resolved using the Laemmli SDS method (Laemmli, 1970). A 15% or 17% polyacrylamide separating gel was prepared and overlaid with a 4% stacking gel in a MiniProtean BioRad gel cassette that was 8.3 cm x 7.3 cm. The components for casting SDS-polyacrylamide gels are listed in Table 2.5. Appropriate volumes of protein samples were loaded along with 5 µl of a Spectra Multicolour Broad Range Protein Ladder. The protein samples were electrophoresed at 120 V for 1 h, using a BioRad PowerPac™ HC High-Current Power Supply and the MiniProtean Tetra Cell (BioRad, South Africa). Gels were fixed by microwaving for 2 min in reverse osmosis (RO) water and removed. After this was repeated twice the gel was stained using a water-based Coomassie Blue stain (8% G-250 dye, 11.1% HCl in MilliQ™ water) and microwaved for 1 min (Schägger, 2006). The gel was incubated in the stain at room temperature for 1 h on a shaking platform and destained with water until it was sufficient to capture an image using the GeneSnap GeneGenius Geldoc scanning system.

The observed molecular weight of proteins were estimated by determining the relative migration distance and correlating the values to a graph of the logarithm of the molecular weights of the ladder marker used.

Following electrophoresis, gels which were not stained were transferred onto a Hybond™-C extra supported nitrocellulose membrane. The membrane, gel and two sheets of blotting paper were soaked in cold Transblot buffer (25 mM Tris, 192 mM glycine, 0.2% SDS, 40% methanol). The Mini-Transblot™ Cassette was assembled with the gel on the cathode side and the nitrocellulose membrane on the side of the anode. To ensure efficient transfer air bubbles between the membrane and gel were removed using a metal roller. The cassette was placed

Table 2.5: **Components of SDS-polyacrylamide gels**

Component	**15%/17% Resolving Gel**	**4% Stacking Gel**
MilliQ™ Water	1025 µl/620 µl	1630 µl
10% (w/v) SDS	53 µl	6.7 µl
4X Resolving Buffer (1.5 M Tris-HCl, pH 8.8)	2500 µl	-
4X Stacking Buffer (0.5 M Tris-HCl, pH 6.8)	-	833 µl
30% (w/v) Acrylamide	5000 µl/5680 µl	433 µl
1% (w/v) Bis-acrylamide	1100 µl	333 µl
10% (w/v) ammonium persulphate	67 µl	67 µl
TEMED	5 µl	2.5 µl
Final Volume	10025 µl	3305.2 µl

in the MiniProtean Tetra Cell, filled with Transblot buffer and an ice pack supplied by the manufacturer. The proteins were transferred at 100 V for 1 h in a cold room.

2.5.5 Immunoblot analysis

The nitrocellulose membranes were washed in Tris buffered saline (TBS- 25 mM Tris-HCl, 500 mM NaCl, pH 7.5) for 5 min and the gels were placed in water-based Coomassie Blue stain (as described in subsection 2.5.4) to ascertain transfer efficiency. The blots were drained and briefly coated in 1 ml Ponceau S (1% Ponceau S, 7% glacial acetic acid) to detect protein transfer. The membranes were placed in MilliQ™ water to rinse off excess dye. The banding of the Spectra Protein Ladder was recorded on the membrane in pencil, which was returned to the water for 10 min to completely remove the Ponceau S stain.

The following protocol (adjusted from Qiagen, 2002) was used to detect the recombinant His-tagged *Pf*MAEBL CCys protein: The membrane was placed in anti-His horse radish peroxidase (HRP)-conjugated blocking buffer (Qiagen, UK) on a shaking platform for 1 h to block the areas of the nitrocellulose membrane that did not contain transferred protein. After removing the blocking buffer the membrane was incubated in a 1:2000 dilution of QIAexpress anti-His HRP-conjugated antibody (Lot no.: 148049181 and 154018594, Qiagen, UK) for 1 h at room temperature with gentle shaking. The membrane was washed twice with shaking in 0.05%

Tween-TBS solution for 5 min each to remove any unbound antibody. This was followed by a rinse in TBS for 5 min.

The following protocol (adjusted from GE Healthcare Life Sciences, 2014) was used to detect the recombinant GST-tagged *Pf*SORTILIN VPS10 protein: The membrane was placed in a 3% bovine serum albumin (BSA)/TBS blocking buffer on a shaking platform for 1 h. After removal of the blocking buffer the membrane was incubated in a 1:12500 dilution of anti-GST HRP-conjugated antibody (Lot no.: 9475836 and 9800159, GE Healthcare Life Sciences, UK) for 1 h at room temperature with gentle shaking. The membrane was washed twice with shaking in 0.1% Tween-TBS solution for 5 min each to remove any unbound antibody. This was followed by a rinse in TBS for 5 min. Recombinant GST was used as a positive control for the immunoblot procedure.

Analysis of the anti-His and anti-GST immunoblots followed the same visualization procedure since they both use HRP-conjugated antibodies. A volume of 2 ml of SuperSignal® West Pico Chemiluminescent Substrate was applied to cover the membrane for 5 min before visualization of the chemiluminescent signal using the GeneSnap GeneGenius Geldoc scanning system. The addition of the luminol substrate and peroxide allows for the HRP enzyme to catalyse the oxidation, thereby releasing energy in the form of light. After visualization the membrane was stained in amido black (10% acetic acid, 25% isopropanol, 0.1% amido black) for 5 min and then destained for 30 min in 10% acetic acid, 10% methanol; 30 min in 10% acetic acid and then photographed using the GeneSnap GeneGenius Geldoc scanning system.

2.6 Recombinant protein purification

Soluble fractions of the extracted *E. coli* proteins were used to purify the recombinant proteins. Purification of His-tagged recombinant fusion proteins was performed using the MagneHis™ Protein Purification System. This system uses paramagnetic precharged Ni particles. The His-tag competitively binds to the Ni-substrate (Promega Corporation, 2013). The purification of

glutathione S-transferase (GST) fusion proteins was carried out using the MagneGST™ Protein Purification System (Promega, USA). The system uses paramagnetic glutathione particles. The GST enzyme binds to its substrate, glutathione, in a lock-and-key manner (Promega Corporation, 2014). When the respective particles are added to a soluble protein fraction containing His/GST-tagged proteins, the fusion proteins bind to the paramagnetic particles. Using a magnetic particle separator, the particles are rinsed to remove unbound proteins. The bound fusion proteins are eluted with a buffer containing a high concentration of imidazole or glutathione, which competes for the Ni particles or glutathione-coated particles, respectively. The protocols for purification were carried out as described by the manufacturer with minor changes, such as 16 h binding at 4 °C (Promega Corporation, 2013, 2014).

2.6.1 Purification of His-tagged proteins

A volume of 5 µl of MagneHis™ particle slurry was washed three times with 1 ml His-binding buffer and resuspended in 100 µl of His-binding buffer in a 2 ml Eppendorf tube. A cell pellet from a 20 ml culture was resuspended in 1.5 ml of His-binding buffer. After sonication, the sample was centrifuged as in section 2.5. A final concentration of 1 mM imidazole (3 µl of a 500 mM solution) was added to the soluble fraction, which was added to the particles. The mixture was incubated at 4 °C overnight on an Intelli-mixer RM 2M Skyline rotating 360° at 25 rpm to allow the recombinant protein to bind to the particles. Beads were separated using the magnetic particle separator and 75 µl of the unbound fraction was kept and diluted 1:1 with His-binding buffer for SDS-PAGE analysis. The remaining unbound fraction was stored at 4 °C to rebind the recombinant protein if the first attempt was not sufficient. The particles were washed in 2 ml 20mM imidazole in His-binding buffer twice on an Intelli-mixer (90°, 99 rpm). An aliquot (150 µl) of each wash was kept for analysis. Samples were eluted using the Intelli-mixer at 30° rotation, 99 rpm for 20 min with 50 mM Na_2HPO_4/NaH_2PO_4 buffer, 150 mM NaCl, 0.5 M, imidazole (pH 7.5).

2.6.2 Purification of GST-tagged proteins

A volume of 5 µl of MagneGST™ particle slurry was washed three times with 1 ml GST binding/wash buffer and resuspended in 100 µl GST binding/wash buffer in a 2 ml Eppendorf tube. A cell pellet from a 20 ml culture was resuspended in 1.5 ml of GST binding/wash buffer. After sonication, the sample was centrifuged as in section 2.5. The soluble fraction was added to the particles. The mixture was incubated at 4 °C overnight on an Intelli-mixer RM 2M Skyline rotating 360° at 25 rpm. Beads were separated and 75 µl of the unbound fraction was kept and diluted 1:1 with GST binding/wash buffer for SDS-PAGE analysis and the remaining unbound fraction was stored at 4 °C, to rebind to the particles if necessary. The particles were washed five times in 2 ml GST binding/wash buffer on an Intelli-mixer (90°, 99 rpm). An aliquot (150 µl) of each wash was kept for analysis. For elutions, 100 µl of GST Elution buffer (500 mM L-Glutathione, 500 mM NaCl, 50 mM Tris-HCl, pH 8.1) was added to the particles. Samples were eluted using the Intelli-mixer at 30° rotation, 99 rpm for 20 min.

Purifications were extrapolated for different culture volumes (i.e multiple 2 ml Eppendorf tubes).

2.7 *In vitro* binding assays

2.7.1 Dialysis of protein samples

All recombinant proteins were dialysed with a Slide-A-Lyzer MINI dialysis unit (ThermoScientific, 2012) which has a molecular weight cut-off of 10 kDa. *Pf*MAEBL CCys (13 kDa) and *Pf*AMA-1 prodomain (11 kDa) were dialysed into His-binding (pH 8.0) or TBS (25 mM Tris-HCl, 150 mM NaCl, pH 7.5) buffer. *Pf*SORTILIN VPS10 (86 kDa) and GST (26 kDa) were dialysed into GST binding/wash (pH 7.2) or TBS (pH 7.5) buffers. Dialysis membranes were first incubated with MilliQ™ water for 2 min to equilibrate the device and then removed. Samples of 100 µl were loaded onto the membrane of dialysis tubes and caps were attached.

The tubes were attached to a Styrofoam floater, which was placed into 50 ml of the respective buffers for 2 h and repeated 3 times, as suggested by the manufacturer's specifications. Samples were placed such that the membrane made contact with the buffer, but care was taken to ensure that the sample was level with, or slightly above the buffer. Samples were collected and stored at 4 °C until the binding assays were performed later on the same day.

All samples were quantified by SDS-PAGE analysis against increasing concentrations of an Albumin standard (made from a 2 mg/ml stock) prior to conducting the binding assays. Densitometry by the GeneSnap GeneGenius Geldoc scanning system was used to determine the amount of each sample available. All assays were set up at constant ratios and not by exact concentrations when repeated. This was due to the variability of final protein yield for each attempt at conducting a binding assay.

2.7.2 Binding assay of immobilized *Pf*MAEBL CCys or *Pf*AMA-1 prodomain to *Pf*SORTILIN VPS10

Individual binding reactions were conducted in 2 ml Eppendorf tubes. For each assay tube 5 µl of MagneHis™ particles were washed and resuspended in 100 µl of His-binding buffer. A constant amount of His-tagged *Pf*MAEBL CCys or *Pf*AMA-1 prodomain (in either ng or pmoles) was immobilized on MagneHis™ particles in each tube by incubation for 1 h at room temperature using an Intelli-mixer (60°; 25 rpm). Using a magnetic separator the unbound fraction was collected and the particles containing reattached *Pf*MAEBL CCys or *Pf*AMA-1 prodomain were washed three times with 500 µl of TBS by shaking (60°; 99 rpm). Increasing amounts of free GST-tagged *Pf*SORTILIN VPS10 (in either ng or pmoles) were then applied to the immobilized His-tagged samples in a total volume of 100 µl of TBS for 1 h at room temperature using an Intelli-mixer (60°; 25 rpm). The unbound fractions were collected and the particles were washed twice in 500 µl TBS to remove unbound *Pf*SORTILIN VPS10 . The particles were then prepared for SDS-PAGE analysis as described in subsection 2.5.4. To account for non-specific

binding, purified GST as well as heat-denatured (boiled for 2 min) GST-tagged *Pf*SORTILIN VPS10 were applied to immobilized His-tagged proteins. A control to determine if there was non-specific binding of *Pf*SORTILIN VPS10 to the MagneHis™ particles was also included.

2.7.3 Binding assay of immobilized *Pf*SORTILIN VPS10 to *Pf*MAEBL CCys or *Pf*AMA-1 prodomain

To verify interactions the assay was repeated (as described in subsection 2.7.2) with GST-tagged *Pf*SORTILIN VPS10 as the immobilized bait protein and *Pf*MAEBL CCys or *Pf*AMA-1 prodomain in solution.

Binding interactions were conducted in individual 2 ml Eppendorf tubes. For each 2 ml Eppendorf tube 5 µl of MagneGST™ particles were washed and resuspended in 100 µl of GST binding/wash buffer. Constant amounts of *Pf*SORTILIN VPS10 (ng or pmoles) were immobilized to the particles by incubation for 1 h at room temperature using an Intelli-mixer (60°; 25 rpm). Using a magnetic separator the unbound fraction was collected and the particles containing reattached *Pf*SORTILIN VPS10 were washed three times with 500 µl of TBS by shaking (60°; 99 rpm). Increasing amounts of each His-tagged protein (ng or pmoles) were then applied to the immobilized *Pf*SORTILIN VPS10 samples in a total volume of 100 µl of TBS for 1 h at room temperature using an Intelli-mixer (60°; 25 rpm). The unbound fractions were collected and the particles were washed twice in 500 µl TBS to remove unbound *Pf*MAEBL CCys or *Pf*AMA-1 prodomain protein. The particles were then prepared for SDS-PAGE analysis as described in subsection 2.5.4. To account for non-specific binding heat-denatured (boiled for 2 min) His-tagged proteins were applied to immobilized GST-tagged *Pf*SORTILIN VPS10 protein. Controls to determine if there was non-specific binding of the His-tagged proteins to the MagneGST™ particles were also included.

2.8 Biopanning using phage display libraries

To identify novel binding partners of *Pf*SORTILIN VPS10 biopanning was used against *P. falciparum* phage display libraries (Lauterbach *et al.*, 2003) according to the Novagen T7Select® System manual (Novagen, 2002; see Appendix C for the phage vector used). *P. falciparum* phage display libraries used for this study were created in the Coetzer lab by Dr D Liebenberg.

2.8.1 Bacterial cultures for biopanning

A BLT5403 *E. coli* culture was started from a glycerol stock in a 100 ml Erlenmeyer flask containing 10 ml M9LB medium (18.7 mM NH_4Cl, 22 mM KH_2PO_4, 22 mM Na_2HPO_4, 0.4% glucose, 22 mM $MgSO_4$, in LB) supplemented with 100 µg/ml ampicillin. The culture was incubated in a shaking incubator at 250 rpm at 37 °C overnight until an $OD_{600\,nm}$ > 3.0 was reached. The BLT5403 culture was used to set up new cultures for phage amplification and plaque assays (subsection 2.8.2) or creating glycerol stocks (subsection 2.8.5).

2.8.2 Phage amplification and plaque assays

To amplify the phage 100 µl of an overnight BLT5403 culture was used to seed 50 ml of M9LB in a 250 ml flask, supplemented with 100 µg/ml ampicillin . This was incubated for 2 h to 3 h at 37 °C at 250 rpm until the $OD_{600\,nm}$ was between 0.6 and 1. The log phase culture was inoculated with glycerol stock scrapings of the three *P. falciparum* phage display libraries and incubated overnight as in subsection 2.8.1. NaCl (at a final concentraion of 0.5 M) was added to the overnight BLT5403/phage culture to allow for the lysis of bacterial cells and the release of individual phage. Phage were isolated by centrifugation of the sample (Eppendorf Centrifuge 5702R) at 3000 x *g* for 20 min at 4 °C. The supernatant containing phage was transferred to a tube and stored at 4 °C until the plaque assay was performed.

A plaque assay was able to determine the viral load or the "titre" of the phage lysates. A series of phage dilutions ranging from 10^6 to 10^9 were prepared, by adding an appropriate volume of the phage lysates to sterile TBS in a final volume of 1 ml. To prepare the sample for phage titration, 100 µl of each dilution was added to 250 µl BLT5403 culture with an $OD_{600\,nm}$ of >1.0. Each sample was then combined with 3 ml of molten top agarose (1% tryptone, 0.5% yeast extract, 85 mM NaCl, 0.6% agarose) and poured onto pre-warmed agar plates, containing 50 µg/ml ampicillin. Once the top agarose mixture had set, the plates were inverted and incubated for 2 h at 37 °C. To calculate the viral titre the plaques on each plate were counted. The following equation was used for each dilution to determine the number of plaque forming units (pfu) per milliliter of phage lysate:

$$\text{Viral Titre (pfu/ml)} = \text{number of plaques X dilution X } 10$$

Where 10 accounts for the 0.1 ml of initial phage plated.

2.8.3 Preparation of *Pf*SORTILIN VPS10 for biopanning

GST-tagged *Pf*SORTILIN VPS10 was purified (subsection 2.6.2) and dialysed (subsection 2.7.1) into GST binding/wash buffer. The sample was stored at 4 °C with 1.5 µl of Protease Inhibitor cocktail Set III until required for biopanning. This protein sample was quantitated using the GeneSnap GeneGenius Geldoc scanning system. An amount of 3 µg of *Pf*SORTILIN VPS10 was rebound to 5 µl of MagneGST™ particles for 1 h in a total volume of 100 µl of GST binding/wash buffer using an Intelli-mixer (60°, 25 rpm). Binding of *Pf*SORTILIN VPS10 to the particles was repeated for each round of biopanning.

2.8.4 Biopanning

After determining the viral titre of the original phage display library (subsection 2.8.2), a 1 x 10^7 phage dilution was made using TBS in a total volume of 500 µl. A volume of 5 µl of washed

MagneGST™ particles was incubated with the phage for 1 h at room temperature with rotation using an Intelli-Mixer (90°, 25 rpm). This step prevented subsequent non-specific interactions by removing any phage that bound to the MagneGST™ particles. A magnetic separator was used to collect the unbound phage fraction.

The GST binding/wash buffer of the immobilized *Pf*SORTILIN VPS10 was removed using a magnetic separator before the unbound phage fraction was added. The phage-protein-particle mixture was incubated for 1 h at room temperature with an Intelli-Mixer (90°, 25 rpm). The MagneGST™ particles were collected and washed five times with 2 ml of 0.05% Tween-TBS by inversion (90°, 90 rpm) on an Intelli-Mixer for 10 min at room temperature. The washed particles were transferred to 50 ml of a log phase BLT5403 culture in a 250 ml flask and incubated in a shaking incubator at 250 rpm at 37 °C overnight to allow for phage amplification. The phage lysate was collected and used in a plaque assay (subsection 2.8.2) to determine the viral titre for the next round of biopanning. Glycerol stocks were prepared for long term storage of the phage lysates as described in subsection 2.8.5. Four rounds of biopanning were performed to enrich for specific protein interactions with the purified recombinant *Pf*SORTILIN VPS10 . After the final round of biopanning, the plaque assay was performed according to subsection 2.8.2.

2.8.5 BLT5403 and bacteriophage glycerol stocks

Once a BLT5403 *E. coli* culture had an $OD_{600\,nm}$ > 3.0, glycerol stocks were prepared by combining 1.5 ml of the culture with 150 µl of 80% sterile glycerol (prepared using MilliQ water and filter sterilised) in an Eppendorf tube (Novagen, 2011). The contents were mixed by inversion and stored at −70 °C. Glycerol stocks of the phage lysates were prepared as described above, including the initial phage before biopanning and the phage lysates after each round of biopanning.

2.8.6 PCR amplification and sequencing of selected phage

Following four rounds of biopanning, individual plaques were selected and used for PCR analysis and preparation of stocks for long term storage (subsection 2.8.5). Isolated plaques were extracted from the top agarose layer of the agar plates using a sterile pipette tip and transferred to a 600 µl PCR tube containing 50 µl 10 mM EDTA (pH 8.0). The tube was vortexed briefly and incubated for 10 min at 65 °C. A PCR reaction was set up using 1 µl of the phage, DreamTaq polymerase (ThermoScientific, South Africa) and the UP and DOWN T7 backbone primers. Primer sequences and PCR conditions are listed in Appendix D. The PCR products were resolved on a 1% agarose gel against a low range DNA standard (1031-100 bp) as described in subsection 2.2.2. The empty cassette product is 237 bp, thus products of 300 bp or larger were considered for further analysis. Selected phage were sent to Inqaba Biotec™ for sequencing. The sequencing reactions were performed on the amplified PCR products using the T7Select-UP-2 primer (see Appendix D for the primer sequence).

Chapter 3

Results

3.1 Identification of protein domains

3.1.1 *Pf*MAEBL CCys domain (PF3D7_1147800)

Previous work by Melanie Wepener (2014) has shown that the CCys domain is sufficient and necessary for apical trafficking of *Pf*MAEBL and thus it was cloned into the pET15-b vector. Models of the domain structure of the protein and the exon/intron gene structure are shown in Figure 3.1 (Blair *et al.*, 2002; Singh *et al.*, 2004; Fu *et al.*, 2005; PlasmoDB, 2018).

The *Pf*MAEBL CCys domain is 114 aa and coded within exon II, so DNA could be used for PCR. The primers that were designed (Appendix A) amplified the 342 bp (114 aa) CCys domain and no other flanking regions, which can be seen in Figure 3.1. Between the two possible ORFs (transcripts PF3D7_1147800.1 and PF3D7_1147800.2), the CCys rich region remained conserved. The difference between the two transcripts lies in the transmembrane domain and the cytoplasmic tail and therefore does not effect the coding sequence of the CCys domain.

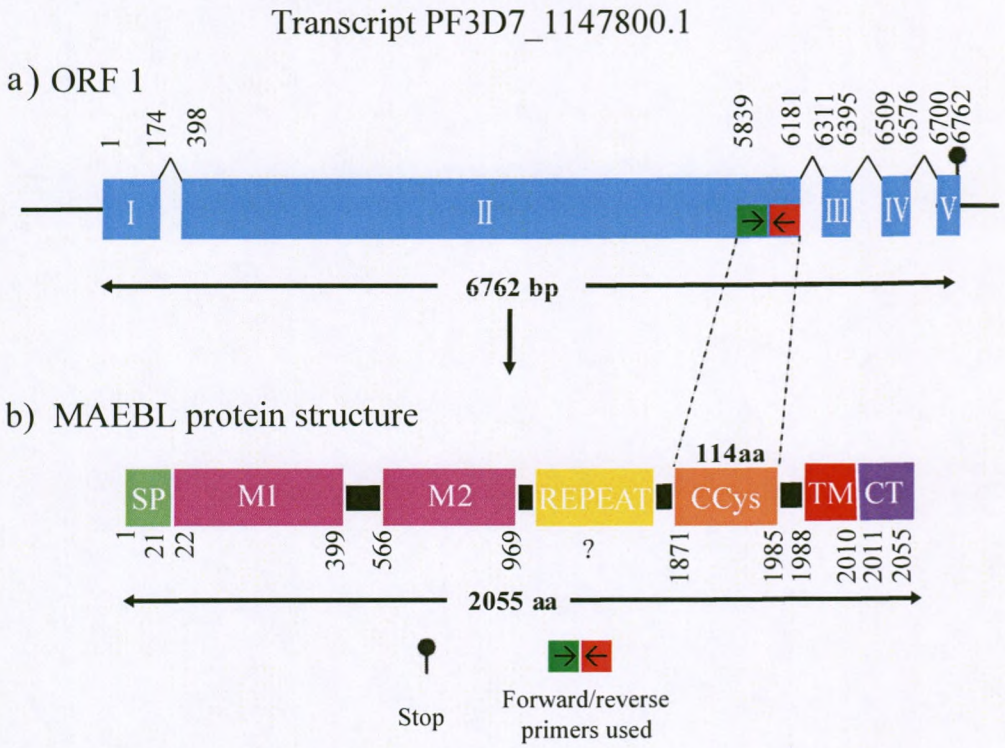


Figure 3.1: **Schematic representation of one transcript of *Pf*MAEBL (ORF 1). a) *Pfmaebl* gene structure and b) *Pf*MAEBL protein structure. SP: signal peptide; M1/M2: MAEBL-specific binding domain; CCys: C-terminal cysteine-rich region; TM: transmembrane domain; CT: cytoplasmic tail. The arrows indicate where the primers bind and what region they amplify. Alternative splicing does not effect the CCys domain. The coding region (dotted line) and size of CCys rich domain are based on Singh *et al.*, 2004; Blair *et al.*, 2002; PlasmoDB, 2018.**

3.1.2 *Pf*SORTILIN VPS10 domain (PF3D7_1451800)

To identify which domain of *Pf*SORTILIN would most likely be involved in apical protein trafficking, it was compared to the *T. gondii* homologue (TGME49_290160). The whole protein BLAST search between *Pf*SORTILIN and *Tg*SORTLR revealed a 50% sequence identity. Smart Protein program (www.smart.embl-heidelberg.de) and Phyre 2 were used to identify a domain of *Pf*SORTILIN. The VPS10 domain of *Pf*SORTILIN was identified and considered because the homologous domains (VPS10) in *T. gondii* were shown to interact with adaptin and clathrin and suggested they play a role in protein sorting and trafficking (Sloves *et al.*, 2012).

