https://dspace.org/

Protein-protein interactions in the erythrocytic stage of Plasmodium falciparum

Smit, Brigit Karen

2009-09-15T08:29:24Z
http://hdl.handle.net/10539/7261

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

PROTEIN-PROTEIN INTERACTIONS IN THE ERYTHROCYTIC STAGE OF *PLASMODIUM FALCIPARUM*

Brigit Karen Smit

A dissertation submitted to the Faculty of Science, University of the Witwatersrand, in fulfillment of the requirements for the degree of Master of Science

Johannesburg, 2008

DECLARATION

I declare that this dissertation is my own, unaided 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 in any other university.

_____

(Signature of candidate)

_____ day of _____ 200____

ABSTRACT

Increasing levels of morbidity and mortality due to malaria are being reported, fuelled by the development of drug-resistant strains of *Plasmodium falciparum*. This indicates that new compounds and drug targets are needed to strengthen the drug arsenal. Concentrating on the discovery of novel drug targets, two *P. falciparum* genes were selected for study, due to their putative importance in major protein-protein interaction pathways facilitating growth and survival within the erythrocytic stage of the parasite. *PFB0150c* codes for a putative protein kinase (PK). Biopanning a phage display library had previously shown that the PK interacted with protein 4.1 and spectrin of the host erythrocyte membrane. The catalytic domain of *PFB0150c* was subcloned with a GST-tag and the recombinant protein was expressed in *E. coli*. A kinase assay using [γ-^{32}P] ATP showed that the recombinant PK phosphorylated exogenous casein. This enzyme is hypothesised to phosphorylate host RBC membrane proteins and facilitate RBC invasion by merozoites. It may also play a role in intraerythrocytic growth and exit of the parasite. The PK may be multifunctional since the catalytic kinase domain makes up only a small portion of the protein. The additional domains could be good drug targets since they have very low sequence homology to human proteins. *PFE1400c* codes for a putative adaptor protein complex-1 β1 subunit presumed to be involved in protein trafficking, which is essential to the survival of malaria parasites. The N-terminal adaptin and C-terminal clathrin adaptor appendage domains were subcloned with a histidine- and a GST-tag, respectively, and were expressed in *E. coli*. The N-terminal adaptin domain was highly conserved, but the smaller C-terminal domain had low homology with the orthologous human protein and has potential as a drug target. Targeting *P. falciparum* invasion proteins is difficult because they are highly redundant. By inhibiting the protein trafficking complex, transport of invasion proteins to the micronemes and rhoptries would be blocked and they would be unable to reach their final destination and carry out their function. Thus invasion of host RBC by these defective merozoites would be prevented. The current study forms the foundation for future work in characterising the structure and function of these *P. falciparum* proteins.

ACKNOWLEDGEMENTS

I would like to thank the South African Malaria Initiative for a Masters bursary, the National Research Foundation for a grant-holder linked bursary and the University of the Witwatersrand for a postgraduate merit award. Thank you to the National Research Foundation and the University of the Witwatersrand Medical Faculty research endowment fund for providing project funding.

INDEX

ABBREVIATIONS **pg 10-11**

CHAPTER 1

INTRODUCTION **pg16-36**

1.1) Malaria: Approaching resistance **pg16-18**

1.2) Heterologous expression of *P. falciparum* proteins **pg18-20**

1.3) Malaria parasite-host protein-protein interactions **pg20-24**

1.3.1) PPIs during invasion pg20-23

1.3.2) Parasite-host PPIs during intraerythrocytic development pg23-24

1.4) What makes a good drug target? **pg24-26**

1.5) Protein kinases **pg26-29**

1.5.1) *P. falciparum* protein kinases pg27-28

1.5.2) *PFB0150c:* A putative PK pg28-29

1.6) Protein trafficking in eukaryotes **pg30-36**

1.6.1) Protein trafficking in *P. falciparum* pg31-34

1.6.2) *PFE1400c:* a putative AP-1 β subunit pg35

1.7) Aims of the project **pg36**

CHAPTER 2

MATERIALS AND METHODS **pg37-70**

2.1) Malaria culturing techniques **pg37-41**

2.1.1) Preparation of culture from frozen stock pg37-38

2.1.2) Heat inactivation of plasma pg38

2.1.3) Washing erythrocytes pg38

2.1.4) Preparation of smears pg39

2.1.5) Calculating parasitaemia pg39

2.1.6) Feeding the culture pg39

2.1.7) Dividing the culture pg40

2.1.8) Synchronisation of the culture – Sorbitol treatment pg40

2.1.9) Freezing of cultures pg41

2.2) DNA experiments **pg41-43**

2.2.1) DNA extraction from *P. falciparum* pg41-42

2.2.2) Determination of DNA concentration and purity pg43

2.3) RNA experiments **pg44-45**

2.3.1) RNA extraction from *P. falciparum* pg44

2.3.2) Determination of RNA concentration and integrity pg45

2.3.3) Reverse transcription pg45

2.4) PCR **pg46-50**

2.4.1) Primer design pg46-47

2.4.2) PCR protocol pg47-49

2.4.3) Restriction digestion of PCR products pg50

2.5) Plasmid preparation **pg51-53**

2.5.1) Extraction of plasmid DNA from DH5α cells pg51-52

2.5.2) Restriction enzyme digestion of plasmids pg52-53

2.5.3) Alkaline phosphatase treatment of linearised plasmids pg53

2.6) Sub-cloning **pg53-57**

2.6.1) Ligation of DNA insert and plasmid pg53-54

2.6.2) Transformation of DH5α cells pg54-55

2.6.3) Transformation of Rosetta 2 (DE3) cells pg55-56

2.6.4) Extraction of plasmid DNA from Rosetta 2 (DE3) cells pg56-57

2.7) Recombinant protein experiments **pg57-67**

2.7.1) Protein induction and expression pg57-58

2.7.2) Protein extraction pg58-59

2.7.3) Protein purification pg59-61

2.7.4) Inclusion body purification pg62

2.7.5) Inclusion body solubilisation pg63

2.7.6) Protein refolding pg64

2.7.7) Laemmli SDS-PAGE pg64-65

2.7.8) Western blotting pg65

2.7.9) Immunoblotting pg65-66

2.7.10) Chemiluminescence pg66-67

2.8) PK assay – *PFB0150c* **pg67-69**

2.8.1) Determination of recombinant PK concentration pg67-68

2.8.2) Kinase assay pg68-69

2.8.3) Densitometric analysis pg69

2.9) Bioinformatics **pg69-70**

CHAPTER 3

RESULTS **pg71-138**

3.1) DNA extraction **pg71-72**

3.2) RNA extraction **pg73-74**

3.3) PCR **pg74-80**

3.3.1) Primer design pg74-79

3.3.1.1) *PFB0150c* pg74-75

3.3.1.2) *PFE1400c* pg76-79

3.3.2) Digestion of PCR products pg79-80

3.4) Plasmid preparation **pg81-83**

3.4.1) Extraction of plasmid DNA pg81

3.4.2) Digestion of plasmids pg82-83

3.5) Results for *P. falciparum* PK **pg83-107**

3.5.1) Verification of PK vector construct pg83-85

3.5.2) PK vector construct sequence data pg85-87

3.5.3) Verification of PK vector construct in Rosetta 2 cells pg88

3.5.4) Laemmli SDS-PAGE of rGST PK pg88-89

3.5.5) Western blot of rGST PK pg90

3.5.6) Immunoblot of rGST PK pg91-93

3.5.7) Purity of rGST PK pg93-94

3.5.8) Concentration of rGST PK pg95-96

3.5.9) Kinase assay pg96-98

3.5.10) Bioinformatic data for *P. falciparum* PK pg99-107

3.5.10.1) Structure of *P. falciparum* PK pg99-100

3.5.10.2) Homology of parasite PK with other proteins pg101-107

3.6) Results for *P. falciparum* AP1 C-terminal **pg108-124**

3.6.1) Verification of AP1 C-terminal vector construct pg108

3.6.2) AP1 C-terminal vector construct sequence data pg109-111

3.6.3) Verification of AP1 C-terminal vector construct in Rosetta 2 cells pg111-112

3.6.4) Laemmli SDS-PAGE of rGST AP1 C-terminal pg112-113

3.6.5) Western blot of rGST AP1 C-terminal pg113-114

3.6.6) Immunoblot of rGST AP1 C-terminal pg114-116

3.6.7) Purity of rGST AP1 C-terminal pg117-119

3.6.8) Bioinformatic data for *P. falciparum* AP1 C-terminal pg119-124

3.6.8.1) Structure of *P. falciparum* AP1 C-terminal pg119-120

3.6.8.2) Homology of *P. falciparum* AP-1 β subunit pg121-124

3.7) Results for *P. falciparum* AP1 N-terminal **pg125-138**

3.7.1) Verification of AP1 N-terminal vector construct pg125

3.7.2) AP1 N-terminal vector construct sequence data pg125-128

3.7.3) Verification of AP1 N-terminal vector construct in Rosetta 2 cells pg129

3.7.4) Laemmli SDS-PAGE of rHis AP1 N-terminal pg129-130

3.7.5) Western blot of rHis AP1 N-terminal pg130-131

3.7.6) Autoradiograph of rHis AP1 N-terminal pg131-136

3.7.6.1) Western blot of rHis AP1 N-terminal extracted from inclusion bodies pg133

3.7.6.2) Autoradiograph of rHis AP1 N-terminal extracted from inclusion bodies pg134-136

3.7.7) Purity of rHis AP1 N-terminal pg136

3.7.8) Bioinformatic data for *P. falciparum* AP1 N-terminal pg136-138

3.7.8.1) Structure of *P. falciparum* AP1 N-terminal pg136-138

3.7.8.2) Homology of *P. falciparum* AP1 N-terminal to other proteins pg138

CHAPTER 4

DISCUSSION **pg139-160**

4.1) *PFB0150c* – A *P. falciparum* PK **pg139-146**
4.1.1) *PFB0150c* protein and mRNA expression pg141-142
4.1.2) Recombinant *PFB0150c* expression in *E. coli* pg142
4.1.3) Classification of *PFB0150c* PK pg143-144
4.1.4) Probable function of *PFB0150c* PK pg144-145
4.1.5) Drug-target potential of *PFB0150c* pg145-146
4.2) *PFE1400c* – A putative AP-1 β subunit **pg147-156**
4.2.1) *PFE1400c* protein and mRNA expression pg150
4.2.2) Recombinant *PFE1400c* expression in *E. coli* pg151-154
4.2.3) Drug-target potential of *PFE1400c* pg154-156
4.3) Future studies **pg156-159**
4.4) Conclusion **pg160**

CHAPTER 5
APPENDIX **pg161-191**

5.1) Malaria culturing techniques **pg161-162**
5.2) Reagents for DNA and RNA experiments **pg162-164**
5.3) Reagents for sub-cloning experiments **pg164-166**
5.4) Reagents for protein experiments **pg166-171**
5.4.1) *Reagents for recombinant protein purification* pg166-167
5.4.2) *Reagents for SDS PAGE analysis of recombinant protein* pg167-169
5.4.3) *Reagents for immunoblot analysis of recombinant protein* pg169-170
5.4.4) *Reagents for inclusion body experiments* pg170-171
5.5) Reagents for PK assay **pg171-172**
5.6) Data for graphs **pg172**
5.7) Vector maps **pg173-174**
5.8) Sequencing data **pg175-178**
5.9) BLAST data **pg179-185**
5.10) Secondary structure data **pg186-188**
5.11) Target-template pair alignments **pg189-191**

REFERENCES **pg192-201**

ABBREVIATIONS

aa – amino acids

ADP – adenosine diphosphate

AP-1 – adaptor protein complex-1

APS – ammonium persulphate

ARF1 – ADP ribosylation factor 1

A-T – adenine-thymine

ATP – adenosine triphosphate

BLAST – Basic Local Alignment Search Tool

BSA – bovine serum albumin

cAMP – cyclic adenosine monophosphate

CDK – cyclin-dependent kinase

cGDP – cyclic guanosine monophosphate

CIP – calf intestinal phosphatase

CSP – circumsporozoite protein

DBL – Duffy binding-like

dsDNA – double-stranded deoxyribonucleic acid

DTT – dithiothreitol

EBA – erythrocyte binding antigens

EBA – erythrocyte binding protein

ER – endoplasmic reticulum

G3PD – glycerol-3-phosphate dehydrogenase

GSK – glycogen-synthase kinase

GST – glutathione S-transferase

Hsc70 – heat shock cognate 70

ICAM-1 – intercellular adhesion molecule 1

KCH_3COO – potassium acetate

LB – Luria broth

MAPK – mitogen-activated PK

MEK – MAP kinase kinase

MHC – major histocompatibility complex

Mr – molecular weight

NaOH – sodium hydroxide

OD – optical density

PBS – phosphate buffered saline

PfEMP1 – *Plasmodium falciparum* erythrocyte membrane protein 1

pI – isoelectric point

PK – protein kinase

PKLS – putative kinase-like sequence

PMSF – phenylmethylsulfonyl fluoride

PPIs – protein-protein interactions

pRBC – parasitised red blood cells

RD – restriction digest

RBC – red blood cells

Rf – retardation factor

SANBS – South African National Blood Services

SDS – sodium dodecyl sulphate

SDS-PAGE – sodium dodecyl sulphate polyacrylamide gel electrophoresis

SERCA – Sarco-endoplasmic reticulum calcium ATPase

TB – terrific broth

TEMED – Tetramethylethylenediamine

TGN – trans-Golgi network

TRAP - thrombospondin-related anonymous protein

tRNAs – transfer ribonucleic acids

TSP – thrombospondin

LIST OF FIGURES

Figure 1. The malaria life cycle (Roberts & Janovy, 1996) *pg17*

Figure 2. Structure of a P. falciparum merozoite (Bannister et al, 2000) *pg21*

Figure 3. Structure of human erythrocyte membrane (Luna & Hitt, 1992) *pg22*

Figure 4. Genomic context of PFB0150c (PlasmoDB 5.4, 2008) *pg29*

Figure 5. Putative trafficking pathways in P. falciparum-infected erythrocytes. (Adapted from Cooke et al, 2004) *pg32*

Figure 6. Classical clathrin-mediated vesicular transport (Adapted from Malaria Parasite Metabolic Pathways, http://sites.huji.ac.il/malaria) *pg34*

Figure 7. Genomic context of PFE1400c (PlasmoDB 5.3, 2008) *pg35*

Figure 8. Agarose gel electrophoresis of P. falciparum DNA *pg71*

Figure 9. Agarose gel of P. falciparum DNA treated with RNase A *pg72*

Figure 10. Agarose gel electrophoresis of P. falciparum RNA *pg73*

Figure 11. PFB0150c diagramatic representation of primer positions *pg74*

Figure 12. DNA sequence of PFB0150c insert (PlasmoDB 5.3, 2008) *pg75*

Figure 13. Protein sequence of rGST PK *pg75*

Figure 14. PFE1400c diagramatic representation of primer positions *pg76*

Figure 15. DNA sequence of PFE1400c clathrin adaptor appendage domain, excluding introns (PlasmoDB 5.3, 2008) *pg77*

Figure 16. Protein sequence of rGST AP1 C-terminal *pg77*

Figure 17. DNA sequence of PFE1400c N-terminal adaptin domain *pg78*

Figure 18. Protein sequence of rHis AP1 N-terminal *pg79*

Figure 19. Digested P. falciparum PCR products resolved on 1% agarose gel *pg80*

Figure 20. Plasmid stocks resolved on 1% agarose gel *pg81*

Figure 21. Digested plasmids resolved on 1% agarose gel *pg82*

Figure 22. Digested PK vector constructs from DH5α cells *pg84*

Figure 23. Amplified PK vector constructs from DH5α cells *pg85*

Figure 24. Section of PK vector construct chromatogram *pg86*

Figure 25. Alignment of PK DNA sequences from 5' sequencing *pg87*

Figure 26. Amplified Rosetta vector constructs resolved on 1% agarose gel *pg88*

Figure 27. Purification of rGST PK protein *pg89*

Figure 28. rGST PK Western blot *pg90*

Figure 29. Immunoblot of rGST PK *pg91*
Figure 30. Densitometric scan of first elution of rGST PK *pg94*
Figure 31. Dried SDS acrylamide gel of kinase assay *pg97*
Figure 32. Autoradiograph of kinase assay gel *pg98*
Figure 33. Putative 3D model of PK kinase domain *pg99*
Figure 34. Sequence homology of PK to 3D template *pg100*
Figure 35. Amplified AP1 C-terminal vector constructs from DH5α cells *pg108*
Figure 36. Section of AP1 C-terminal vector construct chromatogram *pg109*
Figure 37. Alignment of AP1 C-terminal DNA sequences from 5' sequencing *pg111*
Figure 38. Amplified Rosetta vector constructs resolved on 1% agarose gel *pg112*
Figure 39. Purification of rGST AP1 C-terminal protein *pg 113*
Figure 40. rGST AP1 C-terminal Western blot *pg114*
Figure 41. Immunoblot of rGST AP1 C-terminal *pg115*
Figure 42. Densitometric scan of first elution of rGST AP1 C-terminal *pg117*
Figure 43. Densitometric scan of second elution of rGST AP1 C-terminal *pg118*
Figure 44. Predicted 3D model of AP1 C-terminal *pg119*
Figure 45. Sequence homology of AP1 C-terminal to 3D template *pg120*
Figure 46. Digested AP1 N-terminal vector constructs from DH5α cells *pg125*
Figure 47. Section of AP1 N-terminal vector construct chromatogram *pg126*
Figure 48. Alignment of AP1 N-terminal DNA sequences from T7 promoter primer sequencing *pg128*
Figure 49. Digested Rosetta vector constructs resolved on 1% agarose gel *pg129*
Figure 50. Purification of rHis AP1 N-terminal protein *pg130*
Figure 51. rHis AP1 N-terminal Western blot *pg131*
Figure 52. Immunoblot of rHis AP1 N-terminal *pg132*
Figure 53. Immunoblot of rHis AP1 N-terminal extracted from inclusion bodies *pg134*
Figure 54. 3D model of AP1 N-terminal *pg137*
Figure 55. Sequence homology of AP1 N-terminal to 3D template *pg137*
Figure 56A. Catalytic core of PKs *pg139*
Figure 56B. Catalytic core of P. falciparum PK *pg140*
Figure 57. Intraerythrocytic P. falciparum 3D7 photolithographic oligo array data for PFB0150c mRNA expression (Le Roch et al, 2003) *pg142*

Figure 58. *Structure of the AP-1 core. (Heldwein et al, 2004)* *pg147*
Figure 59. *Proposed AP-1 subunit organisation* *pg149*
Figure 60. *Intraerythrocytic P. falciparum 3D7 photolithographic oligo array data for PFE1400c mRNA expression (Le Roch et al, 2003)* *pg150*
Figure 61. *Exon/intron boundary sequences of PFE1400c intron 5* *pg151*
Figure 62. *Alignment of PFB0150c DNA sequences from 3' sequencing* *pg176*
Figure 63. *Alignment of AP1 N-terminal domain DNA sequences from T7 terminal primer sequencing* *pg178*
Figure 64. *BLAST data for P. falciparum PK with Homo sapiens protein* *pg179*
Figure 65. *BLAST data for P. falciparum PK with P. vivax protein* *pg180*
Figure 66. *BLAST data for P. falciparum PK with T. gondii protein* *pg181*
Figure 67. *BLAST data for P. falciparum AP-1 β1 subunit with P. vivax protein* *pg182*
Figure 68. *BLAST data for P. falciparum AP-1 β1 subunit with T. gondii protein* *pg184*
Figure 69. *BLAST data for P. falciparum AP-1 β1 subunit with H. sapiens protein* *pg185*
Figure 70. *Secondary structure prediction of PK* *pg186*
Figure 71. *Secondary structure prediction of AP1 C-terminal* *pg187*
Figure 72. *Secondary structure prediction of AP1 N-terminal* *pg188*
Figure 73. *Target-template pair alignment for PK kinase domain 3D model* *pg189*
Figure 74. *Target-template pair alignment for AP1 C-terminal 3D model* *pg189*
Figure 75. *Target-template pair alignment for AP1 N-terminal 3D model* *pg190*
Figure 76. *Protein sequence of PFB0150c with PEXEL/VTS motif* *pg191*

LIST OF TABLES

TABLE 1: Reaction mixture for PCR using the Expand High Fidelity ^{PLUS} PCR system **pg48**

TABLE 2: Reaction mixture for PCR product restriction digests using Fermentas restriction endonucleases **pg50**

TABLE 3: Reaction mixture for plasmid restriction digests using Fermentas restriction endonucleases **pg52**

TABLE 4: Reaction mixture for plasmid restriction digests using Fermentas FastDigest^{TM} restriction endonucleases **pg53**

TABLE 5: Protein samples collected during extraction and purification **pg61**

TABLE 6: Preparation of diluted BSA standards **pg68**

TABLE 7: Spectrophotometric data for *P. falciparum* genomic DNA **pg72**

TABLE 8: Spectrophotometric data for *P. falciparum* RNA **pg74**

TABLE 9: Spectrophotometric data for digested amplicons **pg80**

TABLE 10: Spectrophotometric data for digested, CIP-treated plasmids **pg83**

TABLE 11: Data used to determine molecular weight of rGST PK **pg92**

TABLE 12: Densitometric data for first elution of rGST PK **pg94**

TABLE 13: Extrapolated protein concentration of rGST PK **pg96**

TABLE 14: Sequence identity of *P. falciparum* PK with other *Plasmodium* proteins **pg102**

TABLE 15: Sequence identity of *P. falciparum* PK with *T. gondii* proteins **pg104**

TABLE 16: PK BLASTP analysis against eukaryotic proteomes **pg105-107**

TABLE 17: Summary of BLASTP analysis **pg107**

TABLE 18 Data used to determine the M_r of rGST AP1 C-terminal **pg116**

TABLE 19: Densitometric data for first elution of rGST AP1 C-terminal **pg117**

TABLE 20: Densitometric data for second elution of rGST AP1 C-terminal **pg118**

TABLE 21: Identity of *P. falciparum* AP1 with other *Plasmodium* proteins **pg121**

TABLE 22: Sequence identity of *P. falciparum* AP1 with *T.gondii* proteins **pg123**

TABLE 23: Data used to determine the M_r of rHis AP1 N-terminal **pg132**

TABLE 24: Data used to determine the M_r of rHis AP1 N-terminal from insoluble fraction **pg135**

TABLE 25: Donor and acceptor sites in non-alternatively spliced exon/intron boundaries of *P. falciparum* AP1 C-terminal **pg152**

CHAPTER 1
INTRODUCTION

1.1) Malaria: Approaching resistance

On an annual basis, over 300 million individuals experience an acute infection of malaria, caused by *Plasmodium falciparum* of the phylum *Apicomplexa*, and more than a million of them perish. The majority of these mortalities occur in sub-Saharan Africa and affect children under the age of five years (*World Health Organisation, 2007*).