The VPS domain sequences were aligned to determine the sequence identity of the homologous domains using the Multiple Sequences Comparison Log-Expectation algorithm because of its fast-speed and high-accuracy outputs (Edgar, 2004). The VPS10 domains of *Pf*SORTILIN and *T. gondii* had a sequence identity of 51% as shown in Figure 3.2. The sequence homology was 73%. This is not unexpected since both species belong to the same phylum (Gardner *et al.*, 2002). *P. falciparum* genes are larger than those of most organisms, but this alignment revealed that the *T. gondii* domains had a longer aa sequence.

```
PfSORTILIN  MKKKLEKQKKLLEYEKGYYS-----NKVKMRNMLF    30
TgSORTLR    MYTSRTPHGGLRRWPLSRSSVSAVASRVLSRSVAF    35

PfSORTILIN  FKSI--------------------------------    34
TgSORTLR    HRRLASFLLSRKKTPLCALSSRPSGLQASSGVQTV    70

PfSORTILIN  ---------------GNKNHLFLVFLLFL--------    48
TgSORTLR    TMAHSFARGRPPAAGRTLAVFLAFLAVLSPSLLPG   105

PfSORTILIN  FFS--------FSCINFTQCQITKKKVSVSEINFDS    76
TgSORTLR    FFSEDAQKRESSVLVANAASYANKKVSVSEVSFDS   140

PfSORTILIN  AVDDVQWCGNNHMTVLVKTVKGRLYRSSDGGKIWT   111
TgSORTLR    HIEDIQWCGTDHRTILLKTRRGRLYRSQDGGKSWT   175

PfSORTILIN  NITGNLSDKENNKSDPSSNHMNEMTTVDLIMVNSV   146
TgSORTLR    EITDLLK----------SSEAATGTVAVDSIIVSPV   201

PfSORTILIN  NKNIVLIIGNQKNHYISEDSGETFRLLNYQNKINF   181
TgSORTLR    DKRVVLIVGSKRNHFISEDSAATFRRLKYKNTIHN   236

PfSORTILIN  WQFHSTKTHWALVSSWTEACYSNDDNSG------E   210
TgSORTLR    FHFHPTRPKYAILSTWTDACYSGSGTASRAQSQQD   271

PfSORTILIN  CMQTLSLTKDLGRTFQLIDIYVVQFNWGDKESHLE   245
TgSORTLR    CNHQLFYTRDLGRSFKLVADYVVQFSWGDKKLGNT   306

PfSORTILIN  DTIYYTRHKNRNGHQQRFSGWSKDVDFVATHNFGK   280
TgSORTLR    DHIFFTQHRGRSGDQPRYGGWSKNVDLMYTPDFGA   341

PfSORTILIN  SVDVLVKQGNKFLISNGYIFVAKLNDVIKQTVNMM   315
TgSORTLR    TITRLVYRGNKFLLSNGYFFVAKVKDAAKQTVSLL   376
```

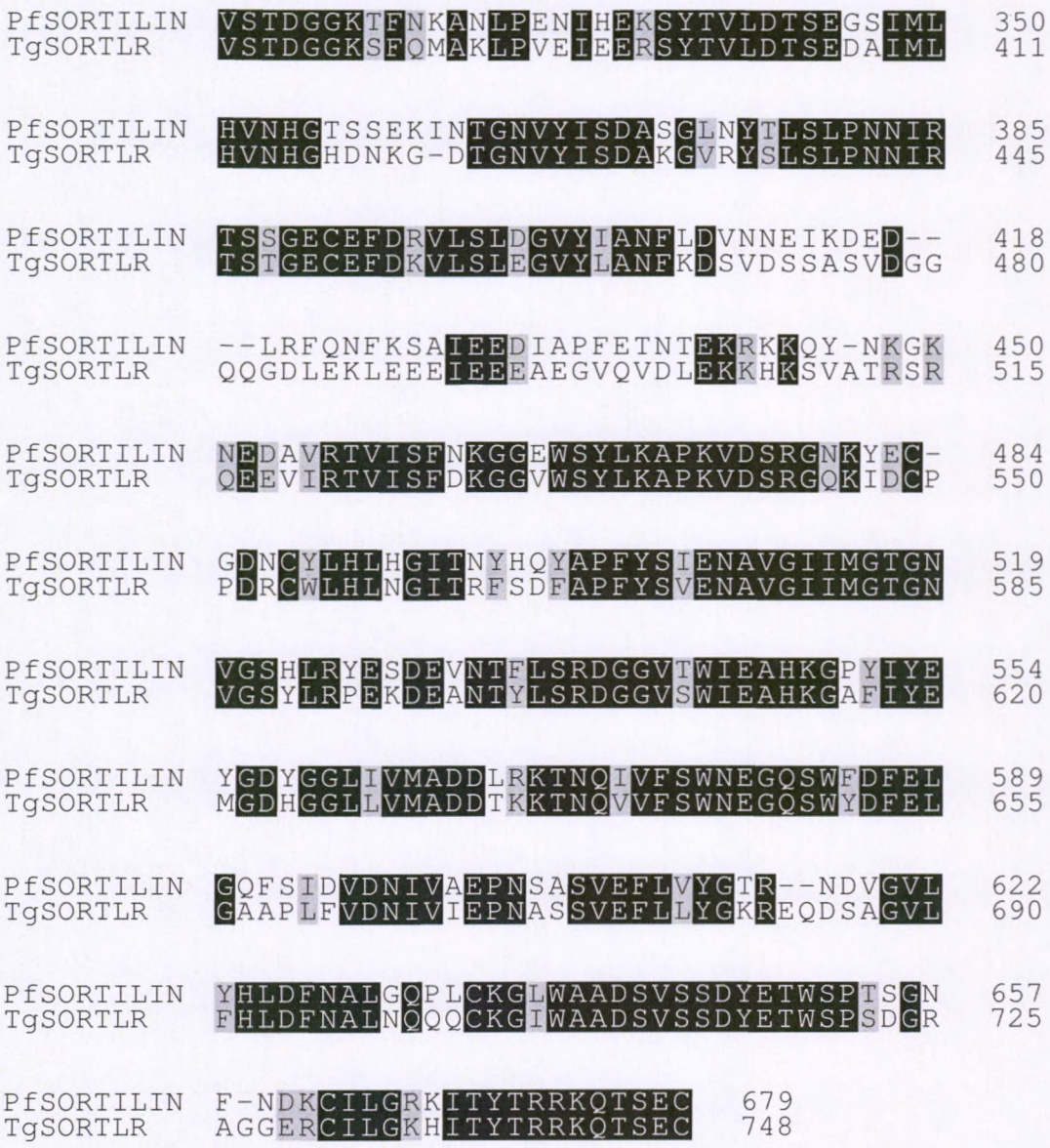


Figure 3.2: **MUSCLE alignment of VPS10 domain of *Pf*SORTILIN and VPS10 domain of *Tg*SORTLR**

aa: identical residues, aa : similar residues, -: gap residues.

A sequence model of *Pf*SORTILIN was made using information gathered from various studies (Sloves *et al.*, 2012; Krai *et al.*, 2014; PlasmoDB, 2018). The *Pf*SORTILIN VPS10 domain is 679 aa (60 kDa) and coded by exon I, therefore to amplify this region genomic DNA could be used. The primers (Appendix A) were designed to amplify a large segment of the VPS10 domain (541 aa) of *Pf*SORTILIN and no other flanking regions (Figure 3.3). The two predicted

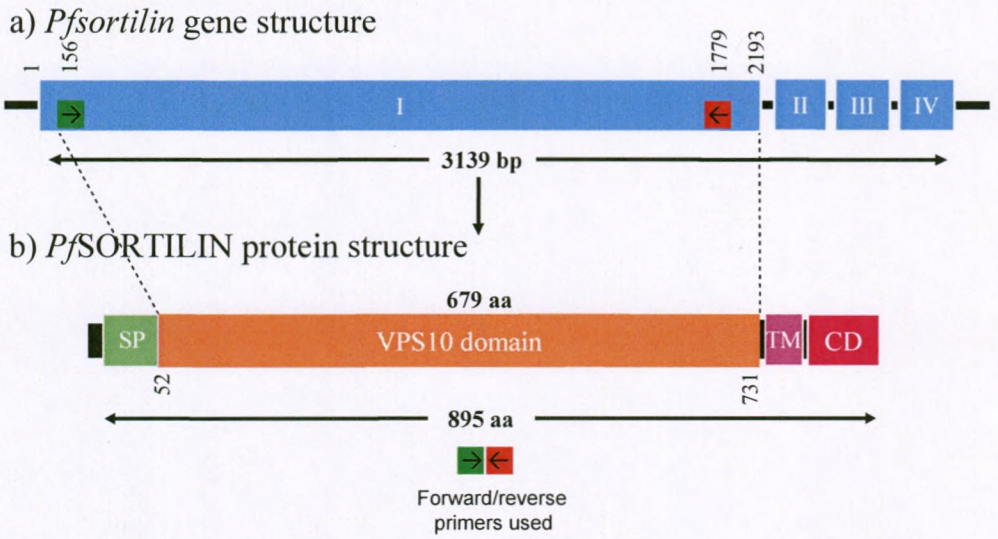


Figure 3.3: **Schematic representation of *Pf*SORTILIN. a) *Pfsortilin* gene structure and b) *Pf*SORTILIN protein structure. SP: signal peptide; TM: transmembrane domain; CD: cytoplasmic domain. The coding region (dotted line) and size of the VPS10 domain are shown. The arrows indicate where the primers bind and the amplified region. This figure was adapted from Sloves *et al.* (2012), Krai *et al.* (2014) and PlasmoDB (2018).**

tertiary structure models show a globular protein structure predominantly comprised of β-sheet formations with a defined central hole/gap (Figure 3.4). The models also revealed that two protruding arms existed which comprised mainly of the signal peptide (N-terminal end) and the transmembrane and cytoplasmic tail domains (C-terminal end).

These domains and a small segment of the C-terminal end (138 aa) of the VPS10 domain were not included in the cloned recombinant construct (Figure 3.4).

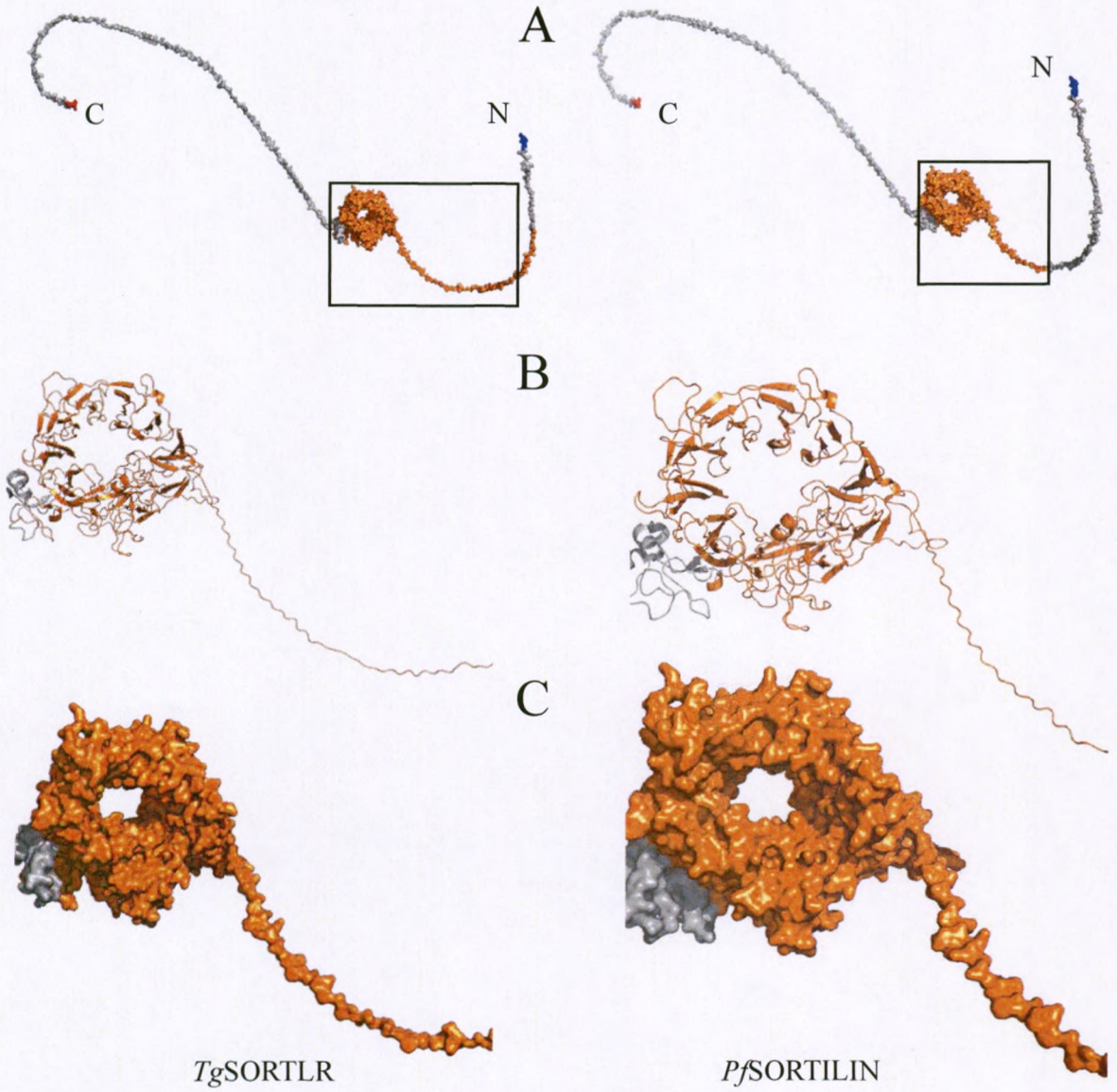


Figure 3.4: **3-D models of the tertiary structure of *Tg*SORTLR (left) and *Pf*SORTILIN (right). A: Surface model of the entire protein, B: ribbon model of VPS10 domain (orange) and C: zoomed in surface model of the cloned VPS10 domain are shown. Both models were based on the crystal structure of the VPS10 domain of human SORTILIN with 62% and 71% of residues modelled at >90% confidence, respectively. Models were generated using the Phyre 2 online tool (www.sbg. bio.ic.ac.uk) and visualized using the PyMOL computer software.**

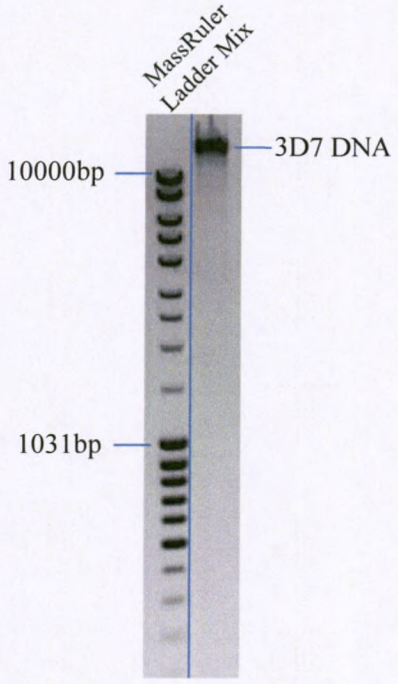


Figure 3.5: **Agarose gel (1%) electrophoresis of 3D7 *P. falciparum* DNA**

3.2 Genomic DNA extraction from *P. falciparum*

From a 5 ml trophozoite-dominant parasite culture, with a parasitaemia of $\approx 5\%$, genomic DNA was extracted and eluted in 60 µl of nuclease-free water. The A_{260} absorbance reflected a yield of 1.75 µg at a concentration of 28 ng/µl (Table 3.1). The A_{260}/A_{280} ratio indicated the DNA was pure (between 1.8-2) and A_{260}/A_{230} confirmed the absence of contaminants such as phenol (between 1.8-2.3).

Table 3.1: **Absorbance values and concentration of genomic DNA of 3D7 parasites**

A_{230}	A_{260}	A_{280}	A_{260}/A_{230}	A_{260}/A_{280}	Yield	[Sample]
1.6	3.6	1.9	2.3	1.9	≈ 1.7 µg	28 ng/µl

The DNA sample was electrophoresed and showed a single high molecular weight band with good integrity (Figure 3.5). Densitometry revealed a similar yield and concentration of the sample compared to the absorbance values. This sample was used for PCR.

3.3 Cloning

3.3.1 PCR amplification

From a 50 µl PCR reaction, 1 µl of each amplicon was loaded and electrophoresed on an agarose gel along with 5 µl of the MassRuler™ Ladder Mix. All PCR products migrated according to their expected sizes (Figure 3.6). In a total reaction volume of 50 µl ≈ 1.5 µg of the *Pf*MAEBL CCys amplicon and ≈ 2.5 µg of the *Pf*SORTILIN VPS10 amplicon were obtained. After the products were cleaned, to remove the PCR buffer and primers, the A_{260} values showed that ≈ 1.3 µg of the *Pf*MAEBL CCys amplicon and ≈ 1.8 µg of the *Pf*SORTILIN VPS10 amplicon were retained in a final volume of 30 µl nuclease-free water.

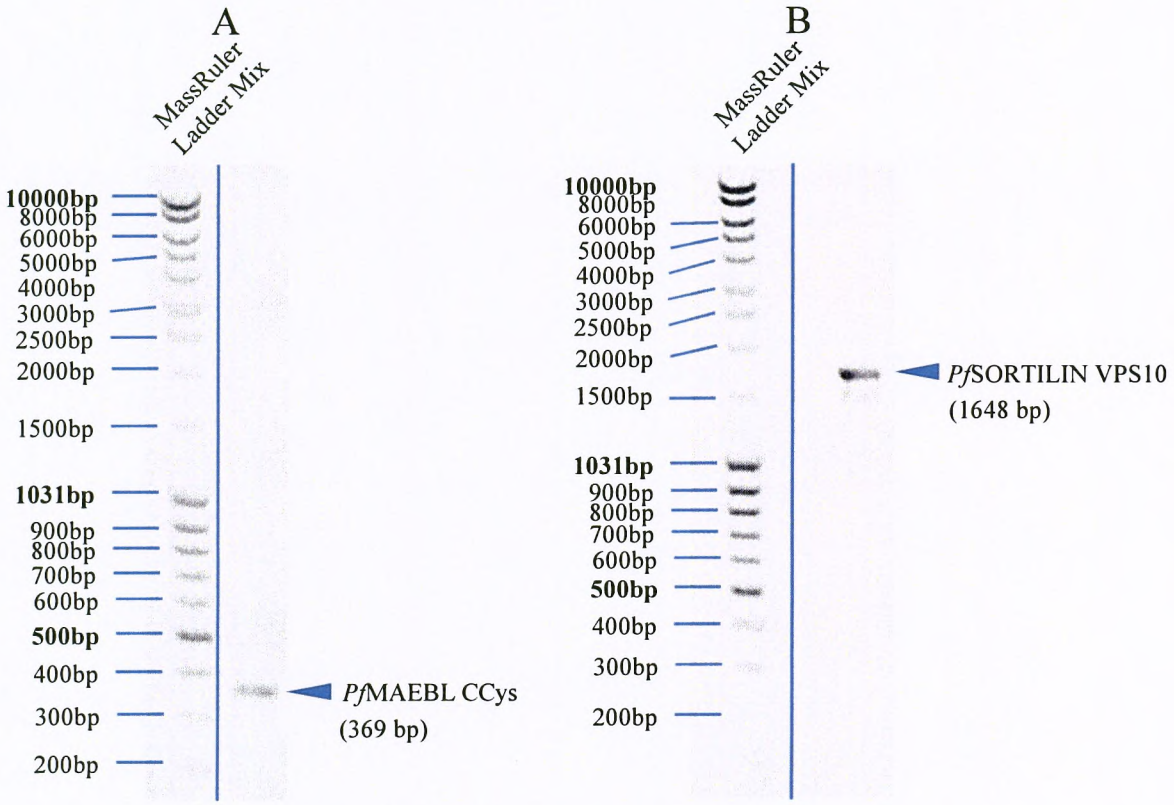


Figure 3.6: **Agarose gel (1%) electrophoresis of PCR amplicons. A: *Pf*MAEBL CCys and B: *Pf*SORTILIN VPS10**

3.3.2 Ligation reactions

The plasmids and amplicons were digested with the appropriate restriction enzymes. All samples (1 µl of each) were electrophoresed and quantitated by densitometry relative to the MassRuler™ Ladder Mix (Figure 3.7).

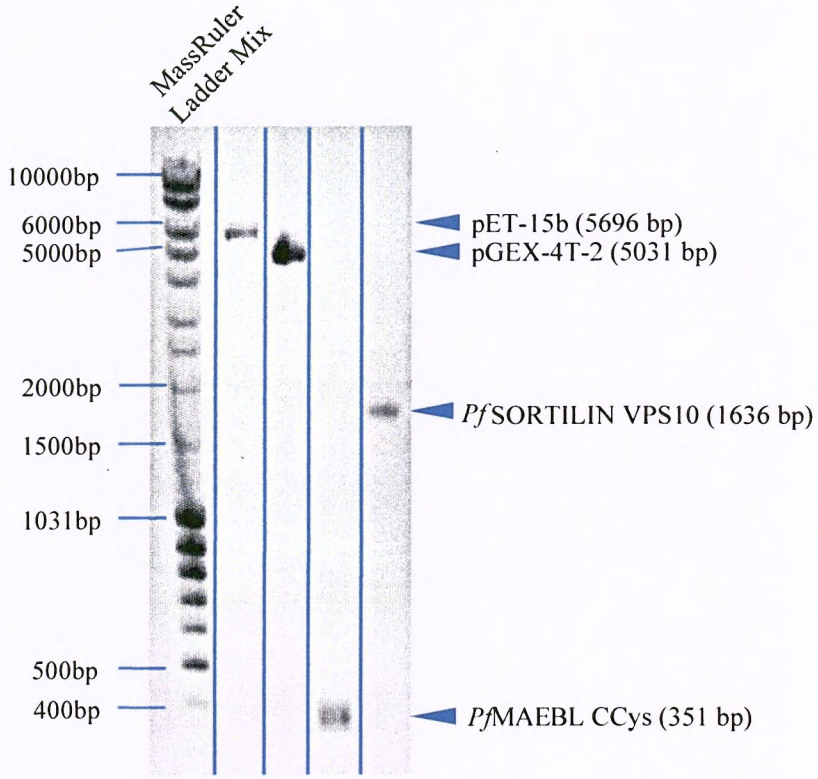


Figure 3.7: **Agarose gel (1%) electrophoresis of digested plasmids, *Pf*MAEBL CCys and *Pf*SORTILIN VPS10 inserts. The first lane is the MassRuler™ Ladder Mix (5 µl) used for quantitation.**

The amounts of each sample required in the ligation reaction for an insert:plasmid molar ratio of 3:1 were calculated in Table 3.2. The total DNA in the ligation reaction did not exceed 200 ng.

Table 3.2: **Ligation reaction insert:plasmid (3:1) molar ratios. All samples are digested.**

Sample	Size (bp)	Amount needed (ng)	[Sample] (ng/µl)	Amount used (µl)
pET-15b	5696	167.4	20.0	8.4
*Pf*MAEBL CCys	351	32.6	27.0	1.2
pGEX-4T-2	5031	100.9	50.0	2.0
*Pf*SORTILIN VPS10	1636	99.1	11.0	9.0

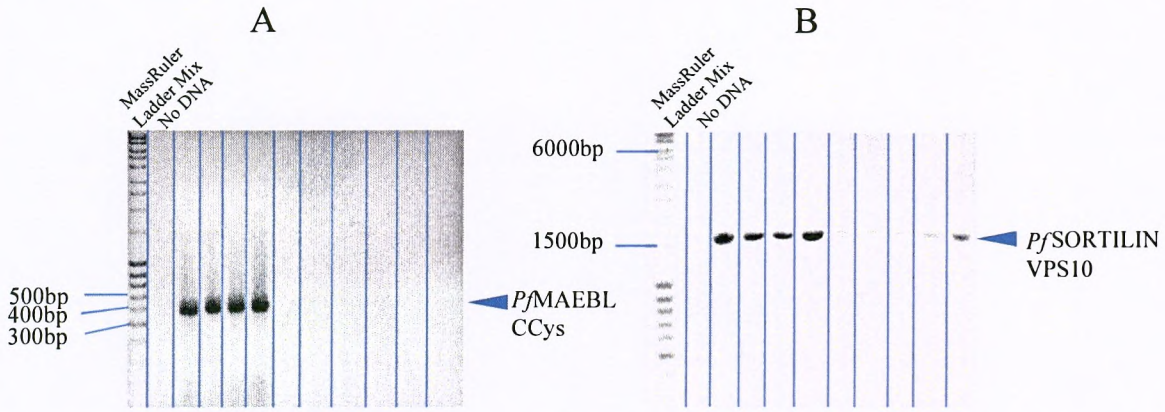


Figure 3.8: **Colony PCR verification of transformed DH5α cells. Agarose gel (1%) of PCR amplicons of A: *Pf*MAEBL CCys and B: *Pf*SORTILIN VPS10 constructs with a negative control for the PCR (no DNA included).**

3.3.3 Recombinant plasmid construct verification

Following ligation and transformation of DH5α cells, plasmid constructs were verified using colony PCR with insert primers as a first line screen and restriction enzyme digestion to confirm the size of the insert. Colony PCRs revealed positive clones in cells containing either *Pf*MAEBL CCys (4/10 colonies) or *Pf*SORTILIN VPS10 (10/10 colonies). Samples containing the inserts migrated as expected (Figure 3.8).

Samples that underwent restriction enzyme digestion are seen in Figure 3.9 and showed that the inserts were the correct size. DNA sequencing confirmed that the inserts contained no PCR-related errors and were in the correct orientation as seen in Figures A.1-A.5 found in Appendix A.

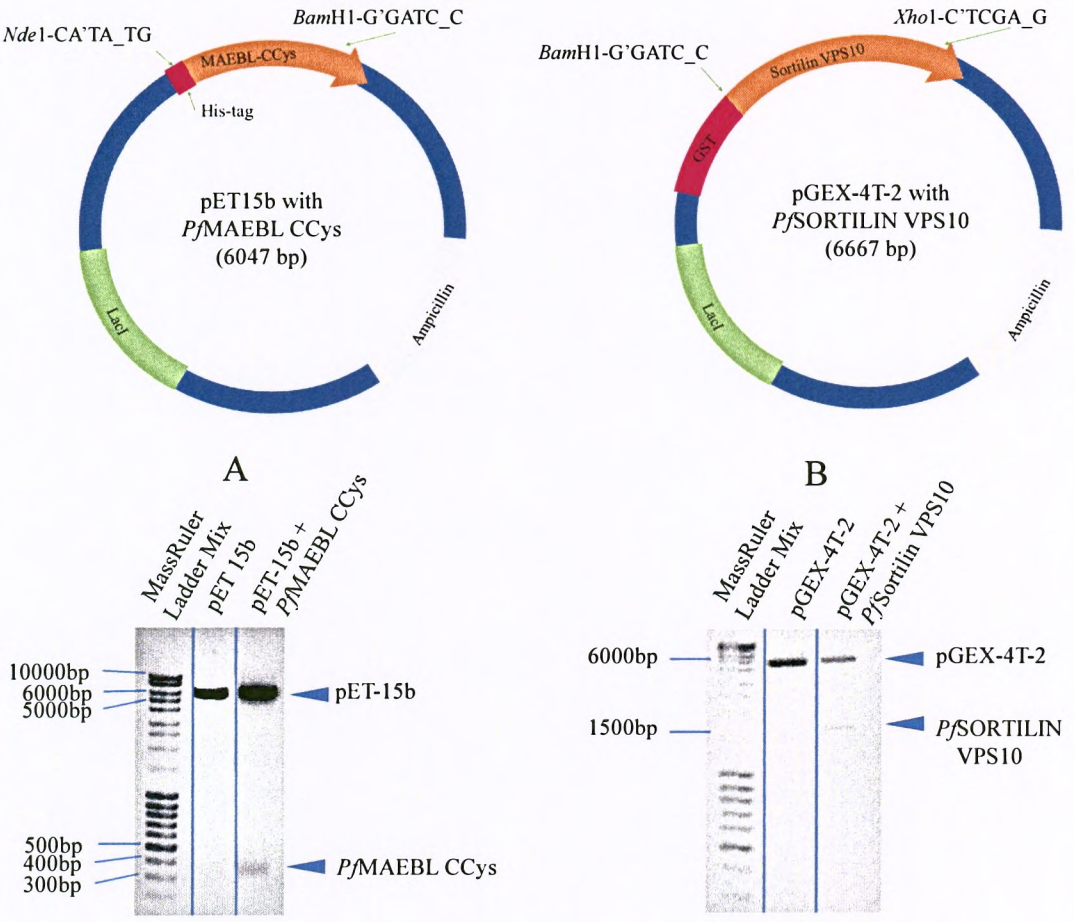


Figure 3.9: **Restriction enzyme digestion of plasmids isolated from clones positive for colony PCR. Digestion of A: pET-15b plasmid and the pET-15b *Pf*MAEBL CCys construct with *Bam*H1 and *Xho*1, B: pGEX-4T-2 plasmid and the pGEX-4T-2 *Pf*SORTILIN VPS10 construct with *Nde*1 and *Bam*H1 after transformation. Vector maps indicate the recognition sites used to clone each amplicon into their respective vectors.**

3.4 Recombinant protein expression

3.4.1 *Pf*MAEBL CCys domain

The *Pf*MAEBL CCys domain (expected size with His-tag: 13 kDa) was first expressed in *E. coli* using Overnight Express™ Instant TB Medium, but expression and solubility were poor so IPTG induction was chosen as the method of expression. By testing Rosetta 2™ (DE3) cell growth conditions at 20 °C and 30 °C for 16 h, expression and solubility of the *Pf*MAEBL CCys

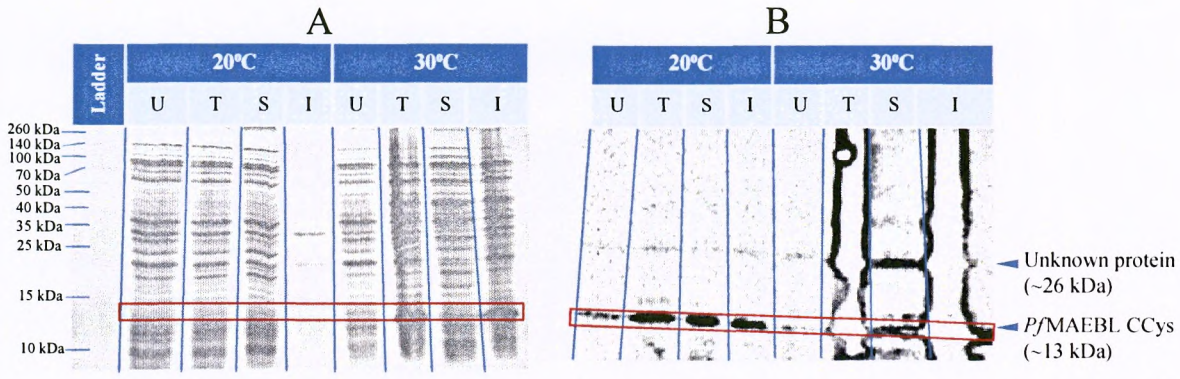


Figure 3.10: **SDS-PAGE (17% acrylamide) and anti-His immunoblot analyses of *Pf*MAEBL CCys using IPTG induction at varying temperatures. A: SDS-PAGE and B: anti-His immunoblot analyses of the expression and solubility of *Pf*MAEBL CCys using IPTG induction at** 20 °C **and** 30 °C. ***E. coli*** **cultures** (20 ml) **were grown for** 16 h. **The red box indicates the protein of interest. An aliquot of** 5 µl **out of a volume** 1.5 ml **of each sample was loaded on the gel. U: uninduced, T: total, S: soluble and I: insoluble fraction of *E. coli* lysates. A Spectra Multicolour Broad Range Protein Ladder (5 µl) was used.**

domain were compared. All fractions and an uninduced control (no IPTG added) were analysed by SDS-PAGE and immunoblot analysis (Figure 3.10). All samples were normalized to have equal $OD_{600\,nm}$ values for comparative analyses.