The life cycle of the *Plasmodium* parasite is divided into an asexual and a sexual phase (*Figure 1*). The former phase is initiated when a female *Anopheles* mosquito carrying the parasite partakes in a human blood meal, thus injecting her anticoagulating saliva into the host along with the infective sporozoites. The parasite enters the bloodstream and subsequently migrates to the liver, where it invades the hepatocytes (*Roberts & Janovy, 1996*). Once safely enveloped in the liver cell, the sporozoite transforms into a feeding trophozoite that ingests the host cell cytoplasm via pinocytosis. Upon maturation, the trophozoite undergoes schizogony to form a schizont which contains multiple daughter nuclei. Further maturation results in the emergence of merozoites which are released from the hepatocytes. This signals the end of the pre-erythrocytic stage and newly liberated merozoites initiate the erythrocytic leg of development by penetrating host red blood cells (RBCs) and invading them (*Roberts & Janovy, 1996*). During the process of invasion, the parasite produces a parasitophorous vacuole membrane (PVM) which surrounds the merozoite. The compartment formed within the membrane is known as the parasitophorous vacuole (PV). The merozoite again metamorphoses into a feeding trophozoite and proceeds to ingest the host cytoplasm and haemoglobin. The parasite flattens, giving rise to the characteristic signet-ring stage (*Haynes, 1993*). Another round of schizogony occurs and the resultant schizont develops into multiple merozoites. Once mature, merozoites secrete proteases that disrupt cell plasma membranes, causing the parasitised erythrocytes to burst, propelling the parasites into the blood stream to infect a fresh set of RBCs (*Glushakova et al, 2005*).

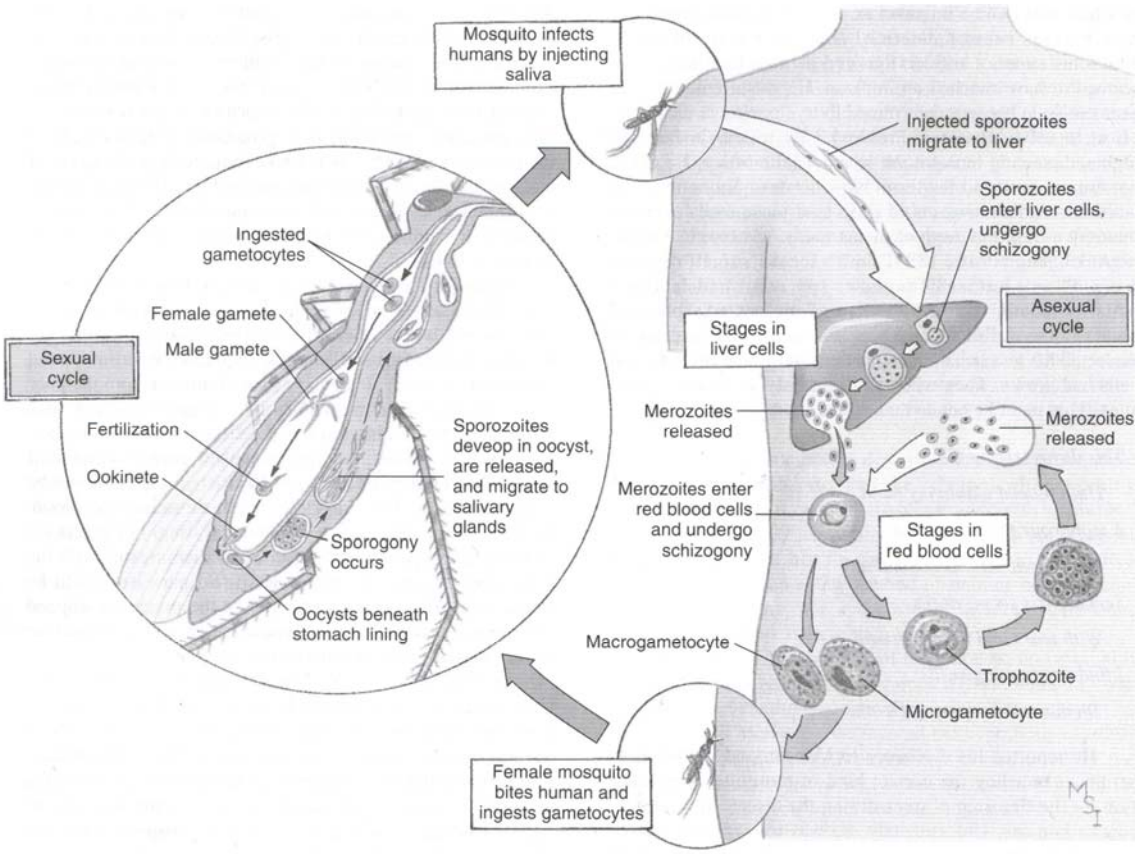


***Figure 1. The malaria life cycle* (*Roberts & Janovy, 1996*)**

The life cycle of P. falciparum consists of a sexual cycle that occurs in the mosquito and an asexual cycle that occurs in the human host.

The asexual cycle repeats itself a number of times. A few merozoites develop into the sexual cycle gametocytes, namely the macro- and microgametocytes. When ingested by a suitable mosquito, the gametocytes slough off the erythrocyte membrane and develop into mature gametes within the insect's gut. The microgamete penetrates and fertilises the macrogamete. A motile ookinete develops and enters the lining of the mosquito's gut. Transformation occurs and an oocyst is formed, maturing to produce multiple sporoblasts. Thousands of sporozoites created from the repeatedly dividing sporoblasts rupture the oocyst and migrate to the salivary glands. Within the channels of the invertebrate glands, *Plasmodium* is in a prime position to infect a new human host when next the insect feeds (*Roberts & Janovy, 1996*).

The latest approximation of the worldwide toll that malaria is taking indicates that the situation is worsening, with increasing levels of morbidity and mortality being reported (*World Health Organisation, 2007*). One of the main reasons for this deterioration is the development of drug-resistant strains of *P. falciparum* against which most conventional drugs – such as amodiaquine, chloroquine and sulfadoxine-pyrimethamine – are ineffective. One group of anti-malarials remains effective against the parasite, namely the artemisinin-derived drugs – dihydroartemisinin, artesunate and artemether (*World Health Organisation, 2007*). These compounds have an endoperoxide function, interacting with haem to generate toxic carbon-centred free radicals. These free radicals disable a number of proposed parasite proteins, one being the enzyme PfATPase6, the only SERCA-type (Sarco-endoplasmic reticulum calcium ATPase-type) Ca^{2+}-ATPase in the *P. falciparum* genome, leading to the death of the parasite (*Eckstein-Ludwig et al, 2003*). No clinical cases of resistance to the artemisinin compounds have occurred to date, but *in vitro* resistance in field isolates has been shown by several research groups (*Afonso et al, 2006; Jambou et al, 2005)*, indicating that new compounds and drug targets are needed to strengthen the drug arsenal in the event that *in vivo* resistance does emerge.

The elucidation of the entire *P. falciparum* genome by Gardner *et al* (*2002*) has facilitated the discovery of potential drug targets in the form of genes that are either specific to the parasite, or have low sequence similarity to human homologues (*Hoffman et al, 2002)*. Using bioinformatics and functional genomics, the *P. falciparum* proteins predicted from the genome can be analysed to determine probable molecular function, cellular localisation and expression levels at each of the life cycle stages. These genes or their protein products can then be exclusively targeted, without disrupting or damaging host functions (*Hoffman et al, 2002)*.

1.2) Heterologous expression of *P. falciparum* proteins

A total of 5 268 predicted proteins are encoded by the parasite genome (*Gardner et al, 2002*) and 65 percent of these have no homology to known proteins and therefore represent a goldmine of potential drug targets (*Florens et al, 2002)*. Unfortunately, the expression and purification of these parasite proteins using an *E. coli* expression system is no easy task, since the recombinant proteins are often insoluble (*Mehlin et al, 2006)*.

This resistance to heterologous expression of soluble proteins stems from the following factors, in decreasing order of significance: the isoelectric point (pI) of the protein; the hypothetical status of the protein and the percentage AT content of the gene.

Mehlin *et al* (*2006*) observed that the insolubility of expressed proteins increased with increasing pI. Expressed proteins that were most soluble had a pI between 3.45 and 6.8, while those that were least soluble had a pI range of 9.6-12.1. At very low pI values (less than 3.45) proteins had deficient expression; however, the small amount that was expressed had a high level of solubility. It was also noted that if the protein was annotated on PlasmoDB (the *P. falciparum* genome database) as hypothetical – meaning that the protein had no known function – the insolubility was significantly higher compared to characterised proteins. The group found no significant correlation between the size of the gene – and subsequently expressed protein – and solubility. The *P. falciparum* genome has approximately 80 percent AT content, resulting in a different codon preference to *E. coli* genes, and this was found to be a significant factor contributing to insolubility when analysed independently. However, multivariate analysis showed that pI was the most important factor for expression of soluble parasite protein, while the percentage AT content had the least effect.

These aforementioned factors may contribute to the expression of *P. falciparum* target proteins as insoluble inclusion bodies (*Mehlin et al, 2006).* Inclusion bodies are aggregations of misfolded proteins that have a low solubility and are therefore not easily extracted (*Kiefhaber et al, 1991*). The formation of these bodies occurs most often when the genes from an evolutionarily superior organism are expressed in a lower organism, as is the case when eukaryotic *P. falciparum* genes are expressed in prokaryotic *E. coli* cells. The microenvironment into which the recombinant protein emerges is foreign in terms of pH and osmolarity; mechanisms and chaperones for folding the eukaryotic protein are also absent in the prokaryotic microenvironment, and these factors may contribute to the production of an insoluble, inactive inclusion body (*Baca and Hol, 2000; Ling Goh et al, 2003*). However, a number of recombinant *P. falciparum* proteins have been successfully expressed in soluble form.

By optimising the time, pH and temperature of induction, as well as eliminating bacterial proteases and using *E. coli* cells that express rare tRNAs, the probability of expressing soluble recombinant proteins is increased (*Sorensen & Mortensen, 2005*). Inclusion bodies can also be treated with high molar concentrations of urea to solubilise the recombinant protein, followed by dialysis to allow renaturation and refolding of the protein which usually – but not always – restores functionality (*Yeo et al, 1997*).

1.3) Malaria parasite-host protein-protein interactions

Protein-protein interactions (PPIs) are of interest when applying molecular methods to combat *P. falciparum* infection, due to the fact that these connections are utilised by the malaria parasite during vital periods of its life cycle. It has been predicted that 516 PPIs occur between *Homo sapiens* and *P. falciparum* (*Dyer et al, 2007*). If some of these fundamental interactions could be disrupted, the parasite would not be able to complete specific stages of development and would perish, thereby alleviating malaria infections in the human population. PPIs within the erythrocytic stage of the *P. falciparum* life cycle are vital, as all the pathogenesis associated with malaria occurs during this time. The invasion of and growth within host red blood cells is therefore an important factor in the disease which, if disrupted, could prevent infection.

1.3.1) PPIs during invasion

The erythrocytic stage is initiated when merozoites come into contact with the red blood cell membrane and specialised secretory organelles – micronemes and rhoptries – release parasitic proteins that facilitate invasion of the host cells. These two types of organelles are located in the apical end of invasive merozoites (*Figure 2*).

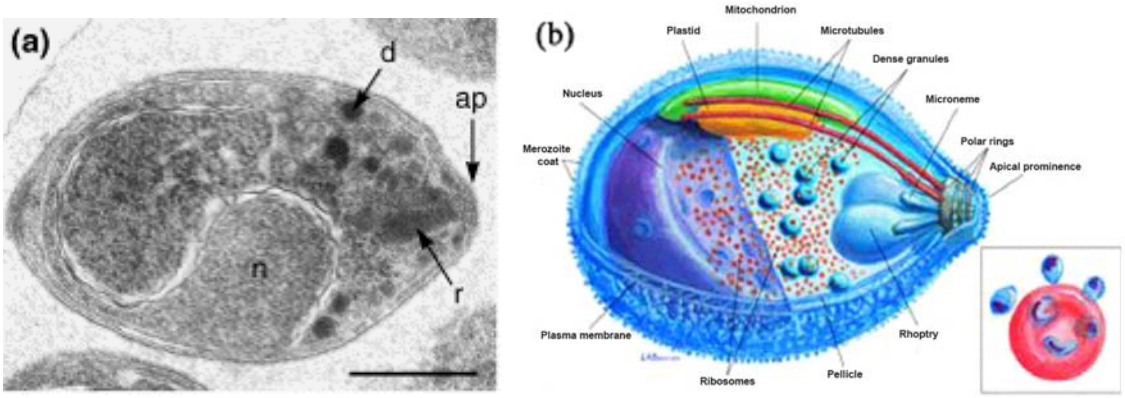


Figure 2. Structure of a P. falciparum merozoite (Bannister et al, 2000)

(a) Electron micrograph: A merozoite from the asexual cycle of P. falciparum showing the apical prominence (ap) with a rhoptry (r), dense granules (d), and an indented nucleus (n). Scale bar =0.5 µm.
(b) 3D structure: The merozoite pellicle is partly cut away to show the internal structure. The micronemes, dense granules and rhoptries are located at the apical end of the merozoite. Inset: relative sizes of merozoites and the RBC being invaded.

Rhoptry-associated and micronemal proteins interact with host proteins on the erythrocyte membrane. The membrane consists of two main components: a lipid bilayer and a skeletal protein network (*Figure 3*). The lipid bilayer is embedded with glycophorins and the anion exchanger, band 3. These transmembrane proteins have sialic acid moieties that extend into the vascular space (*Gallagher, 2006)*. The skeletal protein network consists primarily of spectrin tetramers. Spectrin is a fibrous polypeptide that exists in two isoforms, alpha and beta, which are loosely coiled around one another to form a relaxed helix. Tetramers are formed via the association of two alpha-beta helices which are joined by protein 4.1 to form a protein meshwork. Protein 4.1 stabilises the interaction of spectrin with actin to provide mechanical stability (*Gallagher, 2006).* Phosphorylation of protein 4.1 by protein kinases destabilises the erythrocyte plasma membrane (*Chishti et al, 1994*). This protein is vital to survival of the malaria parasite, as normal development of *P. falciparum* does not occur in red blood cells that lack protein 4.1 (*Chishti et al, 1994*). Band 3 is anchored to the main component of the protein network, namely spectrin, via the 215kD protein, ankyrin. The spectrin network is further stabilised by interactions with the proteins actin, adducin, tropomyosin and tropomodulin (*Gallagher, 2006)*.

Merozoite attachment to the host red blood cell membrane is mediated by a parasite-host PPI between the microneme-derived EBAs and the red cell membrane glycophorins and band 3 (*Mayer et al, 2002*). EBA is part of the Duffy binding-like family of parasitic ligands, which are the primary facilitators of merozoite binding. *Plasmodium* strains that utilise the sialic acid residues of glycophorins as invasion receptors are termed sialic acid-dependent; those that exploit other host receptors are termed sialic acid-independent.

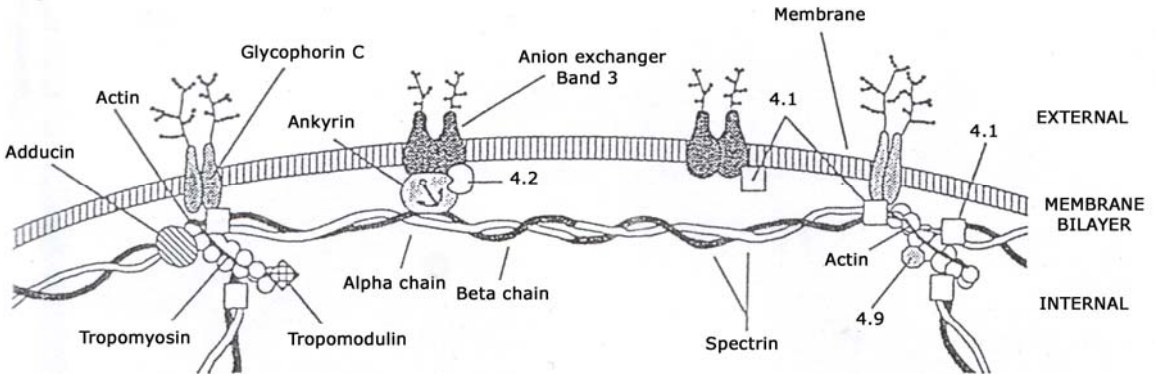


Figure 3. Structure of human erythrocyte membrane (Luna & Hitt, 1992)

The RBC plasma membrane consists of a lipid bilayer, containing polar head groups and non-polar fatty acid tails, and a protein skeletal network of spectrin tetramers joined together by several other proteins. Various accessory proteins link the two layers and stabilise the membrane.

Sialic acid-dependent strains have two main erythrocyte-binding ligands: EBA140, also known as BAEBL, and EBA175. EBA140 recognises glycophorin C, while EBA175 recognises glycophorin A (*Mayer et al, 2002*). By inhibiting the binding of EBA140 to glycophorin C, merozoites were prevented from penetrating host red blood cells (*Maier et al, 2003*). This discovery seemed to point to the possibility of manufacturing a vaccine against this EBA140 strain, but experimentation by Mayer *et al* (*2002*) indicated that multiple polymorphic forms of EBA140 exist, not all of which specifically bind glycophorin C. Because different parasite strains express different EBA isoforms, each with an affinity for different host ligands, the disruption of this PPI to thwart parasite invasion will not be successful as a global intervention (*Mayer et al, 2002*). However, the disruption of transport of these proteins to the micronemes and rhoptries would circumvent this redundancy, as the transport mechanisms in *P. falciparum* are likely to be general and not specific for each EBA (*see Section 1.4: What makes a good drug target?*).

During invasion, various proteins are secreted from the micronemes and rhoptries and interact with host RBC membrane proteins. RhopH3 is one such protein that is found in association with membrane-bound fractions (*Topolska et al, 2003*). This protein interacts with host membrane phospholipids and it has been postulated that RhopH3 causes changes to the inner leaflet of the erythrocyte membrane that facilitate invagination and the formation of the parasitophorous vacuole. Monoclonal antibodies against this protein have inhibited the growth of the parasite and attempts to disrupt the *RhopH3* gene have been unsuccessful (*Topolska et al, 2003*). These reports suggest that RhopH3 has a key role to play in the parasite life cycle (*Topolska et al, 2004*). Invasion proteins secreted by the rhoptries include Rhoptry Associated Proteins (RAP) 1 and 2, and gp76 rhoptry protein. Antibodies to both the RAP proteins prevented invasion of erythrocytes *in vitro.* Furthermore, monkeys immunised with RAP1 had partial immunity against infection (*Topolska et al, 2003*). The gp76 rhoptry protein degraded glycophorin A and band 3, major structural components of the host RBC membrane. This membrane destabilisation by the parasitic serine protease promotes speedy invasion by the merozoite. Proteins involved in the invasion of human erythrocytes that are secreted by the micronemes include *P. falciparum* apical membrane antigen 1 (PfAMA1), which facilitates reorientation of the merozoite such that its apical end is in contact with the erythrocyte cell membrane, along with the erythrocyte binding antigens (EBAs). PfAMA1 performs an essential role in the invasion of erythrocytes, evident from experiments that show inhibition of merozoite invasion in the presence of both monoclonal and polyclonal antibodies directed against PfAMA1 (*Healer et al, 2002*).

1.3.2) Parasite-host PPIs during intraerythrocytic development

Once established within a red blood cell, the parasite secretes various proteins that are transported to the erythrocyte plasma membrane surface (*Marti et al, 2004; Hiller et al, 2004*). These proteins form protrusions on the host cell membrane and are termed ‘knobs’. These parasitic knobs consist primarily of *P. falciparum* erythrocyte membrane protein 1 (PfEMP1) (*Maier et al, 2007*). The various PPIs that PfEMP1 takes part in are responsible for much of the acute pathology of malaria (*Chen et al, 1998; Flick & Chen, 2004*).

Parasitised red blood cells (pRBC) are removed from the host circulation by the binding of PfEMP1 to specific host endothelial cell surface receptors, for example: intercellular adhesion molecule-1 (ICAM-1), a cell surface ligand with a role in leukocyte adhesion and inflammation; CD36, a glycoprotein expressed on leukocytes, platelets and endothelium; and thrombospondin, an extracellular adhesive protein that plays a role in platelet aggregation and angiogenesis. This adhesion to host vasculature prevents the circulation of pRBCs in the blood stream and the subsequent sequestration by the spleen. The CD36 cell surface receptor is also present on dendritic cells – phagocytic cells of the host immune system – and PfEMP1 binding to these cells may initiate apoptosis of the cell and thereby prevent antigen presentation by the class II MHC to T-helper cells (*Flick and Chen, 2004*), contributing to the ineffective host immune response to malaria parasites.

1.4) What makes a good drug target?

A previous approach in identifying malaria proteins suitable for vaccine development and drug targeting concentrated on the EBA invasion proteins after release from micronemes but due to the redundancy of receptor binding specificities, these proteins make poor targets (*Mayer et al, 2002*). PfEMP1 is a protein important for the survival of the parasite, but it does not represent a good drug target as a multigene family of 59 var genes encodes this highly variable protein. Of the 59 genes, only one will be expressed by a single parasite at a time, making population-wide targeting impossible. Malaria proteins that are exposed to the host immune system are therefore not good drug targets, as this confers increased antigenic variation to the proteins so that detection by immune cells is thwarted. By concentrating on malaria proteins that function within the confines of the pRBC – for example, protein kinases (PKs) and trafficking proteins – an effective target for molecular intervention may be elucidated.

PKs have been targeted in various diseases and their inhibition often provides therapeutic benefits. In the case of chronic myeloid leukaemia (CML), the leukaemogenic oncogene *Bcr-Abl* is the causative molecular abnormality of the disease. The product of this oncogene is a constitutively activated tyrosine kinase which functions in numerous signalling pathways, deregulating them to cause the uncontrolled proliferation of myeloid blasts in the blood. The development of Imatinib, an inhibitor of *Bcr-Abl* tyrosine kinase, has proved effective in killing or inhibiting the proliferation of all *Bcr-Abl*-expressing cells (*Druker, 2002*). A number of *P. falciparum* PKs have been targeted by inhibitors with varying degrees of success. Syin *et al* (2001) carried out a study on the catalytic subunit gene of *P. falciparum* cAMP-dependent protein kinase (*Pfpka-c*). This gene is expressed at high levels during the asexual stage of parasite growth, mainly in schizonts. The group determined that the treatment of pRBC in culture with the PKA-C inhibitor H89 halted the growth and development of the parasite. However, due to the high sequence similarity between the human erythrocytic and parasite PKA-C enzymes, it is possible that the RBC PKA-C was being inhibited. Therefore, the growth of the intraerythrocytic parasite could be dependent on the erythrocyte PKA-C, and not the parasitic PKA-C as initially assumed. Another group ascertained that a cGMP-dependent protein kinase (PKG) inhibitor that had antiparasitic activity toward a *Toxoplasma gondii* PKG was also effective in blocking motility and invasion by malaria parasites (*Diaz et al, 2006*). The PKG inhibitor, a trisubstituted pyrrole pyridine known as Compound 1, was active against the *P. falciparum* PKG. It inhibited the growth of chloroquine-resistant and chloroquine-sensitive strains during erythrocytic growth *in vitro*. Compound 1 also delays the onset of parasitemia in *P. berghei* mouse models of infection, although it does not cure the infection (*Diaz et al, 2006*). It should be noted, however, that a large number of PK inhibitors have a rather poor specificity and tend to inhibit other PKs more potently than their presumed targets, rendering some cell-based experiments inaccurate (*Davies et al, 2000; Bain et al, 2003; Bain et al, 2007*).

Trafficking proteins have not been exploited as much as PKs when it comes to drug targeting. Ramya *et al* (*2007*) determined that the mechanism of action of a known antimalarial, 15-Deoxyspergualin (DSG), involved the disruption of protein trafficking. At submicromolar concentrations, DSG kills *P. falciparum* by preventing the transport of nucleus-encoded proteins to the apicoplast during the first asexual cycle, which leads to missegregation and subsequent loss of the apicoplast in the second asexual cycle (*Ramya et al, 2007*). The apicoplast is an organelle specific to some apicomplexan parasites. It is a non-photosynthetic plastid, assumed to be a relic of a photosynthetic organism that was phagocytosed by the host, but not digested. DSG induces delayed death – during the second asexual cycle – by starving the apicoplast of proteins that are indispensable to various metabolic processes in the parasite. This effect was observed in the late trophozoite and schizont stages, where fatty acid biosynthesis was inhibited, killing the malaria parasite. This indicates that the targeting of trafficking proteins within the trans-Golgi network of the parasite could also lead to its demise.

1.5) Protein kinases

PKs are enzymes that catalyse the phosphorylation of proteins within eukaryotic cells. This regulates protein function – activating, modulating or deactivating the molecules – and thereby controls cell behaviour. Of the proteins expressed in an average eukaryotic cell, almost 33 percent contain covalently bound phosphate molecules (*Hubbard & Cohen, 1993*). Approximately 3 percent of all eukaryotic genes code for PKs and these are classified according to structural similarity, as well as parallel substrate specificity and mode of regulation (*Hanks & Hunter, 1995*).

The reaction catalysed by PKs is:

$$\textit{Protein} + \textit{ATP} \xrightarrow{\textbf{PK}} \textit{phosphoprotein} + \textit{ADP}$$

Eukaryotic PKs are divided into seven established groups, with the two main sub-divisions being the protein-serine/threonine kinases and the protein-tyrosine kinases (*Hanks & Hunter, 1995*). The entire complement of PKs encoded in a genome is termed the kinome.

1.5.1) *P. falciparum* protein kinases

Phosphorylation and dephosphorylation processes play an important role in the life cycle of the malaria parasite (*Suetterlin et al, 1991*). This is especially true for the intraerythrocytic stage which is accompanied by a distorted phosphorylation pattern of the host RBC membrane (*Chishti et al, 1994*). This vital stage of the parasite lifecycle is prevented by PK inhibitors (*Ward et al, 2004*; *Anamika et al, 2005*).

According to Ward *et al* (2004) – who identified 65 malaria PK sequences – and Anamika *et al* (*2005*), who identified 99 PKs in the *P. falciparum* genome using various amino acid sequence profile matching algorithms, several of the parasite sequences did not cluster within any of the known eukaryotic PK groups. Furthermore, the highest number of malarial sequences were those involved in the control of cell proliferation, namely the cyclin-dependent- (CDK), mitogen-activated- (MAPK), glycogen-synthase- (GSK) and CDK-like kinases, along with a kinase family that includes PKs A, G and C (AGC). Interestingly, no malarial PK clustered with the tyrosine kinase group; homologues of MEK (MAPK kinase), MEKK (MAPK kinase kinase) and PKC-like kinases were also lacking in the *P. falciparum* genome. A splinter group of 20 PK-related sequences formed a novel family called FIKK, which seems to be restricted to the *Apicomplexa* (*Ward et al, 2004).* According to Schneider & Mercereau-Puijalon (*2005*), even though kinase activity has not been demonstrated in this group, the presence of most of the amino acids necessary for phosphotransfer indicates an enzymatic role. Nunes *et al* (*2007*) provide experimental evidence of kinase activity and transport of some FIKKs to the erythrocyte.

The large divergence in the kinome of *P. falciparum* compared to that of humans suggests that exclusive targeting of parasite enzymes is possible. This is promising as PKs play crucial roles in most cellular processes and thus their targeted inhibition could incapacitate the parasite and prevent disease progression.

1.5.2) ***PFB0150c:*** **A putative PK**

In previous research in Professor T.L. Coetzer's Plasmodium Molecular Research Unit, a *P. falciparum* phage display library was created and bio-panned against a host protein of the erythrocyte plasma membrane, namely protein 4.1 (*Figure 3*).
A host-parasite PPI was found to exist between protein 4.1 and a putative PK encoded by the *PFB0150c* gene (*Lauterbach et al, 2003*). Research by Chishti et al (*1994*) has shown that phosphorylation of protein 4.1 decreases its interaction with the skeletal network, thereby destabilising the erythrocyte plasma membrane. This increased flexibility of the erythrocyte membrane has been postulated to facilitate parasite growth and survival (*Chishti et al, 1994)*. The group determined that the enzyme that phosphorylated protein 4.1 in parasitised erythrocytes was a casein kinase of parasite origin, by showing that phosphorylation of the RBC membrane protein was prevented by casein kinase I and II inhibitors. This casein kinase enzyme could be the product of *PFB0150c,* the gene of interest.

PFB0150c is a *P. falciparum* gene found on chromosome 2. The gene has Crick-orientation, meaning that the coding sequence is found on the antisense (3' to 5') strand of the chromosomal DNA (*Figure 4)*. The gene consists of a single exon and is 7 448bp in length. The PK domain, found at the C-terminal of the protein, contains the catalytic core of the enzyme. *(*www.*plasmodb.org, version 5.4, 2008)*.

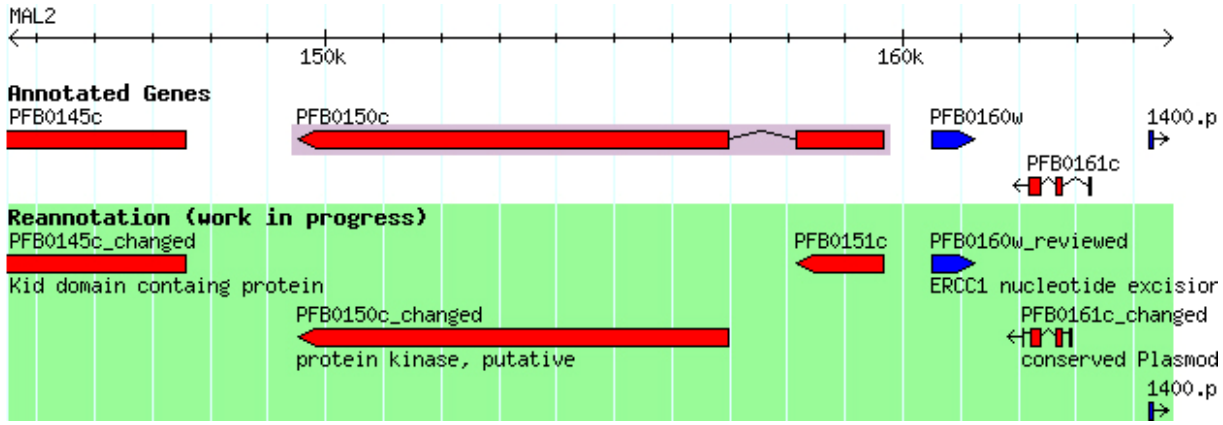


Figure 4. Genomic context of PFB0150c (PlasmoDB 5.4, 2008; www.plasmodb.org)

Crick-orientation genes are represented in red, while Watson-orientation genes are coloured blue. Updates of the gene model for PFB0150c as of February 1, 2008 can be seen in the green area of the diagram.

PFB0150c is described as a hypothetical phosphotransferase of the serine- or threonine-specific kinase subfamily (*Letunic et al, 2006*). The role of this protein as a kinase was inferred from its homology to other eukaryotic PKs (*www.genedb.org*) but its function has not been proved. According to The Gene Ontology Consortium (*2007*) (*http://www.genedb.org/amigo/perl*) there are no data available indicating the cellular location of the *PFB0150c* gene product within the parasite but, as no transmembrane domains or signal peptides are found within the protein, it is unlikely that it is positioned on the parasite plasma membrane. According to the same group, the biological processes that this putative PK is involved in include pathogenesis, protein phosphorylation and proteolysis. The probable molecular function of this enzyme is described as ATP binding, metal-ion binding, metallo-endopeptidase activity, PK activity and protein serine/threonine kinase activity.

1.6) Protein trafficking in eukaryotes

In eukaryotic cells, normal functionality is dependent on the transfer of newly created polypeptides to their correct cellular locations. These sites may occur within the cell's cytoplasm, on the plasma membrane or external to the cell. Protein trafficking in eukaryotes occurs via the classical vesicle-mediated secretory pathway. A hydrophobic sequence at the N-terminal of the proteins acts as a polypeptide signal, directing them into the secretory pathway (*von Heijne, 1985*). Eukaryotic proteins are synthesised by ribosomes and then directed into the endoplasmic reticulum (ER) (*Rodnina et al, 2006*). ER chaperones ensure that the newly-formed proteins are correctly folded and membrane-bound vesicles shuttle the properly-folded polypeptides to the Golgi apparatus. Proteins are released at the cell surface by exocytosis. COP II coat protein functions in the transport of vesicles to the Golgi in what is termed anti-retrograde transport, while COP I returns the vesicles to the ER. These coat proteins are essential for the selection of suitable cargo proteins, as well as the deformation of the plasma membrane to form vesicles for transport (*Duden, 2003)*.

Key molecules in trafficking are the adaptor protein (AP) complexes. AP-1 and AP-2 have been well characterised in mammals, where they are involved in protein sorting and transport to different organelles in the cell (*Ohno, 2006)*. AP complexes are heterotetramers composed of two large adaptin subunits (γ and β), a medium subunit (μ) and a small subunit (σ) (*Heldwein et al, 2004*). Of the four subunits making up the adaptor protein, the μ-subunit is involved in cargo binding via the recognition of particular sequence motifs, while the β-subunit induces the assembly of clathrin triskelions to form a latticed coat (*Brodsky et al, 2001)*. Mammalian AP-1 and -2 complexes are responsible for the formation of clathrin-coated pits and vesicles which are created during endocytosis (*Brodsky et al, 2001*); the former initiates formation of these vesicles at the ER in the trans-Golgi network while the latter is found in plasma membrane clathrin-coated pits. These clathrin-coated vesicles (CCVs) then referee the trafficking of various membrane-bound proteins (*Brodsky et al, 2001)*.

CCVs are made up of three distinct layers. The inner layer contains the transmembrane cargo, which is usually protein in nature. The middle layer consists of AP complexes and links the inner layer to the outer protein layer, which is composed of a stabilising clathrin coat of interlinking triskelia (*Owen et al, 2000; Wang et al, 1995*). Homologues of AP-1 and AP-2 have been found in other organisms (*Brodsky et al, 2001)* including *T. gondii* and *P. falciparum.*

1.6.1) Protein trafficking in *P. falciparum*

Trafficking of proteins by the malaria parasite involves transport of two main types, namely transport of proteins inside the parasite, and export of proteins into the RBC and onto the host RBC membrane. Protein trafficking in *P. falciparum* seems to follow the classical vesicle-mediated secretory pathway described in section 1.6 (*Cooke et al, 2004; Figure 5*). The exact design of the secretory pathway has not been fully elucidated; some groups postulate that it consists of a tubovesicular membrane network that is directly connected to the parasitophorous vacuole and includes a rudimentary Golgi apparatus and ER (*Behari & Haldar, 1994)*.

Templeton and Deitsch (*2005*) maintain that the pathway exists as a conglomeration of Maurer's clefts. In any case, the use of Brefeldin A, a drug that inhibits vesicle-mediated trafficking, has been demonstrated to prevent secretion of most exported malarial proteins (*Cooke et al, 2004).*

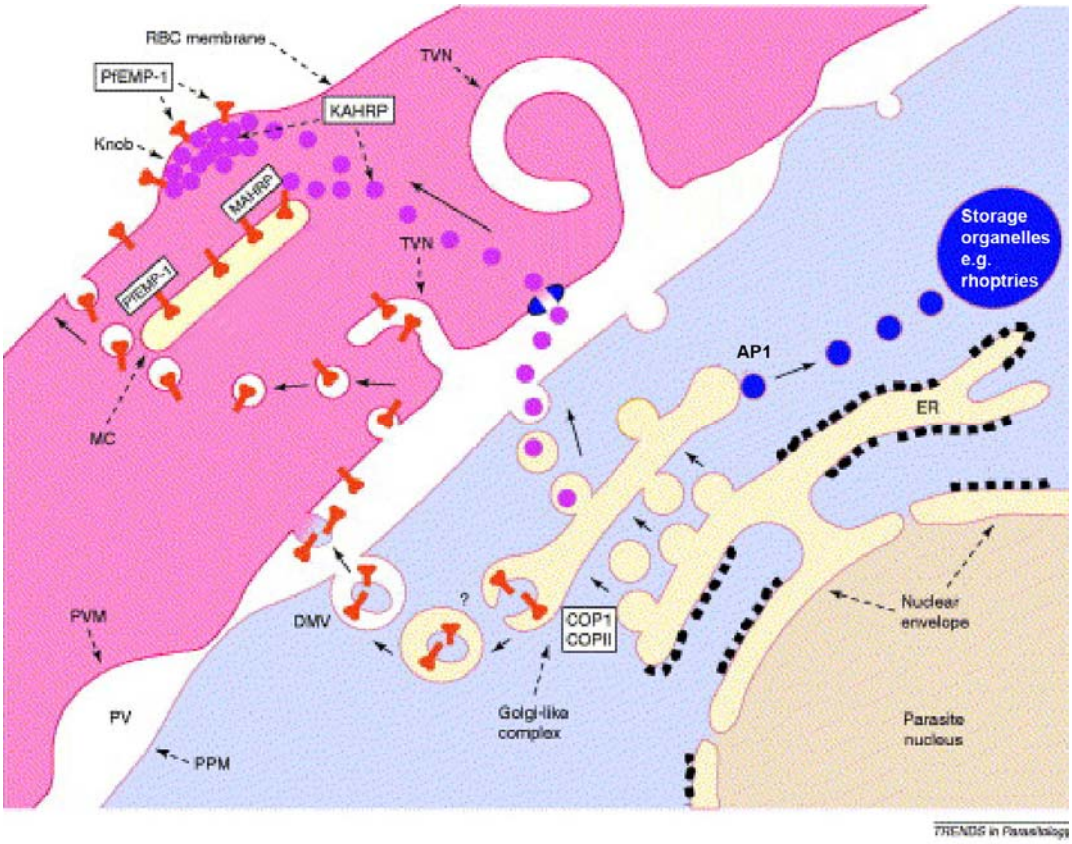


Figure 5. Putative trafficking pathways in P. falciparum-infected erythrocytes.

(Adapted from Cooke et al, 2004)

The classical vesicle-mediated secretory pathway traffics proteins to a number of organelles and cellular domains. PfEMP1 (red), MAHRP and KAHRP (purple) are targeted to the erythrocyte membrane via the PEXEL/VTS motif. For the current study, the pathway of interest is highlighted in blue. It involves AP-1, which facilitates the formation of protein-containing vesicles at the Golgi-like complex for trafficking to storage organelles like the rhoptries and micronemes. COPI and COPII are involved in the transport of proteins between the ER and Golgi-like complex. COP - coat protein; DMV - double membrane-bound vesicle; ER - endoplasmic reticulum; KAHRP - knob-associated histidine-rich protein; MAHRP - membrane-associated histidine-rich protein; MC - Maurer's clefts; PfEMP1 - P. falciparum erythrocyte membrane protein 1; PM - plasma membrane; PPM - parasite PM; PV - parasitophorous vacuole; PVM - PV membrane; RBC - red blood cell; TVN - tubovesicular network.

As can be seen in figure 5, trafficking in the parasite involves the transport of proteins between the ER and Golgi-like complex – mediated by COPI and COPII – and from the trans-Golgi network to other organelles, which is mediated by AP complexes. Various parasite proteins are also exported to the plasma membrane of the host RBC. The exported parasite membrane proteins, knob-associated histidine-rich protein (KAHRP) and PfEMP1 (*Figure 5*), are exposed on the surface of the pRBC. Both soluble and integral membrane proteins are translocated into the ER via an N-terminal signal sequence and are then delivered into the parasitophorous vacuole by way of the Golgi-like complex (*Figure 5*). A second motif, known as the *Plasmodium* export element (PEXEL) or vacuolar transport signal (VTS), targets export proteins for onward transport to the erythrocyte ((*Marti et al, 2004; Hiller et al, 2004*)). Once there, they are exported to the membrane, forming knobby protrusions that are involved in the cytoadherence of pRBC to the host vasculature (*Foley and Tilley, 1998*).

In the apicomplexans AP-1 is involved in sorting of proteins to the micronemes and rhoptries (*Ngô et al, 2003*) as indicated by the blue pathway in figure 5. These organelles secrete numerous proteins involved in the parasitic invasion of host cells (*see Section 1.3.1: PPIs during invasion)*. Rhoptry Associated Membrane Antigen (RAMA) is a protein with the dual roles of RBC invasion and rhoptry creation. RAMA is synthesised in the early trophozoite stage before the rhoptries exist, and is stored within lipid-rich areas of the Golgi-like complex. As maturation progresses, the portions of the Golgi containing RAMA bud to form vesicles that are trafficked to the apical prominence and mature into rhoptries (*Topolska et al, 2003*).

The putative function of AP-1 in malaria transport pathways is outlined in figure 6. AP-1 recruits clathrin molecules to a membrane segment of a post-Golgi endocytic compartment and a clathrin-coated vesicle is formed. The vesicle then surrounds the membrane-bound protein that is to be transported and buds from the endocytic compartment which causes the coat to fall off. The uncoated vesicle is then able to fuse with the rhoptries and micronemes and deliver its contents to these apical organelles (*Ngô et al, 2003)*.

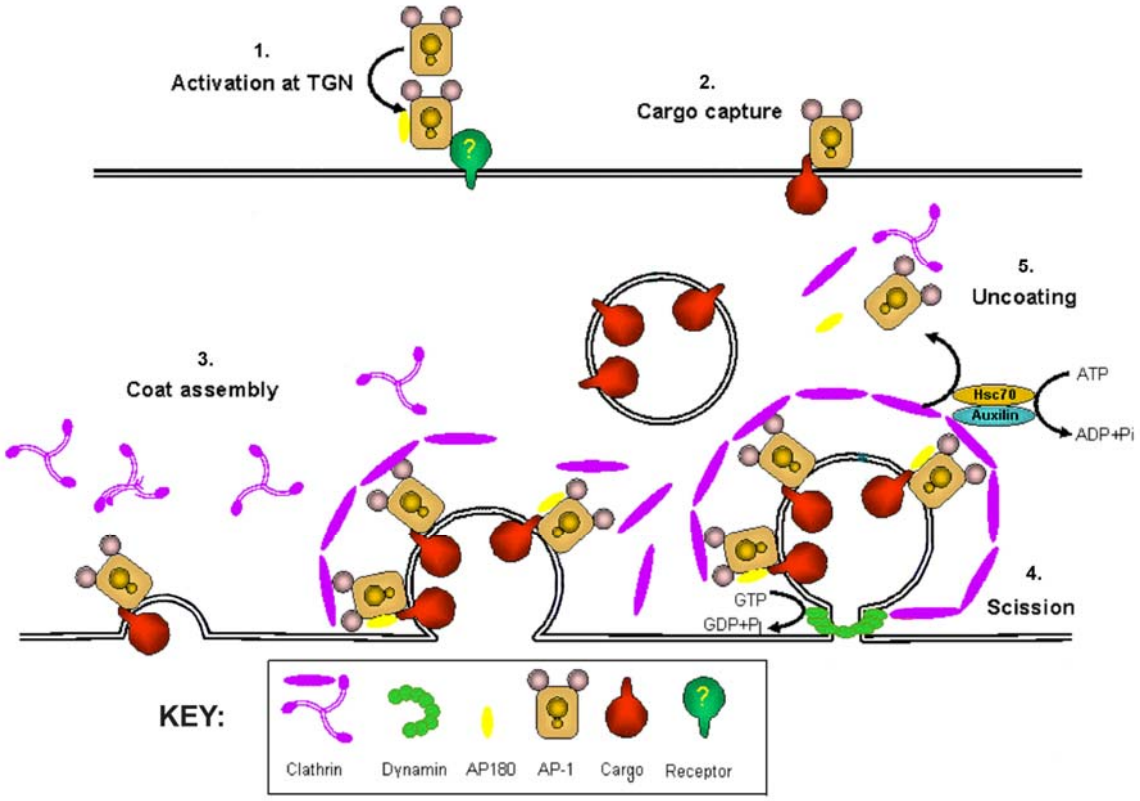


Figure 6. Classical clathrin-mediated vesicular transport (Adapted from Malaria Parasite Metabolic Pathways, http://sites.huji.ac.il/malaria)

1. Activation at the TGN: At the trans-Golgi network (TGN), coat assembly is instigated by adaptor protein-1 (AP-1) which engages a receptor at the plasma membrane.
2. Cargo capture: One end of AP-1 binds to protein cargo.
3. Coat assembly: The other end of AP1 binds the coat component, clathrin. Clathrin triskelion assembly results in the formation of hexagonal and pentagonal cages that cause membrane distortion.
4. Scission: The resulting exvagination is pinched off via the action of dynamin.
5. Uncoating: As the vesicle buds off from the TGN compartment, uncoating occurs via the action of heat shock cognate 70 (Hsc70) and auxillin with the concomitant hydrolysis of ATP. The vesicle is then transported to its destination and fuses with its destination compartment's plasma membrane, releasing the protein contents.

By targeting proteins like AP-1 involved in vesicle-formation within the trafficking pathway, proteins involved in attachment to and invasion of RBCs could be prevented from reaching the secretory organelles, and the erythrocyte invasion process could thus be prevented.

1.6.2) *PFE1400c:* a putative AP-1 β subunit

The second gene that will be investigated, namely *PFE1400c,* found on chromosome 5, is 3 714bp in length and contains seven exons (*Figure 7*).

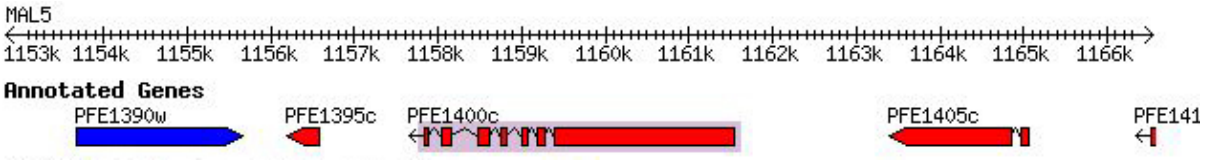


Figure 7. Genomic context of PFE1400c (PlasmoDB 5.3, 2007; www.plasmodb.org)

Crick-orientation genes are represented in red, while Watson-orientation genes are coloured blue.