At 20 °C more *Pf*MAEBL CCys is present in the soluble fraction than at 30 °C. Considering that the experiments were conducted under identical conditions (i.e. cell lysis), it could be assumed that the difference observed was mainly due to protein solubility. In the immunoblot it was noted that all fractions contained a larger second band ($\approx$ 26 kDa), approximately twice the size, which could be a dimerized form of the protein. It should be noted that protein samples were treated with β-mercaptoethanol, which should break disulphide bonds in this Cys rich protein. However it may be possible that a covalent bond formed between the two monomers giving rise to the larger protein. The band was most pronounced at 30 °C in the soluble fraction and least pronounced at 20 °C.

Solubility was tested further (Figure 3.11) at 20 °C and different incubation periods (3 h, 6 h, 8 h and 12 h). As the incubation period increased so did the amount of soluble *Pf*MAEBL CCys protein. All samples contained the previously seen $\approx$ 26 kDa protein (Figure 3.11) and another

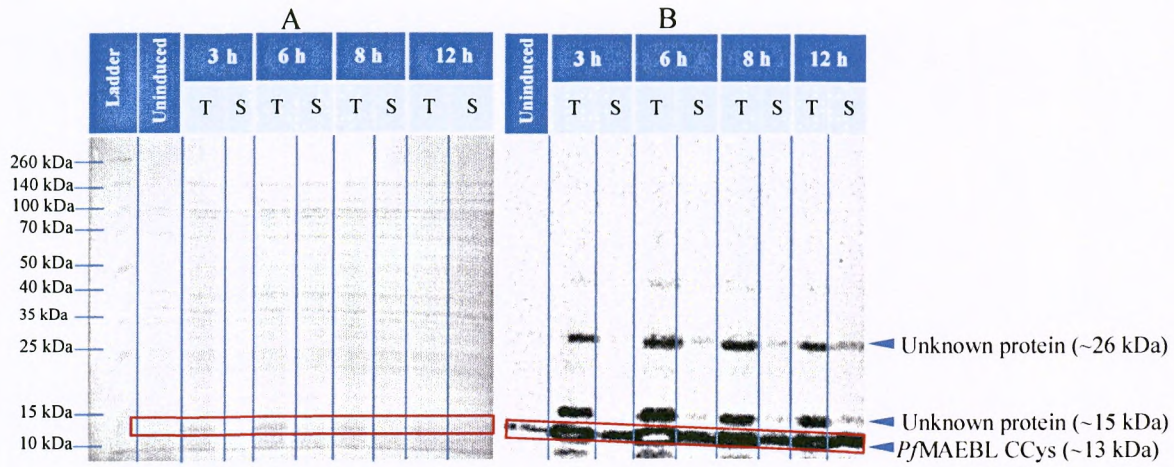


Figure 3.11: **SDS-PAGE (17% acrylamide) and immunoblot analyses *Pf*MAEBL CCys using IPTG induction at varying incubation periods. A: SDS-PAGE and B: immunoblot analyses of the expression and solubility of *Pf*MAEBL CCys using IPTG induction at** 20 °C **for** 3 h, 6 h, 8 h **and** 12 h. **The red box indicates the protein of interest. An aliquot** 1.5 µl **out of a volume of** 1.5 ml **of each sample was loaded on the gel. T: total and S: soluble fractions of *E. coli* lysates. A Spectra Multicolour Broad Range Protein Ladder (5 µl) was used.**

unknown band at ≈ 15 kDa. The immnoblot showed leaky expression of *Pf*MAEBL CCys under uninduced conditions.

The anti-His HRP-conjugated antibody bound to both unknown proteins. Minimal amounts of these proteins were present in the soluble fractions but seemed to increase incubation time since the largest amounts were detected at 12 h. Such minimal amounts would not have posed major issues when attempting to purify *Pf*MAEBL CCys protein. However, to limit the presence of the larger proteins, a shorter incubation period of 6 h was chosen for further experiments.

3.4.2 *Pf*SORTILIN VPS10 domain

Expression and solubility of *Pf*SORTILIN VPS10 (expected size with GST-tag: 86 kDa) using the Overnight Express™ Instant TB Medium and IPTG induction were compared by SDS-PAGE and immunoblot analysis (Figure 3.12). *E. coli* Rosetta 2™ (DE3) cell cultures for each method were grown at 37 °C for 16 h. Recombinant GST was expressed from the empty pGEX-4T-2 vector and used as a positive control. Immunoblot analysis showed similar expression of

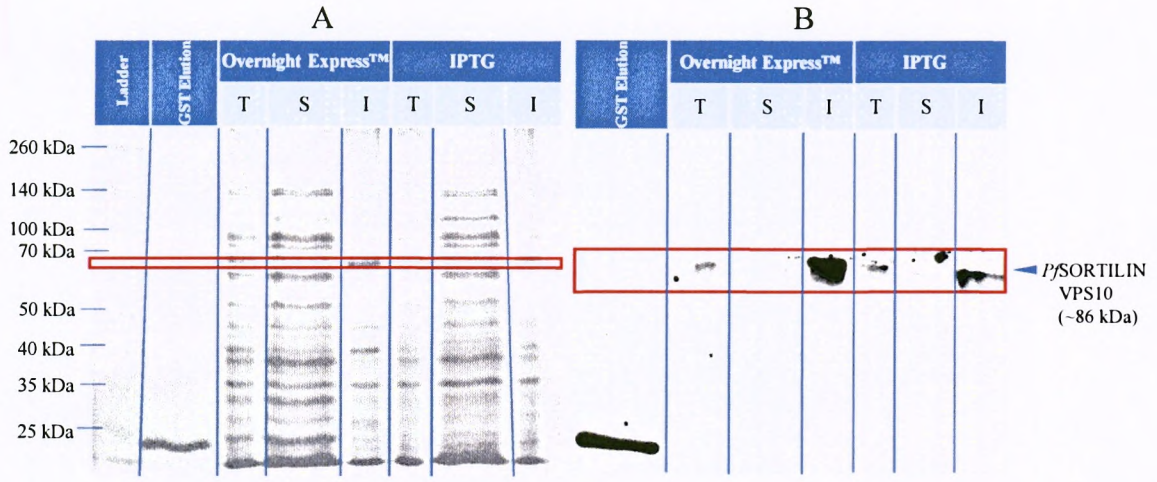


Figure 3.12: **SDS-PAGE (15% acrylamide) and anti-GST immunoblot analyses of *Pf*SORTILIN VPS10 using Overnight Express™ or IPTG induction. A: SDS-PAGE and B: anti-GST immunoblot analyses of *Pf*SORTILIN VPS10 (red box) using Overnight Express™ or IPTG induction. Cultures (20 ml) were grown for 16 h at 37 °C and the T: total, S: soluble and I: insoluble fractions of *E. coli* lystates were analysed between the two methods of expression. An aliquot representing 12 µl (total volume: 1.5 ml) of each fraction was loaded on the gel. A purified GST elution (by glutathione affinity chromatography) was included as an immunoblot positive control. A Spectra Multicolour Broad Range Protein Ladder (5 µl) was used.**

*Pf*SORTILIN VPS10 in Overnight Express™ Instant TB Medium and IPTG induction in the total fractions (Figure 3.12). Almost no protein was seen in the soluble fractions when using either method. This was also true for both insoluble fractions which contained the majority of the total expressed *Pf*SORTILIN VPS10 protein.

The solubility of *Pf*SORTILIN VPS10 induced with IPTG was tested at 20 °C and 30 °C for 16 h as well as at 20 °C for 3 h, 6 h, 8 h and 12 h (Figure 3.13 and Figure 3.14). All total fractions were normalized to have equal $OD_{600\,nm}$ values for comparative purposes.

SDS-PAGE analysis revealed higher expression at 30 °C but the protein was largely insoluble at both temperatures. Immunoblot analysis revealed that almost no soluble protein was expressed when incubated at 30 °C, but at 20 °C a small amount was soluble. This incubation temperature was chosen for further experiments. SDS-PAGE analysis showed that there was a gradual increase in the amount of protein expressed as the incubation period increased. Soluble protein existed at all time periods tested. Considering that all samples were normalized to have

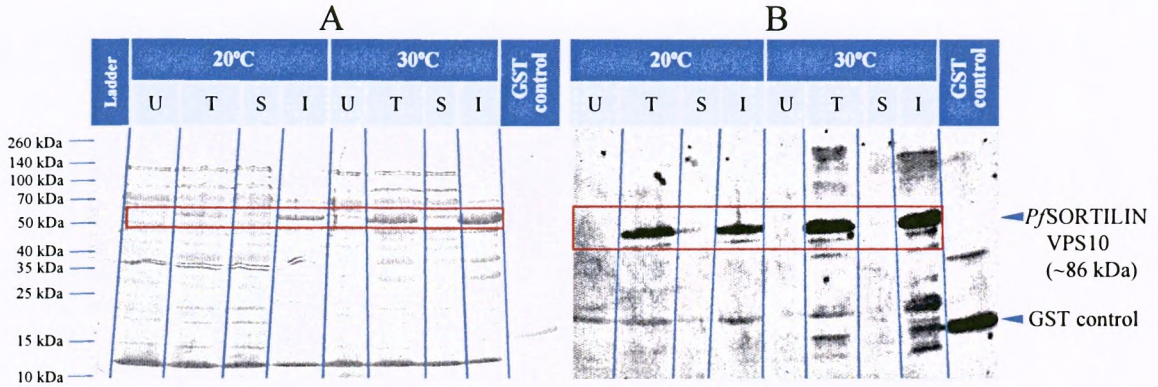


Figure 3.13: **SDS-PAGE (15% acrylamide) and anti-GST immunoblot analyses of *Pf*SORTILIN VPS10 using IPTG induction at varying temperatures. A: SDS-PAGE and B: anti-GST immunoblot analyses of *Pf*SORTILIN VPS10 using IPTG induction at 20 °C and 30 °C. Cultures (20 ml) were grown for 16 h. The red box indicates the protein of interest. An aliquot of 10 µl out of a volume of 1.5 ml of each sample was loaded on the gel. U: uninduced, T: total, S: soluble and I: insoluble fractions of *E. coli* lysates. A Spectra Multicolour Broad Range Protein Ladder (5 µl) was used.**

equal $OD_{600\,nm}$ values, there was no noticeable increase in the amount of soluble protein with an increasing period of incubation, so it was decided to use a period of 3 h incubation.

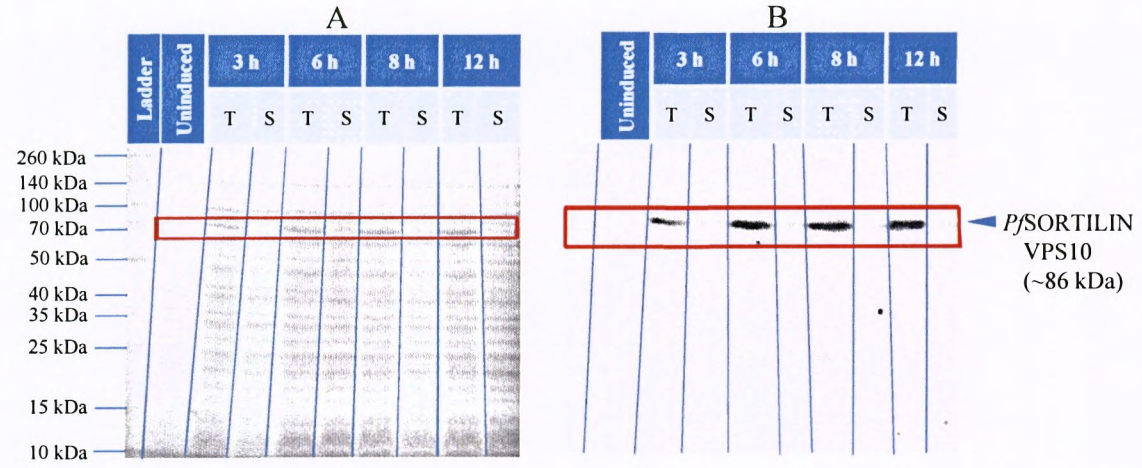


Figure 3.14: **SDS-PAGE (15% acrylamide) and anti-GST immunoblot analyses of *Pf*SORTILIN VPS10 using IPTG induction at varying incubation periods. A: SDS-PAGE and B: anti-GST immunoblot analyses of *Pf*SORTILIN VPS10 using IPTG induction at 20 °C for 3 h, 6 h, 8 h and 12 h. The red box indicates the protein of interest. An aliquot of 1.5 µl out of a volume of 1.5 ml of each sample was loaded on the gel. T: total and S: soluble fractions of *E. coli* lysates. A purified GST elution (by glutathione affinity chromatography) was included as an immunoblot positive control. A Spectra Multicolour Broad Range Protein Ladder (5 µl) was used.**

In all experiments the GST-tagged *Pf*SORTILIN VPS10 protein migrated at ≈67kDa (calculated using the protein ladder as a reference) compared to its theoretical molecular weight of ≈86kDa. Migration of *P. falciparum* proteins at unexpected sizes is not uncommon (Mehlin *et al.*, 2006). Proteins are denatured by binding to SDS, facilitating their migration according to molecular weight. Several factors can influence this binding resulting in a shift in migration. Normally low complexity regions result in more SDS binding, causing the protein to migrate faster and appear to have a smaller molecular weight (Mehlin *et al.*, 2006). The segment of the *Pf*SORTILIN VPS10 domain that was cloned does not contain a low complexity region (Figure 3.15) which therefore does not explain the aberrant migration.

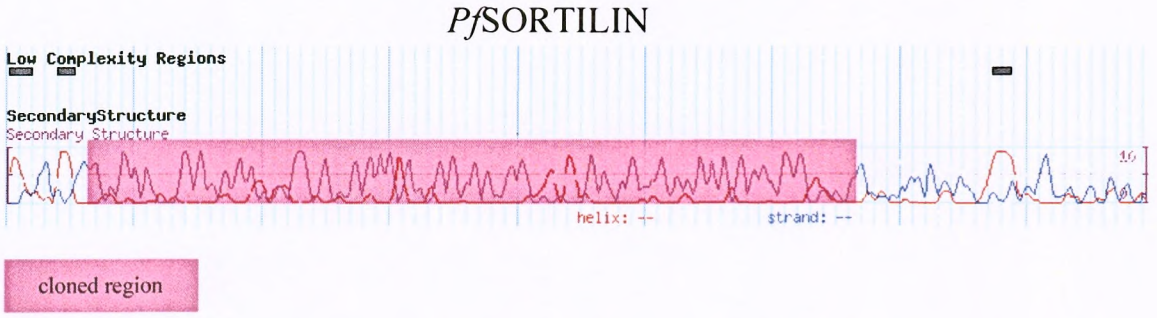


Figure 3.15: **Low complexity regions of the *Pf*SORTILIN VPS10 protein. The highlighted region represents the cloned sequence of *Pf*SORTILIN VPS10 . The grey boxes above show where the low complexity regions are found. Low complexity regions can be found in the signal peptide and transmembrane domain of this protein. Image adapted from PlasmoDB (2018).**

3.4.3 Final recombinant protein expression conditions

To obtain an adequate amount of protein for downstream applications such as *in vitro* binding assays the following growth conditions in Table 3.3 were used based on experiments described in subsection 3.4.1 and subsection 3.4.2. Conditions for *Pf*AMA-1 prodomain were optimized by Churchyard (Churchyard, 2017).

Table 3.3: **Final growth conditions for expression of soluble recombinant proteins**

Protein	Method	Culture volume	Flask size	Temperature	Time
*Pf*MAEBL CCys	IPTG	6X 20 ml	250 ml	20 °C	6 h
*Pf*AMA-1 prodomain	Overnight Express™	2X 20 ml	250 ml	20 °C	10 h
*Pf*SORTILIN VPS10	IPTG	16X 100 ml	1000 ml	20 °C	3 h
GST control	Overnight Express™	2X 20 ml	250 ml	20 °C	10 h

3.5 Recombinant protein purification

3.5.1 *Pf*MAEBL CCys domain

The purification of the His-tagged *Pf*MAEBL CCys protein, expressed under the conditions stated in Table 3.3, is shown in Figure 3.16.

Three other protein bands were present in the total and insoluble fractions. Not all of the *Pf*MAEBL CCys protein bound to the MagneHis™ particles as indicated by the unbound fraction. Some protein was lost after 2 washes with 20 mM imidazole buffer. Most of the protein was eluted in elution 1 with 50 µl of 500 mM imidazole buffer. When the 3 elutions (50 µl each) were combined the yield was ≈6 µg from 120 ml bacterial culture with a >75% purity and a concentration of 40 ng/µl based on densitometry of the bands on the gel. A substantial amount of protein was still bound to the MagneHis™ particles. The protein was dialysed into His-binding buffer and also electrophoresed on the gel, which indicated that some protein was lost (≈ 900 ng).

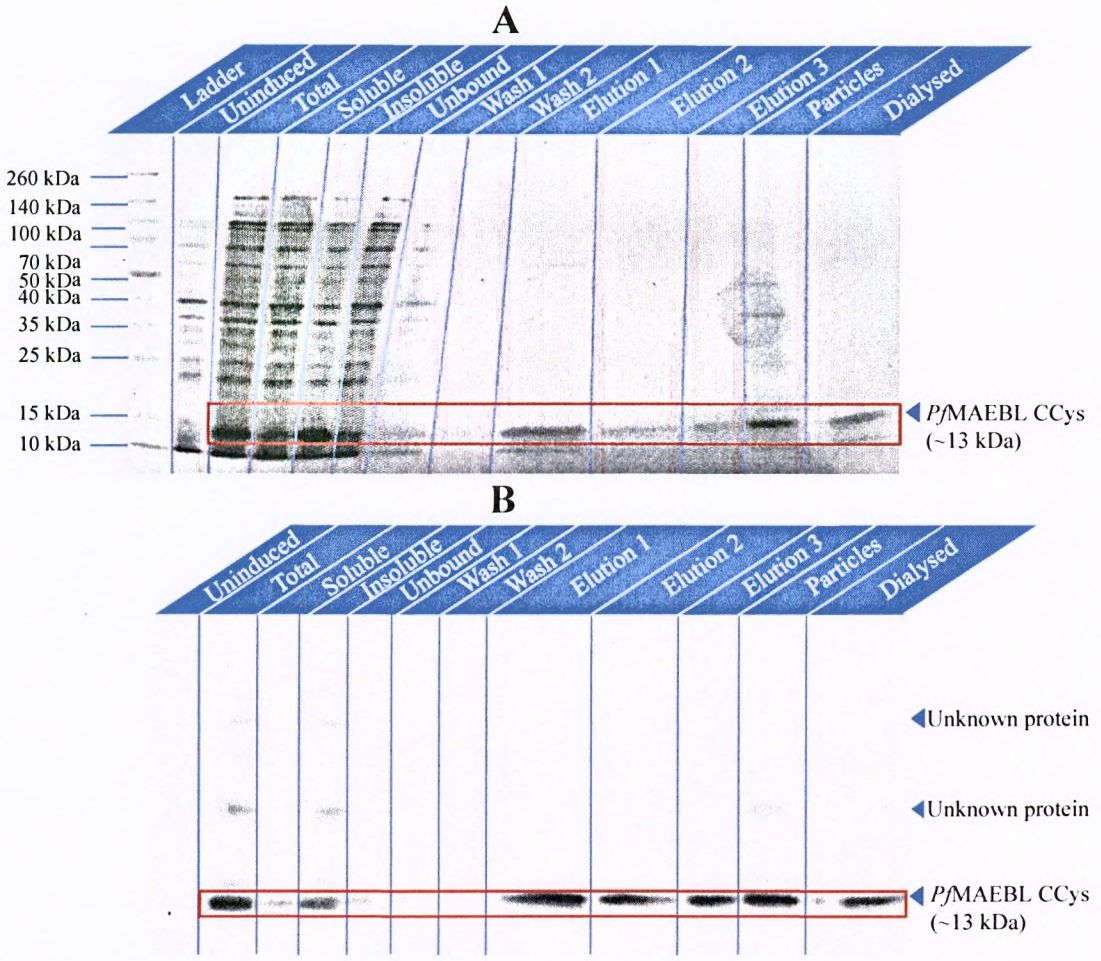


Figure 3.16: **Purification of His-tagged *Pf*MAEBL CCys protein. A: SDS-PAGE (17% acrylamide) and B: anti-His immunoblot analyses of the purification of *Pf*MAEBL CCys on the MagneHis^{TM} particles after IPTG induction at** 20 °C **for** 6 h**. The blue arrows show unknown proteins and the red box shows the protein of interest. Aliquots out of a total volume were loaded as follows: Uninduced/Total/Soluble/Insoluble/Unbound-** 2 µl/1.5 ml**, Washes-** 20 µl/2 ml**, Elutions/Particles-** 6 µl/50 µl**, Dialysed-** 6 µl/132 µl**. The predicted molecular weight of *Pf*MAEBL CCys is 13 kDa, but the protein migrated at ≈ 12 kDa. A Spectra Multicolour Broad Range Protein Ladder (5 µl) was used.**

3.5.2 *Pf*AMA-1 prodomain

The expression and purification conditions of His-tagged *Pf*AMA-1 prodomain were determined by Churchyard (2017, Figure 3.17). The purified *Pf*AMA-1 prodomain protein migrated at ≈16kDa compared to its theoretical molecular weight of ≈11kDa (Figure 3.17). Most non-specific bands were washed away in the first wash. Some protein was lost in both washes. The

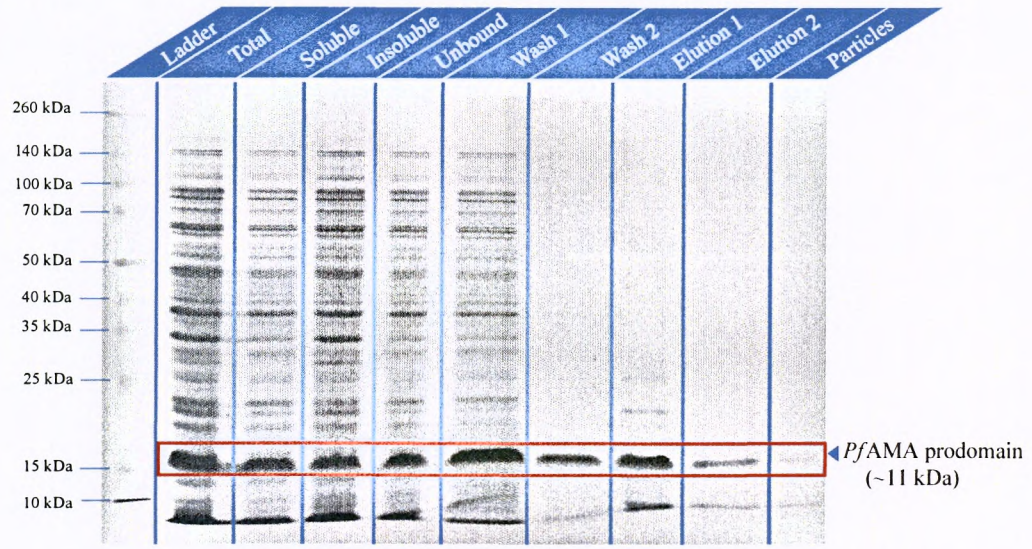


Figure 3.17: **Purification of His-tagged *Pf*AMA-1 prodomain protein. SDS-PAGE (15% acrylamide) analysis of the purification of *Pf*AMA-1 prodomain after induction with Overnight Express™ Instant TB Medium for** 16 h **at** 20 °C**. The red box shows the protein of interest. Aliquots out of a total volume were loaded as follows: Uninduced/Total/Soluble/Insoluble/Unbound-** 2 µl/1.5 ml**, Washes-** 20 µl/2 ml**, Elutions/Particles-** 6 µl/50 µl**. A Spectra Multicolour Broad Range Protein Ladder (5 µl) was used. The predicted molecular weight of *Pf*AMA-1 prodomain is 11 kDa however the protein migrated at ≈ 16 kDa.**

first elution of *Pf*AMA-1 prodomain had a purity of >75%. A small amount of protein was still bound to the MagneHis™ particles. A yield of ≈65 µg of *Pf*AMA-1 prodomain from 100 ml bacterial culture at a concentration of ≈320 ng/µl was obtained.

3.5.3 *Pf*SORTILIN VPS10 domain

The GST-tagged *Pf*SORTILIN VPS10 protein was expressed under the conditions stated in Table 3.3 and purified using MagneGST™ particles (Figure 3.18). SDS-PAGE and immunoblot analysis revealed the presence of several smaller bands in the elutions. This could be due to premature translation termination of *Pf*SORTILIN VPS10. Since the GST-tag is at the N-terminal of the recombinant protein, prematurely terminated proteins will still be recognized by the MagneGST™ particles and the anti-GST conjugated antibody. The smallest band seen is most likely to be endogenous GST expressed by the bacteria. Immunoblot analysis revealed

no leaky expression of the protein. A large amount of *Pf*SORTILIN VPS10 was lost after the first wash using the GST buffer, but 5 washes were performed in an attempt to improve purity. The last wash showed no residual *Pf*SORTILIN VPS10 protein. The majority of the protein was eluted in elution 1 (50 µl of a 500 m M glutathione buffer), however some protein was still bound to the particles. When both elutions were combined the yield was ≈3 µg from 1.4 l bacterial cultures with a >50% purity at a concentration of 32 ng/µl.