The gene has Crick-orientation, meaning that the coding sequence is found on the antisense strand. This gene codes for the putative adaptor protein complex-1 (AP-1) β1 subunit. This subunit is potentially involved in the formation of clathrin-coated vesicles for protein trafficking within the parasite, ensuring that proteins reach their final cellular destinations (*Cooke et al, 2004*).

1.7) Aims of the project

The aim of this project is to express, extract and purify recombinant *P. falciparum* proteins that participate in vital protein-protein interactions within the parasite and with the host erythrocyte membrane. Two proteins have been chosen: a PK, the product of *PFB0150c*, and an AP-1 β1 subunit, the product of *PFE1400c*, that will be expressed as two domains – the clathrin adaptor appendage domain and the adaptin N-terminal domain. The specific objectives of the project are as follows:

- culture the 3D7 strain of *P. falciparum*
- design primers and amplify the genes of interest using PCR
- subclone the genes into expression vectors
- express the recombinant proteins in competent *E. coli* cells
- purify the recombinant proteins using affinity separation
- analyse the structural and biochemical characteristics of the expressed proteins

This project will form the groundwork for further studies involving the inhibition of these proteins by gene knockout or knockdown to determine whether they are feasible drug targets.

CHAPTER 2
MATERIALS AND METHODS

All solutions were made using ultrapure laboratory grade water filtered using the Milli-Q™ Water System (*Millipore Corporation, USA*). Sterile methods were employed for all culturing techniques, and solutions were sterilised by autoclaving or filtration. The methods for preparing the solutions in each section appear in the Appendix.

2.1) Malaria culturing techniques

The *P. falciparum* 3D7 strain was cultured in a sterile hood using an amended method of Trager and Jensen (*1976*), the amendment being the exclusion of the candle jar procedure to obtain the correct gas mixture for culturing**.** The parasitised red blood cells (pRBC) were decanted into sterile culture flasks (*Nunc, Germany)* and maintained in complete medium at 37°C in a gas mixture of 5.5% CO_2, 2.75% O_2 and 91.25% N_2 (*Afrox, South Africa*). The life cycle stage, percentage parasitaemia and health of the parasites were determined on a daily basis via preparation of smears and examination under a microscope (*Zeiss Axiostar, Germany*). Complete medium was aspirated and replaced on a daily basis. Cultures were maintained at approximately 5% parasitaemia and when this level exceeded 10%, cultures were divided or excess cells were aspirated. Fresh washed erythrocytes were added to maintain the haematocrit at 5%.

2.1.1) Preparation of culture from frozen stock

The preparation of cultures from frozen stock was carried out according to the method of Haeggström and Schlichtherle (*2004*). Tubes of cryo-frozen *P. falciparum*-infected erythrocytes were removed from liquid nitrogen. The tubes were placed into a 37°C water bath for 5 minutes or until thawed. The stock was transferred to 15ml Nunc tubes and 100µl 12% NaCl was added in a drop-wise fashion for each millilitre of stock. The tube was gently swirled and left to stand at room temperature for 5 minutes. Nine volumes of 1.6% NaCl were added and gently mixed.

The tube was centrifuged using an Eppendorf centrifuge 5702R (*Eppendorf, Germany*) at 250x g at 4°C for 5 minutes and the supernatant was aspirated. The pellet was resuspended in 9 volumes of 0.9% NaCl/0.2% glucose solution and centrifuged at 250x g for 5 minutes. The supernatant was aspirated and the remaining pellet was made up to 250µl with freshly washed erythrocytes. This solution was transferred to a 50ml culture flask and made up to 5ml using complete medium with 20% plasma. This plasma-enriched medium was used for the first week to initiate culture growth; thereafter medium with 10% plasma was used. The flask was gassed with 5.5% CO_2, 2.75% O_2 and 91.25% N_2 for 30 seconds and sealed tightly. The flask was incubated at 37°C for 48hrs.

2.1.2) Heat inactivation of plasma

Aliquots of frozen AB plasma (*SANBS*, *South Africa*) were thawed at room temperature, followed by heat-inactivation for 2 hours at 56°C. The inactivated plasma was transferred to sterile 50ml Nunc tubes. The tubes of plasma were centrifuged (*Centrifuge 5702R; Eppendorf, Germany*) at 750x g at 20°C for 10 minutes. 10ml or 45ml of the plasma was aliquotted into 15ml or 50ml Nunc tubes respectively and stored at –70°C.

2.1.3) Washing erythrocytes

Fresh blood from volunteers was collected in 6ml acid citrate dextrose tubes (*Becton Dickinson Biosciences, USA*) ((Ethics number M03-11-06; University of the Witwatersrand; Committee for Research on Human Subjects (medical)) and centrifuged (*MSE Coolspin centrifuge; Fisons Scientific, UK*) at 400x g for 10 minutes at 4°C. The plasma and buffy coat were aspirated in a sterile hood. Two volumes of PBS were added to the remaining erythrocytes and the solution was mixed and centrifuged as before. This process was repeated twice, after which the erythrocytes were resuspended in 1 volume of incomplete medium. Washed erythrocytes were stored in sterile 15ml tubes at 4°C for 5 days after which the unused cells were discarded and a fresh batch was prepared.

2.1.4) Preparation of smears

The culture flask was gently tilted so that the medium was displaced and the pRBC were concentrated on the bottom of the flask. A pipette was used to aspirate 10µl of the erythrocytes and this was transferred onto a glass microscope slide. The drop of erythrocytes was thinly smeared across the slide and allowed to air dry. The smear was then stained using the Rapindiff staining kit (*Global Diagnostics, South Africa*). The slide was air dried and viewed at 1000x magnification under oil immersion using a microscope (*Zeiss Axiostar, Germany*).

2.1.5) Calculating parasitaemia

The infected erythrocytes and uninfected erythrocytes were counted in at least 5 different fields containing more than 100 erythrocytes per field.
The percentage parasitaemia for each field was calculated using the following formula:

$$\% \text{ Parasitaemia} = \frac{\#\text{ Infected erythrocytes}}{\#\text{ Infected erythrocytes} + \#\text{ uninfected erythrocytes}} \times 100$$

An average percentage parasitaemia was calculated.

2.1.6) Feeding the culture

The flask was tilted gently to collect the medium in a corner. Using a Pasteur pipette and pump, the stale medium was aspirated. Fresh complete medium, warmed up to 37°C, was added to the flask. The volume of complete medium added was 5ml for small 50ml-volume flasks, 15ml for medium 80ml-volume flasks and 30ml for large 175ml-volume flasks. The flask was gassed with 5.5% CO_2, 2.75% O_2 and 91.25% N_2 for 30 seconds and sealed tightly. The flask was incubated at 37°C in an incubator. The medium was changed once a day, at approximately the same time each day.

2.1.7) Dividing the culture

A culture was divided when the parasitaemia was greater than 10% or when a culture needed to be continued when using the majority for DNA/RNA extraction. 5ml of complete medium was decanted into the small flask and washed erythrocytes were added to a final haematocrit of 5%. A 1ml aliquot of pRBC was taken from the culture flask and added to the new flask to give an approximate parasitaemia of 2%. Flasks were gassed and incubated at 37°C in an incubator.

2.1.8) Synchronisation of the culture – Sorbitol treatment

Synchronisation of *P. falciparum* erythrocytic stages in culture was achieved using the method determined by Lambros and Vanderberg (*1979*). This method involves the lysis of pRBCs containing trophozoites and schizonts, due to the increased permeability of the RBC membrane to sorbitol during these later stages. The influx of sorbitol causes a concomitant influx of water and the osmotically fragile pRBCs burst. The only pRBCs that will remain are those containing the early ring-stages of the parasite.
A 5% D-Sorbitol solution was prepared and sterilised by filtration using a Millex GP 0.22µm syringe-driven filter unit (*Millipore, Ireland*) and stored at 4°C. A smear was prepared to verify that parasites were mainly in the early ring stage with a parasitaemia greater than 10%. The culture was transferred to a 50ml Nunc tube and centrifuged (*Centrifuge 5702R; Eppendorf, Germany*) at 250x g for 5 minutes at 20°C. The supernatant was aspirated and 10 volumes of D-sorbitol were added to the pRBC pellet. The pellet was resuspended and left to stand at room temperature for 30 minutes. The solution was centrifuged as before and the supernatant was removed. The pellet was resuspended in complete medium and washed erythrocytes were added to a haematocrit of 5%. The flask was gassed with 5.5% CO_2, 2.75% O_2 and 91.25% N_2 for 30 seconds and sealed tightly. The flask was incubated at 37°C.

2.1.9) Freezing of cultures

Cultures were frozen when parasites were mainly in the ring stage and parasitaemia was 10% or more. The culture was centrifuged (*Centrifuge 5702R; Eppendorf, Germany*) at 250x g for 5 minutes at 20°C and the supernatant was removed. The packed cells were resuspended in a 1:1 ratio with 60% glycerol solution. The suspension was transferred to sterile cryotubes in 1ml aliquots and left to settle for 5 minutes. The tubes were then placed in liquid nitrogen.

2.2) DNA experiments

2.2.1) DNA extraction from *P. falciparum*

DNA extraction was carried out on parasites in the late stages of the life cycle (trophozoites and/or schizonts) from cultures with 10-15% parasitaemia according to the methods of Schlichtherle and Wahlgren (*2004*). The culture was transferred from 175ml-volume large culture flasks, containing 30ml of culture, to sterile 50ml Nunc tubes and centrifuged (*Centrifuge 5702R; Eppendorf, Germany*) at 750x g for 5 minutes at 4°C. The supernatant was removed and the pellet was washed in PBS followed by centrifugation at 750x g for 5 minutes at 4°C. The supernatant was removed and the pRBC pellet was resuspended in 1ml PBS in a 15ml Nunc tube. 10µl of a 5% saponin (*USB Corporation, USA*) solution was added for every millilitre of pRBC. The suspension was left at room temperature for 3-5 minutes to allow the erythrocytes to lyse. This was followed by centrifugation at 750x g for 5 minutes at 4°C. The supernatant was decanted. PBS was added to a volume twice that of the pellet volume to resuspended the pellet. This mixture was centrifuged at 750x g for 5 minutes at 4°C and the pellet was washed three times with brief vortexing. 1mg of PCR-grade recombinant proteinase K (*Roche Applied Science, Germany*) was added to 1ml of lysis buffer. 250µl lysis buffer/proteinase K solution was diluted with 750µl water and added to the parasite pellet, which was resuspended by gentle aspiration.

Proteinase K is an endopeptidase that cleaves peptide bonds, thus promoting cell lysis and inactivating endogenous nucleases.

The sample was incubated for 3 hours at 37°C, with mixing every hour using a pipette to promote parasite lysis. After 3 hours, the solution was aspirated with a syringe and 21 gauge needle to ensure that parasites were completely lysed. DNA was extracted from the crude cell lysate using the phenol extraction and ethanol precipitation method (*Moore & Dowhan, 2002*). An equal volume of 1:1 phenol/chloroform mixture was added. The phenol (*Sigma-Aldrich Inc, USA*) used had a pH of 8.2 and was saturated with 10mM Tris HCl (pH 8.0) and 1mM EDTA. The solution was mixed well by inverting the tube several times. The sample was then centrifuged (*Centrifuge 5415R; Eppendorf, Germany*) at 18 000x g for 5minutes at 4°C. 80-90% of the aqueous phase was collected and transferred to a new Eppendorf tube. 500µl TE buffer was added to the remaining organic phase. This mixture was centrifuged at 18 000x g for 5 minutes at 4°C. The aqueous phase was collected and added to the previously collected aqueous phase. An equal volume of chloroform was added, followed by centrifugation at 18 000x g for 5 minutes at 4°C. The aqueous phase was collected. 10µl of RNase A (*Fermentas, Europe*), supplied in 50mM Tris-HCl (pH 7.4) and 50% glycerol, was added per ml of aqueous phase (final concentration = 100ug/ml) followed by incubation at 37°C for 30 minutes. The phenol/chloroform extraction was repeated. DNA from the aqueous phase was precipitated by adding 0.1 volumes of 3M sodium acetate (pH 5.2) and 2.5 volumes 100% ice-cold ethanol. This solution was placed at -70°C for 30 minutes and then centrifuged at 18 000x g for 30 minutes at 4°C. The supernatant was removed and 70% ice-cold ethanol was added to the pellet. This was followed by centrifugation at 18 000x g for 5 minutes at 4°C. The supernatant was removed and the pellet was air-dried for 15 minutes to ensure that the ethanol had evaporated. The pellet was resuspended in 20µl nuclease-free water (*Promega, USA*) and stored at 4°C overnight and at -70°C thereafter.

2.2.2) Determination of DNA concentration and purity

DNA concentration was determined using a Nanodrop ND-1000 spectrophotometer (*Nanodrop Technologies Inc, USA*) according to the following equation**:**

$$[DNA] \text{ in } \mu g/\mu l = A_{260} \times 50 \times DF \times 0.001$$

Where: DF = dilution factor

50 = extinction coefficient of 1µg/ml dsDNA

0.001 = conversion factor for ml to µl

The DNA purity was determined using the A_{260}/A_{280} value. A value of 1.8 indicates a pure sample. If the A_{260}/A_{280} reading is less than 1.8, contamination with proteins, salt, ethanol or aromatic substances like phenol may be present. An A_{260}/A_{280} reading greater than 2 indicates the presence of RNA in the sample. Samples with an A_{260}/A_{280} reading between 1.5 and 1.9 were considered pure enough for further use. Integrity and purity of the DNA samples were determined via electrophoresis on a 0.8% mini agarose gel (10x8cm). The agarose used was D-1 LE agarose (*Hispanagar, Spain*). The gel was cast into a horizontal mini-gel kit model # MGU-200T (*CBS Scientific, USA*). The gel was electrophoresed for 1 hour at 65V (*Electrophoresis power supply EPS 301; Amersham pharmacia biotech, Sweden*) in 1x TAE buffer. 2.5µl of 10µg/µl ethidium bromide was added to the anode chamber of the gel system. Ethidium bromide binds to nucleic acids and fluoresces orange under UV light. DNA was visualised under UV light and photographed using GeneSnap version 6.05 image acquisition software (*Syngene, UK*).

2.3) RNA experiments

2.3.1) RNA extraction from *P. falciparum*

Parasitised erythrocytes were harvested from a large culture flask with 10-15% parasitaemia in the late stages of development. The cultures were centrifuged using an Eppendorf centrifuge 5702R (*Eppendorf, Germany*) at 750x g for 4 minutes at 4°C. RNA was extracted according to the Tri-Reagent (*Sigma-Aldrich Inc, USA*) manufacturer's protocol. The supernatant was removed and 5ml of Tri-Reagent was added for every 0.5 ml erythrocytes. The solution was shaken and incubated at 37°C for 5 minutes. Chloroform was added to a volume 0.2 times that of the original Tri-Reagent volume. The solution was vigorously vortexed for 15 seconds and placed at room temperature for 3 minutes. Centrifugation at 1 600x g at 4°C for 35 minutes was then carried out. The upper aqueous phase was carefully transferred into 2ml Eppendorf tubes. Cold isopropanol to 0.5 times the original Tri-Reagent volume was added to precipitate the RNA. The solution was mixed by inversion and placed on ice for 2 hours. The solution was centrifuged (*Centrifuge 5415R; Eppendorf, Germany*) at 18 000x g for 30 minutes at 4°C. The supernatant was removed and 800µl of cold 70% ethanol was added to the RNA pellet. Centrifugation at 18 000x g for 5 minutes at 4°C followed. The ethanol was removed and the pellet was air dried for 15 minutes at room temperature. 20µl of nuclease-free water (*Promega, USA*) was added and the RNA was stored at -70°C.

2.3.2) Determination of RNA concentration and integrity

RNA concentration was determined using a Nanodrop ND-1000 spectrophotometer (*Nanodrop Technologies Inc, USA*) according to the following equation:

$$[RNA] \text{ in } \mu g/\mu l = A_{260} \times 40 \times DF \times 0.001$$

Where: DF = dilution factor

40 = average extinction coefficient of 1µg/ml ribonucleic acids

0.001 = conversion factor for ml to µl

RNA purity was ascertained according to the A_{260}/A_{280} value obtained. A value equal to 2 indicates a pure sample of RNA. If the ratio is less than 2 then protein contamination may be present. Samples with A_{260}/A_{280} values between 1.7 and 2 were considered pure enough for RT-PCR. RNA integrity was determined via electrophoresis on a 1% agarose gel in TAE buffer. The gel was cast, electrophoresed and visualised as described in section 2.2.2.

2.3.3) Reverse transcription

RNA was reverse transcribed using the SuperScript ™ III RNase H Reverse Transcriptase kit (*Invitrogen Life Technologies, UK*) to produce cDNA. The manufacturer's protocol was followed. The following components were added to a nuclease-free PCR tube: 1µl oligo $(dT)_{12\text{-}18}$, 1µg total RNA, 4µl 2.5mM dNTP, nuclease-free water to 13µl. The mixture was heated to 65°C for 5 minutes - to ensure that any RNA secondary structures were resolved - and then incubated on ice for 1 minute. The following reagents were added: 4µl 5X first strand buffer, 1µl 0.1M DTT, 1µl RNaseOUT and 1µl SuperScript ™ III. The solution was gently mixed and incubated at 55°C for 60 minutes. The enzyme was inactivated by heating to 70°C for 15 minutes. The cDNA was then amplified to produce the *PFE1400c* clathrin adaptor appendage domain insert.

2.4) PCR

2.4.1) Primer design

Primers were designed based on DNA sequences obtained from the *Plasmodium* genome database, PlasmoDB version 5.3 (*2007*). Using Integrated DNA Technologies SciTools Oligo Analyzer 3.0 (*www.idtdna.com)* the primers were analysed to ensure that minimal hairpins and primer-dimers would be formed. Recognition sites for the restriction endonucleases *Bam*HI, *Xho*I and *Nde*I were included to facilitate ligation into either the pGEX-4T-2 (*Amersham Biosciences, UK*) or pET-15b (*Novagen, USA*) vector (*Vector maps, Appendix section 5.8*). Primers were synthesised by Inqaba Biotec (*South Africa*) and provided in lyophilised form. The lyophilised primers were reconstituted in TE buffer to a final concentration of 100µM and stored at -20°C.

The following primers were used, with the bases in red indicating the restriction endonuclease recognition site; the bases in bold indicate the extra bases added to facilitate enzyme binding and activity at the site, and the arrows show the exact site of cleavage:

PFB0150c PK domain:

Forward primer (with *Bam*HI recognition site)
5'- **CGC** **G↓GA TCC** GAT GAA AAG GAT GGA TAT G -3'

Reverse primer (with *Xho*I recognition site)
5'- **CCG** **C↓TC GAG** **TCA** TTT CTG GGA TTG TTC AGT -3'

The restriction endonuclease recognition sites (red) are included to facilitate insertion into the pGEX-4T-2 vector. The PCR product produced using these primers was 1 515bp in size.

PFE1400c adaptin N-terminal domain:

Forward primer (with *Nde*I recognition site)
5'- **CGG TTT CCA↓ TAT G**AT GTC TGA TTT ACG CTA CTT -3'

Reverse primer (with *Bam*HI recognition site)
5'- **CGC G↓GA TCC** ATG ATC ATT ATT ATT ATC AGA A -3'

The restriction endonuclease recognition sites (red) are included to facilitate insertion into the pET-15b vector. The PCR product produced using these primers was 1 800bp in size.

PFE1400c clathrin adaptor appendage domain:

Forward primer (with *Bam*HI recognition site)
5'- **CGC G↓GA TCC** TCT TCG GAT GAA TTC AAT AA -3'

Reverse primer (with *Xho*I recognition site)
5'- **CCG C↓TC GAG TCA** CTG AGT TAC ACT TAA GGA AAA -3'

The restriction endonuclease recognition sites (red) are included to facilitate insertion into the pGEX-4T-2 vector. The PCR product produced using these primers was 723bp in size.

2.4.2) PCR protocol

The PCR was carried out using the Expand High Fidelity ^{PLUS} PCR system (*Roche Applied Science, Germany*). The high-fidelity DNA polymerase ensures that the DNA sequence is amplified with high specificity and accuracy, and is six-fold more accurate than *Taq* DNA polymerase alone. The reaction mixture was prepared as indicated in Table 1.

TABLE 1: Reaction mixture for PCR using the Expand High Fidelity ^{PLUS} PCR system

REAGENT	VOLUME (µl)	FINAL CONCENTRATION
PCR grade water	Add to 50	
Expand HiFi ^{PLUS} reaction buffer 5X ; Vial 2	10	1.5mM $MgCl_2$
dNTP (10mM)	1	200µM
Forward primer (20µM)	1	0.4µM
Reverse primer (20µM)	1	0.4µM
Expand HiFi ^{PLUS} enzyme blend; Vial 1	0.5	2.5U
Genomic DNA	0.8	125ng

The SuperScript ^{TM} manufacturer's protocol recommended 2µl of cDNA be added to the reaction mixture for PCR. The reaction tubes were placed in a Mastercycler gradient thermocycler (*Eppendorf, Germany*). Two successive annealing temperatures were used for each amplicon. Annealing temperatures were calculated to be 3°C lower than the melting temperatures, which were supplied by the manufacturer (*Inqaba Biotec, South Africa*). A range of temperatures (within 6°C of the calculated melting temperature) were tested to optimise the reaction conditions for each DNA sequence to be amplified. The first annealing temperature was calculated for the *P. falciparum* sequence primer without the added bases making up the restriction site. The second annealing temperature took into account the extra bases. The following programs produced the best results and were used for subsequent PCRs:

PFB0150c PK domain:

Initial denaturation: 94°C for 2 minutes
Denaturation: 94°C for 1 minute
Primer annealing: 52.2°C for 1 minute, 5 cycles; 59.2°C for 1 minute, 29 cycles
Extension: 72°C for 1 minute
Final extension: 72°C for 5 minutes

PFE1400c adaptin N-terminal domain:

Initial denaturation: 94°C for 2 minutes
Denaturation: 94°C for 1 minute
Primer annealing: 48.5°C for 1 minute, 5 cycles; 56.9°C for 1 minute, 29 cycles
Extension: 72°C for 1 minute
Final extension: 72°C for 5 minutes

PFE1400c clathrin adaptor appendage domain:

Initial denaturation: 94°C for 2 minutes
Denaturation: 94°C for 1 minute
Primer annealing: 46.1°C for 1 minute, 5 cycles; 52°C for 1 minute, 29 cycles
Extension: 72°C for 1 minute
Final extension: 72°C for 5 minutes

The PCR products were purified using the phenol extraction and ethanol precipitation method as previously described (*section 2.2.1*) to change the PCR buffer. For every 50µl of PCR product, 50µl of nuclease-free water was added. Following the extraction, the dried pellet was resuspended in 40µl of nuclease-free water. Aliquots of the purified PCR products were resolved on a 1% agarose gel and the concentration of the samples was determined on a Nanodrop ND-1000 spectrophotometer (*Nanodrop Technologies Inc, USA*) as described for DNA (*section 2.2.2*).