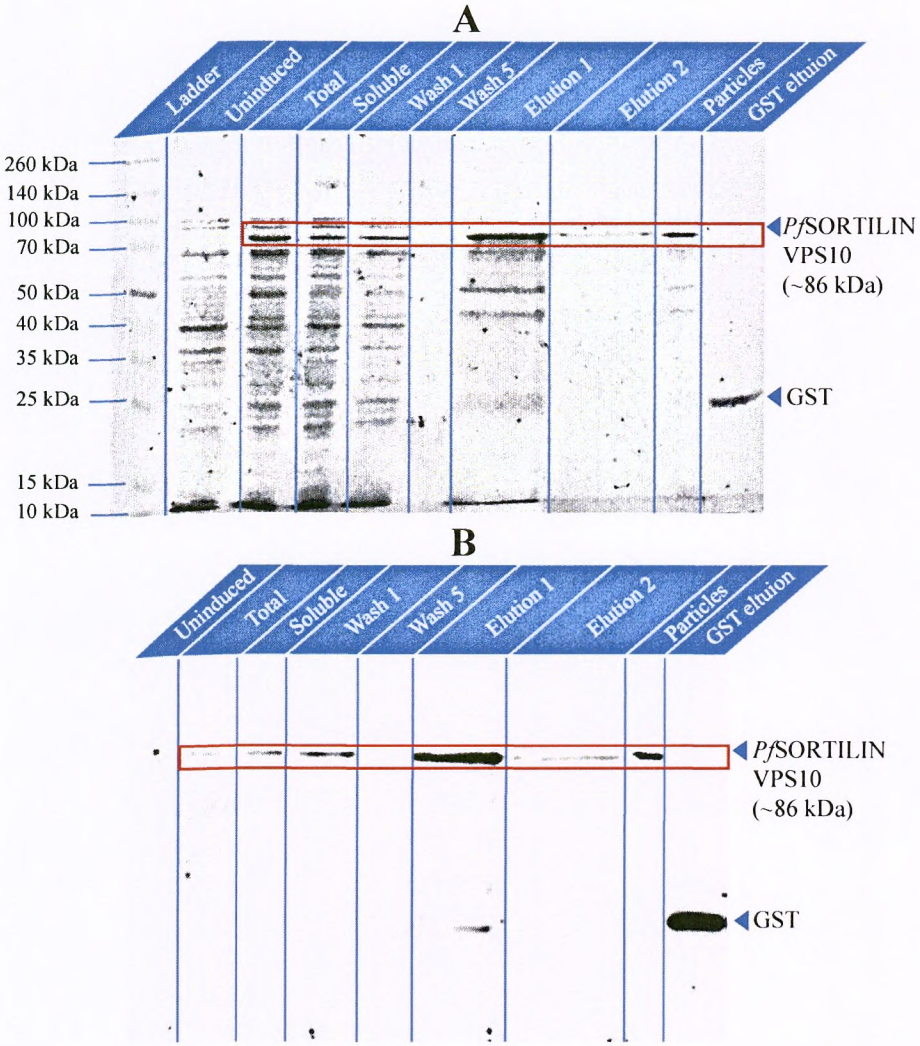


Figure 3.18: **Purification of *Pf*SORTILIN VPS10 protein. A: SDS-PAGE (15% acrylamide) and B: anti-GST immunoblot analyses of the purification of *Pf*SORTILIN VPS10 using MagneGST^{TM} particles after IPTG induction at 20 °C for 3 h. The red box indicates the protein of interest and the blue arrow shows the GST control. Aliquots of a total volume were loaded as follows: Uninduced/Total/Soluble/Unbound- 2 µl/1.5 ml, Washes- 20 µl/2 ml, Elutions/particles- 13 µl/50 µl, GST control-2 µl/50 µl. The predicted molecular weight of *Pf*SORTILIN VPS10 is 86 kDa, but the protein migrated at ≈ 67 kDa. A Spectra Multicolour Broad Range Protein Ladder (5 µl) was used.**

3.6 *In vitro* binding assays

To identify an interaction between *Pf*SORTILIN VPS10 and *Pf*MAEBL CCys or *Pf*AMA-1 prodomain, *in vitro* binding assays were performed. This would indicate that *Pf*SORTILIN

may be an escorter for a variety of invasion proteins which localize to different apical organelles. All purified proteins were dialysed into relevant buffers depending on whether the protein was immobilized or used in-solution. In all assays that were performed, GST-tagged *Pf*SORTILIN VPS10 was the limiting protein as it always gave the lowest yield.

3.6.1 Quantitation by densitometry

Prior to conducting an assay, all protein samples were quantitated by densitometry. Not all quantitation gels are shown but an example is provided (Figure 3.19). Protein samples were electrophoresed, along with increasing amounts of BSA (100 ng to 500 ng), on a 15% acrylamide gel. Densitometry was used to prepare a BSA standard curve and to estimate the total amounts and concentrations of proteins (Figure 3.20 and Table 3.4).

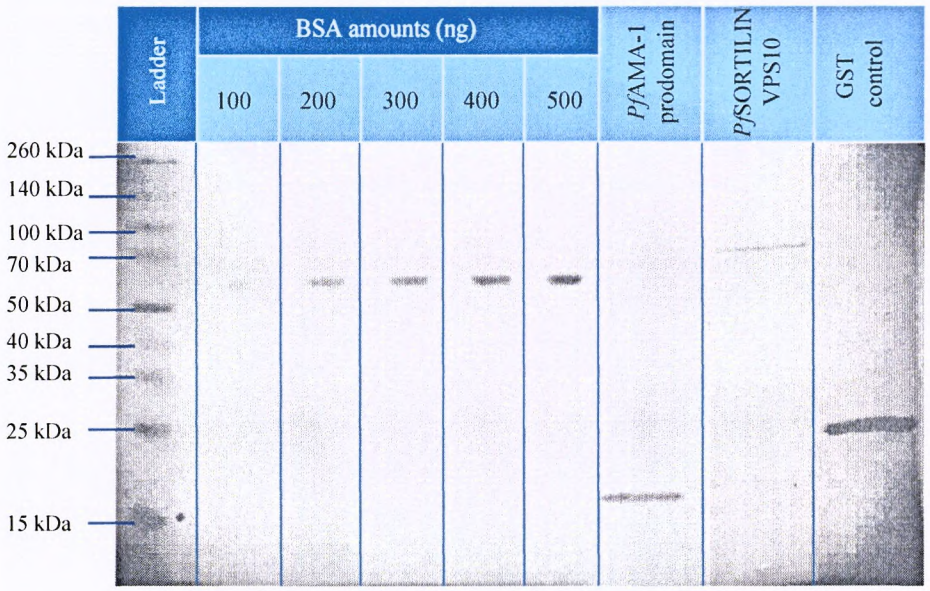


Figure 3.19: **SDS-PAGE (15% acrylamide) of proteins used for *in vitro* binding assay of immobilized *Pf*AMA-1 prodomain and *Pf*SORTILIN VPS10 . Aliquots of 6 µl of each sample were electrophoresed. BSA samples (100 ng to 500 ng) were loaded as standards for densitometry. A Spectra Multicolour Broad Range Protein Ladder (5 µl) was used.**

Table 3.4: **Final yield and concentration of proteins for binding assay of *Pf*AMA-1 prodomain and *Pf*SORTILIN VPS10 based on densitometry of SDS polyacrylamide gel (Figure 3.19).**

Sample	Peak volume	ng amount	Total volume (µl)	[Sample] (ng/µl)	Total yield (ng)
BSA	1355	100	2	-	-
BSA	2803	200	4	-	-
BSA	4720	300	6	-	-
BSA	7279	400	8	-	-
BSA	8166	500	10	-	-
*Pf*AMA-1 prodomain	5054	**324**	160	54	8640
*Pf*SORTILIN VPS10	1937	**124**	160	21	3296
GST control	16701	1072	179	160	28587

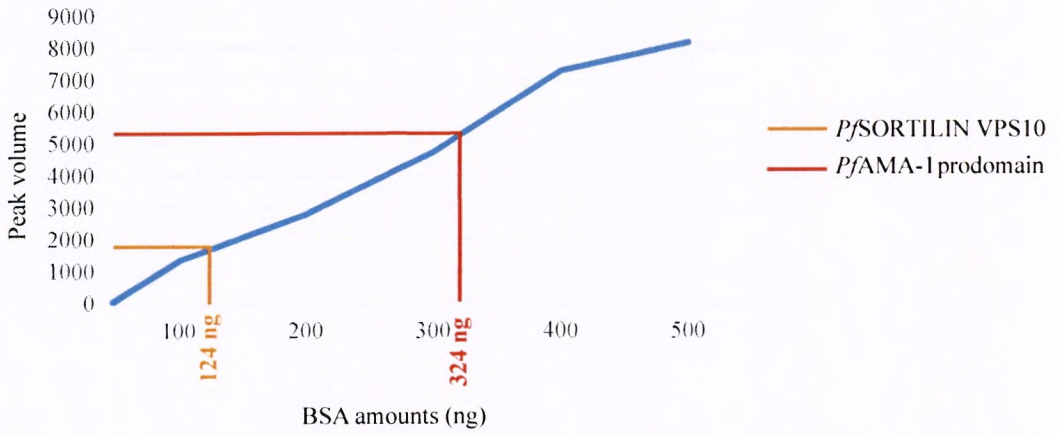


Figure 3.20: **Standard curve of BSA (ng) versus peak volume. The peak volumes of *Pf*AMA-1 prodomain and *Pf*SORTILIN VPS10 (Table 3.4) were used to determine the amounts of these proteins.**

3.6.2 Immobilized *Pf*MAEBL CCys vs *Pf*SORTILIN VPS10

Purified His-tagged *Pf*MAEBL CCys was dialysed into His-binding buffer (pH 8) and purified GST-tagged *Pf*SORTILIN VPS10 and GST control were dialysed into TBS (25 mM Tris-HCl, 150 mM NaCl, pH 7.5). In this assay constant amounts of 150 ng of His-tagged *Pf*MAEBL CCys were immobilized onto MagneHis^{TM} particles and incubated with increasing amounts (100 ng to 400 ng) of free GST-tagged *Pf*SORTILIN VPS10 or GST (Figure 3.21-A).

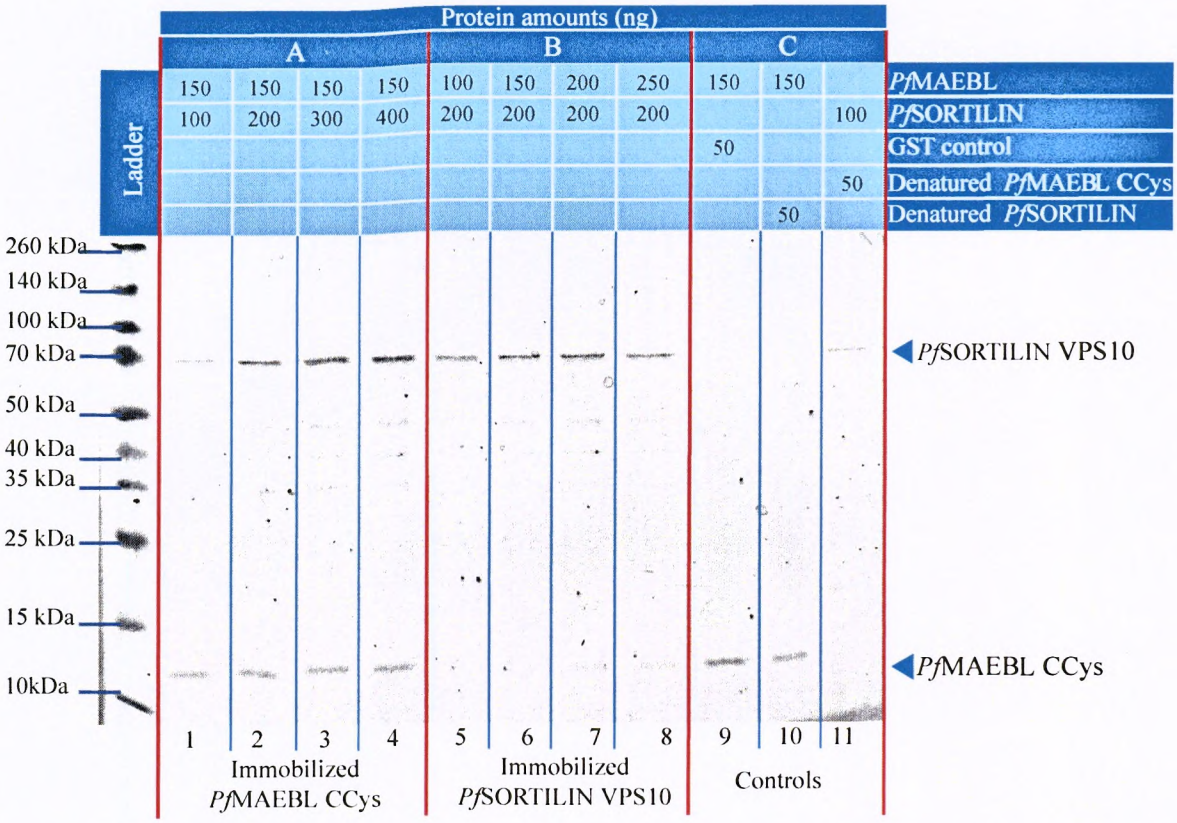


Figure 3.21: **SDS-PAGE (15% acrylamide) analysis of *in vitro* binding assay of His-tagged *Pf*MAEBL CCys and GST-tagged *Pf*SORTILIN VPS10. Entire particle samples were loaded. A (Lane 1-4): Immobilized *Pf*MAEBL CCys (ng) with increasing amounts of free *Pf*SORTILIN VPS10 , B (Lane 5-8): Immobilized *Pf*SORTILIN VPS10 (ng) with increasing amounts of free *Pf*MAEBl CCys, Lane 9: Immobilized *Pf*MAEBL CCys with free GST, Lane 10: Immobilized *Pf*MAEBL CCys with denatured *Pf*SORTILIN VPS10 , Lane 11: Immobilized *Pf*SORTILIN VPS10 with denatured *Pf*MAEBl CCys. A Spectra Multicolour Broad Range Protein Ladder (10 µl) was used.**

As seen in the SDS-polyacrylamide gel (Figure 3.21), densitometry analysis (Table 3.5) and the binding curve (Figure 3.22) there was a dose-dependent association between the proteins that approached saturation at 400 ng. *Pf*SORTILIN VPS10 bound to *Pf*MAEBL CCys with a stoichiometry of approximately 1:3.2. When free GST was incubated with immobilized *Pf*MAEBL CCys (Figure 3.21-C; Lane 9) no binding occurred, implying that the GST-tag did not facilitate the binding of *Pf*SORTILIN VPS10 to *Pf*MAEBL CCys. Denatured *Pf*SORTILIN VPS10 also did not bind to immobilized *Pf*MAEBL CCys which suggested that there was no non-specific binding (Figure 3.21-C; Lane 10).

Table 3.5: **Densitometry analysis of the *in vitro* binding assay between immobilized His-tagged *Pf*MAEBL CCys and free GST-tagged *Pf*SORTILIN VPS10. A constant amount of *Pf*MAEBL CCys (150 ng, 11.4 pmol) was incubated with increasing amounts of *Pf*SORTILIN VPS10.**

SORTILIN added		SORTILIN bound			
ng	**pmol**	**Peak volume**	**ng**	**pmol**	**pmol/pmol MAEBL**
100	1.1	551	90	1.00	0.09
200	2.2	1374	225	2.48	0.22
300	3.3	1779	291	3.21	0.28
400	4.4	1968	290	3.55	0.31

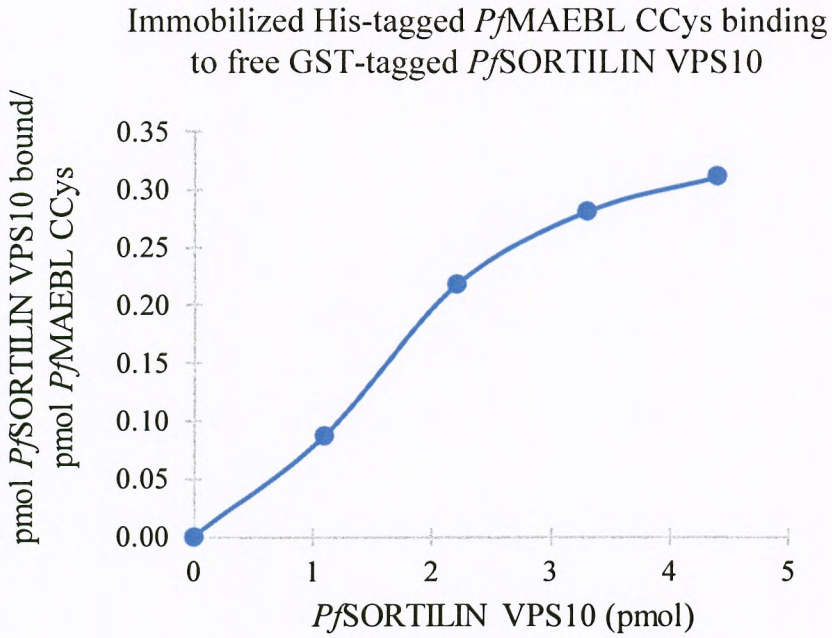


Figure 3.22: **Graph showing the dose-dependent binding of immobilised His-tagged *Pf*MAEBL CCys to free GST-tagged *Pf*SORTILIN VPS10. Densitometry analysis (Table 3.5) was used to quantitate the amount of bound *Pf*SORTILIN VPS10. Pmol amounts were used to determine the stoichiometry of the binding.**

3.6.3 Immobilized *Pf*SORTILIN VPS10 vs *Pf*MAEBL CCys

To determine whether the binding seen in subsection 3.6.2 can also be seen with immobilized *Pf*SORTILIN VPS10 , the following assay was performed. Purified GST-tagged *Pf*SORTILIN VPS10 was dialysed into GST binding/wash buffer (pH 7.2) and purified His-tagged *Pf*MAEBL

CCys and GST into TBS (pH 7.5). This assay used constant amounts of 200 ng of *Pf*SORTILIN VPS10 immobilized on MagneGST™ particles which were incubated with increasing amounts 100 ng to 250 ng of free *Pf*MAEBL CCys protein.

The samples from this assay were electrophoresed on the same SDS polyacrylamide gel as the previous assay (Figure 3.21-B). Densitometry analysis (Table 3.6) indicated that as the amount of free *Pf*MAEBL CCys increased so did its binding to immobilized *Pf*SORTILIN VPS10 (Figure 3.21-B; Lane 5-8). At the lowest concentrations of free *Pf*MAEBL CCys (100 ng to 150 ng) very faint bands can be seen, however, clear binding to immobilized *Pf*SORTILIN VPS10 can be seen at 200 ng to 250 ng of free *Pf*MAEBL CCys. This indicated dose-dependent binding which had not reached saturation (Figure 3.23), so the stoichiometry could not be determined. A control showed that there was no non-specific binding of denatured *Pf*MAEBL CCys to immobilized *Pf*SORTILIN VPS10 (Figure 3.21-C, Lane 11).

Table 3.6: **Densitometry analysis of the *in vitro* binding assay between immobilized GST-tagged *Pf*SORTILIN VPS10 and free His-tagged *Pf*MAEBL CCys. A constant amount of *Pf*SORTILIN VPS10 (200 ng, 2.3 pmol) was incubated with increasing amounts of *Pf*MAEBL CCys.**

MAEBL added		MAEBL bound			
ng	**pmol**	**Peak volume**	**ng**	**pmol**	**pmol/pmol SORTILIN**
100	7.6	87	9.77	0.74	0.32
150	11.4	198	22.23	1.69	0.73
200	15.2	393	44.13	3.35	1.44
250	19	562	63.11	4.80	2.06

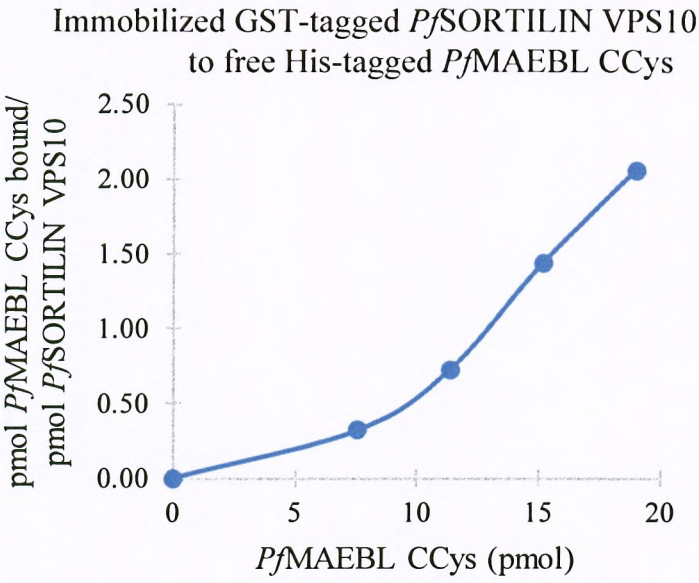


Figure 3.23: **Graph showing the binding of immobilised GST-tagged *Pf*SORTILIN VPS10 to free His-tagged *Pf*MAEBL CCys. Densitometry analysis (Table 3.6) was used to quantitate the amount of bound *Pf*MAEBL CCys.**

3.6.4 Immobilized *Pf*AMA-1 prodomain vs *Pf*SORTILIN VPS10

To determine if *Pf*SORTILIN could be an escorter for microneme proteins, the interaction between *Pf*SORTILIN VPS10 and *Pf*AMA-1 prodomain was investigated. Purified His-tagged *Pf*AMA-1 prodomain was dialysed into His-binding buffer (pH 8) and purified GST-tagged *Pf*SORTILIN VPS10 and GST control were dialysed into TBS (pH 7.5).

In this assay constant amounts of 480 ng of His-tagged *Pf*AMA-1 prodomain were immobilized on MagneHis™ particles and incubated with increasing amounts (160 ng to 480 ng) of free GST-tagged *Pf*SORTILIN VPS10 or a GST control.

Whole particle samples were electrophoresed (Figure 3.24) and densitometry was used to analyse the binding of *Pf*AMA-1 prodomain to *Pf*SORTILIN VPS10 (Table 3.7).

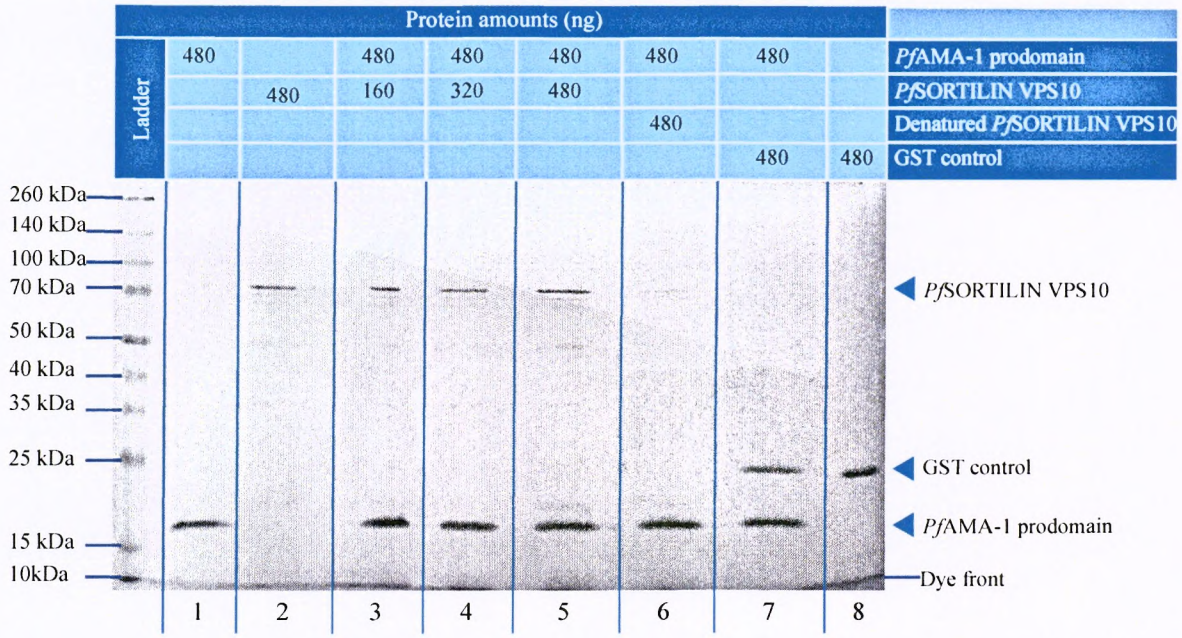


Figure 3.24: **SDS-PAGE (15% acrylamide) analysis of *in vitro* binding assays of *Pf*AMA-1 prodomain vs *Pf*SORTILIN VPS10 . A Spectra Multicolour Broad Range Protein Ladder (5 µl) was used.**

Table 3.7: **Densitometry analysis of the *in vitro* binding assay between immobilized His-tagged *Pf*AMA-1 prodomain and free GST-tagged *Pf*SORTILIN VPS10. A constant amount of *Pf*AMA-1 prodomain (480 ng, 43.6 pmol) was incubated with increasing amounts of *Pf*SORTILIN VPS10.**

SORTILIN added		SORTILIN bound			
ng	pmol	Peak volume	ng	pmol	pmol/pmol AMA
160	1.71	605	163	1.80	0.04
320	3.52	1245	336	3.70	0.08
480	5.29	1789	483	5.32	0.12

As seen in the SDS-polyacrylamide gel (Figure 3.24), densitometry analysis (Table 3.7) and the binding curve (Figure 3.25) there was a linear association between the proteins, which never reached saturation. Virtually all of the free GST control bound to *Pf*AMA-1 prodomain. This implies that the GST-tag of recombinant *Pf*SORTILIN VPS10 facilitated the binding to the His-tagged *Pf*AMA-1 prodomain protein.

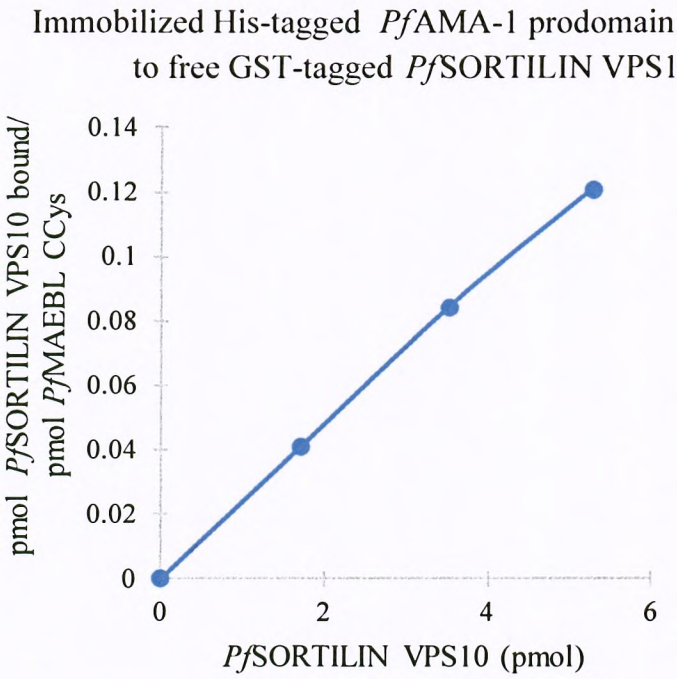


Figure 3.25: **Graph showing the dose-dependent binding of immobilised His-tagged *Pf*AMA-1 prodomain to free GST-tagged *Pf*SORTILIN VPS10. Densitometry analysis (Table 3.7) was used to quantitate the amount of bound *Pf*SORTILIN VPS10.**

3.6.5 Immobilized *Pf*SORTILIN VPS10 vs *Pf*AMA-1 prodomain

To assess the role of the GST tag in the interaction between *Pf*SORTILIN VPS10 vs *Pf*AMA-1 prodomain the following assay was performed. Purified GST-tagged *Pf*SORTILIN VPS10 and GST control were dialysed into GST binding/wash buffer (pH 7.2) and purified His-tagged *Pf*AMA-1 prodomain into TBS (pH 7.5). Constant amounts of 480 ng of GST-tagged *Pf*SORTILIN VPS10 were immobilized to MagneGST™ particles and incubated with increasing amounts (160 ng to 480 ng) of free His-tagged *Pf*AMA-1 prodomain. Total particle samples were electrophoresed (Figure 3.26). No binding between the two proteins was observed. A control of denatured *Pf*AMA-1 prodomain also showed no binding.

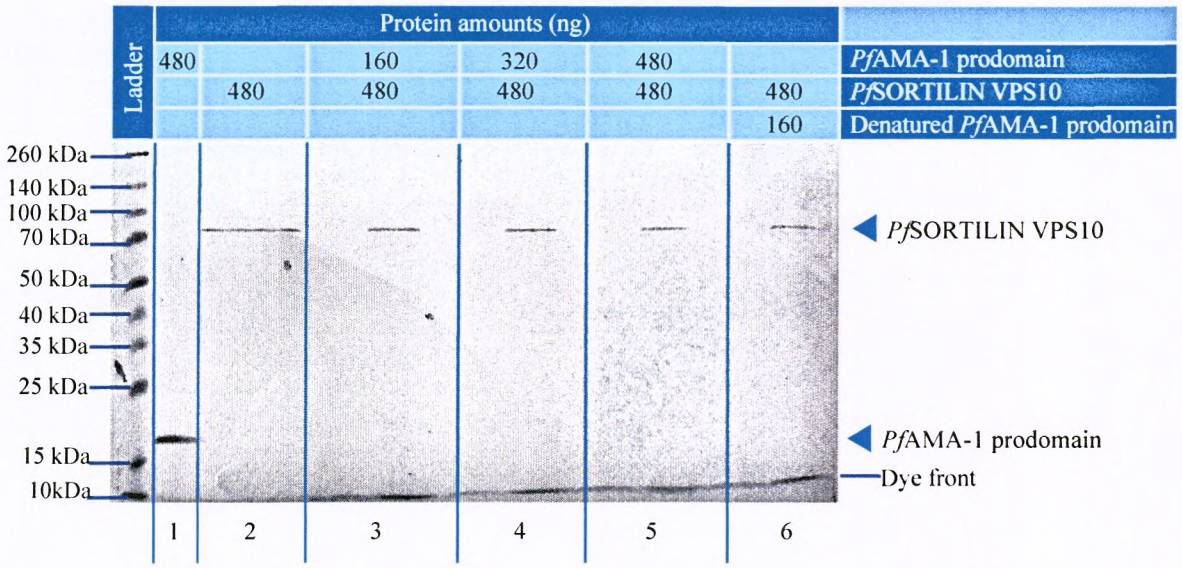


Figure 3.26: **SDS-PAGE (15% acrylamide) of *in vitro* binding assays of *Pf*SORTILIN VPS10 vs *Pf*AMA-1 prodomain . A Spectra Multicolour Broad Range Protein Ladder (5 µl) was used.**

In this assay the GST-tag of *Pf*SORTILIN VPS10 would not be accessible to the free *Pf*AMA-1 prodomain since it is bound to the MagneGST™ particles. These results correlate with the previous assay and confirm that *Pf*SORTILIN VPS10 and *Pf*AMA-1 prodomain do not interact.

3.7 Biopanning *Pf*SORTILIN VPS10 using phage display library

To identify novel binding partners of the putative escorter *Pf*SORTILIN, the purified recombinant protein was biopanned against mixed stage *P. falciparum* phage display libraries as outlined in Figure 3.27. To enrich for specific interactions, the biopanning was repeated for 4 consecutive rounds.

After the fourth round of biopanning the resultant plaques were visible as defined, non-overlapping clear areas on a lawn of *E. coli* on the titre plates as shown in Figure 3.28. Individual plaques

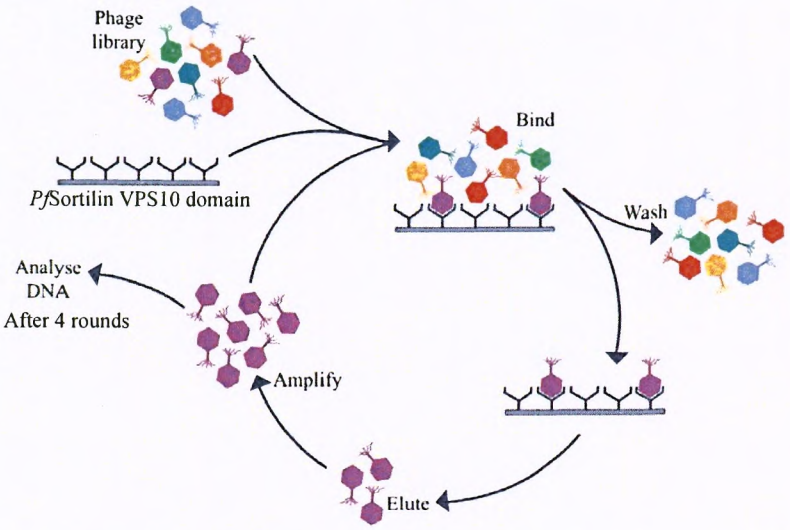


Figure 3.27: **A schematic depicting biopanning of *Pf*SORTILIN VPS10 against a *P. falciparum* phage display library. This image was adapted from Churchyard (2017).**

(60) were chosen from all plates and used for PCR and agarose gel assessment to determine the size of the *P. falciparum* DNA insert in each phage.

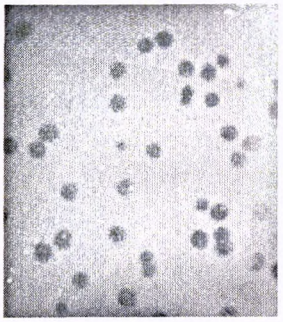


Figure 3.28: **A segment of a titering plate after four rounds of biopanning. A 10^6 dilution of a *Pf*-specific phage display library plated after four rounds of biopanning against GST-tagged *Pf*SORTILIN VPS10 protein. The high number of plaques indicated successful phage amplification.**

An empty phage cassette is 237 bp and none of the plaques were smaller than 250 bp, proving that the phage did contain an insert. Most inserts were just below 300 bp (Figure 3.29-A/B). Multiple similar size inserts suggested strong enrichment of these phage from the original starting library. The 10 largest PCR products were selected for DNA sequencing to identify the unknown peptides (Figure 3.29-C).

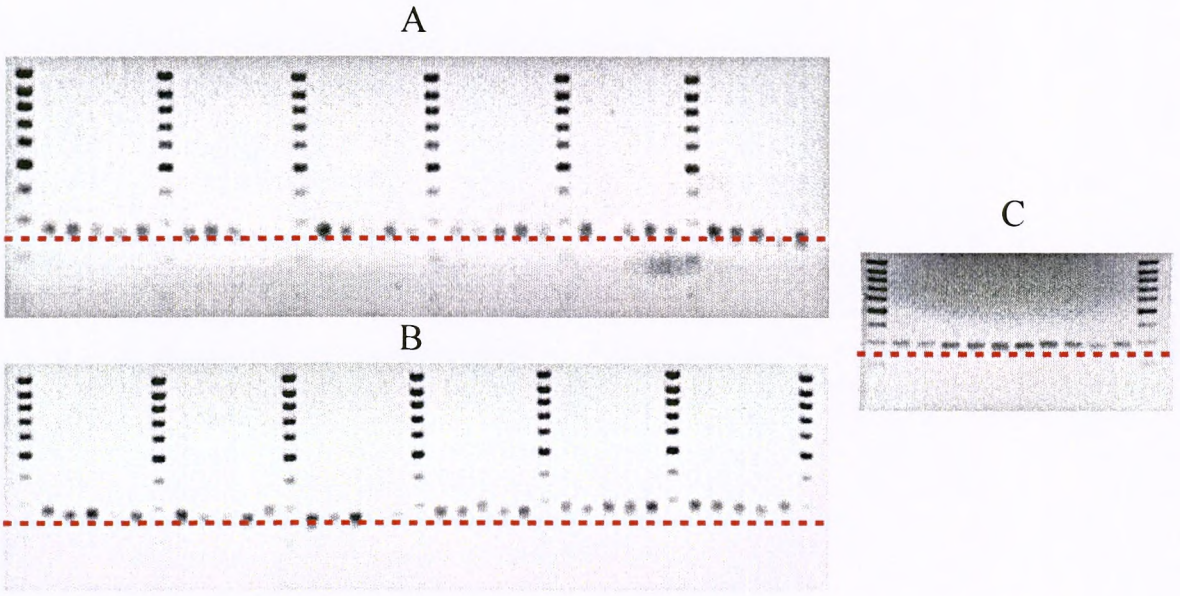


Figure 3.29: **PCR of the cDNA inserts present in the phage isolated after four rounds of biopanning against *Pf*SORTILIN VPS10 . Agarose gels (1%) showing A/B: 60 randomly selected phage and C: 10 PCR products selected for DNA sequencing. The low-range MassRuler^{TM} DNA Ladder was used. The red dotted line indicates the 250 bp marker.**

DNA sequencing analysis revealed small inserts which were ≈50 bp (Figure 3.30). All 10 sequences were the same with long stretches of A nucleotides.

The reading frame of each sequence was determined by identifying the location of the *EcoR*I restriction site (which was used to insert the *Pf*-specific DNA sequences into the phage). All the in-frame DNA sequences coded for strings of Lys residues, which did not correlate to any peptide segment available in the PlasmoDB database.

EcoR1 site

↓

GAATTCAAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 50

↑

Start of P. falciparum sequence

Figure 3.30: **DNA sequence of one phage insert which bound to the *Pf*SORTILIN VPS10 domain**

Chapter 4

Discussion

The *P. falciparum* parasite has an extensive range of proteins involved in many different processes which allows for its invasion into a host and further development (Pease *et al.*, 2013). The final destination of the individual proteins varies significantly from intracellular compartments such as the apicoplast or apical organelles to extracellular targets such as the PV or the host cell. The export of proteins from the parasite to its RBC host as been well characterized (Gilson *et al.*, 2017), however, little is known about the internal protein sorting and trafficking mechanisms.

An important component of the *P. falciparum* proteome is the invasion proteins which are localized in the apical organelles, the rhoptries and micronemes, and facilitate invasion through many ligand-receptor interactions (Beeson *et al.*, 2016). It is estimated that more than 50 proteins are involved in erythrocyte invasion (Kats *et al.*, 2008). The parasite is able to switch between many different invasion pathways and processes, rendering invasion proteins, like EBL-DBPs, non-essential and therefore unlikely drug targets (Cortés, 2008). However, the mechanisms governing invasion protein sorting and trafficking may be similar among these proteins which are all destined for either the micronemes or the rhoptries. Disruption of this putative common pathway may prevent the invasion proteins from reaching their target destinations and thus the parasite will be unable to invade the host.

4.1 Protein domain selection

A multifunctional escorter has been identified in *T. gondii*, an apicomplexan closely related to *P. falciparum*. It has been shown that *Tg*SORTLR is essential not only for trafficking to, but also for biogenesis of micronemes and rhoptries (Sloves *et al.*, 2012). A *P. falciparum* homologue, *Pf*SORTILIN, has been speculated to sort specific cargo from the cis-Golgi to the endosomal network, with subsequent recycling to the Golgi mediated by the plasmodial retromer cargo-selective complex (Krai *et al.*, 2014).

The current study hypothesized that *Pf*SORTILIN interacts with target domains of cargo invasion proteins and may be a common escorter of these proteins to both the micronemes and rhoptries. The CCys domain of *Pf*MAEBL and the prodomain of PfAMA-1 were chosen because they are needed for correct trafficking to the rhoptries and micronemes, respectively (Wepener, 2014; Churchyard, 2017). These domains were expressed as recombinant proteins and their interaction with *Pf*SORTILIN VPS10 was determined by conducting *in vitro* binding assays.

This study also attempted to identify novel binding partners of *Pf*SORTILIN VPS10 by biopanning it against a *P. falciparum* phage display library. No binding partners were identified.

4.2 Challenges with recombinant protein expression

Protein expression is a tightly regulated and complex process. It requires RNA to be transcribed from DNA, which is then processed into mRNA which is translated into proteins on ribosomes. These events occur separately in eukaryotes whereas in prokaryotes they are integrated (Berg *et al.*, 2002).

4.2.1 Expressing eukaryotic proteins in a prokaryotic system

Prokaryotic organisms, like *E. coli*, are one of the preferred hosts for recombinant protein expression because they offer i) an inexpensive form of carbon source needed for growth, ii) fast biomass accumulation and iii) a simple process that can be upscaled. However, *E. coli* lack post-translational machinery and may also produce inactive protein due to inclusion body formation (Sahdev *et al.*, 2007). These drawbacks create many challenges when attempting to express foreign proteins like those found in *P. falciparum*. These include poor expression and solubility due to the A/T richness of the *Plasmodium* genome, protein disorder, large size of the proteins and the presence of low complexity regions (Birkholtz *et al.*, 2008).

4.2.2 Expression levels and solubility

The expression and solubility of His-tagged *Pf*MAEBL CCys and GST-tagged *Pf*SORTILIN VPS10 recombinant proteins were investigated in this study, whereas conditions for *Pf*AMA-1 prodomain had already been determined by A. Churchyard (2017).

Solubility profiles display the predicted solubility per amino acid position (Figure 4.1). The algorithm considers physico-chemical properties of amino acid sequences, such as, disorder, coil and hydrophilicity propensities which best discriminate between soluble and insoluble proteins. Pro, Glu, Ser, Lys and Gln are the most disorder-prone and associated with solubility whereas, Asn, Asp, Gly and His are the most coil-prone and associated with insolubility (Center for Genomic Regulation, 2012). The algorithm predicted that *Pf*SORTILIN VPS10 and *Pf*AMA-1 prodomain would have a high degree of solubility which is illustrated by the black area in Figure 4.1 and Table 4.1. The solubility profile of *Pf*MAEBL CCys predicted a high degree of insolubility.

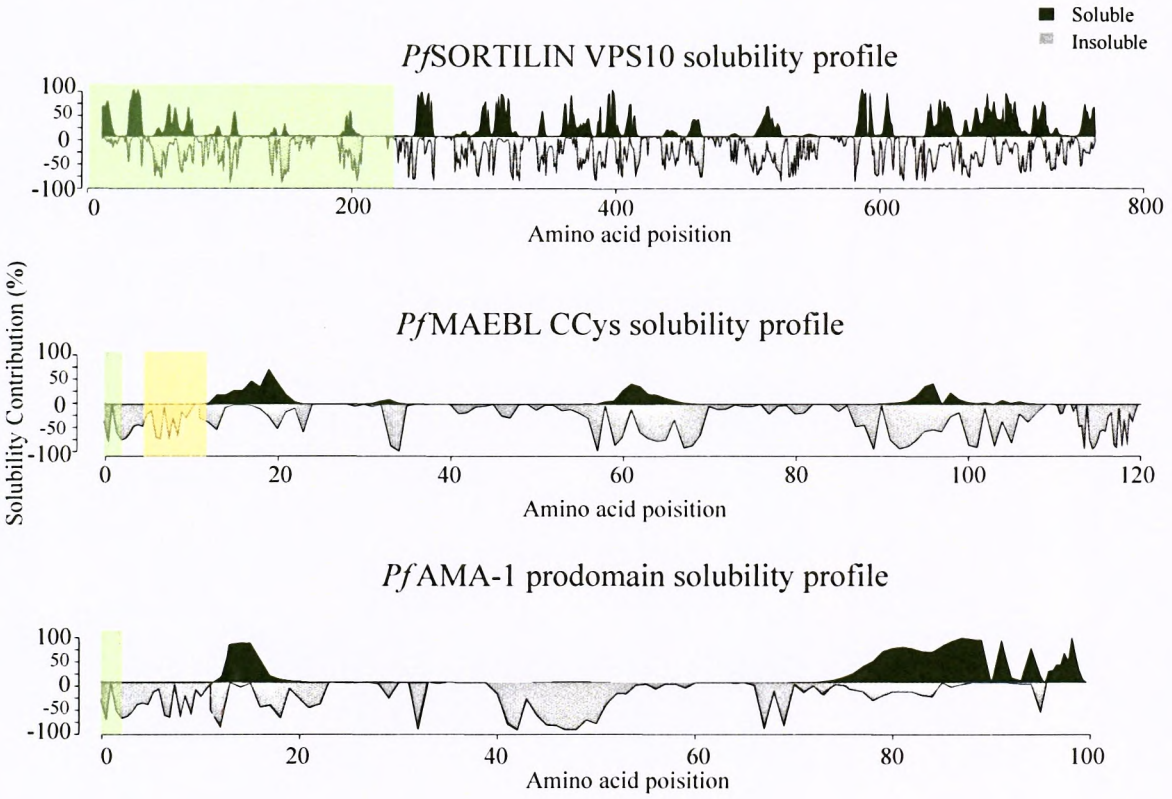


Figure 4.1: **Predicted solubility profiles of recombinant *Pf*SORTILIN VPS10 , *Pf*MAEBL CCys and *Pf*AMA-1 prodomain. This diagram indicates the probability of solubility as per the amino acid composition of each recombinant protein. Yellow represents a low complexity region. Green areas show the N-terminal fusion tags, GST or His. These profiles were generated by the ccSOL omics online program (http://s.tartaglialab.com/ page/ccsol_group).**

Table 4.1: **Summary of properties of recombinant proteins generated in this study (with and without tag). Solubility was predicted with the ccSOL omics online program (http://s.tartaglialab.com/ page/ccsol_group) and pI was determined using the ProtParam online tool (web.expasy.org/compute_pi).**

Protein	Predicted solubility %	pI
*Pf*SORTILIN VPS10	72%	5.48
GST-tagged *Pf*SORTILIN VPS10	98%	5.67
*Pf*MAEBL CCys	22%	8.36
His-tagged *Pf*MAEBL CCys	51%	8.38
*Pf*AMA-1 prodomain	79%	4.51
His-tagged *Pf*AMA-1 prodomain	17%	5.45

The fusion tags, highlighted in Figure 4.1, should increase solubility of the expressed protein (GE Healthcare Life Sciences, 2014; Qiagen, 2002) and it can be seen in that the GST-tag contributes more to the solubility of a protein than the His-tag. Table 4.1 shows that the predicted solubility of *Pf*SORTILIN VPS10 and *Pf*MAEBL CCys increased with the addition of their respective recombinant tag, whereas *Pf*AMA-1 promain's predicted solubility decreased from 79% to 17% with the His-tag. These predictions are said to be ≈78% accurate (Center for Genomic Regulation, 2012), however, they did not correlate with the actual solubility levels, indicating the unreliability of the algorithm.

All proteins predominantly expressed in an insoluble form, presumably due to the development of inclusion bodies. Smaller proteins are expected to have higher expression levels which was true for *Pf*MAEBL CCys (13 kDa) and *Pf*AMA-1 prodomain (11 kDa). Proteins larger than 40 kDa tend to aggregate and misfold into inclusion bodies, which may explain the low yield of the *Pf*SORTILIN VPS10 protein (86 kDa) (Mehlin *et al.*, 2006). The isolectric point (pI) of a protein also influences its expression. Proteins with pI values that do not range between 6-8 often elicit poor expression/solubility (Mehlin *et al.*, 2006). The pI values of the recombinant proteins in this study did not fall within this range (Table 4.1), however, the challenge was not in expression of these proteins but rather the amount of soluble protein obtained. The recombinant tags negligibly changed the pI values.

Other characteristics of proteins can contribute to their expression and solubility in *E. coli* bacteria. The region of *Pf*MAEBL CCys cloned for this study contained a low complexity region which is known to contribute to insolubility of a protein (Mehlin *et al.*, 2006; Figure 4.1). In addition, the domain has 8 Cys residues (7% of the domain) and the disulphide bonds may contribute to poor expression and solubility (De-Marco, 2009). *Pf*SORTILIN VPS10 does not contain a low complexity region and the 5 Cys residues which constitute <1% of the entire protein, is unlikely influence solubility. The large size of this protein may contribute to the insolubility (Mehlin *et al.*, 2006).

After attempting to improve the yield of soluble recombinant protein, the solubility levels were

still poor and multiple flasks of bacterial cultures were needed to obtain enough protein for the downstream applications in this study. Large scale *E. coli* cultures were investigated (>200 ml), however this yielded smaller amounts of soluble protein, which was likely due to incomplete lysis of the cell pellet. The ratio of the volume of the cell pellet to the volume of lysis buffer was suboptimal leading to inefficient sonication and disruption of the cells.

The amount of soluble protein could be increased by treating inclusion bodies with high molar concentrations of urea to solubilise the recombinant protein, followed by dialysis to allow renaturation and oxidative refolding of the protein which usually restores functionality (De-Marco, 2009), but further tests to confirm this would be necessary.

Several other strategies to improve solubility could be considered. These include the use of different tags, such as maltose-binding protein, or codon optimisation to suit the codon preference of *E. coli* (Kapust and Waugh, 1999; Birkholtz *et al.*, 2008). Instead of codon-optimization, Rosetta 2™ (DE3) cells, which contain a plasmid which encode for rare tRNAs, were used in this study to minimize codon bias (Mehlin *et al.*, 2006; Birkholtz *et al.*, 2008). Expression could also be attempted in eukaryotic insect cells or in a cell free wheat germ system. The latter would be able to express proteins without the need for codon optimization and has been successful in expressing *Pf*AMA-1 (Tsuboi *et al.*, 2008; Birkholtz *et al.*, 2008).

4.3 Unexpected migration of recombinant proteins on SDS-PAGE

GST-tagged *Pf*SORTILIN VPS10 migrated at a molecular size much smaller than expected throughout this study (Table 4.2). This anomaly is not uncommon for *P. falciparum* proteins. The presence of more Lys residues are likely to cause more SDS to bind to a protein so it will migrate faster and appear to have a smaller molecular weight (MW) (Tompa, 2003). The *Pf*SORTILIN VPS10 protein along with its GST-tag contained 56 Lys residues which make up

7% of the whole protein (Figure 4.2) and may explain the apparent MW difference. *Pf*MAEBL CCys contained 17 and *Pf*AMA-1 prodomain contained 2 Lys residues which may account for the slight difference in the observed MWs (Table 4.2).

Table 4.2: **Summary of molecular weights of recombinant proteins generated in this study. Observed molecular weights were determined by a graph of logarithmic MW versus distance migrated on SDS-PAGE. MW: molecular weight**

Protein	Expected MW	Observed MW
GST-tagged *Pf*SORTILIN VPS10	86 kDa	≈67 kDa
His-tagged *Pf*MAEBL CCys	13 kDa	≈12 kDa
His-tagged *Pf*AMA-1 prodomain	11 kDa	≈16 kDa

MSPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDG
DVKLTQSMAIIRYIADKHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDFL
SKLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKR
IEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDLVPRGSPGIHRDMTVLVKTVKGRL
YRSSDGGKIWTNITGNLSDKENNKSDPSSNHMNEMTTVDLIMVNSVNKNIVLIIGNQKNHYI
SEDSGETFRLLNYQNKINFWQFHSTKTHWALVSSWTEACYSNDDNSGECMQTLSLTKDLGR
TFQLIDIYVVQFNWGDKESHLEDTIYYTRHKNRNGHQQRFSGWSKDVDFVATHNFGKSVDV
LVKQGNKFLISNGYIFVAKLNDVIKQTVNMMVSTDGGKTFNKANLPENIHEKSYTVLDTSE
GSIMLHVNHGTSSEKINTGNVYISDASGLNYTLSLPNNIRTSSGECEFDRVLSLDGVYIANFL
DVNNEIKDEDLRFQNFKSAIEEDIAPFETNTEKRKKQYNKGKNEDAVRTVISFNKGGEWSYL
KAPKVDSRGNKYECGDNCYLHLHGITNYHQYAPFYSIENAVGIIMGTGNVGSHLRYESDEV
NTFLSRDGGVTWIEAHKGPYIYEYGDYGGLIVMADDLRKTNQIVFSWNEGQSWFDFELGQF
SIDVDNIVAEPNSASVEFLVYGTRNDVGVLYHLDFNA

Figure 4.2: **Lys residues in GST-tagged *Pf*SORTILIN VPS10. The sequence highlighted in yellow is the GST-tag and all Lys residues are indicated in red.**

4.4 *In vitro* binding assays determine interactions of *Pf*SORTILIN VPS10

To determine the escorter properties of *Pf*SORTILIN, binding assays were conducted with *Pf*MAEBL, a rhoptry protein and *Pf*AMA-1, a microneme protein.

4.4.1 *Pf*SORTILIN VPS10 interacts with *Pf*MAEBL CCys

Immobilized *Pf*MAEBL CCys bound to *Pf*SORTILIN VPS10 with a stoichiometry of 3.2:1, but in the reciprocal experiment the stoichiometry could not be determined. This interaction was dose-dependent, specific and not facilitated by the GST-tag of *Pf*SORTILIN VPS10 . This is the first description of an interaction between these two proteins. Previously *Pf*RAMA has been shown to interact with *Pf*RAP1and *Pf*SORTILIN, indicating that *Pf*RAMAand *Pf*RAP1are likely to form a complex (RAP complex) associated with *Pf*SORTILIN that is then translocated to the rhoptries from the trans-Golgi (Richard *et al.*, 2009; Hallée *et al.*, 2018). The results in the current study support the possibility that *Pf*SORTILIN VPS10 is a rhoptry protein escorter.

4.4.2 *Pf*SORTILIN VPS10 does not interact with *Pf*AMA-1 prodomain

This study is the first investigation of the interactions of *Pf*SORTILIN VPS10 with a microneme invasion protein. The *in vitro* assays revealed that immobilized *Pf*AMA-1 prodomain bound to *Pf*SORTILIN VPS10 but the GST-tag of *Pf*SORTILIN VPS10 also bound. When the GST-tag was used to immobilize *Pf*SORTILIN VPS10 there was no binding to free *Pf*AMA-1 prodomain. These results imply that *Pf*SORTILIN VPS10 is unlikely to be an escorter of *Pf*AMA-1 prodomain to the micronemes. Other domains of *Pf*AMA-1 could be investigated, such as the ectodomain, which makes up the hydrophobic binding pocket that binds to *Pf*RON2 on the surface of the merozoite where it facilitates re-orientation and tight junction formation (Harris *et al.*, 2005; Bai *et al.*, 2005). The role of the binding pocket in the trafficking of *Pf*AMA-1 to the micronemes is unknown.

4.5 Comparison between the *Pf*MAEBL CCys domain and *Pf*AMA-1 prodomain

The protein sequence of *Pf*MAEBL CCys was aligned to that of *Pf*AMA-1 prodomain using the Clustal alignment tool (www.ebi.ac.uk/Tools/msa/clustalo/) to determine if there was any similarity between the domains. These domains have a 10% sequence identity (Figure 4.3), but there is no common sequence between these proteins that *Pf*SORTILIN VPS10 would have recognized.

```
PfMAEBLCCys        MSN-------NNMKGNNNDIIYTKLDVEEY    23
PfAMA-1prodomain   YQNSDVYRPINEHREHPKEYEYPLHQEHTY    36

PfMAEBLCCys        KKRDVNETRE----------------------   33
PfAMA-1prodomain   QQEDSGEDENTLQHAYPIDHEGAEPAPQEQ    66

                                  ↓          ↓
PfMAEBLCCys        ----KIIKISKKNMCNNDFSSKYCDYMKDN    59
PfAMA-1prodomain   NLFSSIEIVERSNYMGNPWTEYMAKYDIEE    96

                        ↓        ↓↓      ↓
PfMAEBLCCys        ISSGTCSNEERKSLCCSISDFCLKYFDHNS    89
PfAMA-1prodomain   VHG--------------------------    99

                        ↓                 ↓
PfMAEBLCCys        NKYYDCTKIEFADPLYRCFKKKEFS   114
PfAMA-1prodomain   -------------------------    99
```

Figure 4.3: **ClustalW alignment of *Pf*MAEBL CCys and *Pf*AMA-1 prodomain**

aa: identical residues, aa: similar residues, -: gap residues. Cys residues are indicated by arrows. Alignment was performed using the ClustalW2 tool from www.ebi.ac.uk/Tools.

Both the 3D model of *Pf*MAEBL and the crystal structure of *Pf*AMA-1 have a globular conformation, even though the *Pf*MAEBL model had a 40% correlation with the crystal structure of *Tg*AMA-4, which is quite low. The cloned domains used in this study are highlighted in blue

in Figure 4.4. *Pf*AMA-1 prodomain is not as exposed as *Pf*MAEBL CCys which could also explain why only the latter bound to *Pf*SORTILIN VPS10.