2.4.3) Restriction digestion of PCR products

Double digests on the purified PCR products were carried out using Fermentas (*Europe*) restriction endonucleases. The reaction mixtures were prepared - to a final volume of 20µl - as recommended by the Fermentas DoubleDigest™ engine (*www.fermentas.com/ doubledigest*), outlined in Table 2. Briefly, double digestion with *Bam*HI and *Xho*I was carried out in 2X Tango™ buffer at 37°C for 18 hours. Incubation times of 2, 6 and 18 hours were tested, with the latter resulting in optimum digestion. Under these conditions *Bam*HI has an activity of 50-100% and star activity is ameliorated; *Xho*I has 100% activity under these conditions. Star activity is the non-specific cleavage of DNA by restriction enzymes when reaction conditions are not optimised. Double digestion with *Bam*HI and *Nde*I was carried out in 2X Tango™ buffer at 37°C for 18 hours. These reaction conditions result in enzyme activity of 50-100% for both *Bam*HI and *Nde*I.

TABLE 2: Reaction mixture for PCR product restriction digests using Fermentas restriction endonucleases

PCR PRODUCT	RESTRICTION ENZYME 1	RESTRICTION ENZYME 2	2X BUFFER	MILLI-Q WATER
PFB0150c Kinase (1µg)	*Bam*HI - 2µl (20U)	*Xho*I - 1µl (10U)	10X Tango™ buffer - 4µl	Make up to 20µl
PFE1400c Clathrin (1µg)	*Bam*HI - 2µl (20U)	*Xho*I - 1µl (10U)	10X Tango™ buffer - 4µl	Make up to 20µl
PFE1400c Adaptin (1µg)	*Bam*HI - 2µl (20U)	*Nde*I - 2µl (20U)	10X Tango™ buffer - 4µl	Make up to 20µl

Aliquots of the samples were resolved on a 1% agarose gel as previously described (*section 2.2.2*) to check the size of the DNA inserts (in case of star activity). After restriction digestion, DNA inserts were purified using phenol extraction and ethanol precipitation as previously described (*section 2.2.1*). The concentrations of DNA inserts were determined using a Nanodrop ND-1000 spectrophotometer (*Nanodrop Technologies Inc, USA*).

2.5) Plasmid preparation

2.5.1) Extraction of plasmid DNA from DH5α cells

Plasmids, namely pGEX-4T-2 (*Amersham Biosciences, UK*) and pET-15b (*Novagen, USA*) were extracted from Subcloning Efficiency™ DH5α™ chemically competent cells (*Invitrogen, USA*) using the GenElute Plasmid Miniprep Kit (*Sigma-Aldrich Inc, USA*) according to the manufacturer's protocol and using kit reagents. The method relies on an SDS-based lysis of *E. coli* cells, followed by the adsorption of plasmid DNA onto silica in the presence of high salt concentrations. A spin-wash step removes any contaminants and the bound DNA is eluted.

DH5α cells from glycerol stock were inoculated into sterile 100ml Erlenmeyer flasks containing 10ml Luria broth, which contains nutrients and a buffer to provide optimum growth conditions for the cells. The cell cultures were grown overnight, incubated in an orbital shaker at 37°C. The cells were harvested by centrifuging (*Beckman model J2-21 centrifuge; Beckman Coulter, USA*) at 3 500x g for 10 minutes at room temperature. The supernatant was discarded and cells were resuspended in 200µl Resuspension solution. 200µl of alkaline Lysis solution was added to the cells and the tubes were inverted gently to mix. The solution was allowed to clear for 5 minutes and then 350µl of Neutralisation solution was added. The solution was gently mixed and transferred to 2ml Eppendorf tubes. The debris was pelleted by centrifuging (*Centrifuge 5415R; Eppendorf, Germany*) at 18 000x g for 10 minutes. Binding columns were prepared by adding 500µl Column Preparation Solution and centrifuging at 18 000x g for 1 minute. The flow-through was discarded and the cleared lysate was transferred into the Eppendorf tube containing the binding column. Centrifugation at 18 000x g for 1 minute followed and the flow-through was discarded. 750µl of Wash Solution was added to the column, followed by centrifugation at 18 000x g for 1 minute. The flow-through was discarded and the binding column was centrifuged again to remove any excess ethanol.

The binding column was transferred to a new collection tube and 100µl of Elution solution was added. The collection tube was centrifuged at 18 000x g for 1 minute and the plasmid DNA was eluted. The concentration of plasmid DNA was determined using a Nanodrop ND-1000 spectrophotometer (*Nanodrop Technologies Inc, USA*)**.** Aliquots of the samples were resolved on a 1% agarose gel as previously described (*section 2.2.2*) to check the size and integrity of the plasmids. Samples were stored at -20°C until required.

2.5.2) Restriction enzyme digestion of plasmids

Plasmids were subjected to double restriction digests, using Fermentas (*Europe*) restriction endonucleases, to prepare them for ligation with *P. falciparum* DNA inserts. The reaction mixtures were prepared as outlined in Table 3, and previously described (*section 2.4.3*).

TABLE 3: Reaction mixture for plasmid restriction digests using Fermentas restriction endonucleases

PLASMID	RESTRICTION ENZYME 1	RESTRICTION ENZYME 2	2X BUFFER	MILLI-Q WATER
pGEX-4T-2 (1µg)	*Bam*HI - 2µl (20U)	*Xho*I - 1µl (10U)	10X Tango™ buffer - 4µl	Make up to 20µl
pET-15b (1µg)	*Bam*HI - 2µl (20U)	*Nde*I - 2µl (20U)	10X Tango™ buffer - 4µl	Make up to 20µl

The solutions were incubated at 37°C for 18 hours. After restriction digestion, plasmids were purified using the phenol-chloroform extraction and ethanol precipitation method as previously described (*section 2.2.1*). The concentration of plasmid DNA was determined using a Nanodrop ND-1000 spectrophotometer (*Nanodrop Technologies Inc, USA*) and aliquots were resolved on a 1% agarose gel as before (*section 2.2.2*) to check that the plasmids had been linearised.

Fermentas FastDigest™ restriction endonucleases were also utilised, which only required a 1 hour incubation time. The reaction mixtures were prepared as follows in Table 4.

TABLE 4: Reaction mixture for plasmid restriction digests using Fermentas FastDigest™ restriction endonucleases

PLASMID	RESTRICTION ENZYME 1	RESTRICTION ENZYME 2	1X BUFFER	MILLI-Q WATER
pGEX-4T-2 (1µg)	FastDigest™ *Bam*HI - 1µl (0.25U)	FastDigest™ *Xho*I - 1µl (0.5U)	10X Tango™ buffer - 2µl	Make up to 20µl
pET-15b (1µg)	FastDigest™ *Bam*HI - 1µl (0.25U)	FastDigest™ *Nde*I - 1µl (0.5U)	10X Tango™ buffer - 2µl	Make up to 20µl

2.5.3) Alkaline phosphatase treatment of linearised plasmids

After purification the restriction digested plasmids were treated with calf intestinal phosphatase (CIP) (*Roche Applied Science, Germany*) to prevent self-ligation. For every 20µl of insert, 1µl of CIP and 2µl of 10X CIP buffer was added. The solution was incubated at 37°C for 1 hour. The purification process was then repeated. The concentration of plasmid DNA was determined using a Nanodrop ND-1000 spectrophotometer (*Nanodrop Technologies Inc, USA*) (*section 2.2.2*).

2.6) Subcloning

2.6.1) Ligation of DNA insert and plasmid

Ligations were carried out using the DNA Ligation Kit (*Roche Applied Science, Germany*). A 1:3 molar ratio of plasmid to DNA insert, with the total amount of DNA not exceeding 200ng, was added to 2µl 5X DNA dilution buffer and made up to 10µl with nuclease-free water.

Next, 10µl of T4 DNA Ligation buffer was added to this solution and mixed. Lastly, 1µl of T4 DNA Ligase was added and the solution was incubated at 16°C for 30 minutes. Thereafter, 5µl of ligated DNA solution was used for transformation, while the remainder was stored at -20°C.

2.6.2) Transformation of DH5α cells

50µl aliquots of Subcloning Efficiency™ DH5α™ chemically competent cells (*Invitrogen, USA*) were thawed on ice. The plasmid-DNA insert solution was incubated at 70°C for 10 minutes to heat-inactivate the T4 DNA Ligase enzyme. 5µl of the ligated plasmid-DNA insert solution was added to the DH5α cells and incubated on ice for 30 minutes. A control reaction contained the plasmid with no insert. The suspension was incubated at 37°C for 20 seconds to heat shock the cells. The cells were placed on ice for 2 minutes. 500µl of Luria broth was added and the solution was placed in an orbital shaker at 37°C for 1 hour. Aliquots of 50µl, 100µl and 350µl were spread in a sterile manner onto agar plates. The agar plates were made by autoclaving 1L of Luria broth with 15g of Biolab bacteriological agar (*Merck, Germany*), cooling the solution and adding 1ml of 100mg/ml ampicillin (*Roche Applied Science, Germany*). This solution was poured into 85mm culture plates (*Costar, USA*). 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-bearing, ampicillin-resistant bacterial colonies will proliferate.

The next morning, colonies were compared to the control plate to check for any self-ligated vectors. Five single colonies were removed from the plate using a pipette and placed in 10µl Milli-Q water. 5µl of this solution was used in a PCR reaction. Primers specific for the *P. falciparum* insert were used. The bacteria were first lysed by heating at 94°C for 5 minutes in a thermocycler and PCR reagents were added as outlined in Table 1 (*section 2.4.2*). The PCR products were resolved on a 1% agarose gel (*section 2.2.2*) to check that the inserts were present.

The remaining 5µl of the positive colonies was added to BD Falcon™ round bottom tubes (*Becton Dickinson Biosciences, USA*) containing 2ml Luria broth with 20µl of 100mg/ml ampicillin and placed in a 37°C orbital shaker overnight. The next day, stock solutions were made by adding 500µl of the transformed DH5α cells to 500µl of 60% sterile glycerol. The stocks were stored at -70°C. The plasmids were extracted from the remaining cells using the GenElute Plasmid Miniprep Kit (*Sigma-Aldrich Inc, USA*) as described in section 2.5.1. 50µl of the plasmid was digested for two hours with the relevant restriction endonucleases (*section 2.5.2*). Whole plasmids and digested plasmids were resolved on a 1% agarose gel (*section 2.2.2*) to verify the presence of the insert.

2.6.3) Transformation of Rosetta 2 (DE3) cells

Rosetta 2 (DE3) cells (*Novagen, USA)* supply tRNAS for seven rare codons that are not usually present in *E. coli* cells, which enhances the expression of eukaryotic proteins. 20µl aliquots of Rosetta 2 (DE3) cells in were thawed on ice. 1µl of the recombinant plasmid was added to the competent cells and gently mixed by inversion. The cells were incubated on ice for 30 minutes, followed by a heat shock at 37°C for 30 seconds. The solution was immediately placed on ice for 2 minutes. 500µl of Luria broth was added, which was placed in an orbital shaker at 37°C for 1 hour. 50µl and 450µl aliquots were sterily spread onto agar plates treated with ampicillin (100mg/ml) and chloramphenicol (50mg/ml) and placed in an incubator at 37°C overnight. 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. The next morning, colonies were compared to the control plate to check for any self-ligated vector. Ten single colonies were removed using a pipette and placed into 10µl Milli-Q water. 5µl of this solution was used in a PCR reaction as described in section 2.4.2. The PCR products were resolved on a 1% agarose gel (*section 2.2.2*) to check that the inserts were present. Stock solutions were prepared and stored as in section 2.6.2. The plasmids were extracted from the remaining cells using the manual plasmid preparation method (*section 2.6.4*).

Half of the plasmid solution was digested for two hours with the relevant restriction endonucleases (*section 2.5.2*). Whole plasmids and digested plasmids were resolved on a 1% agarose gel (*section 2.2.2*) to verify the presence of the insert.

2.6.4) Extraction of plasmid DNA from Rosetta 2 (DE3) cells

The extraction of plasmid DNA from Rosetta 2 (DE3) cells was performed by utilising the alkaline lysis method as described by Bimboim & Doly (*1979*) with modern enhancements by Ehrt & Schnappinger (*2003*). This manual method resulted in better yields of plasmid than the kit. The method is based on the disparity between the denaturation of chromosomal and plasmid DNA, which allows the separation of the two. Both the high molecular weight chromosomal DNA and the plasmid denature on addition of an SDS/ NaOH solution, but the latter remains double-stranded. The solution is neutralised by the addition of KCH_3COO, which causes the chromosomal DNA to form an insoluble mass while the plasmid remains soluble. The precipitated chromosomal DNA forms a complex with the SDS and potassium, and is removed via centrifugation along with the cellular proteins. The plasmid DNA is retrieved by ethanol precipitation. Rosetta 2 (DE3) cells were grown overnight in 2ml Luria broth, 20µl 100mg/ml ampicillin and 20µl 50mg/ml chloramphenicol. The overnight growths were transferred into 2ml Eppendorf tubes and centrifuged (*Centrifuge 5415R; Eppendorf, Germany*) at 18 000x g for 1 minute at 4°C. The supernatant was discarded and 100µl ice cold lysis buffer was added to resuspend the cell pellet. The samples were incubated at room temperature for 5 minutes. 200µl of fresh 0.2M NaOH/1% SDS solution was added to each sample and mixed by inversion. 150µl ice-cold, fresh KCH_3COO (pH 4.8) was added to each sample, followed by incubation on ice for 5 minutes. Samples were centrifuged at 18 000x g for 10 minutes at room temperature. The supernatants were removed to Eppendorf tubes and 500µl of a 1:1 phenol/chloroform solution was added. Samples were centrifuged at 18 000x g for 5 minutes at 4°C. The upper aqueous phase was collected and 2.5 volumes of 100% ice cold ethanol was added, followed by incubation at room temperature for 5 minutes.

Samples were centrifuged at 18 000x g for 10 minutes at 4°C. The supernatants were discarded and the pellets were air dried for 20 minutes.
Once the pellets were transparent and dry, 100µl TE buffer and 1µl RNase A (*Fermentas, Europe*) were added to each pellet and incubated at 37°C for 30 minutes. 40µl fresh KCH_3COO (pH 4.8) was added, followed by 260µl of nuclease-free water. 500µl of a 1:1 phenol/chloroform solution was added and mixed. Samples were centrifuged at 18 000x g for 5 minutes at 4°C.The aqueous phase was collected and 250µl chloroform was added and mixed by inversion. Samples were then centrifuged at 18 000x g for 5 minutes at 4°C. The aqueous phase was collected and 1ml of 100% ice cold ethanol was added. The samples were incubated at -70°C for 15 minutes. Samples were centrifuged at 20 000x g for 15 minutes at 4°C. The 100% ethanol was discarded and the pellet was washed with 1ml 70% ethanol. The pellets were air dried at room temperature for 20 minutes, resuspended in 40µl nuclease-free water and stored at 4°C. Vector constructs were sequenced by Inqaba Biotech using pGEX sequencing primers (*Amersham Biosciences, UK*) for recombinant GST-tagged proteins and T7 sequencing primers for recombinant His-tagged protein. The 5' pGEX Sequencing Primer binds nucleotides 869-891 while the 3' pGEX Sequencing Primer binds nucleotides 1 019-997. The T7 promoter sequencing primer has the following sequence: 5'-TAA TAC GAC TCA CTA TAG GG-3'. The T7 terminator sequencing primer has the following sequence: 5'-GCT AGT TAT TGC TCA GCG G-3'

2.7) Recombinant protein experiments

2.7.1) Recombinant protein induction and expression

Expression of recombinant proteins was induced using the Overnight Express™ Autoinduction System (*Novagen, USA)*. 10µl of glycerol stock containing transformed Rosetta 2 (DE3) cells was transferred into 1ml of Luria broth with 10µl ampicillin (100mg/ml) and 10µl chloramphenicol (50mg/ml). The samples were placed in an orbital shaker at 37°C and incubated overnight.

The next day the optical density (OD) was checked at a wavelength of 600nm using a Du® -65 Spectrophotometer (*Beckman, UK*). The OD indicates the amount of bacterial cells present in the broth and therefore the approximate phase of growth that has been reached. If the OD was less than 0.4, the samples were incubated for a longer period of time to allow the bacterial cells to proliferate further and reach log phase. If the OD was between 0.4 and 0.6 the samples were used for expression. 200µl of the overnight grow was added to 20ml Overnight Express™ instant TB medium along with 20µl ampicillin (100mg/ml) and 20µl chloramphenicol (50mg/ml) in a 250ml Erlenmeyer flask. The flask was placed on a shaking platform for 18 hours at room temperature. After 18 hours the OD was checked at a wavelength of 600nm. If the OD was between 1.3 and 1.5, the bacterial cells were estimated to have reached the stationary phase and protein extraction was carried out.

2.7.2) Recombinant protein extraction

Protein extraction was carried out using BugBuster™ Protein Extraction Reagent (*Novagen, USA*), which utilises a mixture of non-ionic detergents to disrupt the cell wall of *E. coli* without denaturing soluble proteins. The cell culture was transferred to 50ml Beckman tubes and centrifuged (*Beckman model J2-21 centrifuge; Beckman Coulter, USA*) at 3 500x g for 10 minutes. The supernatant was removed and the remaining pellet was placed in a -70°C freezer for 20 minutes. 1ml of BugBuster™ Protein Extraction Reagent and 1µl of Protease Inhibitor Cocktail Set III (*Novagen,* USA) were added to resuspend each frozen pellet. The cell suspension was transferred to an Eppendorf tube and placed on a shaking platform at room temperature for half an hour. The insoluble fraction was removed by centrifugation (*Centrifuge 5415; Eppendorf, Germany*) at 18 000x g for 20 minutes. The supernatant was transferred to a fresh Eppendorf tube and stored at 4°C. 150µl aliquots were taken at specific stages of the extraction to evaluate protein expression.

An aliquot of overnight express culture was taken to analyse total cell protein content; an aliquot of supernatant after BugBuster™ protein extraction was taken to analyse the soluble protein fraction; the pellet after BugBuster™ protein extraction was resuspended in 1ml TE buffer and a 150µl aliquot was used to analyse the insoluble protein fraction (*table 5*). Each 150µl sample was added to 40µl 5x suspension solution, 4µl β-mercaptoethanol and 5µl bromophenol blue solution. The samples were boiled for 2 minutes and stored at -20°C prior to electrophoresis (*section 2.7.7*).

2.7.3) Recombinant protein purification

The purification of glutathione S-transferase (GST) fusion proteins was carried out using the MagneGST™ Protein Purification System (*Promega, USA*). The system utilises paramagnetic particles which have glutathione immobilised on their surface. The GST enzyme specifically binds to its substrate, glutathione, in a lock-and-key manner. On addition of the particles to a soluble protein fraction containing GST fusion proteins, the fusion proteins bind to the immobilised glutathione. 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 glutathione, which competes with the glutathione-coated particles for binding of the GST fusion proteins.

150µl aliquots of sample from various steps of the purification were taken to be analysed via SDS PAGE (*table 5*). The manufacturer's protocol was followed; for every 20ml of original culture, 20µl of MagneGST™ particles were resuspended and transferred to a 2ml Eppendorf tube. 20µl of particles have a binding capacity of 50µg fusion protein. The tube was placed in a magnetic particle separator (*Roche Applied Science, Germany*) and the particles were captured by the magnet. The supernatant was carefully removed and 100µl of MagneGST™ binding/wash buffer was added to the particles. The particles were resuspended by pipetting and the tube was placed in the magnetic stand. The supernatant was removed and the wash step was repeated twice. After the final wash the particles were gently resuspended in 100µl MagneGST™ binding/wash buffer.

The soluble protein fraction (*section 2.7.2*) was added to the particles along with 1µl Protease Inhibitor Cocktail Set III (*Novagen,* USA). The suspension was incubated at 4°C overnight with gentle mixing on an Intelli-mixer RM 2M Skyline (*ELMI Ltd, Latvia*) to prevent the beads from settling. The manufacturer's protocol recommended incubation at 4°C for 30 minutes to facilitate binding of the tagged recombinant protein to the beads, but incubation overnight resulted in better yields of eluted protein.

The next day the tube was placed in the magnetic particle separator. The supernatant was removed and stored to be used for rebinding if the recombinant protein yield was low. 1ml of MagneGST™ binding/wash buffer was added to the particles and incubated at room temperature for 5 minutes with occasional mixing. The tube was placed in the magnetic particle separator and the supernatant was removed. The washing step was repeated a total of three times. After the final wash, 50µl of elution buffer was added to the sample and incubated at room temperature on the Intelli-mixer for 15 minutes. The tube was placed in the magnetic particle separator and the eluted protein was removed to a clean tube. A second elution was carried out and the beads were then stripped of protein by adding 50µl of 1% SDS in binding/wash buffer. To each 50µl elution volume, 13µl 5x suspension solution, 1.5µl β-mercaptoethanol and 1.7µl bromophenol blue solution was added. The samples were boiled for 2 minutes and stored at -20°C prior to electrophoresis (*section 2.7.7*).

The purification of histidine (His)-tagged proteins was carried out using HIS-Select™ Magnetic Agarose Beads (*Sigma-Aldrich Inc, USA*). The beads consist of nickel bound to a chelate, which is covalently attached to magnetic 6% beaded agarose via a linker. The system exploits the high metal affinity of histidine. On addition of the particles to a soluble protein fraction containing His-tagged proteins, binding of the amino acid tag to the immobilised nickel occurs. Using a magnetic particle separator, the particles are rinsed and the bound fusion protein is eluted with a buffer containing a high concentration of imidazole – a histidine analog – which competes with the nickel-coated particles for binding of the histidine fusion proteins.

The manufacturer's protocol was followed. The magnetic bead suspension was mixed until uniformly suspended. 20µl aliquots were then transferred into Eppendorf tubes. 20µl of particles have a binding capacity of 150µg fusion protein. The beads were equilibrated in 250µl Equilibration/Wash buffer by washing three times. The supernatant containing the soluble protein fraction (*section 2.7.2*) was added to the beads along with 1µl Protease Inhibitor Cocktail Set III (*Novagen, USA*). The solution was incubated at 4°C overnight with gentle mixing on the Intelli-mixer. The manufacturer's protocol recommended incubation at 4°C for 30 minutes to facilitate binding of the tagged recombinant protein to the beads, but incubation overnight resulted in better yields of eluted protein.

The next day the tubes were placed in the magnetic separator and the same elution protocol for the GST-tagged proteins was followed, using buffers specific for the HIS-Select™ system. Table 5 illustrates the complete set of protein samples gathered at each stage of extraction and purification.

TABLE 5: Protein samples collected during extraction and purification

PROTEIN SAMPLE	TOTAL VOLUME (ml)	ALIQUOT VOLUME & SOURCE
Total protein, induced	20 (*E. coli* culture)	150µl of cell culture from Overnight Express™ Autoinduction System overnight grow
Total protein, uninduced	20 (*E. coli* culture)	150µl of cell culture from LB overnight grow
Total insoluble protein	1	150µl protein pellet resuspended in TE buffer after BugBuster™ treatment
Total soluble protein	1	150µl of protein after BugBuster™ treatment
Unbound fraction	1	150µl of supernatant after overnight binding
1st Elution	0.05	50µl of eluted protein
2nd Elution	0.05	50µl of eluted protein
Stripped beads	0.05	50µl of bound protein

2.7.4) Inclusion body purification

In the event that the expressed protein was insoluble, inclusion body purification was carried out using BugBuster™ Protein Extraction Reagent (*Novagen, USA*). Washing the insoluble pellet containing the inclusion bodies with a 1:10 diluted solution of BugBuster™, which contains non-ionic detergents, effectively purifies the pellet and removes any contaminating proteins.

The manufacturer's protocol was followed. Protein was extracted as described in section 2.7.2. The insoluble cell debris – containing inclusion bodies – was pelleted by centrifugation (*Centrifuge 5415; Eppendorf, Germany*) at 18 000x g for 20 minutes. The pellet was washed in 1ml of BugBuster™. The mixture was pipetted and vortexed to obtain an even suspension. Thorough resuspension is crucial to obtaining a high purity inclusion body preparation, free of contaminating proteins. A final concentration of 1KU/ml rLysozyme™ (*Novagen, USA*) was added from a stock of 130KU/µl (i.e 7.7µl of rLysozyme™ stock was added to 1ml of the BugBuster™ suspension.) The solution was mixed by gentle vortexing and incubated at room temperature for 5 minutes. 6 volumes of a 1:10 diluted BugBuster™ reagent was added to the suspension and mixed by vortexing for 1 minute. The suspension was centrifuged at 18 000x g for 20 minutes at 4°C to collect the inclusion bodies. The supernatant was removed and the inclusion bodies were resuspended in half the original culture volume (i.e. for 20ml grow, resuspend in 10ml) of 1:10 diluted BugBuster™ reagent. The mixture was vortexed and centrifuged (*Beckman model J2-21 centrifuge; Beckman Coulter, USA*) at 3 500x g for 20 minutes at 4°C. This step was repeated twice. The supernatant was removed and the pellet of purified inclusion bodies was stored at -70°C.

2.7.5) Inclusion body solubilisation

High molar concentrations of guanidine hydrochloride were used to denature the inclusion bodies, thus solubilising the recombinant protein. The purified inclusion bodies from a 20ml culture were defrosted, resuspended in 2ml 6M guanidine hydrochloride buffer and incubated on a shaker at 4°C overnight. Once fully dissolved, the solution was transferred to an Eppendorf tube and centrifuged (*Centrifuge 5415R; Eppendorf, Germany*) at 30 000x g for 20 minutes. The supernatant containing the denatured recombinant protein was carefully transferred to a fresh Eppendorf tube. A_{280nm} values were obtained from analysis of the supernatant using a Du® -65 Spectrophotometer (*Beckman, UK*). Samples were calibrated against an aliquot of 6M guanidine hydrochloride buffer. Protein concentration was estimated using the following equation (*Simonian & Smith, 2003)*:

$$A_{280nm} = ec\mathit{l}$$

$$\textit{Therefore } c = \frac{A_{280nm}}{e\mathit{l}} \text{ x dilution factor}$$

Where e = protein extinction coefficient ($M^{-1}.cm^{-1}$)

c = concentration (mol/L)

l = light path length (cm^{-1})

The protein extinction coefficient was obtained using the Recombinant Protein Solubility Prediction tool *(University of Oklahoma, School of Chemical Engineering and Materials Science, http://www.biotech.ou.edu/).* The recombinant protein was stored in 100µl aliquots at -20°C.

2.7.6) Recombinant protein refolding

Following denaturation, recombinant proteins were dialysed against a native buffer to remove the guanidine hydrochloride and allow renaturation and refolding of the protein, thus restoring functionality. A 100µl aliquot of denatured protein was thawed and an equal volume of 10mM DTT was added, reducing the guanidine hydrochloride buffer concentration to 3M. The solution was incubated on a shaker at 4°C for 1 hour to ensure that all disulphide bonds were reduced. 20µl aliquots of 1x refolding buffer were added to the sample on ice every 12 minutes over the period of an hour. This decreased the guanidine hydrochloride buffer concentration to 2M. 100µl of protein sample was pipetted into a Slide-A-Lyzer® mini dialysis unit (*Pierce, USA)* and placed inside an Eppendorf tube filled with 1ml of refolding buffer. The sample was dialysed at room temperature for 5 minutes. This process was repeated 10 times, which was the maximum amount of dialysis before recombinant adaptin precipitated. The guanidine hydrochloride was completely removed from the recombinant protein solution by this time. The recombinant protein was purified from the solution using HIS-Select™ Magnetic Agarose Beads (*Sigma-Aldrich Inc, USA*) as described in section 2.7.3.

2.7.7) Laemmli SDS-PAGE

Protein samples were resolved using the Laemmli SDS-PAGE method (*Laemmli, 1970*). A 12% polyacrylamide SDS separating gel was prepared and overlaid with a 4% stacking gel in a Mighty Small II SE250 gel cassette (*Hoefer Scientific Instruments, USA*) that was 8x10cm in size. Volumes of protein sample loaded were as follows: total cell extract and insoluble fraction - 10µl; soluble fraction and eluted protein fraction - 20µl; unbound fraction - 20µl. 5µl of red blood cell membrane proteins (4-8µg) was used as a marker to estimate the relative size of proteins. The marker was prepared from human red blood cells by Kubendran Naidoo, Plasmodium Molecular Research Unit, Wits Medical School.

The protein samples underwent separation at 15mA per gel for 1 hour, followed by 20mA per gel for a further hour, using a Mighty Slim™ SX250 power supply (*Hoefer Scientific Instruments, USA*). During electrophoresis, gels were cooled to 4°C using a Labcon CPE 50 circulator (*Labcon, South Africa*). Gels were stained in 0.05% Coomassie blue stain overnight and destained in 10% acetic acid/10% methanol solution for 2 hours. The gels were destained further in 10% acetic acid overnight.

2.7.8) Western blotting

Following electrophoresis, two pieces of blotting paper and a piece of Hybond™-c extra supported nitrocellulose membranes (*Amersham Biosciences, UK*) were soaked in transblot buffer along with the polyacrylamide gels. The blotting cassettes (*Hoefer Scientific Instruments, USA*) were assembled with the gel on the cathode side and the nitrocellulose membrane on the side of the anode. The cassette was placed in the Western blotting chamber (*TE Series Transphor electrophoresis unit; Hoefer Scientific Instruments, USA)* which was filled with 1.5L of transblot buffer. The proteins were transferred at 35V, 90mA (*Model 200/2.0 power supply; Biorad, USA*) at 4°C overnight on a magnetic stirrer.

2.7.9) Immunoblotting

The nitrocellulose membranes were washed in TBS for 5 minutes and the gels were placed in Coomassie blue stain to ascertain transfer efficiency. The blots were drained and coated for 1 minute in 1ml Ponceau S (*Sigma-Aldrich Inc, USA*) to detect protein transfer. The blots were placed in Milli-Q water to rinse off excess dye. The banding of the RBC membrane marker was recorded on the membrane, which was returned to the water for 30 minutes to completely remove the Ponceau S stain.
For genes that were subcloned into the pGEX-4T-2 vector the recombinant protein contained a GST tag and the following Amersham Biosciences (*UK*) protocol was followed to detect the recombinant protein:

Once the membrane was sufficiently rinsed it was placed in a 3% BSA/TBS solution on a shaking platform for 1 hour to block the areas of the nitrocellulose membrane that did not contain transferred protein. This would prevent the antibodies from binding to the membrane due to its positive charge. The blot was washed five times in 0.5% Tween-TBS solution for 5 minutes each to remove any unbound protein, followed by a rinse in TBS for 5 minutes. The membrane was incubated in a 1:100 000 dilution of anti-GST horse radish peroxidise (HRP)-conjugated primary antibody (*Amersham Biosciences, UK*) for 1 hour at room temperature. The membrane was washed five times in 0.5% Tween-TBS solution for 5 minutes each to remove any unbound antibody. This was followed by a rinse in TBS for 5 minutes.

For the gene that was subcloned into the pET-15b vector the recombinant protein contained a histidine tag and the following Qiagen (*UK*) protocol was followed: Once the membrane was sufficiently rinsed it was placed in anti-His HRP-conjugated blocking buffer (*Qiagen, UK*) on a shaking platform for 1 hour to block the areas of the nitrocellulose membrane that did not contain transferred protein. The blot was washed twice in 0.5% Tween-TBS solution for 10 minutes each to remove any unbound protein, followed by a rinse in TBS for 10 minutes. The membrane was incubated in a 1: 2 000 dilution of QIAexpress anti-His HRP-conjugated antibody (*Qiagen, UK*) for 1 hour at room temperature. The membrane was washed twice in 0.5% Tween-TBS solution for 10 minutes each to remove any unbound antibody. This was followed by a rinse in TBS for 10 minutes.

2.7.10) Chemiluminescence

GST- and His-tagged proteins immobilised on the nitrocellulose membrane and complexed with HRP-conjugated antibodies were visualised using luminescent substrates (*Sasse & Gallagher, 2003*) in the SuperSignal® West Pico chemiluminescent substrate kit (*Pierce, USA)*. On addition of the peroxide substrate and luminol, the HRP enzyme catalyses the oxidation of the latter, thereby releasing energy in the form of light. Exposing x-ray film to the membrane records the luminescent band.

The manufacturer's protocol was followed. Operating in a darkroom, the working solution was prepared by adding equal parts of stable peroxide solution and luminol/enhancer solution provided in the kit. The membrane was incubated in working solution for 5 minutes. The excess solution was drained off the membrane, which was covered with clear plastic. The blot was placed in an x-ray film cassette with the protein side facing upward. A piece of CP-G plus medical x-ray film (*Agfa, Germany*) was placed over the blot and exposed for 60 seconds. Subsequent pieces of film were exposed for 5 minutes and 10 minutes, respectively. The film was developed using manual x-ray developer diluted 1:4 with water (*Axim, South Africa*). It was rinsed in water and finally fixed in Perfix high speed x-ray fixer (*Champion photochemistry, South Africa*) diluted 1:4 with water, followed by drying at room temperature for 30 minutes. The nitrocellulose membrane was placed in amido black stain (*AppliChem, Germany*) to detect the proteins on the membrane. To remove background dye, the membrane was destained in amido black destain.

2.8) PK assay

2.8.1) Determination of recombinant PK concentration

The Coomassie Plus – The Better Bradford™ Assay Kit (*Pierce, USA*) was used to determine the recombinant kinase concentration. The kit consists of Coomassie Plus Reagent and bovine serum albumin (BSA) standard ampoules containing 1ml each of 2mg/ml BSA stock solution. The method is based on the spectrophotometric shift from 465nm to 595nm that occurs when Coomassie blue dye adsorbs to arginine, histidine and the aromatic amino acids making up the protein. The absorbance data can be plotted against the known BSA concentrations to construct a standard curve from which unknown sample concentrations can be extrapolated.
The manufacturer's procedure was followed. Diluted BSA standards were prepared as follows in Table 6.

TABLE 6: Preparation of diluted BSA standards

SAMPLE	VOL. OF MagneGST™ ELUTION BUFFER (µl)	VOLUME OF BSA (µl)	AMOUNT OF BSA (µg)
Blank	10	0	0
Standard 1	10	1	2
Standard 2	10	2	4
Standard 3	10	4	8
Standard 4	10	6	12
Standard 5	10	8	16

10µl of recombinant PK was transferred to an Eppendorf tube. The Coomassie Plus reagent was mixed by inversion immediately before use. The amount needed was transferred to a test tube and equilibrated to room temperature. 1.5ml of Coomassie Plus reagent was added to each sample and the solution was mixed well. The samples were incubated at room temperature for 10 minutes. The Du® - 65 Spectrophotometer (*Beckman, UK*) was set to 595nm and calibrated against an aliquot of MagneGST™ elution buffer. The A_{595} of all samples was measured. A standard plot of A_{595} to standard protein amount (µg) was constructed. The amount of recombinant PK in the unknown sample was extrapolated from the graph.

2.8.2) Kinase assay

Recombinant PK was tested for activity using a standard kinase assay (*Sefton & Shenolikar, 1996*; *Wang & Roach, 1993*). The method is based on the principle of assaying the PK by the transfer of radiolabel from 2.5µCi [γ-^{32}P] ATP (*PerkinElmer, USA*), with a specific activity of 3 000Ci/mmol, to the following exogenous substrates: 10µg histone protein 1 (*Merck, Germany*), 5µg myelin basic protein (*Sigma-Aldrich Inc, USA*) and 2µg bovine casein (*Merck, Germany*). Sodium fluoride and β-glycerophosphate (*Calbiochem, Germany*) were added to inhibit phosphatases. Sodium fluoride inhibits serine/threonine phosphatases, while β-glycerophosphate inhibits tyrosine phosphatases (*Schaefer et al, 1996*).

The assay was carried out in a reaction volume of 30µl, consisting of 0.5 - 3µg recombinant protein in a kinase reaction mix. The Eppendorf containing the reaction solutions was incubated at 30°C for 45 minutes and the reaction was terminated by addition of 8µl 5x suspension solution, 0.8µl β-mercaptoethanol and 1µl bromophenol blue solution. The solution was boiled for 5 minutes. The samples were loaded onto a 12% SDS polyacrylamide gel and electrophoresed, followed by staining in Coomassie blue overnight (*section 2.7.7*). The gel was destained and soaked in gel drying solution for 2 hours. The gel was dried in a Model 224 Gel Slab Dryer (*Biorad, USA*) for 2 hours at 80°C, which was attached to a Vacuum Pump XF54 230 50 (*Millipore Corporation, USA*) with a maximum vacuum of 84KPa. The dried gel was transferred to an x-ray cassette with intensifying screens and exposed to CP-G plus medical x-ray film (*Agfa, Germany*) for 12 hours at -70°C. The film was developed as for chemiluminescence (*section 2.7.10*).

2.8.3) Densitometric analysis

Densitometric analysis was performed on Coomassie-stained SDS-PAGE gels using a GS300 transmittance/reflectance scanning densitometer (*Hoefer, USA*). The data were integrated using the Hoefer Electrophoresis data reduction software and presented as recombinant protein peak areas.

2.9) Bioinformatics

The solubility of recombinant proteins was determined using a recombinant protein solubility prediction tool designed by the University of Oklahoma, School of Chemical Engineering and Materials Science (*www.biotech.ou.edu*). The ExPasy ProtParam tool was used to obtain the molecular weight and pI of recombinant proteins (*http://au.expasy.org*). Sequence alignments were carried out using the EMBOSS Pairwise Alignment Algorithm tool of the European Bioinformatics Institute *(www.ebi.ac.uk).* The protein secondary structure was determined using the Jnet secondary structure prediction algorithm (*Cuff and Barton, 2000*).

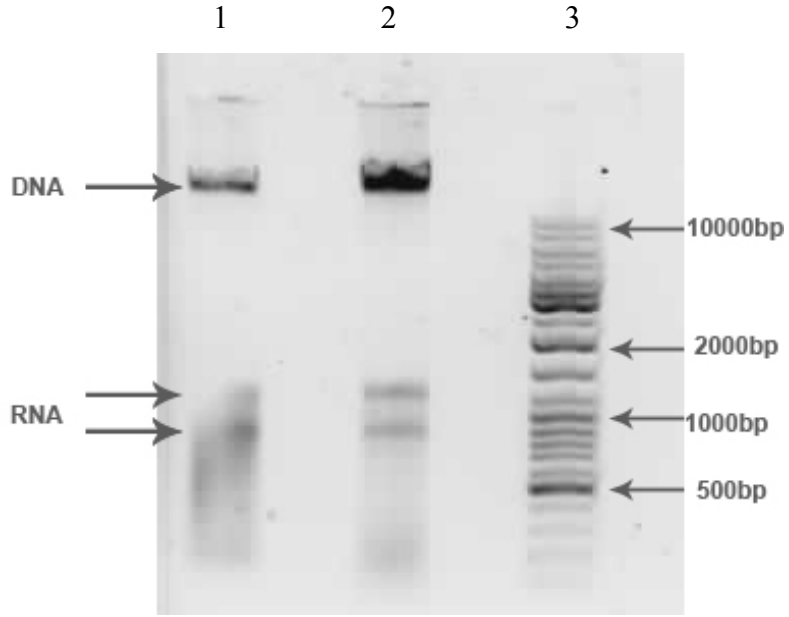
Using the Swiss Model Template Library (*http://swissmodel.expasy.org*), three dimensional models of the *P. falciparum* protein domains, without the recombinant tags, were predicted (*Arnold et al, 2006; Kopp & Schwede, 2004; Schwede et al, 2003; Guex and Peitsch, 1997).* This comparative homology modeling server reliably generates a 3D model of the target from its amino acid sequence, provided that a template sharing sequence identities of more than 40 percent with the target is present in the library. If target-template pairs have less than 40 percent identity, errors in sequence alignment algorithms may occur. Homology of *P. falciparum* proteins with other *Plasmodium* species proteins was achieved using the Basic Local Alignment Search Tool (BLAST) on PlasmoDB (*www.plasmoDB. org*). *T gondii* proteins with homology were found using the BLAST on ToxoDB (*www.toxoDB.org*). Homologous *H. sapiens* proteins were discovered using the Ensembl Multi-BLAST View tool (*www.ensembl.org*)

CHAPTER 3

RESULTS

3.1) DNA extraction

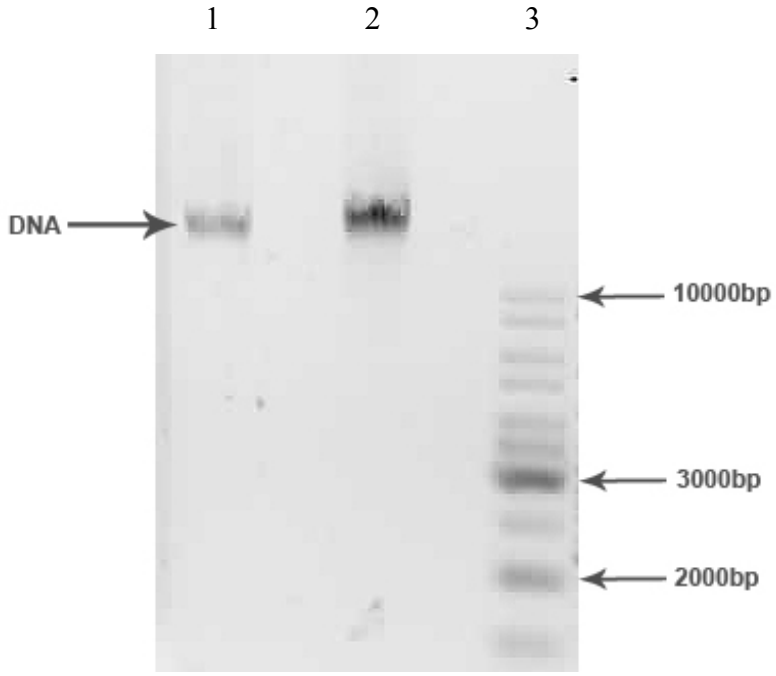
The *P. falciparum* DNA samples were electrophoresed and the integrity was good with a single high molecular weight band. However, some RNA coprecipitated with the DNA (*Figure 8*), resulting in A_{260}/A_{280} values greater than 2. Pure DNA has an A_{260}/A_{280} value between 1.8 and 2. If this ratio is less than 1.8, the preparation is contaminated with proteins and aromatic substances (*Hoffmann-Rohrer & Kruchen, 2006*).



1 – DNA 1
2 – DNA 2
3 – MassRuler™ DNA ladder mix

Figure 8. Agarose gel electrophoresis of P. falciparum DNA

To remedy the RNA contamination, DNA was treated with RNase A. This resulted in improved A_{260}/A_{280} values (*Table 7)*. To confirm that the RNA had been removed and the DNA had retained structural integrity, electrophoresis was repeated (*Figure 9).*



1 – DNA 1
2 – DNA 2
3 – MassRuler^{TM} DNA ladder mix

Figure 9. Agarose gel of P. falciparum DNA treated with RNase A

Of the two DNA samples obtained after RNase A treatment, DNA 1 (*Table 7*) had the best purity, but a low concentration. The low recovery of DNA was probably due to the short duration (5 minutes) of centrifugation during precipitation (*Zeugin & Hartley, 1985*). The second sample, DNA2, was centrifuged for 30 minutes at the precipitation stage, which dramatically improved the yield (*Zeugin & Hartley, 1985*). DNA 2 was slightly less pure, but was used for all PCR experiments due to the higher concentration.

TABLE 7: Spectrophotometric data for *P. falciparum* genomic DNA

SAMPLE	CONCENTRATION (ng/µl)	A_{260}/A_{280}
DNA 1	201.29	2.14
DNA 2	199.13	2.06
DNA 1 (RNase A-treated)	42.8	1.81
DNA 2 (RNase A-treated)	161.56	1.74

3.2) RNA extraction

RNA integrity was determined using agarose gel electrophoresis and spectrophotometric data. Intact *P. falciparum* ribosomal RNA resolves as three distinct bands of 28s, 18s and 5.8s, with sizes of 4 104 bases, 2 149 bases and 121 bases, respectively (*www.plasmo-DB.org, 2008; Daily et al, 2004*), as seen in figure 10. RNA was considered to be degraded if the expected bands appeared as a smear within the gel.

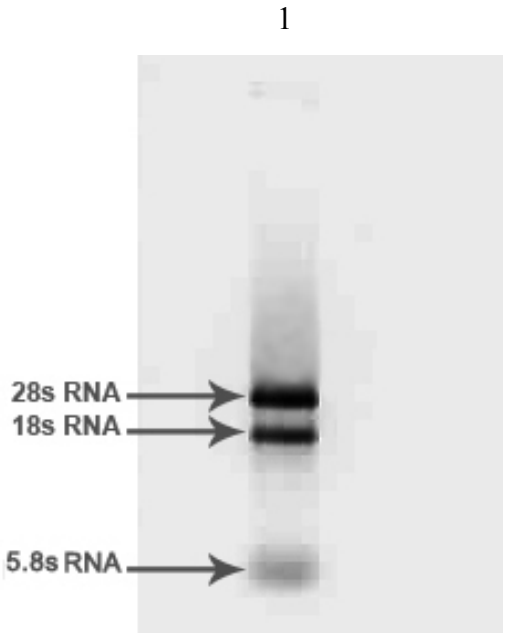


Figure 10. Agarose gel electrophoresis of P. falciparum RNA

The RNA showed good integrity with minimal smearing. No RNA marker was available and sizes are presumptive based on information from Daily et al (2004). 28s RNA has a size of 4 104 bases. 18s RNA is 2 149 bases and 5.8s RNA is 121 bases in length.

An A_{260}/A_{280} value greater than or equal to 2 for RNA indicates a pure sample, while a value less than 2 indicates carry over of protein or phenol (*Hoffmann-Rohrer & Kruchen, 2006*). Sample RNA 1 (*Table 8*) was used for RT-PCR experiments.

TABLE 8: Spectrophotometric data for *P. falciparum* RNA

SAMPLE	CONCENTRATION (ng/µl)	A_{260}/A_{280}
RNA 1	323.57	2.14

3.3) PCR

3.3.1) Primer design

3.3.1.1) *PFB0150c*

Primers were designed for three domains from the two *P. falciparum* genes coding for putative proteins. In the case of *PFB0150c* – the putative PK – the entire gene was too large to amplify and express, so only the PK domain was chosen for subcloning. This functional region of the enzyme containing the catalytic core is 1 515bp in length and spans nucleotides 149 527 to 151 042 (*Figure 11*).