It could be speculated that the hydrophobic residues of *Pf*MAEBL CCys, which are illustrated in Figure 4.4, may facilitate the binding of *Pf*MAEBL CCys to *Pf*SORTILIN VPS10. The binding pocket of *Pf*MAEBL is unknown and structural studies would be needed to ascertain which domain contributes to its formation. The M1/M2 domains of *Pf*MAEBL make up the majority of the protein, hence it would be of interest to investigate these domains as well. The hydrophobic binding pocket of *Pf*AMA-1 is illustrated in Figure 4.4 and includes a part of the ectodomain. The prodomain is a very small component of the whole protein and does not contribute to the formation of the binding pocket.

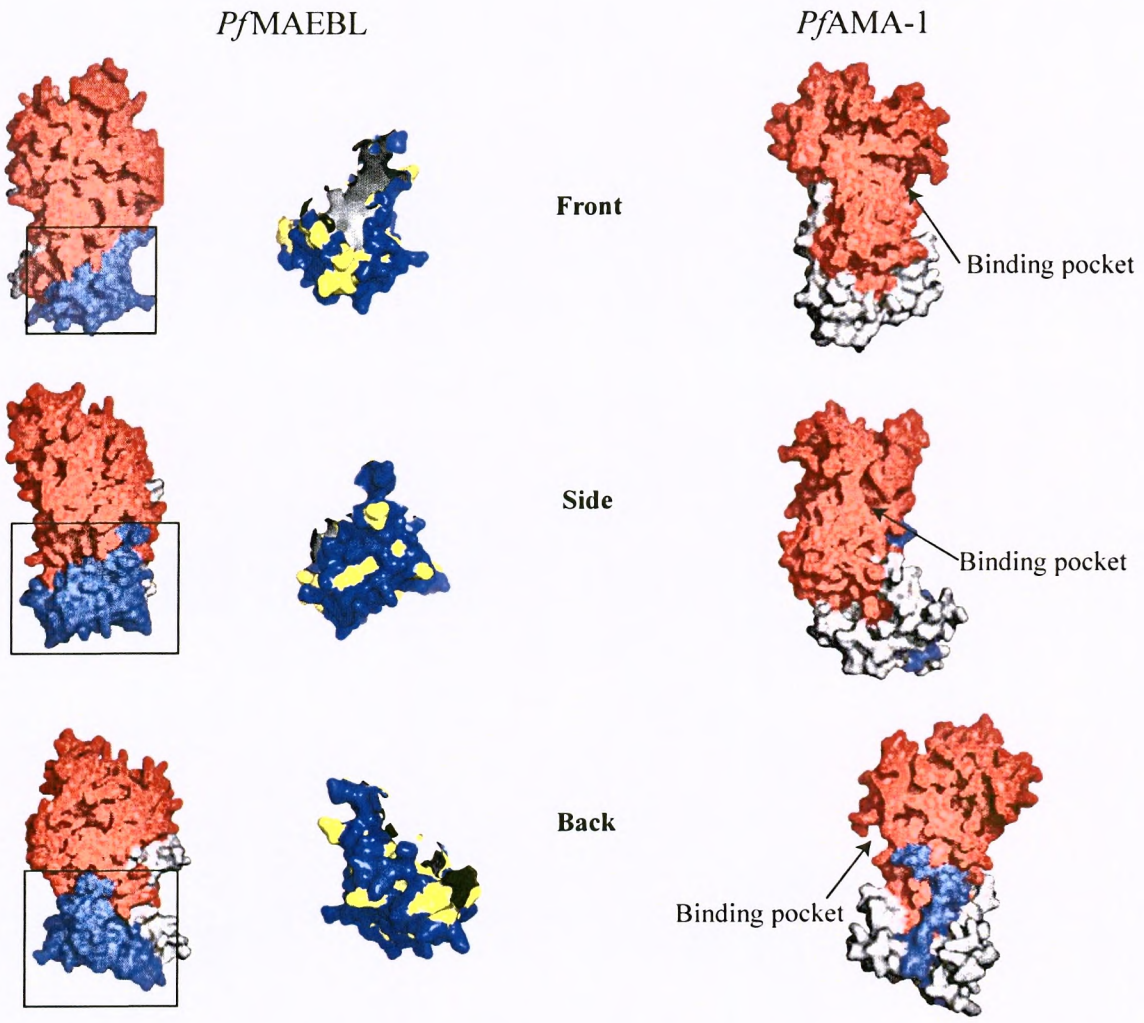


Figure 4.4: **3D surface model of *Pf*MAEBL and crystal structure of *Pf*AMA-1 (Bai *et al.*, 2005). The red residues indicate the M1/M2 domains and ectodomain of *Pf*MAEBL and *Pf*AMA-1, respectively. The blue residues indicate the CCys domain and prodomain of *Pf*MAEBL and *Pf*AMA-1, respectively. The *Pf*RON2 binding pocket of *Pf*AMA-1 is shown by an arrow. The hydrophobic residues of *Pf*MAEBL CCys are indicated in yellow. The 3D model of *Pf*MAEBL was based on the crystal structure of *Tg*AMA-4 with 40% of residues modelled at >90% confidence. The model was generated using the Phyre 2 online tool (www.sbg. bio.ic.ac.uk) and structures were visualized using the PyMOL computer software.**

4.6 Utility of the phage display library

Biopanning experiments conducted in this study were unable to identify binding partners of *Pf*SORTILIN VPS10. Sequencing revealed many identical small inserts composed of strings of Lys residues. The positively charged peptides most likely promoted non-specific ionic interactions during biopanning. The A/T bias of the *P. falciparum* genome was responsible for this enriched Lys dominance (Lys codons are AAA and AAG). During creation of the library the mRNA was isolated using oligodT, which would also interact with long stretches of A nucleotides in ribosomal or other non-coding sequences. The phage libraries will thus express some unnatural peptides on their surfaces, such as Lys peptides (Lanzillotti and Coetzer, 2008).

The 3-D surface model generated by the Phyre2 programme showed a variety of positively and negatively charged residues exposed on the outer surface of the recombinant PfSORTILIN VPS10. A group of negatively charged residues was observed in a protruding segment of the recombinant domain (Figure 4.5) and many negative residues are also concentrated in the central gap of the model, which could provide contact points for the Lys peptides.

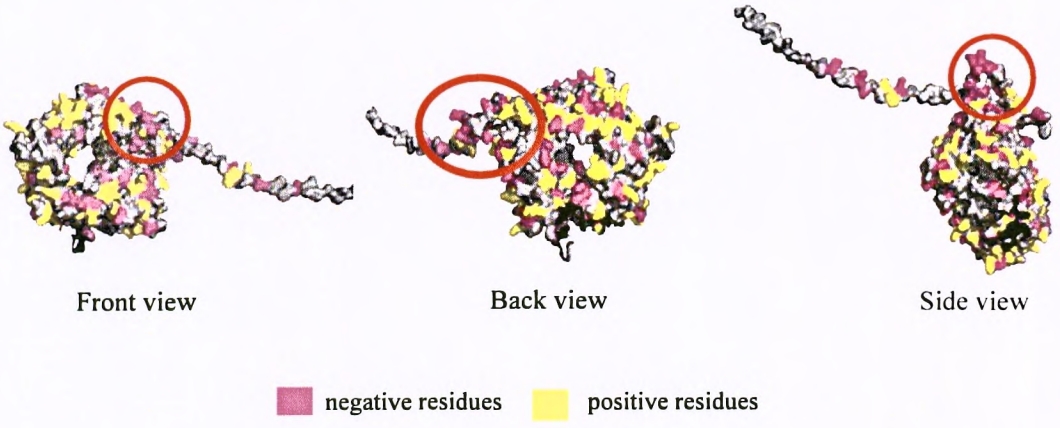


Figure 4.5: **3D surface model of *Pf*SORTILIN VPS10 showing positive and negative residues. The front, back and side views are shown. A protruding segment (red circle) contains mainly negatively charged residues (Asp/Glu). Model generated from Phyre 2 online tool (www.sbg. bio.ic.ac.uk).**

It should also be noted that the phage display library, which was generated in the Coetzer lab, was made with a mixed-stage parasite population. Unsuccessful binding could have been because the peptides did not represent schizont-specific inserts, such as the invasion proteins, which are predominantly expressed at this stage (Le Roch *et al.*, 2003). It would be of interest to develop stage-specific libraries for future studies.

4.7 *Pf*SORTILIN as a rhoptry protein escorter

4.7.1 Escorter Model

Invasion proteins are synthesized on ribosomes and translocated from the ER to the cis-Golgi network by their signal peptides (Kats *et al.*, 2008). The exact trafficking of these proteins to their respective secretory organelles is still not completely understood, however escorter models have been described.

4.7.1.1 Rhoptry escorter model

Results from this study showed that *Pf*MAEBL CCys binds directly to *Pf*SORTILIN VPS10 suggesting *Pf*SORTILIN could escort *Pf*MAEBL to the rhoptries in the same way as the *Pf*RAP complex (Figure 4.6-1/2). Rhoptry proteins aggregate into lipid rafts on the Golgi membrane which facilitates their association with an escorter (Topolska *et al.*, 2003). *Pf*SORTILIN is a transmembrane protein, with the VPS10 domain contained within the lumen of the Golgi apparatus (Krai *et al.*, 2014). It binds to *Pf*RAMA, which associates with *Pf*RAP1and this whole complex buds off from the trans-Golgi network into CCVs. The cytoplasmic tail of *Pf*SORTILIN is proposed to associate *Pf*AP1μ in the cytoplasm, which facilitates the movement of CCVs to the rhoptries where they fuse and the contents are delivered to the rhoptry lumen (Kibria *et al.*, 2015; Hallée *et al.*, 2018).

4.7.1.2 Microneme escorter model

The trafficking of proteins to the micronemes is less well elucidated. This study is the first to investigate the interaction of *Pf*SORTILIN VPS10 with a micronemal protein, *Pf*AMA-1. There was no binding between *Pf*SORTILIN VPS10 and *Pf*AMA-1 prodomain, which supports results from another study that revealed that *Pf*SORTILIN and *Pf*AMA-1 do not co-localize during any stage of the parasite in the RBC (Hallée *et al.*, 2018).

Micronemal proteins therefore adopt a different trafficking process, which most probably requires a microneme-specific escorter that interacts with a microneme-specific AP complex in the cytoplasm (Figure 4.6-3).

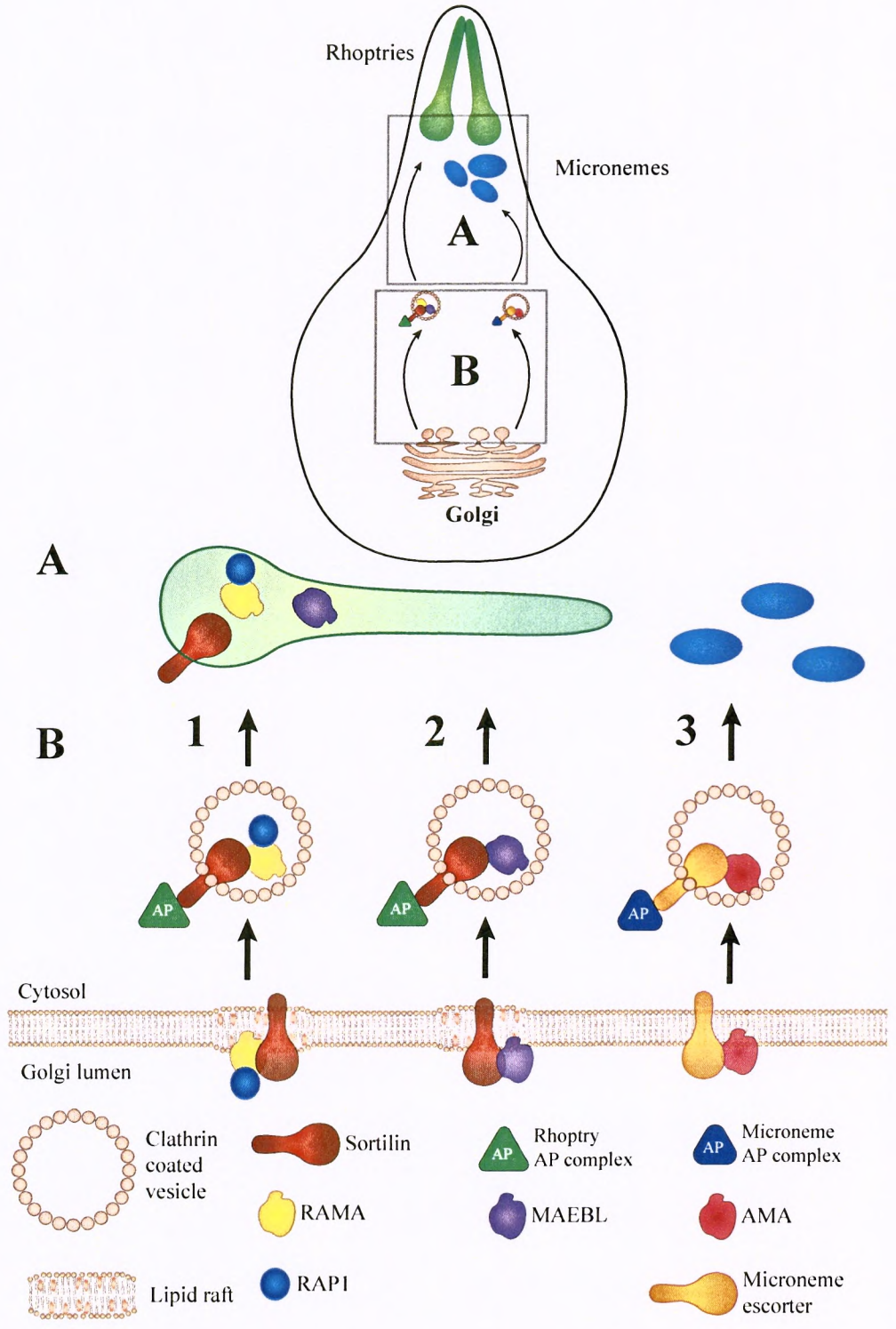


Figure 4.6: **Model of *Pf*SORTILIN VPS10 as a rhoptry protein escorter. A: Trafficking of proteins from the CCV to the rhoptries or micronemes. B: Trafficking of proteins from the Golgi into CCV. Trafficking models of 1: *Pf*RAP1 complex, 2: *Pf*MAEBL and 3: *Pf*AMA-1. MAEBL: merozoite adhesive erythrocyte binding-like protein, RAMA: rhoptry-associated membrane antigen, RAP1: rhoptry associated protein 1, AMA-1: apical membrane antigen-1, AP: adaptor protein.. Adapted from models by Richard *et al.* (2009), Counihan *et al.* (2013) and Kibria *et al.* (2015).**

The information gained from this study contributes to our knowledge of apical organelle trafficking in the malaria parasite. *Pf*SORTILIN is known to be an escorter of the *Pf*RAP1 complex (Hallée *et al.*, 2018) and appears to also facilitate the trafficking of *Pf*MAEBL. It is speculated that *Pf*SORTILIN is a common rhoptry protein escorter, which may be targeted using novel therapeutic strategies in combating malaria.

4.8 Future work

Further work can be done to support the hypothesis that *Pf*SORTILIN is a rhoptry protein escorter. Proteins could be expressed on a larger scale in a cell free wheat germ system. To improve the purity of the recombinant proteins, additional methods such as size exclusion or ion exchange chromatography can be used. With a larger amount of protein the interaction between *Pf*MAEBL/*Pf*AMA-1 and *Pf*SORTILIN can be determined using biophysical methods such as surface plasmon resonance or isothermal titration calorimetry, which would provide more quantitative data. The proposed interaction between *Pf*AP1μ and *Pf*SORTILIN should be investigated. To identify novel binding partners of *Pf*SORTILIN a schizont-specific stage *P. falciparum* phage display library could be created.

The interaction of the full length proteins, as opposed to domains, should be investigated. Soluble expression of these large recombinant proteins may be problematic, so *in vivo* methods can be employed. A transgenic parasite line expressing *Pf*SORTILIN tagged with a fluorescent molecule, such GFP, can be generated with the CRISPR-Cas9 system. Pull down assays can be conducted with this line using an antibody against GFP and novel binding partners can be identified by mass spectrometry. The *Pfsortilin* gene can be knocked-out using either the CRISPR-Cas9 system or double homologous crossover recombination. If the gene is essential no viable parasites will be produced. If the knock-out is not lethal, this line can be used to investigate the effect of the lack of this protein on the parasite. Episomal plasmids expressing

fluorescently-tagged recombinant invasion proteins, such as *Pf*MAEBL, can be used to transfect these parasites and determine what effect the absence of *Pf*SORTILIN will have on the trafficking of these proteins.

Appendix A

Primers, DNA and protein Sequences

A.1 PCR primers

The red underlined sequence show the primer sequences from the *P. falciparum* gene sequence. Green bases show the enzyme restriction sites with cleavage sites shown by an arrow (↓). Bold bases are the **extra bases** added to facilitate enzyme recognition. Red highlighted bases show the insertion of a stop codon in the reverse primer. T_m values were determined using the Integrated DNA Technologies SciTools Oligo Analyzer 3.1.

***Pfmaebl CCys* Sense primer**

MAEBL CCys-*Nde*1: 5'-**GCGAGGCAGG** CA↓TATGTCG ATT AAT AAC ATG AAA-3'

T_m of whole primer = 61.6 °C

T_m of *Pf*-sequence = 45.2 °C

***Pfmaebl CCys* Anti-sense primer**

MAEBL CCys-*Bam*H1: 5'-**GCGAG** G↓GATCC TCA TGA AAA TTC CTT TTT TTT AA-3'

T_m of whole primer =59.5 °C

T_m of *Pf*-sequence = 41.3 °C

***Pfsortilin VPS10* Sense primer**

SORTILIN VPS10-*Bam*H1: 5'-**GCGAG** G↓GATCCATG ACA GTT TTA GTA AAA AC-3'

T_m of whole primer = 59.3 °C

T_m of *Pf*-sequence = 43.5 °C

***Pfsortilin VPS10* Anti-sense primer**

SORTILIN VPS10-*Xho*1:5'-**GCGAG** C↓TCGAG TCA AGC ATT AAA ATC TAG ATG AT-3'

T_m of whole primer = 61 °C

T_m of *Pf*-sequence = 47.8 °C

A.2 Primers used for construct sequencing

Primers used for sequencing of *Pf*MAEBL CCys and *Pf*SORTILIN VPS10 constructs:

pET-15-b T7 promoter primer

5'-CCTATAGTGAGTCGTATTA-3'

pGEX-4T-2 Sense primer

5'-GGGCTGGCAAGCCACGTTTGGTG-3'

pGEX-4T-2 Anti-sense primer

5'-CCGGGAGCTGCATGTGTCAGAGG-3'

A.3 Sizes of vectors and *P. falciparum* inserts

Table A.1: **Sizes of vectors and *P. falciparum* inserts**

Sample	Size (bp)
pET-15b	5708
*Pf*MAEBL CCys *Pf*-sequence amplified	342
*Pf*MAEBL CCys amplicon	369
*Pf*MAEBL CCys insert	351
*Pf*MAEBL CCys recombinant construct	6047
pGEX-4T-2	5056
*Pf*SORTILIN VPS10 *Pf*-sequence amplified	1623
*Pf*SORTILIN VPS10 amplicon	1648
*Pf*SORTILIN VPS10 insert	1636
*Pf*SORTILIN VPS10 recombinant construct	6667

A.4 Protein sequences

The protein sequences of the *Pf*MAEBL CCys and *Pf*SORTILIN VPS10 domains which were cloned into the the pET-15b and pGEX-4T-2 vectors respectively, are shown in the figures below:

A.4.1 *Pf*MAEBL CCys

```
                                                          ↓
MSNNNMKGNNNDIIYTKLDVEEYKKRDVNETREKIIKISKKNMCNNDFSS    50

  ↓             ↓          ↓↓     ↓               ↓
KYCDYMKDNISSGTCSNEERKSLCCSISDFCLKYFDHNSNKYYDCTKIEF   100

       ↓
ADPLYRCFKKKEFS   114
```

Figure A.1: **Protein sequence of recombinant *Pf*MAEBL CCys**

Sequence was obtained from PlasmoDB (2018). Cys residues are indicated by arrows.

A.4.2 *Pf*AMA-1 prodomain

```
NYWEHPYQNSDVYRPINEHREHPKEYEYPLHQEHTYQQEDSGEDENTLQH    50

AYPIDHEGAEPAPQEQNLFSSIEIVERSNYMGNPWTEYMAKYDIEEVHG    99
```

Figure A.2: **Protein sequence of recombinant *Pf*AMA-1 prodomain**

Sequence was obtained from PlasmoDB (2018).

A.4.3 *Pf*SORTILIN VPS10

```
MTVLVKTVKGRLYRSSDGGKIWTNITGNLSDKENNKSDPSSNHMNEMTTV   50

DLIMVNSVNKNIVLIIGNQKNHYISEDSGETFRLLNYQNKINFWQFHSTK  100

THWALVSSWTEACYSNDDNSGECMQTLSLTKDLGRTFQLIDIYVVQFNWG  150

DKESHLEDTIYYTRHKNRNGHQQRFSGWSKDVDFVATHNFGKSVDVLVKQ  200

GNKFLISNGYIFVAKLNDVIKQTVNMMVSTDGGKTFNKANLPENIHEKSY  250

TVLDTSEGSIMLHVNHGTSSEKINTGNVYISDASGLNYTLSLPNNIRTSS  300

GECEFDRVLSLDGVYIANFLDVNNEIKDEDLRFQNFKSAIEEDIAPFETN  350

TEKRKKQYNKGKNEDAVRTVISFNKGGEWSYLKAPKVDSRGNKYECGDNC  400

YLHLHGITNYHQYAPFYSIENAVGIIMGTGNVGSHLRYESDEVNTFLSRD  450

GGVTWIEAHKGPYIYEYGDYGGLIVMADDLRKTNQIVFSWNEGQSWFDFE  500

LGQFSIDVDNIVAEPNSASVEFLVYGTRNDVGVLYHLDFNA   541
```

Figure A.3: **Protein sequence of recombinant *Pf*SORTILIN VPS10**

Sequence was obtained from PlasmoDB.

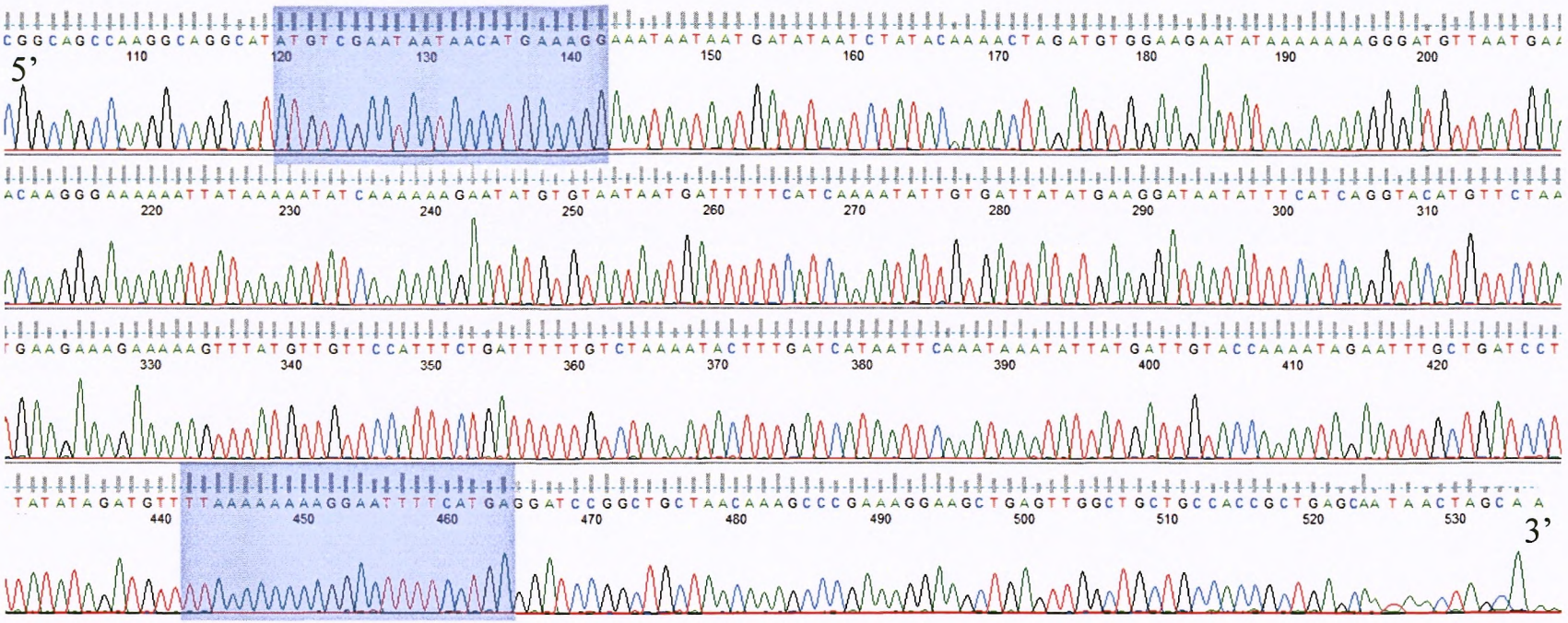


Figure A.4: **The chromatogram of the pET-15-b-*Pf*MAEBL CCys construct from colony 1. The vector construct was sequenced with the T7 promoter primer. The insert sequence starts at 120bp and ends at 465bp, with the forward and reverse primers highlighted in blue.**

```
120    ATGTCGAATAATAACATGAAAGGAAATAATAATGATATAATCTATACAAAACTAGATGTG  180
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5621   ATGTCGAATAATAACATGAAAGGAAATAATAATGATATAATCTATACAAAACTAGATGTG  5681

181    GAAGAATATAAAAAAAGGGATGTTAATGAAACAAGGGAAAAAATTATAAAAATATCAAAA  241
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5682   GAAGAATATAAAAAAAGGGATGTTAATGAAACAAGGGAAAAAATTATAAAAATATCAAAA  5741

242    AAGAATATGTGTAATAATGATTTTTCATCAAAATATTGTGATTATATGAAGGATAATATT  301
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5742   AAGAATATGTGTAATAATGATTTTTCATCAAAATATTGTGATTATATGAAGGATAATATT  5801

181    TCATCAGGTACATGTTCTAATGAAGAAAGAAAAAGTTTATGTTGTTCCATTTCTGATTTT  361
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5802   TCATCAGGTACATGTTCTAATGAAGAAAGAAAAAGTTTATGTTGTTCCATTTCTGATTTT  5861

362    TGTCTAAAATACTTTGATCATAATTCAAATAAATATTATGATTGTACCAAAATAGAATTT  421
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5862   TGTCTAAAATACTTTGATCATAATTCAAATAAATATTATGATTGTACCAAAATAGAATTT  5867

422    GCTGATCCTTTATATAGATGTTTTAAAAAAAAGGAATTTTCATGA  465
       |||||||||||||||||||||||||||||||||||||||||||
5868   GCTGATCCTTTATATAGATGTTTTAAAAAAAAGGAATTTTCATGA  5912
```

Figure A.5: **Alignment of *Pf*MAEBL CCys sequence from 5' sequencing. The sequence obtained from cloning (red) was aligned with the original sequence of *Pf*MAEBL CCys which was obtained from the PlasmoDB database (blue). The underlined sections indicate the forward and reverse primer positions.**