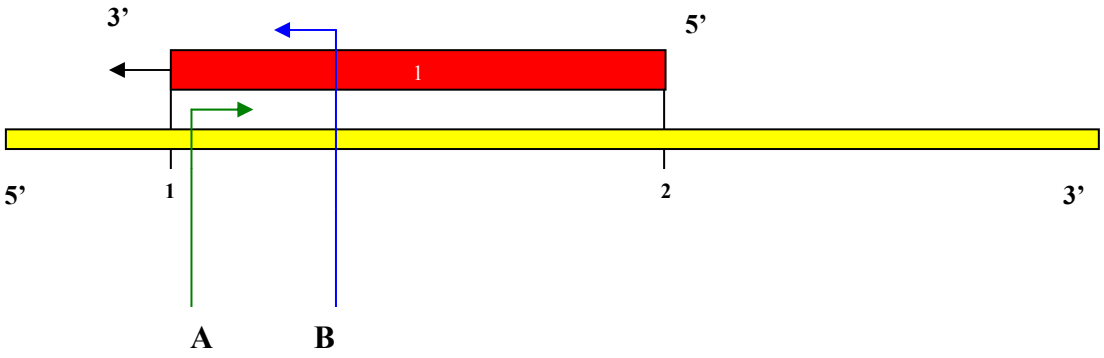


Figure 11. PFB0150c diagramatic representation of primer positions

P. falciparum chromosome 2 is represented by the yellow bar. The PFB0150c gene is represented in red and has Crick orientation. It consists of a single exon 7 448bp in length. The gene codes for a 352.8kD protein with a pI of 7.21. The forward primer is green, while the reverse primer is blue. Base pair positions on chromosome 2 are as follows: 1 - 149 524; 2 - 156 971; A - 149 527; B - 151 042

The DNA sequence of the insert chosen for amplification is represented in figure 12. The PK domain is central to the amplicon.

GATGAAAAGGATGGATATGAAGAAATGAATGGGGGAGATAAGAATGAAGAAATGAATGGGGGAGATAAGAATGAAGA
AATGAATGTGGGAGATAAAAATGGAGGAATAAATGAGGAACATAAAAATGAAGGAATAAATGAGGAACATAAGGATG
AACTAATAAATAAGGAACATAAAAACGAGCGAATAAATGAGGAACATAAAAACGAACGAATAAATGAGGAACATAAA
AATGAAGGAATAAATGAGGAACATAAAAATGAAGGAATAAATGAGGAACATAAAAACGAACGAATAAATGAGGAACA
TAAAAATGAAGGAATAAATAAACTGACCTATCATAATATGAATAAAAATAATATTTCAAATGAAAATAATTATAATG
ATGACGATTCTTATGATGAAGATAATTTGGTATCCCTGAAGATAATAAACTTAAAATATTTAAGTAAAAAGAATAGT
TTAAAAAACATTTTGAGAGAAGTAAATTTTTTAAAAATGTGTGAACATCCAAATGTAGTAAAATATTTCGAATCTTT
TTTTTGGCCTCCTTGTTATTTAGTTATTGTGTGTGAATATTTATCAGGAGGAACATTATATGATTTATATAAAAATT
ATGGTAGAATATCAGAAGATCTTTTAGTATATATCTTAGATGATGTATTAAATGGTTTAAATTATTTACATAATGAA
TGTAGTTCACCACTTATACATAGAGATATAAAACCAACAAATATCGTTCTTTCCAAAGATGGTATAGCTAAGATAAT
TGATTTTGGTTCTTGTGAAGAATTGAAAAATAGTGATCAGTCTAAAGAATTAGTGGGTACTATATATTATATATCAC
CTGAAATATTGATGAGAACTAATTATGATTGTTCATCTGATATATGGTCATTGGGTATTACAATATATGAAATTGTT
TTATGTACCTTACCATGGAAAAGAAATCAATCATTTGAAAATTATATAAAAACCATAATTAATTCATCACCAAAAAT
TAACATAACAGAAGGATATAGTAAACACTTATGTTATTTTGTTGAGAAGTGTTTACAAAAGAAACCTGAGAACAGAG
GAAATGTGAAAGATTTATTAAATCATAAATTTTTGATTAAAAAGAGGTATATTAAAAAGAAACCTAGTTCTATATAT
GAAATAAGAGATATATTAAAAATATATAATGGTAAAGGTAAAACAAATATCTTCCGAAATTTTTTTAAGAACCTTTT
TTTCTTCAATGATAAGAATAAAAAAAAAAAACCAAATAAAATGATCAGTTCCAAATCCTGTGATGCAGAAATGTTCT
TTGAACAGTTAAAAAGGGAAAATTTTGATTTTTTTGAAATTAAATTAAAAGATGATGAAAATAGTAGATCCTTGAAT
ACGTTTAATATAAATATATCTAAAGAAAGAGACGACATATCATATTCTTCTTTAAATTTGGAAAAAATCAAAGAACA
CAGTCTCAATATGGTAGCATCTGTTGTCGGGACTGAACAATCCCAGAAA

Figure 12. DNA sequence of PFB0150c insert (PlasmoDB 5.3, 2007)

Primer positions are indicated in blue and start at base pair position 149 527 of chromosome 2, ending at base pair 151 042; the red, underlined segment denotes the PK domain.

The protein sequence was predicted to be most soluble with a GST tag, as opposed to a His tag, and was annotated rGST PK (*Figure 13*). A 63.2 percent chance of insolubility was predicted for this recombinant protein, along with a molecular weight of 84.8kD. The calculated pI of the rGST PK insert is 5.88.

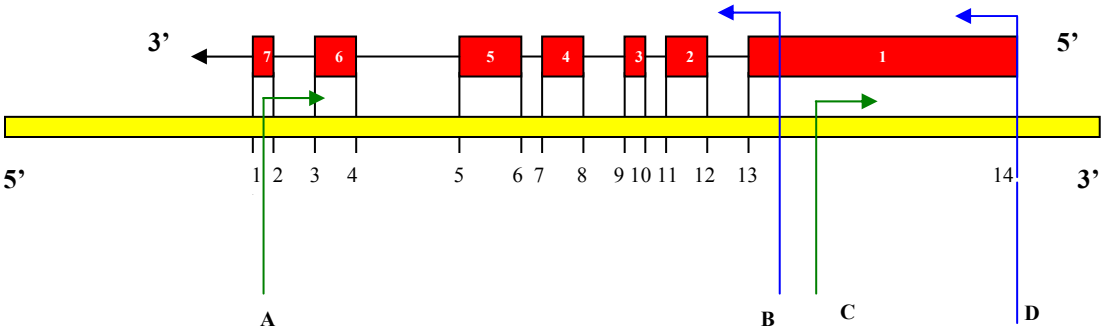
MSPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGDVKLTQSMAIIRYIAD
KHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFM
LYDALDVVLYMDPMCLDAFPKLVCFKKRIEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDEKDGYEEMNGG
DKNEEMNGGDKNEEMNVGDKNGGINEEHKNEGINEEHKDELINKEHKNERINEEHKNERINEEHKNEGINEEHKNEG
INEEHKNERINEEHKNEGINKLTYHNMNKNNISNENNYNDDDSYDEDNLVSLKIINLKYLSKKNSLKNILREVNFLK
MCEHPNVVKYFESFFWPPCYLVIVCEYLSGGTLYDLYKNYGRISEDLLVYILDDVLNGLNYLHNECSSPLIHRDIKP
TNIVLSKDGIAKIIDFGSCEELKNSDQSKELVGTIYYISPEILMRTNYDCSSDIWSLGITIYEIVLCTLPWKRNQSF
ENYIKTIINSSPKINITEGYSKHLCYFVEKCLQKKPENRGNVKDLLNHKFLIKKRYIKKKPSSIYEIRDILKIYNGK
GKTNIFRNFFKNLFFFNDKNKKKKPNKMISSKSCDAEMFFEQLKRENFDFFEIKLKDDENSRSLNTFNINISKERDD
ISYSSLNLEKIKEHSLNMVASVVGTEQSQK

Figure 13. Protein sequence of rGST PK

The entire PK protein is 2 485 amino acids in length. The rGST PK insert is 723 amino acids in length. It consists of the GST tag, (green) which has 220 amino acids, and the P. falciparum protein which has 503 amino acids including the PK domain (red) which spans amino acids 97-372. The amino acid composition of the P. falciparum sequence is: Alanine (A) 1.8%; Arginine (R) 3.3%; Asparagine (N) 9.4%; Aspartic Acid (D) 6.5%; Cysteine (C) 1.9%; Glutamine (Q) 1.5%; Glutamic Acid (E) 10.1%; Glycine (G) 5.4%; Histidine (H) 2.9%; Isoleucine (I) 8.3%; Leucine (L) 9.3%; Lysine (K) 11.2%; Methionine (M) 2.5%; Phenylalanine (F) 4.0%; Proline (P) 3.3%; Serine (S) 6.2%; Threonine (T) 2.5%; Tryptophan (W) 1%; Tyrosine (Y) 5.3%; Valine (V) 3.6%.

3.3.1.2) *PFE1400c*

Primers were designed for the two domains of interest on the *PFE1400c* gene, namely the clathrin adaptor appendage domain and the adaptin N-terminal domain. The clathrin adaptor appendage domain primers delineated an amplicon 723bp in length (*Figure 14, A and B*), spanning nucleotides 1 157 840 to 1 159 584. As the sequence included six introns, the primers were designed within exons one and seven and the insert was amplified from reverse transcribed RNA. The *PFE1400c* N-terminal adaptin domain primers defined an amplicon spanning nucleotides 1 159 748 to 1 161 547, resulting in a 1 800bp-sized product (*Figure 14, C and D*). The insert was amplified directly from DNA.



Base pair positions on chromosome 5:

1 – 1 157 834	**10 – 1 159 065**
2 – 1 157 893	**11 – 1 159 195**
3 – 1 158 038	**12 – 1 159 276**
4 – 1 158 226	**13 – 1 159 393**
5 – 1 158 334	**14 – 1 161 547**
6 – 1 158 608	**A – 1 157 840**
7 – 1 158 744	**B – 1 159 584**
8 – 1 158 814	**C – 1 159 748**
9 – 1 159 004	**D – 1 161 547**

Figure 14. PFE1400c diagramatic representation of primer positions

P. falciparum chromosome 5 is represented by the yellow bar. The PFE1400c gene is represented in red and has Crick orientation. It consists of seven exons. The gene is 2 778bp in size, including introns, and codes for a 105.6kD protein with a pI of 5.83. The forward primers are green, while the reverse primers are blue.

The DNA sequence of the clathrin adaptor appendage domain chosen for amplification is represented in figure 15.

TCTTCGGATGAATTCAATAATGATATCGATGATGCAGATGATAGTAAAAAATCTATGGACTTGATAGGATTAAATGA
TGATGAAAGCAAACCCCAAAAAACAATTCCACCGGTCAAAATGGTTCAGGTATTATCATCAGAGGATGCTGGTTTGA
AGGGACAGACGGGTCTCTCCATTTTTGCCTCAATAAATCGTATTGACAGAAAAATTCAACTCAAAATATCAGTAACA
AATCAAACCCAGAACGAAATCGTAGTATCAGGAGTTCAAATAAATAAAAATTCATTTGGATTATCGTCACCGAACAA
TTTGGATGTACAGAATATTGGCTTTGGAGAAACAAAAGAAATGCTCATATATTTAATTCCAAATACGTTAAATTCGA
ATACACCTCCAGCAACCCCCTTATTTTTACAAGTTGCCATAAGAACAAATTTAGATATATTTTATTTCAATGTACCG
TACGACATTTTTGTCGTTTTTGTCGAAAATTTTCATATGGAAAAAGACATATTTAAAAAGAAATGGCAAATTATAGA
AGAAGCAAAAGAGGAATTATATTATTTTGCTTGTATAACCACAAACAATCTAGTAATATTAAGTGAAGTCACAATTC
AACCTGAGAAAAAGAACGTGAAATTGTGCATTCGTACAGATTCCTCATCTGTGATACCACTTTATAAGCTTTTATTT
GTTAAGGCA**TTTTCCTTAAGTGTAACTCAG**

Figure 15. DNA sequence of PFE1400c clathrin adaptor appendage domain, excluding introns (PlasmoDB 5.3, 2007)

A 723bp DNA sequence was amplified from cDNA. Primer positions are indicated in blue, with the forward primer starting at base pair number 1 157 840 on chromosome 5, and the reverse primer ending at base pair 1 159 584.

The protein sequence translated from this amplicon was predicted as most soluble with a GST tag, and was annotated rGST AP1 C-terminal (*Figure 16*). rGST AP1 C-terminal has a predicted insolubility of 56.9 percent and the molecular weight is 53.2kD. The calculated pI is 5.4.

MSPILGYWKIKGLVQPTRLLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGDVKLTQSMAIIRYIAD
KHNMLGGCPKERAEISMLEGAVLDIRYGVSRIAYSKDFETLKVDFLSKLPEMLKMFEDRLCHKTYLNGDHVTHPDFM
LYDALDVVLYMDPMCLDAFPKLVCFKKRIEAIPQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDNNSSDEFNNDI
DDADDSKKSMDLIGLNDDESKPQKTIPPVKMVQVLSSEDAGLKGQTGLSIFASINRIDRKIQLKISVTNQTQNEIVV
SGVQINKNSFGLSSPNNLDVQNIGFGETKEMLIYLIPNTLNSNTPPATPLFLQVAIRTNLDIFYFNVPYDIFVVFVE
NFHMEKDIFKKKWQIIEEAKEELYYFACITTNNLVILSEVTIQPEKKNVKLCIRTDSSSVIPLYKLLFVKAFSLSVT
QT

Figure 16. Protein sequence of rGST AP1 C-terminal

The entire AP-1 β complex is 925 amino acids in length. The AP1 C-terminal protein sequence is 464 amino acids in length. It consists of the 220 amino acid GST tag (green) and the parasite AP1 C-terminal of 244 amino acids. The amino acid composition of the P. falciparum sequence is: Alanine (A) 3.9%; Arginine (R) 2.8%; Asparagine (N) 5.6%; Aspartic acid (D) 7.3%; Cysteine (C) 1.3%; Glutamine (Q) 3.7%; Glutamic acid (E) 6.5%; Glycine (G) 4.7%; Histidine (H) 1.5%; Isoleucine (I) 8.0%; Leucine (L) 10.6%; Lysine (K) 8.8%; Methionine (M) 2.8%; Phenylalanine (F) 5.0%; Proline (P) 5.2%; Serine (S) 6.5%; Threonine (T) 4.5%; Tryptophan (W) 1%; Tyrosine (Y) 4.3%; Valine (V) 6.0%.

The DNA sequence of the adaptin N-terminal domain chosen for amplification is represented in figure 17.

ATGTCTGATTTACGCTACTTTCAAACGACTAAAAAAGGAGAAATCCATGAGCTTAAAGAA
GAATTACATTCTTCTCATAAGGAAAAAAAAAAGGAAGCTATTAAAAAGATTATTGCTGCT
ATGACTGTTGGGAAAGATGTCTCGACATTATTTTCTGATGTTGTGAATTGCATGCAAACA
TCGAATATAGAATTAAAGAAGCTTGTTTATTTATACGTTATAAATTATGCTAAGGTACAA
CCTGAACTAGCTATTTTAGCTGTTAATACATTTCGAAAAGATTCCTCAGACCCAAATCCC
CTTATAAGAGCTTTAGCTATTCGAACAATGGGATGTATACGACTAGAACAAATAACAGAA
TATTTAATCGAACCATTGAGGAGATGTTTAAAAGATGAAGATCCATACGTAAGAAAAACA
GCTGTTATATGTATAGCCAAATTATATGATATATCACCGAAATTAGTAGAAGAAGAAGGT
TTTATAGATACGCTTTTAGATATATTAGATGATAATAATGCCATGGTTGTTGCTAATGCT
GTTATATCGTTAACGGATATTTGTGAGAATTCAAACAAGAGCATATTAAAAGATGTTATA
AACAAAGATGAAAATAATGTGAATAAACTTTTAAATGCTATTAATGAATGTGTAGAATGG
GGACAAGTTTTTATTTTAGATGCATTAGTTTTATATGAACCTAAAACTAGTAAAGATGCT
GAACGTGTTTTAGAAAGAATATTACCAAGATTATCACATGCTAATTCAGCAGTAGTTCTA
TCATCTATTAAAGTTATCTTATGCTTATTAGATAAAATCAATGATAAAGAATTTATTAAA
AATGTACATAAGAAATTAAGCCCATCTTTAGTCACACTTTTATCTGCGGAACCAGAAATT
CAATATATTGCATTAAGAAATATTAATTTAATAACACAGAAATTACCCAACATGCTCTCT
GATAAAATCAATATGTTTTTCTGTAAATATAATGAACCTGCTTATGTAAAAATGGAAAAA
CTTGATATTATTATAAGACTTGTATCAGATAAAAATGTAGACCTTGTTCTATATGAATTA
AAAGAATATTCTACAGAAGTAGATGTCGAATTTGTTAAAAAAAGTGTAAGAGCTATAGGT
AGCTGTGCAATAAAATTACCACAATCAAGTGAAAAATGTATTAATATATTATTAGATTTA
ATAGATACCAAAATTAATTATGTTATACAAGAATGTATAGTTGTAATTAAAGATATTTTT
AGAAAATATCCAAATAAATATGAAAGTATAATAACCATTCTATGTGAAAATCTAGAATCT
TTAGATGAATCAAATGCAAAAGCATCCTTAATATGGATAATAGGAGAATATGTAGAACGT
ATTGATAATGCTGATGAATTAATAGATTCCTTTTTAGAAAATTTTAGTGATGAACCATAT
AATGTTCAATTACAAATTCTAACGGCTAGCGTCAAATTATTTTTAAAATGTTCAAAAAAC
ACAAAAGATATTATAACCAAAGTCTTAAAATTATCCACAGAAGAAAGTGATAATCCTGAT
TTAAGAGATAGGGCTTATATTTATTGGAGATTGTTATCTAAAAATATTGATGTTGCCAAA
AAAATTGTACTTGCCGATAAACCCCCAATACAAGAAGAGAATAAAATTACTGATACCAAA
GTATTAAATAAATTAATTAAAAATATATCCATGTTATCATCTGTATATCATAAATTACCA
GAAACTTTTATATCAAAAAAAAATTCCTATTCCTTAAATTCTGATAATAATAATGATCAT

Figure 17. DNA sequence of PFE1400c N-terminal adaptin domain

A 1 800bp DNA sequence was obtained from PlasmoDB version 5.3, 2007. Primer positions are indicated in blue, with the forward primer starting at position 1 159 748 on chromosome 5 and the reverse primer ending at base pair number 1 161 547.

The polypeptide derived from this amplicon was predicted to be most soluble with a histidine tag, and was annotated rHis AP1 N-terminal (*Figure 18*). rHis AP1 N-terminal has a predicted insolubility of 50.4 percent. The calculated molecular weight of the recombinant protein is 72.2kD and the pI is 6.3.

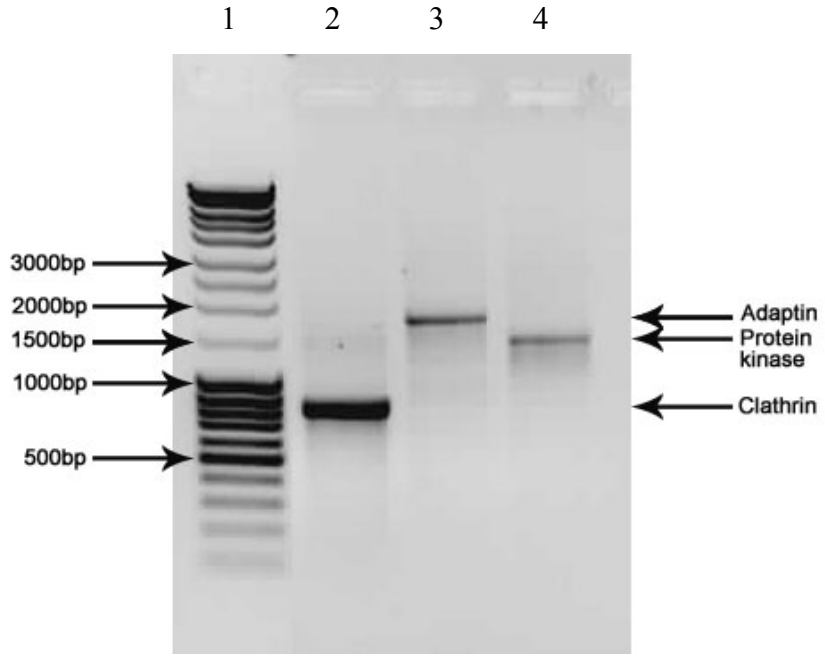
HHHHHHSSGLVPRGSMSDLRYFQTTKKGEIHELKEELHSSHKEKKKEAIKKIIAAMTVGKDVSTLFSDVVNCMQT
SNIELKKLVYLYVINYAKVQPELAILAVNTFRKDSSDPNPLIRALAIRTMGCIRLEQITEYLIEPLRRCLKDEDP
YVRKTAVICIAKLYDISPKLVEEEGFIDTLLDILDDNNAMVVANAVISLTDICENSNKSILKDVINKDENNVNKL
LNAINECVEWGQVFILDALVLYEPKTSKDAERVLERILPRLSHANSAVVLSSIKVILCLLDKINDKEFIKNVHKK
LSPSLVTLLSAEPEIQYIALRNINLITQKLPNMLSDKINMFFCKYNEPAYVKMEKLDIIIRLVSDKNVDLVLYEL
KEYSTEVDVEFVKKSVRAIGSCAIKLPQSSEKCINILLDLIDTKINYVIQECIVVIKDIFRKYPNKYESIITILC
ENLESLDESNAKASLIWIIGEYVERIDNADELIDSFLENFSDEPYNVQLQILTASVKLFLKCSKNTKDIITKVLK
LSTEESDNPDLRDRAYIYWRLLSKNIDVAKKIVLADKPPIQEENKITDTKVLNKLIKNISMLSSVYHKLPETFIS
KKNSYSLNSDNNNDHMQDDHYDDDDYDKDNHVL

Figure 18. Protein sequence of rHis AP1 N-terminal

The protein sequence is 633 amino acids in length. It consists of the His tag (green) which is 6 amino acids in length, the vector sequence (orange) which is 9 amino acids in length and the parasite AP1 N-terminal of 618 amino acids. The amino acid composition of the P. falciparum sequence is: Alanine (A) 6.0%; Arginine (R) 3.3%; Asparagine (N) 7.1%; Aspartic acid (D) 6.5%; Cysteine (C) 2.1%; Glutamine (Q) 2.1%; Glutamic acid (E) 7.9%; Glycine (G) 1.4%; Histidine (H) 2.1%; Isoleucine (I) 10.1%; Leucine (L) 11.7%; Lysine (K) 10.3%; Methionine (M) 1.4%; Phenylalanine (F) 2.2%; Proline (P) 3.2%; Serine (S) 7.6%; Threonine (T) 3.9%; Tryptophan (W) 0.5%; Tyrosine (Y) 3.5%; Valine (V) 7.1%.