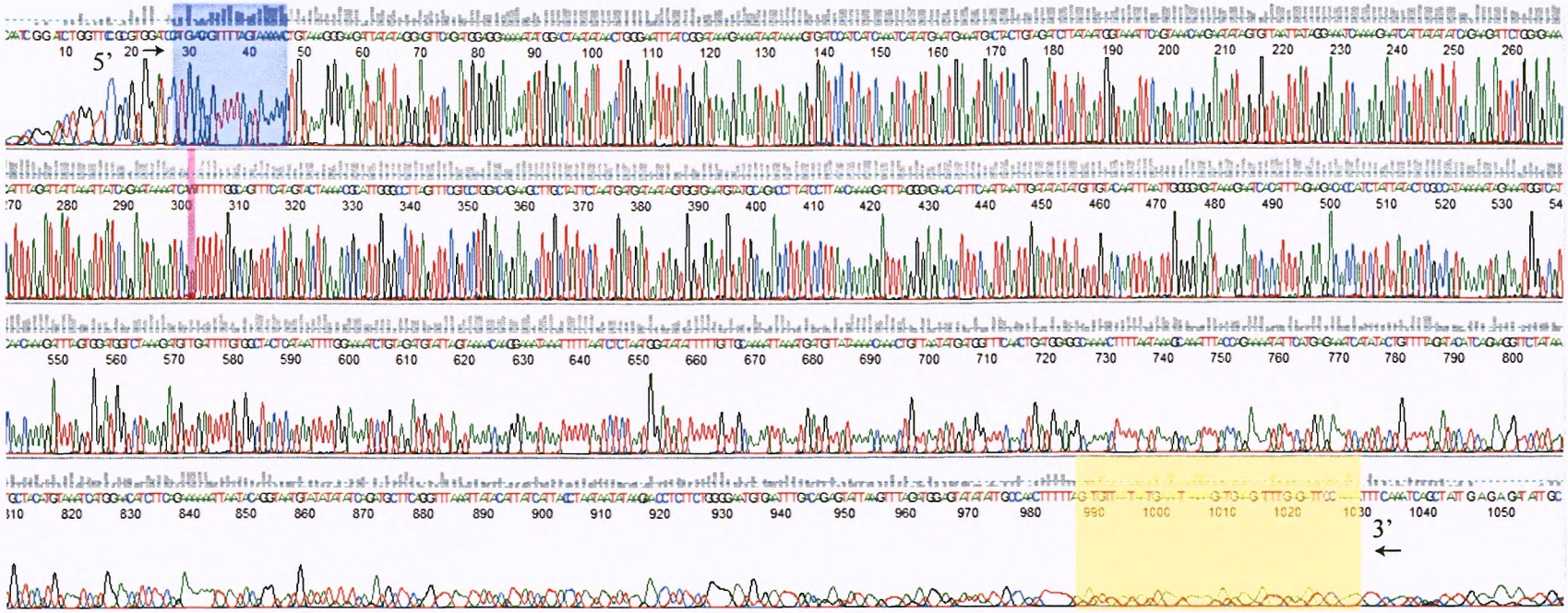


Figure A.6: **Forward sequence of *Pf*SORTILIN VPS10 vector construct. A vector construct from *Pf*SORTILIN VPS10 colony 1A was sequenced with pGEX-4T-2 sense primer, highlighted in blue. The overlapping region of the forward and reverse sequences (Figure A.7) is highlighted in yellow. W indicates a weak A/T which is highlighted in pink. Manual inspection revealed an A signal, which is the correct nucleotide in the gene sequence. The insert sequence used for alignment verification was from 28 bp to 1030 bp (arrow).**

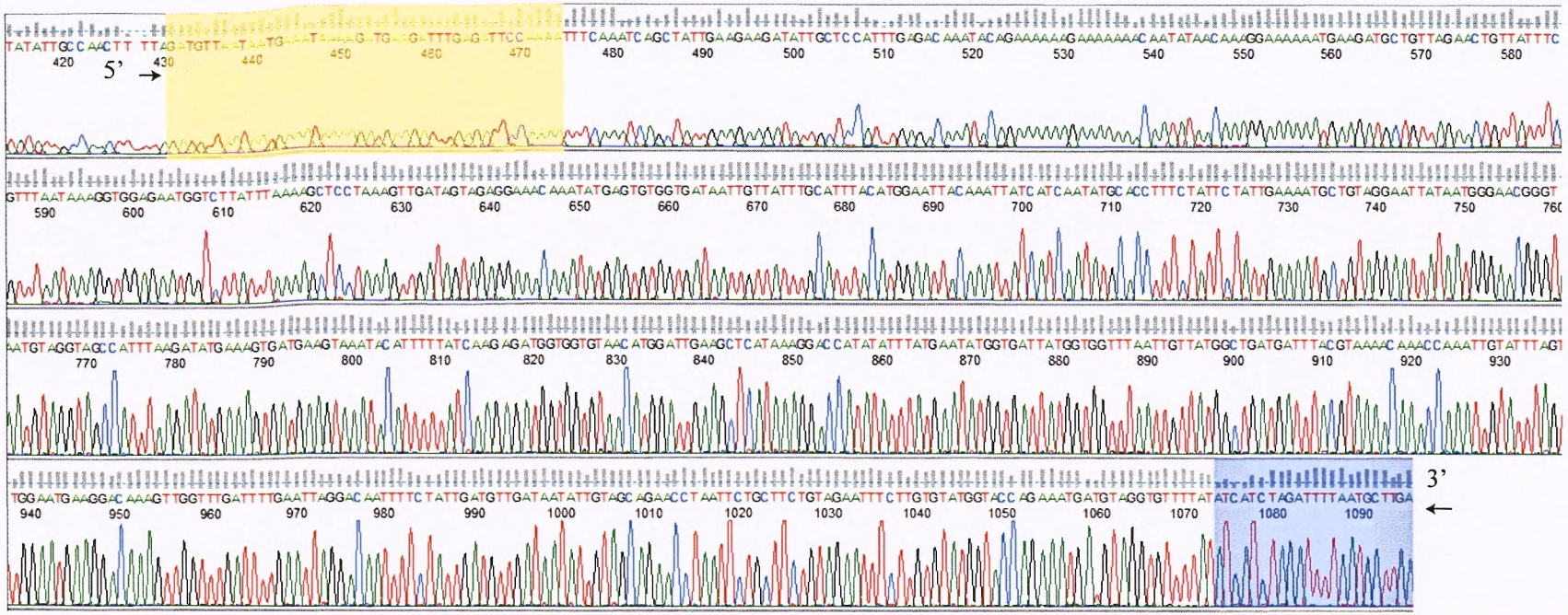


Figure A.7: **Reverse sequence of *Pf*SORTILIN VPS10 vector construct. A vector construct from *Pf*SORTILIN VPS10 colony 1A was sequenced with pGEX-4T-2 anti-sense primer, highlighted in blue. The overlapping region of the forward (Figure A.6) and reverse sequences is highlighted in yellow. The insert sequence used for alignment verification was from 431 bp to 1096 bp (arrow).**

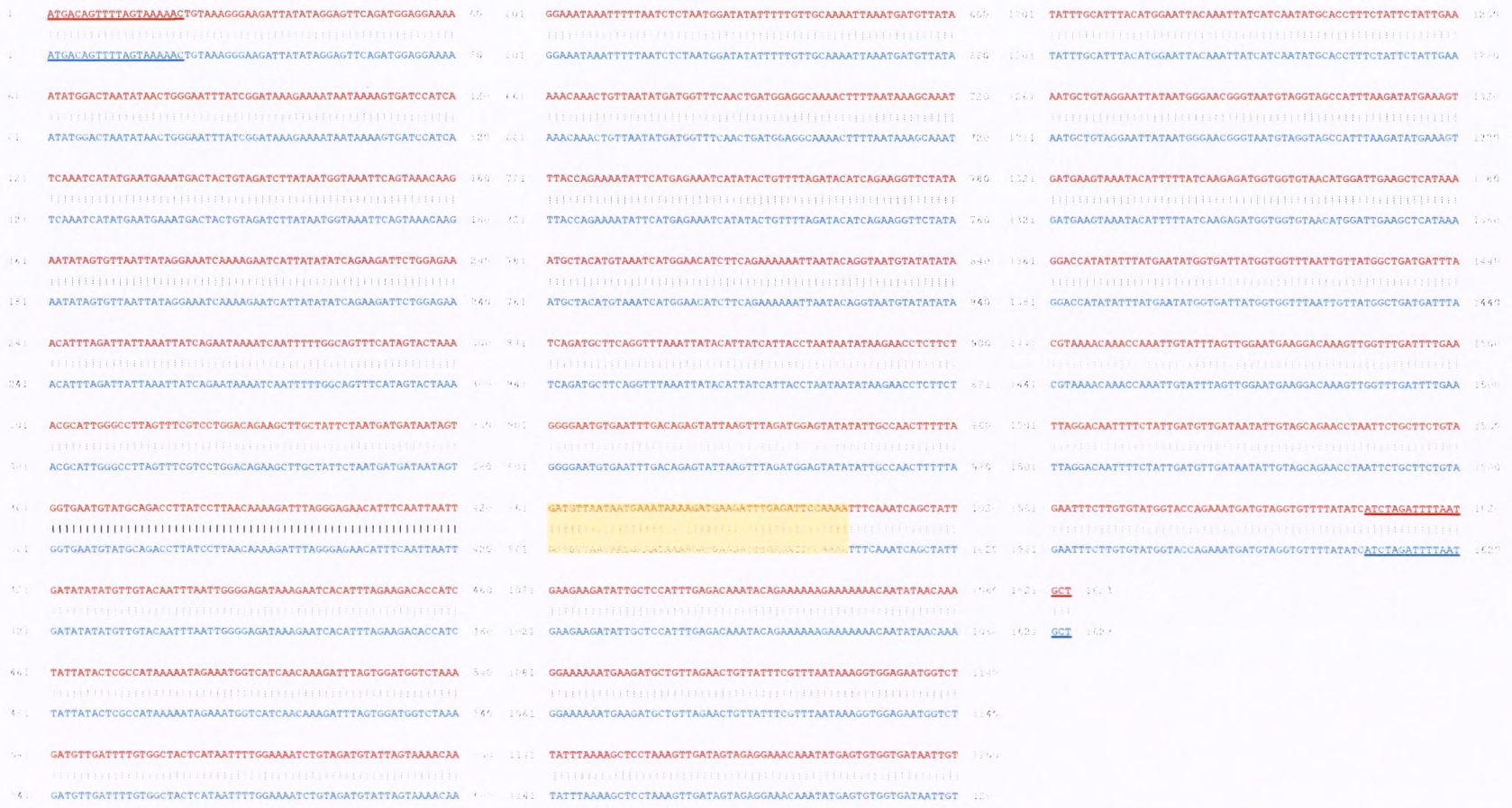


Figure A.8: **Alignment of *Pf*SORTILIN VPS10 construct sequence from 5' sequencing. The sequence obtained from cloning (red) was aligned with the original sequence for *Pf*SORTILIN VPS10 which was obtained from the PlasmoDB database (blue). The underlined sections indicate the forward and reverse primer positions. Highlighted in yellow are the overlapping sequences obtained using either the pGEX-4T-2 sense and anti-sense primer.**

Appendix B

Commercial chemical reagents

1,4-piperazinediethanesulphonic acid (Sigma-Aldrich Corporation, USA)

Acetic acid (Merck, Germany)

Acrylamide (Sigma-Aldrich Corporation, USA)

Agarose (SeaKem® LEAgarose) (Lonza, USA)

Albumax (Gibco, UK) Stored at 4 °C

Amido black (AppliChem, Germany)

Ammonium chloride (NH_4Cl) (Saarchem (Pty) Ltd., RSA)

Ammonium persulfate (Merck, Germany)

Ampicillin (Roche Applied Science, Germany)

Anti-GST HRP-conjugated antibody (Lot no.: 9475836 and 9800159, GE Healthcare Life Sciences, UK) Stored at –20 °C

Anti-His HRP-conjugated antibody (Lot no.: 148049181 and 154018594, Qiagen, UK) Stored at 4 °C

Bacteriological agar (Merck, Germany)

Bis-acrylamide (Sigma-Aldrich Corporation, USA)

Bromophenol blue (ThermoScientific, South Africa)

BSA (Pierce,USA)

Chloramphenicol (Sigma-Aldrich Corporation, USA)

Chloroform (Merck, Germany)

DH5α *E. coli* cells (Invitrogen Ltd, UK) Stored at –70 °C

Di-Sodium hydrogen orthophosphate ($Na_2HPO_4 \cdot 2\,H_2O$) (Saarchem, South Africa)

DNAseI (ThermoScientific, USA)

DreamTaq polymerase (ThermoScientific, South Africa)

D-sorbitol, SigmaUltra (Sigma-Aldrich Corporation, USA)

EDTA (Ethylenediaminetetraacetic acid) (Merck, Germany)

Fast AP™ (ThermoScientific, USA) Stored at –20 °C

Fast Digest™ enzymes (ThermoScientific, USA) Stored at –20 °C

Fast Digest™ Green buffer (10X) (ThermoScientific, USA) Stored at 4 °C

Coosmassie blue G-250 dye (ThermoScientific, South Africa)

Gentamycin sulphate (Sigma-Aldrich Corporation, USA) Stored at 4 °C

Glycerol (Merck, Germany)

Glycine (Sigma-Aldrich Corporation, USA)

HEPES (Sigma-Aldrich Corporation, USA)

HRP-conjugated blocking buffer (Qiagen, UK)

Hypoxanthine (Sigma-Aldrich Corporation, USA)

IPTG (Promega,USA)

Isopropanol (Merck/ACE, South Africa)

MagneGST™ Protein Purification System (Promega, USA) Stored at 4 °C

MagneHis™ Protein Purification System (Promega, USA) Stored at 4 °C

Magnesium sulphate ($MgSO_4$) (Merck, Germany)

MassRuler™ DNA Ladder Mix (Fermentas, Canada)

Methanol (Merck/ACE, South Africa)

Molecular grade ethanol (Merck, South Africa)

Nuclease-free water (Inqaba, South Africa)

pET-15b plasmid (Novagen, USA) Stored at –70 °C

Overnight Express™ Instant TB Medium (Novagen, USA) Petri dishes (Costar, USA)

pGEX-4T-2 plasmid (Amersham Biosciences, UK) Stored at –70 °C

Phenol (Sigma-Aldrich Corporation, USA)

Ponceau S (Sigma-Aldrich Corporation, USA)

Phosphate buffered saline (PBS) tablets (Sigma, South Africa)

Phusion® Flash High-Fidelity PCR Master Mix (ThermoScientific, USA) Stored at –20 °C

Potassium chloride (KCl) (Sigma-Aldrich Corporation, USA)

Potassium dihydrogen phosphate (KH_2PO_4) (Saarchem, South Africa)

Potassium phosphate (K_2HPO_4) (BDH, UK)

Protease Inhibitor cocktail Set III (Calbiochem®, USA) Stored at –20 °C

Rapid™ Ligation Kit (Roche Applied Science, Germany)

Rapi-Diff Stain Set (Clinical Sciences Diagnostics CC, South Africa)

RNaseA (Fermentas, Europe) Stored at –20 °C

Rosetta 2™ *E. coli* cells (Novagen, UK) Stored at –70 °C

RPMI 1640 with L-Glutamine with 25 mM HEPES (Gibco, UK) Stored at 4 °C

Sodium acetate (CH_3COONa) (Hopkin & Williams, United Kingdom)

Sodium bicarbonate ($NaHCO_3$) (Saarchem, South Africa)

Sodium chloride (NaCl) (Saarchem, South Africa)

Sodium dihydrogen phosphate(NaH_2PO_4) (Saarchem, South Africa)

Sodium dodecyl sulphate (SDS) (Merck, Germany)

Sodium hydroxide (NaOH) (Merck, Germany)

Spectra Multicolour Broad Range Protein Ladder (ThermoScientific, USA)

Subcloning Efficiency™ DH5α cells (Invitrogen, USA)

Sucrose (Sigma-Aldrich Corporation, USA)

SuperSignal® West Pico Chemiluminescent Substrate (ThermoScientific,USA) Stored at 4 °C

TEMED (Sigma-Aldrich Corporation, USA) Stored at 4 °C

Trisaminomethane (Tris) (Roche, Germany)

Tryptone Pancreatic Digest of Casein (Merck, Germany)

Tween-20 (Sigma-Aldrich Corporation, USA)

Yeast extract (Oxoid, UK)

Appendix C

Equipment and consumables

Acid citrate dextrose tubes (Becton Dickinson Biosciences, USA)

Centrifuge 5702R (Eppendorf, Germany)

Centrifuge J2-21 (Beckman Coulter, USA)

Centrifuge tubes- 50 ml and 15 ml (Greiner Bio-one, Germany)

Cryotubes (TPP, Switzerland)

Eppendorf microcentrifuge tubes- 2 ml and 1,5 ml (Eppendorf, Germany)

Falcon™ round bottom tubes- 14 ml (Becton Dickinson Biosciences, USA)

Glass microscope slide (Lasec, South Africa)

GeneSnap GeneGenius Geldoc scanning system and software (Syngene, UK)

Hybond™-C extra supported nitrocellulose membrane(Amersham Biosciences, UK)

Intelli-mixer RM 2M Skyline (ELMI Ltd, Latvia)

Mastercycler gradient thermocycler (Eppendorf, Germany)

Axiostar Microscope (Zeiss, Germany)

MilliQ™ Water System (Millipore Corporation, USA)

MiniProtean BioRad gel cassette (BioRad, South Africa)

MiniProtean tetra cell (BioRad, South Africa)

MiniTransblot™ cassette (BioRad, South Africa)

Nanodrop ND-1000 spectrophotometer (ThermoScientific, USA)

Nuclease-free microcentrifuge tubes (QSP, USA)

NucleoSpin® Plasmid DNA purification Kit (Macherey-Nagel, Germany)

NucleoSpin® Gel and PCR Clean-up Kit (Macherey-Nagel, Germany)

Orbital Shaker (Labotec, South Africa)

Orbital bench top shaker PR250 (Hoefer Inc., USA)

Pasteur pipette (Lasec, South Africa)

PowerPac™ HC High-Current Power Supply (BioRad, South Africa)

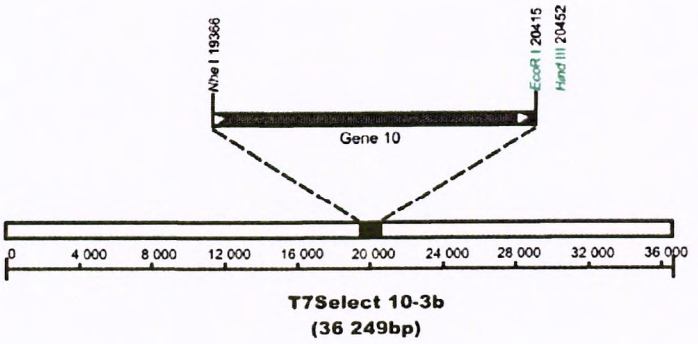
Sonoplus HD 3400 sonicator (Bandelin, Germany)

Sterile culture flasks- all sizes (Nunc, Denmark)

Appendix D

P. falciparum phage display library and PCR amplification of phage isolates

A



B

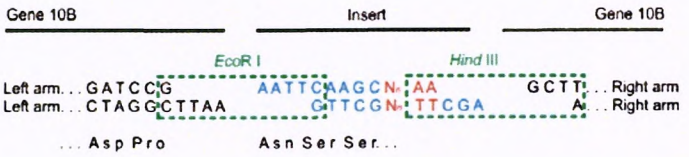


Figure D.1: **Schematic representation of the T7Select10-3b vector which was used to generate the *P. falciparum* phage display library from mRNA. Libraries were created by generating cDNA which was converted to dsDNA (with linkers on either end of the strand) and ligated into the vector as shown above. A: dsDNA was inserted into gene 10 of the T7Select10-3b vector using the *Eco*RI/*Hind*III restrictions sites. B: The sequence of Gene 10 after cloning the *P. falciparum* inserts is shown in more detail. Sequences in blue are provided by appended *Eco*RI/*Hind*III directional cloning linkers, whereas the nucleotides in red are provided by parasite inserts (N_n = *P. falciparum* cDNA sequence). The adenosine and thymine residues (red) are provided by the 3' end of the parasite cDNA strand. Image from Lanzillotti (2007).**

Table D.1: **Primers used for sequencing of bacteriophage that bound to *Pf*SORTILIN VPS10**

Vector	Primer name	Primer direction	Primer sequence (5' to 3')	Predicted T_M (°C)
T7 phage	T7Select-UP-2	Forward	GCT AAC TTC CAA GCG GAC CA	62.5
	T7Select-DOWN	Reverse	GCT AGT TAT TGC TCA GCG GT	60.4

Table D.2: **PCR conditions for amplifying bacteriophage after biopanning**

	Conditions	
Initial denaturation	94 °C for 1 min	
Denaturation	94 °C for 50 s	Repeat for 35 cycles
Primer annealing	50 °C for 60 s	
Extension	72 °C for 60 s	
Final Extension	72 °C for 6 min	

Appendix E

Vector maps

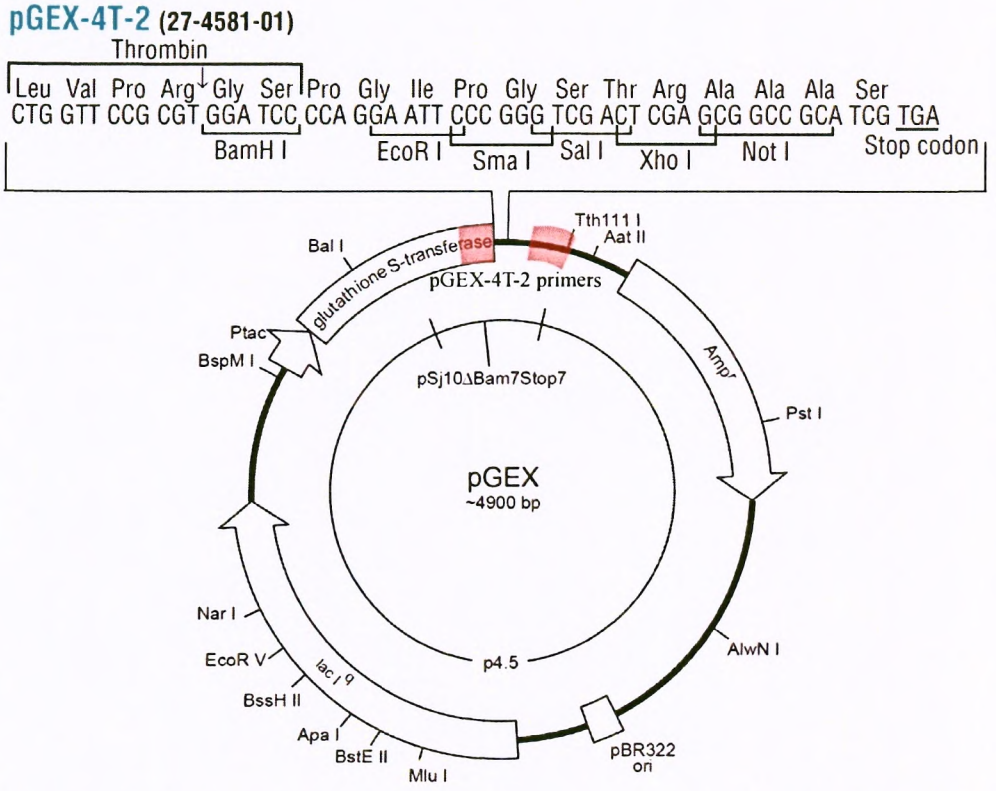


Figure E.1: **The pGEX-4T-2 plasmid. pGEX-4T-2 vector is a low copy number plasmid used for the expression of GST-tagged proteins. The sequence encodes: the ampicillin resistance gene β-lactamase (for positive selection); the pBR322 origin of replication; and the lacI repressor. *Bam*H1 and *Xho*I restriction endonuclease sites were used for directional insertion of the amplicons into the pGEX-4T-2 backbone, which fused a 5' glutathione S-transferase (GST) tag onto the gene fragment. The sites where the sense and anti-sense primers bound for sequencing of the *Pf*SORTILIN VPS10 construct are indicated in red. Adapted from Healthcare (2009).**

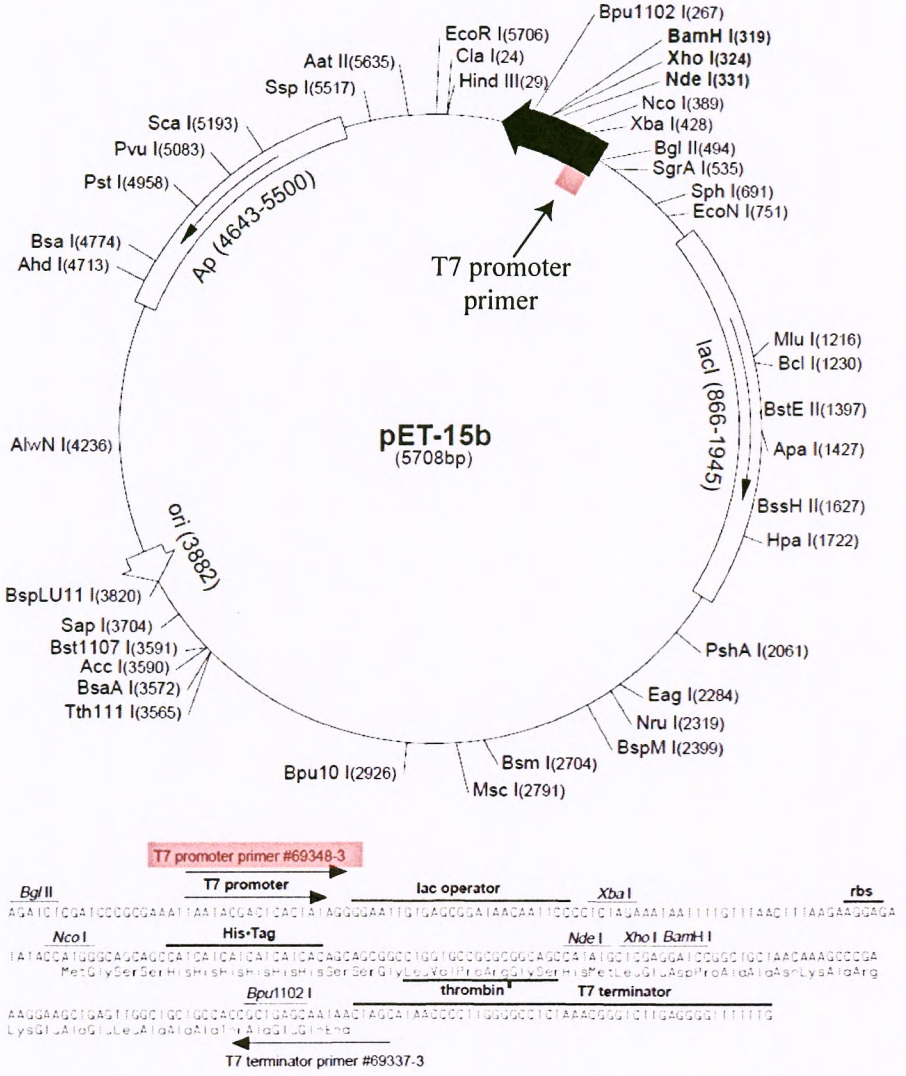


Figure E.2: **The pET-15b plasmid. pET-15b vector is a low copy number plasmid and is used for the expression of His-tagged proteins. The sequence encodes: the ampicillin resistance gene β-lactamase (for positive selection); the pBR322 origin of replication; and the lacI repressor. *Nde*I and *Bam*HI restriction endonuclease sites were used for directional insertion of the amplicons into the pET backbone, which fused a 5' 6xHis tag onto the gene fragment. The site where the T7 promoter primer bound, which was used for sequencing of the *Pf*MAEBL CCys construct, is shown in red. The image was adapted from Novagen (2011).**

Appendix F

Ethics clearance certificate

R14/49 Miss Serena Shunmugam

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

CLEARANCE CERTIFICATE NO. M171088

NAME: (Principal Investigator)	Miss Serena Shunmugam
DEPARTMENT:	Molecular Medicine and Haematology University of the Witwatersrand
PROJECT TITLE:	The Human Trafficking of MAEBL (Merozoite Adhesive Erythrocyte Binding-like Protein) in the Human Malaria Parasite, Plasmodium
DATE CONSIDERED:	Adhoc
DECISION:	Approved unconditionally
CONDITIONS:	Sub-Study under Primary Study (M130569)
SUPERVISOR:	Prof Theresa Coetzer
APPROVED BY:	Professor P. Cleaton-Jones, Chairperson, HREC (Medical)
DATE OF APPROVAL:	10/11/2017

This clearance certificate is valid for 5 years from date of approval. Extension may be applied for.

DECLARATION OF INVESTIGATORS

To be completed in duplicate and **ONE COPY** returned to the Research Office Secretary on the 3rd floor, Phillip Tobias Building, Parktown, University of the Witwatersrand. I/We fully understand the conditions under which I am/we are authorised to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated, from the research protocol as approved, I/we undertake to resubmit to the Committee. **I agree to submit a yearly progress report**. The date for annual re-certification will be one year after the date of convened meeting where the study was initially reviewed. In this case, the study was initially review October and will therefore be due in the month of October each year. Unreported changes to the application may invalidate the clearance given by the HREC (Medical).

Principal Investigator Signature — Date: 17 November 2017

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

Appendix G

Turn-it-in Plagiarism report

599340:MScSerenaShunmugam2018maincompressed.pdf

ORIGINALITY REPORT

12%	5%	6%	8%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to University of Witwatersrand Student Paper	5%
2	Submitted to Higher Education Commission Pakistan Student Paper	1%
3	Submitted to University of Kent at Canterbury Student Paper	1%
4	Jimenez-Ruiz, Elena, Juliette Morlon-Guyot, Wassim Daher, and Markus Meissner. "Vacuolar protein sorting mechanisms in apicomplexan parasites", Molecular and Biochemical Parasitology, 2016. Publication	1%
5	eprints.gla.ac.uk Internet Source	<1%
6	studentsrepo.um.edu.my Internet Source	<1%
7	Luis Eduardo Rodriguez, Hernando Curtidor, Mauricio Urquiza, Gladys Cifuentes, Claudia	<1%

References

Adams, J. H., Blair, P. L., Kaneko, O., *et al.* (2001). An expanding ebl family of Plasmodium faclciparum. *Trends Parasitol.* 17(6), pp. 297–299.

Adams, J. H., Kim, B, Simt, L., *et al.* (1992). A family of erythrocyte binding proteins of malaria parasites. *Microbiology* 89(15), pp. 7085–7089.

Alberts, B., Johnson, A., Lewis, J., *et al.* (2002). "Intracellular Compartments and Protein Sorting". In: *Mol. cell Biol.* 4th. New York: Garland Science. Chap. 12, pp. 615–657.

Andersen, J. L., Schrøder, T. J., Christensen, S., *et al.* (2014). Identification of the first small-molecule ligand of the neuronal receptor sortilin and structure determination of the receptor-ligand complex. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 70(2), pp. 451–460.

Bai, T., Becker, M., Gupta, A., *et al.* (2005). Structure of AMA1 from Plasmodium falciparum reveals a clustering of polymorphisms that surround a conserved hydrophobic pocket. *Proc. Natl. Acad. Sci. U. S. A.* 102(36), pp. 12736–12741.

Bannister, L. H., Hopkins, J. M., Dluzewski, A. R., *et al.* (2003). Plasmodium falciparum apical membrane antigen 1 (PfAMA-1) is translocated within micronemes along subpellicular microtubules during merozoite development. *J. Cell Sci.* 116, pp. 3825–3833.

Bannister, L. and Dluzewski, A. (1990). The ultrastructure of red cell invasion in malaria infections: a review. *Blood Cells* 16(3), pp. 287–297.

Beeson, J., Drew, D. R., Boyle, M., *et al.* (2016). Merozoite surface proteins in red blood cell invasion, immunity and vaccines against malaria. *FEMS Microbiol. Rev.* 40(3), pp. 343–372.

Berg, J. M., Tymoczko, J. L., and Stryer, L. (2002). "Eukaryotic Transcription and Translation Are Separated in Space and Time". In: *Biochemistry*. 5th. New York, NY: W H Freeman. Chap. 12.

Besteiro, S., Michelin, A., Poncet, J., *et al.* (2009). Export of a Toxoplasma gondii Rhoptry Neck Protein Complex at the Host Cell Membrane to Form the Moving Junction during Invasion. *PLoS Pathog.* 5(2), e1000309.

Bhanot, P., Frevert, U., Nussenzweig, V., *et al.* (2003). Defective sorting of the thrombospondin-related anonymous protein (TRAP) inhibits Plasmodium infectivity. *Mol. Biochem. Parasitol.* 126(2), pp. 263–273.

Birkholtz, L. M., Blatch, G., Coetzer, T. L., *et al.* (2008). Heterologous expression of plasmodial proteins for structural studies and functional annotation. *Malar. J.* 7(1), pp. 197–217.

Birnboim, H. C. and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* 7(6), pp. 1513–1523.

Blackburn, G., Gait, M., Loakes, D, *et al.* (2006). "Spectroscopic properties of Nucleosides and Nucleotides". In: *Nucleic Acids Chem. Biol.* Ed. by G. Blackburn. 3rd. Cambridge: The Royal Society of Chemistry. Chap. 2, p. 428.

Blair, P. L., Kappe, S. H. I., Maciel, J. E., *et al.* (2002). Plasmodium falciparum MAEBL is a unique member of the ebl family. *Mol. Biochem. Parasitol.* 122(1), pp. 35–44.

Blomqvist, K. (2008). "Thawing of glycerolyte-frozen parasites with NaCl". In: *Methods Malar. Res.* 5th. BioMalPar and MR4, p. 15.

Bonifacino, J. S. and Lippincott-Schwartz, J. (2003). Coat protein: shaping membrane transport. *Nat. Rev.* 4(5), pp. 409–413.

Boyle, M., Wilson, D., and Beeson, J. (2013). New approaches to studying Plasmodium falciparum merozoite invasion and insights into invasion biology. *Int. J. Parasitol.* 43(1), pp. 1–10.

Cabrera, A., Herrmann, S., Warszta, D., *et al.* (2012). Dissection of minimal sequence requirements for rhoptry membrane targeting in the malaria parasite. *Traffic* 13(10), pp. 1335–1350.

Center for Genomic Regulation (2012). *ccSOL*. URL: `http://s.tartaglialab.com/page/ccsol{\_}group` (visited on 02/16/2018).

Child, M. A., Harris, P. K., Collins, C. R., *et al.* (2013). Molecular Determinants for Subcellular Trafficking of the Malarial Sheddase PfSUB2. *Traffic* 14(10), pp. 1053–1064.

Churchyard, A. (2017). Trafficking of Plasmodium falciparum invasion proteins to the parasite micronemes. PhD thesis. The University of the Witwatersrand, pp. 48–49.

Cortés, A. (2008). Switching Plasmodium falciparum genes on and off for erythrocyte invasion. *Trends Parasitol.* 24(11), pp. 517–524.

Counihan, N. A., Kalanon, M., Coppel, R. L., *et al.* (2013). Plasmodium rhoptry proteins: Why order is important. *Trends Parasitol.* 29(5), pp. 228–236.

Cowman, A., Crabb, B., Alexander Maier, C., *et al.* (2008). "Preparation of P.falciparum genomic DNA". In: *Methods Malar. Res.* Ed. by K Moll, I Ljungström, H Perlmann, *et al.* BioMalPar and MR4, p. 298.

Cowman, A. F. and Crabb, B. S. (2006). Invasion of Red Blood Cells by Malaria Parasites. *Cell* 124(4), pp. 755–766.

Cowman, A. F., Berry, D., and Baum, J. (2012). The cellular and molecular basis for malaria parasite invasion of the human red blood cell. *J. Cell Biol.* 198(6), pp. 961–971.

Culleton, R. and Kaneko, O. (2010). Erythrocyte binding ligands in malaria parasites: Intracellular trafficking and parasite virulence. *Acta Trop.* 114(3), pp. 131–137.

De-Marco, A. (2009). Strategies for successful recombinant expression of disulfide bond-dependent proteins in Escherichia coli. *Microb. Cell Fact.* 8(26).

DeSimone, T. M., Jennings, C. V., Bei, A. K., *et al.* (2009). Cooperativity between Plasmodium falciparum adhesive proteins for invasion into erythrocytes. *Mol. Microbiol.* 72(3), pp. 578–589.

Di Cristina, M., Spaccapelo, R., Soldati, D., *et al.* (2000). Two conserved amino acid motifs mediate protein targeting to the micronemes of the apicomplexan parasite Toxoplasma gondii. *Mol. Cell. Biol.* 20(19), pp. 7332–7341.

Duraisingh, M. T., Maier, A. G., Triglia, T., *et al.* (2003). Erythrocyte-binding antigen 175 mediates invasion in Plasmodium falciparum utilizing sialic acid-dependent and -independent pathways. *PNAS* 100(8), pp. 4796–4801.

Edeling, M. A., Smith, C., and Owen, D. (2006). Life of a clathrin coat: insights from clathrin and AP structures. *Nat. Rev. Mol. Cell Biol.* 7(1), pp. 32–44.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32(5), pp. 1792–1797.

Fewell, S. W. and Brodsky, J. L. (2009). "Entry into the Endoplasmic Reticulum: Protein Translocation, Folding and Quality Control". In: *Traffick. Insid. Cells*. New York, NY: Springer. Chap. 7, pp. 119–142.

Fu, J., Saenz, F., Reed, M., *et al.* (2005). Targeted disruption of MAEBL in Plasmodium falciparum. *Mol Biochem Parasitol.* 141(1), pp. 113–117.

Gardner, M. J., Hall, N., Fung, E., *et al.* (2002). Genome sequence of the human malaria parasite Plasmodium falciparum. *Nature* 419(6906), pp. 498–511.

GE Healthcare Life Sciences (2014). *GST Gene Fusion System Handbook*. 1-12. Buckinghamshire, UK, pp. 68–69.

Ghai, M., Dutta, S., Hall, T., *et al.* (2002). Identification, expression, and functional characterization of MAEBL, a sporozoite and asexual blood stage chimeric erythrocyte-binding protein of Plasmodium falciparum. *Mol. Biochem. Parasitol.* 123(1), pp. 35–45.

Ghoneim, A., Kaneko, O., Tsuboi, T., *et al.* (2007). The Plasmodium falciparum RhopH2 promoter and first 24 amino acids are sufficient to target proteins to the rhoptries. *Parasitol. Int.* 56(1), pp. 31–43.

Gilberger, T. W., Thompson, J. K., Triglia, T., *et al.* (2003). A novel erythrocyte binding antigen-175 paralogue from Plasmodium falciparum defines a new trypsin-resistant receptor on human erythrocytes. *J. Biol. Chem.* 278(16), pp. 14480–14486.

Gilson, P. R., Chisholm, S. A., Crabb, B. S., *et al.* (2017). Host cell remodelling in malaria parasites: a new pool of potential drug targets. *Int. J. Parasitol.* 47(2-3), pp. 119–127.

Grabski, A., Mehler, M., and Drott, D. (2005). The Overnight Express Autoinduction System: High-density cell growth and protein expression while you sleep. *Nat. Methods* 2(3), pp. 233–235.

Hallée, S., Boddey, J. A., Cowman, A. F., *et al.* (2018). Evidence that the Plasmodium falciparum Protein Sortilin Potentially Acts as an Escorter for the Trafficking of the Rhoptry-Associated Membrane Antigen to the Rhoptries. *mSphere* 3(1), e00551–17.

Harris, K. S., Casey, J. L., Coley, A. M., *et al.* (2005). Binding hot spot for invasion inhibitory molecules on Plasmodium falciparum apical membrane antigen 1. *Infect. Immun.* 73(10), pp. 6981–9. URL: `http://www.ncbi.nlm.nih.gov/pubmed/16177378http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1230972`.

Hodder, A. N., Crewther, P. E., Matthew, M. L.S. M., *et al.* (1996). The Disulfide Bond Structure of Plasmodium Apical Membrane Antigen-1. *J. Biol. Chem.* 271(46), pp. 29446–29452.

Hoppe, H. C., Ngô, H. M., Yang, M., *et al.* (2000). Targeting to rhoptry organelles of Toxoplasma gondii involves evolutionarily conserved mechanisms. *Nat. Cell Biol.* 2(7), pp. 449–456. URL: `http://www.nature.com/articles/ncb0700{\_}449`.

Howell, S. A., Withers-Martinez, C., Kocken, C. H., *et al.* (2001). Proteolytic processing and primary structure of Plasmodium falciparum apical membrane antigen-1. *J. Biol. Chem.* 276(33), pp. 31311–31320.

Huang, S. (2011). And Now There Are Five: A New Player in Intracellular Trafficking Pathways. *PLoS Biol.* 9(10), e1001173.

Integrated DNA Technologies (2016). *Integrated DNA Technologies SciTools Oligo Analyzer 3.1*. URL: `https://euidtdna.com/calc/analyzer` (visited on 05/01/2016).

Jimenez-Ruiz, E., Morlon-Guyot, J., Daher, W., *et al.* (2016). Vacuolar protein sorting mechanisms in apicomplexan parasites. *Mol. Biochem. Parasitol.* 209(2), pp. 18–25.

Kappe, S., Noe, A., Fraser, T., *et al.* (1992). A family of erythrocyte binding proteins of malaria parasites. *Proc. Natl. Acad. Sci. U. S. A.* 89(15), pp. 7085–7089.

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(8), pp. 1668–1674.

Kats, L. M., Black, C. G., Proellocks, N. I., *et al.* (2006). Plasmodium rhoptries: how things went pear-shaped. *Trends Parasitol.* 22(6), pp. 269–276.

Kats, L. M., Cooke, B. M., Coppel, R. L., *et al.* (2008). Protein trafficking to apical organelles of malaria parasites - Building an invasion machine. *Traffic* 9(2), pp. 176–186.

Keeley, A. and Soldati, D. (2004). The glideosome: a molecular machine powering motility and host-cell invasion by Apicomplexa. *Trends Cell Biol.* 14(10), pp. 528–532.

Keen, J. H., Willingham, M. C., and Pastan, I. H. (1979). Clathrin-coated vesicles: isolation, dissociation and factor-dependent reassociation of clathrin baskets. *Cell* 16(2), pp. 303–312.

Kibria, K. M. K., Rawat, K., Klinger, C. M., *et al.* (2015). A role for adaptor protein complex 1 in protein targeting to rhoptry organelles in Plasmodium falciparum. *BBA - Mol. Cell Res.* 1853(3), pp. 699–710.

Krai, P., Dalal, S., and Klemba, M. (2014). Evidence for a Golgi-to-endosome protein sorting pathway in Plasmodium falciparum. *PLoS One* 9(2), e89771.

Laemmli, U. K. (1970). SDS-page Laemmli method. *Nature* 227(5259), pp. 680–685.

Lambros, C and Vanderberg, J. P. (1979). Synchronization of Plasmodium falciparum Erythrocytic Stages in Culture. *Am. Soc. Parasitol.* 65(3), pp. 418–420.

Lanzillotti, R. (2007). Identification of protein-protein interactions between plasmodium falciparum and the human erythrocyte membrane protein 4.1. PhD thesis. The Univeristy of the Witwatersrand, p. 148.

Lanzillotti, R. and Coetzer, T. L. (2008). Phage display: a useful tool for malaria research? *Trends Parasitol.* 24(1), pp. 18–23.

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).

Le Roch, K. G., Zhou, Y., Blair, P. L., *et al.* (2003). Discovery of Gene Function by Expression Profiling of the Malaria Parasite Life Cycle. *Science* 301(5639), pp. 1503–1508.

Leete, T. H. and Rubin, H. (1996). Malaria and the cell cycle. *Parasitol. Today* 12(I), pp. 442–444.

Li, X., Chen, H., Oo, T., *et al.* (2004). A co-ligand complex anchors Plasmodium falciparum merozoites to the erythrocyte invasion receptor band 3. *J. Biol. Chem.* 279(7), pp. 5765–5771.

Macherey-Nagel (2017). *PCR clean-up Gel extraction User manual NucleoSpin Gel and PCR Clean-up Manual.*

McMahon, H. T. and Mills, I. G. (2004). COP and clathrin-coated vesicle budding: Different pathways, common approaches. *Curr. Opin. Cell Biol.* 16(4), pp. 379–391. URL: `https://pdfs.semanticscholar.org/084d/16c9230c0808c7aaa15a4b1a3b7315468a32.pdf`.

Medicines for Malaria Venture (2016). *Parasite lifecycle*. URL: `http://www.mmv.org/malaria-medicines/parasite-lifecycle` (visited on 02/05/2016).

Mehlin, C., Boni, E., Buckner, F. S., *et al.* (2006). Heterologous expression of proteins from Plasmodium falciparum: Results from 1000 genes. *Mol. Biochem. Parasitol.* 148(2), pp. 144–160.

Miller, L. H., Ackerman, H. C., Su, X., *et al.* (2013). Malaria biology and disease pathogenesis: insights for new treatments. *Nat. Med.* 19(2), pp. 156–167.

Mitchell, G. H., Thomas, A. W., Margos, G., *et al.* (2004). Apical Membrane Antigen 1, a Major Malaria Vaccine Candidate, Mediates the Close Attachment of Invasive Merozoites to Host Red Blood Cells. *Infect. Immun.* 72(1), pp. 154–158.

Moll, K. (2008). "Stockholm Sorbitol Method". In: *Methods Malar. Res.* 5th. BioMalPar and MR4. Chap. 2, pp. 12–13.

Moore, D. and Dowhan, D. (2002). "Purification and Concentration of DNA from Aqueous Solutions". In: *Curr. Protoc. Mol. Biol.* Ed. by I. John Wiley and Sons. John Wiley {&} Sons Inc. Chap. 2, pp. 2.1.1–2.1.10.

NanoDrop Technologies Inc (2006). *ND-1000 Spectrophotometer V3.3 User ' s Manual.*

Ngô, H. M., Yang, M., Paprotka, K., *et al.* (2003). AP-1 in Toxoplasma gondii mediates biogenesis of the rhoptry secretory organelle from a post-Golgi compartment. *J. Biol. Chem.* 278(7), pp. 5343–5352.

Nielsen, M. S., Madsen, P, Christensen, E. I., *et al.* (2001). The sortilin cytoplasmic tail conveys Golgi-endosome transport and binds the VHS domain of the GGA2 sorting protein. *EMBO J.* 20(9), pp. 2180–90. URL: `http://www.ncbi.nlm.nih.gov/pubmed/8861945http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC452140http://www.ncbi.nlm.nih.gov/pubmed/11331584http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC125444`.

Novagen (2002). *Novagen T7Select System Manual.*

– (2011). *pET-15b Vector map insert.*

– (2016). *Novagen Competent Cells Manual.* Darmstadt, Germany.

Paul, A., Egan, E., and Duraisingh, M. (2015). Host-parasite interactions that guide red blood cell invasion by malaria parasites. *Curr. Opin. Hematol.* 22(3), pp. 220–226.

Pearse, B. M. (1976). Clathrin: a unique protein associated with intracellular transfer of membrane by coated vesicles. *Proc. Natl. Acad. Sci. U. S. A.* 73(4), pp. 1255–1259.

Pease, B. N., Huttlin, E. L., Jedrychowski, M. P., *et al.* (2013). Global Analysis of Protein Expression and Phosphorylation of Three Stages of Plasmodium falciparum Intraerythrocytic Development.

Phillips, M. A., Burrows, J. N., Manyando, C., *et al.* (2017). Malaria. *Nat. Rev.* 3(17050).

PlasmoDB (2018). *PlasmoDB.* URL: `http://plasmodb.org/plasmo/`.

Promega Corporation (2013). *MagneHis ™ Protein Purifi cation System Techincal Manual.*

– (2014). *MagneGST™ Protein Purifi cation System Technical Manual.* Madison, USA.

Qiagen (2002). *QIAexpress® Detection and Assay Handbook.* October, pp. 14; 16; 22–24; 26–27; 34–35; 41–46; 86–87.

Raymond, C. K., Howald-Stevenson, I, Vater, C. A., *et al.* (1992). Morphological classification of the yeast vacuolar protein sorting mutants: evidence for a prevacuolar compartment in class E vps mutants. *Mol. Biol. Cell* 3(12), pp. 1389–402.

Reiss, M, Viebig, N, Brecht, S, *et al.* (2001). Identification and characterization of an escorter for two secretory adhesins in Toxoplasma gondii. *J. Cell Biol.* 152(3), pp. 563–78.

Remarque, E. J., Faber, B. W., Kocken, C. H., *et al.* (2008). Apical membrane antigen 1: a malaria vaccine candidate in review. *Trends Parasitol.* 24(2), pp. 74–84.

Richard, D., Kats, L. M., Langer, C., *et al.* (2009). Identification of rhoptry trafficking determinants and evidence for a novel sorting mechanism in the malaria parasite plasmodium falciparum. *PLoS Pathog.* 5(3), e1000328.

Riglar, D. T., Richard, D., Wilson, D. W., *et al.* (2011). Super-Resolution Dissection of Coordinated Events during Malaria Parasite Invasion of the Human Erythrocyte. *Cell Host Microbe* 9(1), pp. 9–20.

Robinson, M. S. (2004). Adaptable adaptors for coated vesicles. *Trends Cell Biol.* 14(4), pp. 167–174.

Rodriguez, L. E., Urquiza, M, Ocampo, M, *et al.* (2000). Plasmodium falciparum EBA-175 kDa protein peptides which bind to human red blood cells. *Parasitology* 120(3), pp. 225–235.

Rosano, G. L. and Ceccarelli, E. A. (2014). Recombinant protein expression in Escherichia coli: Advances and challenges. *Front. Microbiol.* 5, p. 172.

Rothman, J. E. and Wieland, F. T. (1996). Protein Sorting by Transport Vesicles. *Science* 272(5259), pp. 227–234.

Sahdev, S., Khattar, S. K., and Saini, K. S. (2007). Production of active eukaryotic proteins through bacterial expression systems: a review of the existing biotechnology strategies. *Mol. Cell. Biochem.* 307(1-2), pp. 249–264.

Sakura, T., Yahata, K., and Kaneko, O. (2013). The upstream sequence segment of the C-terminal cysteine-rich domain is required for microneme trafficking of Plasmodium falciparum erythrocyte binding antigen 175. *Parasitol. Int.* 62(2), pp. 157–64.

Schägger, H. (2006). Tricine–SDS-PAGE. *Nat. Protoc.* 1(1), pp. 16–22.

Sebastian, R., Díaz, E., Ayala, G., *et al.* (2006). Studying endocytosis in space and time by means of temporal Boolean models. *Pattern Recognit.* 39(11), pp. 2175–2185.

Silvie, O., Mota, M. M., Matuschewski, K., *et al.* (2008). Interactions of the malaria parasite and its mammalian host. *Curr. Opin. Microbiol.* 11(4), pp. 352–359.

Sim, B., Chitnis, C., Wasniowska, K, *et al.* (1994). Receptor and ligand domains for invasion of erythrocytes by Plasmodium falciparum. *Science* 264(5167), pp. 1941–1944.

Singh, N, Preiser, P, Renia, L, *et al.* (2004). Conservation and developmental control of alternative splicing in maebl among malaria parasites. *J. Mol. Biol.* 343(3), pp. 589–599.

Sloves, P, Delhaye, S, Mouveaux, T, *et al.* (2012). Toxoplasma Sortilin-like Receptor Regulates Protein Transport and Is Essential for Apical Secretory Organelle Biogenesis and Host Infection. *Cell Host Microbe* 11(5), pp. 515–527.

Srinivasan, P., Beatty, W. L., Diouf, A., *et al.* (2011). Binding of Plasmodium merozoite proteins RON2 and AMA1 triggers commitment to invasion. *PNAS* 108(32), pp. 2–7.

Stornaiuolo, M., Lotti, L. V., Borgese, N., *et al.* (2003). KDEL and KKXX retrieval signals appended to the same reporter protein determine different trafficking between endoplasmic reticulum, intermediate compartment, and Golgi complex. *Mol. Biol. Cell* 14(3), pp. 889–902.

Tham, W., Healer, H., and Cowman, A. (2011). Erythrocyte and reticulocyte binding-like proteins of Plasmodium falciparum. *Trends Parasitol.* 28(1), pp. 23–30.

ThermoScientific (2012). *Slide-A-Lyze MINI Dialysis Devices manual.*

– (2014). *Thermo Scientific Phusion Flash High-Fidelity PCR Master Mix Manual.*

Tomavo, S., Slomianny, C., Meissner, M., *et al.* (2013). Protein Trafficking through the Endosomal System Prepares Intracellular Parasites for a Home Invasion. *PLoS Pathog.* 9(10), e1003629.

Tompa, P. (2003). *Intrinsically unstructured proteins evolve by repeat expansion.*

Topolska, A. E., Lidgett, A., Truman, D., *et al.* (2003). Characterization of a Membrane-associated Rhoptry Protein of Plasmodium falciparum. *J. Biol. Chem.* 279(6), pp. 4648–4656.

Trager, W and Jensen, J (1976). Human Malaria Parasites in Continuous Culture. *Science* 193(4254), pp. 673–675.

Treeck, M., Struck, N. S., Haase, S., *et al.* (2006). A conserved region in the EBL proteins is implicated in microneme targeting of the malaria parasite Plasmodium falciparum. *J. Biol. Chem.* 281(42), pp. 31995–32003.

Treeck, M., Zacherl, S., Herrmann, S., *et al.* (2009). Functional Analysis of the Leading Malaria Vaccine Candidate AMA-1 Reveals an Essential Role for the Cytoplasmic Domain in the Invasion Process. *PLoS Pathog.* 5(3). Ed. by M. Blackman, e1000322.

Tsuboi, T., Takeo, S., Iriko, H., *et al.* (2008). Wheat germ cell-free system-based production of malaria proteins for discovery of novel vaccine candidates. *Infect. Immun.* 76(4), pp. 1702–1708.

Tuteja, R. (2007). Malaria - An overview. *FEBS J.* 274(18), pp. 4670–4679.

Tyler, J. S., Treeck, M., and Boothroyd, J. C. (2011). Focus on the ringleader: the role of AMA1 in apicomplexan invasion and replication. *Trends Parasitol.* 27(9), pp. 410–420.

Urquiza, M., Suarez, J. E., Cardenas, C., *et al.* (2000). Plasmodium falciparum AMA-1 erythrocyte binding peptides implicate AMA-1 as erythrocyte binding protein. *Vaccine* 19(4-5), pp. 508–513.

Weiss, G., Gilson, P., Taechalertpaisarn, T, *et al.* (2015). Revealing the Sequence and Resulting Cellular Morphology of Receptor-Ligand Interactions during Plasmodium falciparum Invasion of Erythrocytes. *PLoS Pathog.* 11(2), pp. 1–25.

Wepener, M. (2014). Identification of trafficking determinants needed for the targeting of apical organelle proteins in Plasmodium falciparum. Dissertation. The University of the Witwatersrand, pp. 1–138.

WHO (2017). *World malaria report 2017.* Tech. rep., pp. 1–160.

Withers-Martinez, C., Jean, L., and Blackman, M. J. (2004). Subtilisin-like proteases of the malaria parasite. *Mol. Microbiol.* 53(1), pp. 55–63.

Wolfe, B. L. and Trejo, J. (2007). Clathrin-Dependent Mechanisms of G Protein-coupled Receptor Endocytosis. *Traffic* 8(5), pp. 462–470.

Woo, Y. H., Ansari, H., Otto, T. D., *et al.* (2015). Chromerid genomes reveal the evolutionary path from photosynthetic algae to obligate intracellular parasites. *Elife* 4, e06974.

Yang, A. S. P., Lopaticki, S., O'Neill, M. T., *et al.* (2017). AMA1 and MAEBL are important for Plasmodium falciparum sporozoite infection of the liver. *Cell. Microbiol.* 19(9), e12745.