3.3.2) Digestion of PCR products

PCR products were digested with their respective endonucleases to ensure that the correct ‘sticky ends’ were generated for ligation into the vectors. The amplicon sizes were estimated against a DNA ladder after restriction digestion, as observed in figure 19. The adaptin and protein kinase inserts migrated to their estimated correct positions in line with the DNA ladder markers. The clathrin insert, however, did not align as expected. Though the annotated size of clathrin is 723bp, it consistently moved to a position between the 800bp and 900bp markers. This occurred with all clathrin inserts and indicated that the sequence was longer than expected.



1 – MassRuler^{TM} DNA ladder mix
2 – Clathrin
3 – Adaptin
4 – PK

Figure 19. Digested P. falciparum PCR products resolved on 1% agarose gel

PCR products were digested with BamHI, and either XhoI or NdeI restriction endonucleases and purified, followed by electrophoresis. The sizes of the PCR products are as follows: clathrin – 723bp; PK – 1 515bp; adaptin – 1 800bp.

The concentration and purity of digested amplicons are shown in table 9. The low purity of the PK amplicon sample may have been due to contamination with proteins, salt, ethanol or phenol.

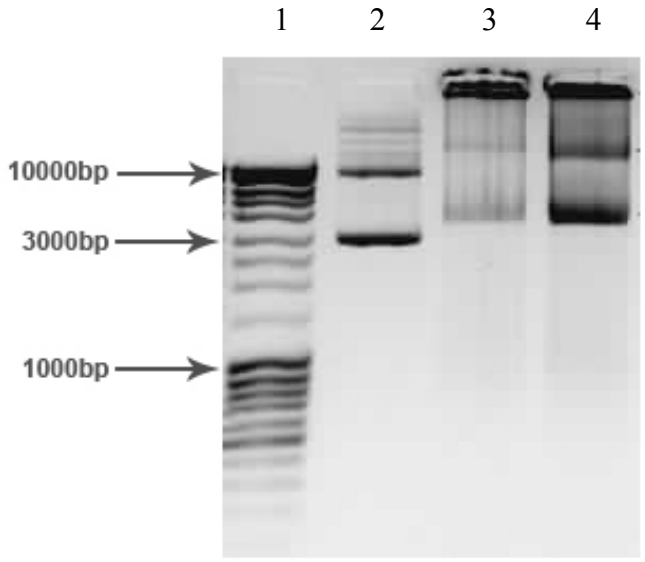
TABLE 9: Spectrophotometric data for digested amplicons

AMPLICON	CONCENTRATION (ng/µl)	A_{260}/A_{280}
Adaptin	100	1.82
Clathrin	62	1.74
PK	98	1.52

3.4) Plasmid preparation

3.4.1) Extraction of plasmid DNA

Plasmids were extracted from DH5α *E. coli* cells and 10µl aliquots of the samples were electrophoresed on agarose gel to determine their integrity. Native plasmids resolve into three major bands representing the various conformations of the circular DNA. This was the case when extracted samples were electrophoresed (*Figure 20*), indicating that the plasmids were not degraded.



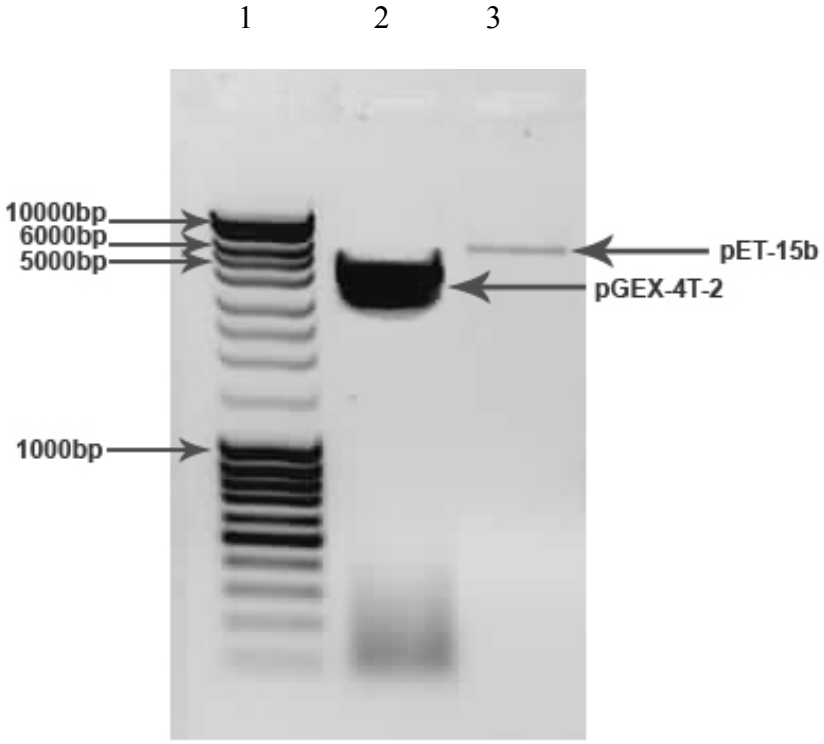
1 – MassRuler™ DNA ladder mix
2 – pGEX-4T-2 (0.6µg)
3 – pET-15b (1.5µg)
4 – pET-15b (2.7µg)

Figure 20. Plasmid stocks resolved on 1% agarose gel

Plasmids extracted from DH5α cells were intact and resolved into three main bands representing the native plasmid conformations.

3.4.2) Digestion of plasmids

Plasmids were digested with their respective endonucleases to linearise their structure in preparation for ligation with the digested PCR amplicons. The pGEX-4T-2 plasmid, which encodes a 5' GST tag, was digested with *Bam*HI and *Xho*I. The pET-15b plasmid, which encodes a 5' hexahistidine tag, was digested with *Bam*HI and *Nde*I. A single band of the correct size was observed for both the pGEX-4T-2 and pET-15b digested samples (*Figure 21*). The yield and concentration for pGEX-4T-2 were consistently greater than for pET-15b. This can be observed in figure 21, where the band for pGEX-4T-2 is much broader and darker than that for pET-15b, indicating a higher concentration of the former.



1 – MassRuler™ DNA ladder mix
2 – pGEX-4T-2
3 – pET-15b

Figure 21. Digested plasmids resolved on 1% agarose gel

Plasmids were digested and purified, followed by treatment with calf intestinal phosphatase (CIP) to prevent self-ligation of the linearised vectors. 10µl of each sample was loaded. The plasmids had the following sizes: pET-15b – 5 708bp; pGEX-4T-2 – 4 900bp.

Spectrophotometric data confirmed that the concentration of the pGEX-4T-2 sample was much higher than that of the pET-15b sample (*Table 10*). Both samples were of high purity, as seen by the A_{260}/A_{280} ratios. These samples were used for subsequent ligation reactions.

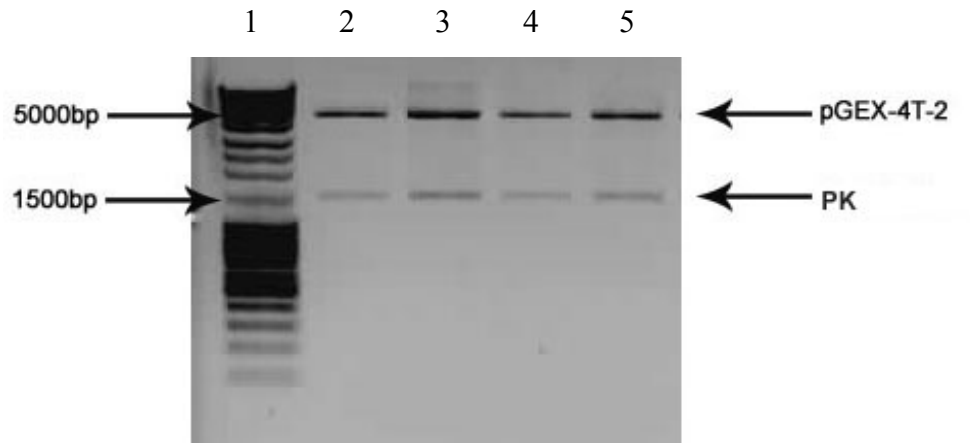
TABLE 10: Spectrophotometric data for digested, CIP-treated plasmids

SAMPLE	CONCENTRATION (ng/µl)	A_{260}/A_{280}
pET-15b	201	2
pGEX-4T-2	1 246	1.94

3.5) Results for *P. falciparum* PK

3.5.1) Verification of PK vector construct

After vector-insert ligation and transformation of *E. coli* DH5α cells, restriction digests and PCR were carried out on plasmids extracted from the transformed colonies to ensure that the correct inserts were present. Colonies 1 to 5 contained the PK insert, as can be seen in lanes 2 to 5 of figure 22. Linear pGEX-4T-2 resolves at 4 900bp while the PK insert migrates to 1 515bp.

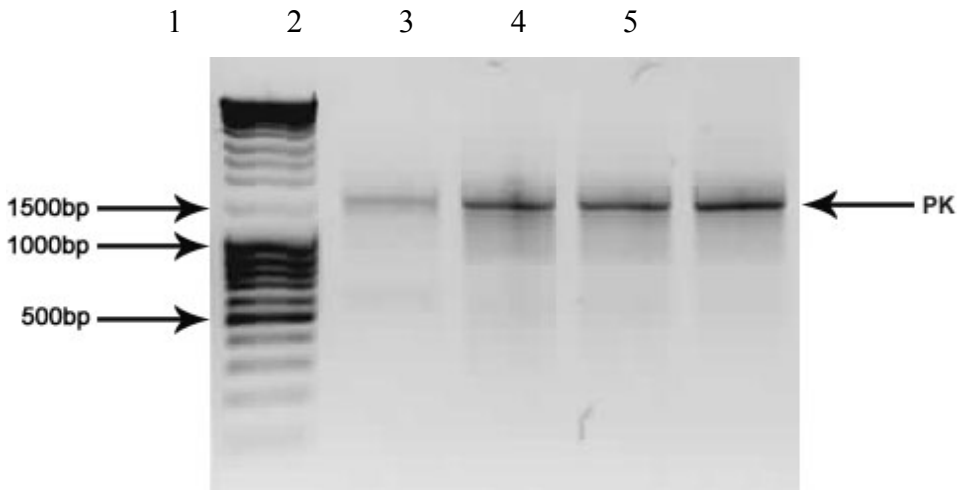


1 – MassRuler^{TM} DNA ladder mix
2 - 5 – Digested vector constructs from colonies 1-4

Figure 22. Digested PK vector constructs from DH5α cells

The PK insert was excised by digestion with the same restriction endonucleases used for the plasmid preparation (section 3.4.2) and was resolved on a 1% agarose gel.

The presence of the insert was also verified by PCR of plasmids extracted from transformed colonies using insert-specific primers. Three positive colonies are shown in figure 23, lanes 3, 4 and 5. Amplified PK from genomic *P. falciparum* DNA was used as a control to verify the size of the inserts at 1 515bp.



1 – MassRuler™ DNA ladder mix
2 – Amplified *P. falciparum* DNA (control)
3 – Amplified vector construct from colony 2
4 – Amplified vector construct from colony 5
5 – Amplified vector construct from colony 10

Figure 23. Amplified PK vector constructs from DH5α cells

The presence of the PK insert was verified in plasmids extracted from DH5α colonies 2, 5 and 10 which were resolved on 1% agarose gel. The insert size was estimated against the DNA ladder as 1 515bp, which was the expected size.

3.5.2) PK vector construct sequence data

The chromatogram for the 5' section of the PK DNA sequence (*Figure 24*) was well resolved with definite, high amplitude peaks and low baseline noise, indicating a high quality sequence at the 5' end. Towards the 3' end of the sequence the peaks overlapped and were lower in amplitude. This is a normal occurrence as the sequencing gel loses resolution and is difficult to read. The insert was therefore also sequenced from the 3' end to obtain reliable sequence data for the entire insert.

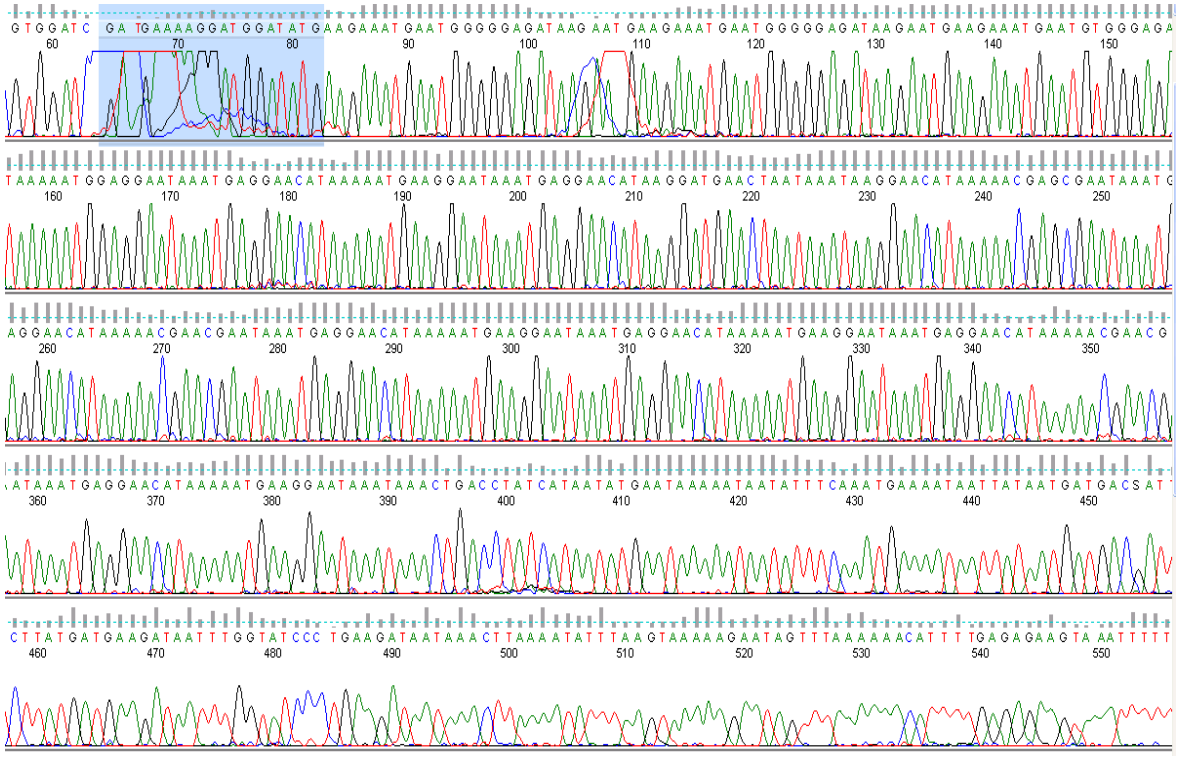


Figure 24. Section of PK vector construct chromatogram

A vector construct from colony 10 was sequenced with a 5' pGEX 4T-2 primer. The P. falciparum sequence starts at base pair 64, highlighted in blue, with the forward primer 5'-GAT GAA AAG GAT GGA TAT G -3'.

To confirm that the subcloned sequence was identical to the original sequence obtained from the PlasmoDB website, the two sequences were aligned (*Figure 25*).

```
7450 GATGAAAAGGATGGATATGAAGAAATGAATGGGGGAGATAAGAATGAAGA   7499
     |||||||||||||||||||||||||||||||||||||||||||||||||||
   1 GATGAAAAGGATGGATATGAAGAAATGAATGGGGGAGATAAGAATGAAGA     50

7500 AATGAATGGGGGAGATAAGAATGAAGAAATGAATGTGGGAGATAAAAATG   7549
     |||||||||||||||||||||||||||||||||||||||||||||||||||
  51 AATGAATGGGGGAGATAAGAATGAAGAAATGAATGTGGGAGATAAAAATG    100

7550 GAGGAATAAATGAGGAACATAAAAATGAAGGAATAAATGAGGAACATAAG   7599
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 101 GAGGAATAAATGAGGAACATAAAAATGAAGGAATAAATGAGGAACATAAG    150

7600 GATGAACTAATAAATAAGGAACATAAAAACGAGCGAATAAATGAGGAACA   7649
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 151 GATGAACTAATAAATAAGGAACATAAAAACGAGCGAATAAATGAGGAACA    200

7650 TAAAAACGAACGAATAAATGAGGAACATAAAAATGAAGGAATAAATGAGG   7699
     |||||||||||||||||||||||||||||||||||||||||||||||||||
 201 TAAAAACGAACGAATAAATGAGGAACATAAAAATGAAGGAATAAATGAGG    250
```

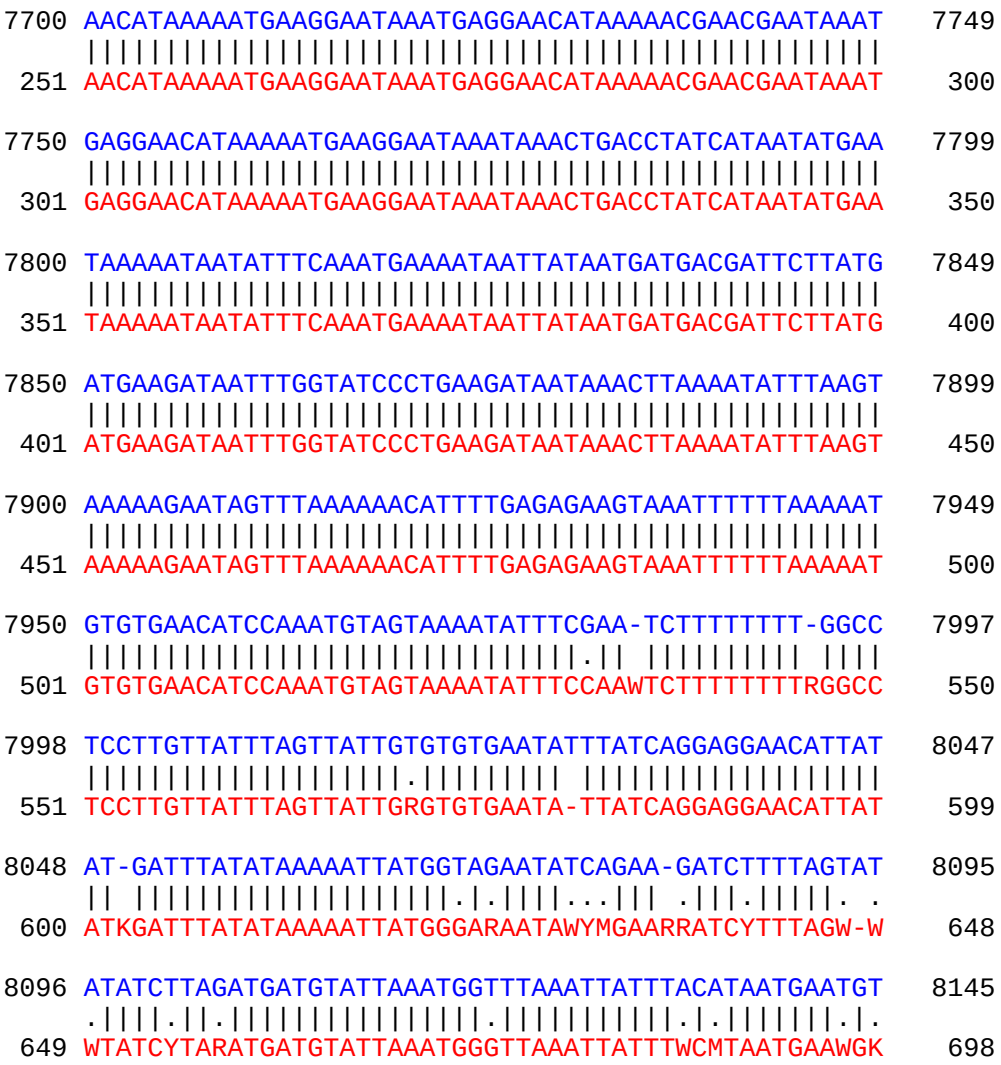


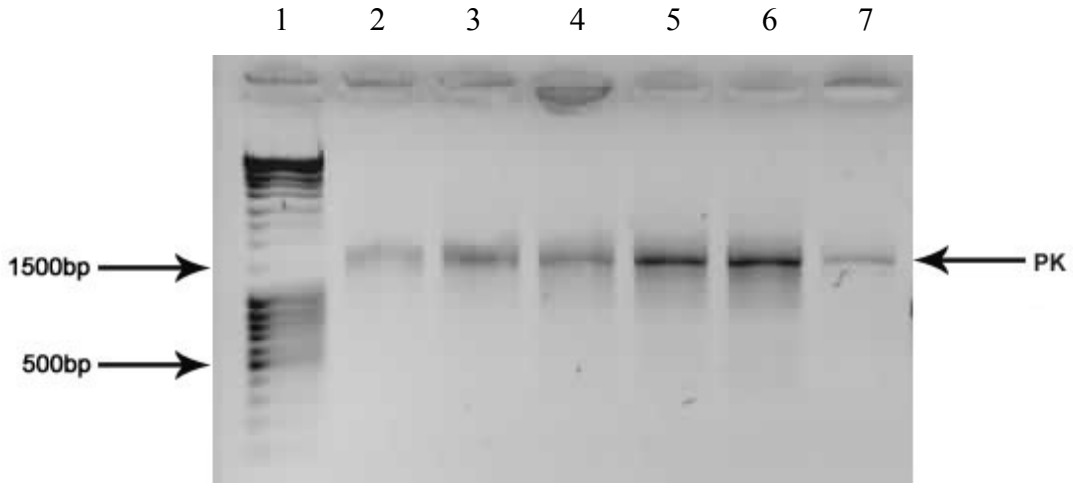
Figure 25. Alignment of PK DNA sequences from 5' sequencing

The sequence obtained from subcloning (red) was aligned with the original sequence obtained from the PlasmoDB database (blue). The underlined sections indicate the forward primer position 5'-GAT GAA AAG GAT GGA TAT G -3'. Key to ambiguous bases: W = T and A overlap in sequence; M = multiple overlap of peaks in sequence; R = A and G overlap; Y = C and T overlap. A dot between base pairs indicates a non-alignment, while a dash indicates a missing base.

Discrepancies towards the 3' end were resolved by sequencing the reverse strand (*Appendix, section 5.8*) which completes the entire 1 515bp sequence correctly, indicating that there is 100 percent sequence identity between the subcloned sequence and the original DNA sequence from PlasmoDB.

3.5.3) Verification of PK vector construct in Rosetta 2 cells

Rosetta 2 (DE3) cells were transformed with the PK vector constructs and transformation was verified by digestion of plasmids extracted from the cells (*not shown*) and by PCR with *P. falciparum* gene specific primers (*Figure 26*). The presence of PCR products with a length of 1 515bp indicated that the recombinant insert was present and that the transformation had been successful. This can be seen in lanes 2 - 6 of figure 26.



1 – Mass- Ruler™ DNA ladder mix
2-6 – Amplified vector constructs, colonies 1-5
7 – Amplified *P. falciparum* DNA (control)

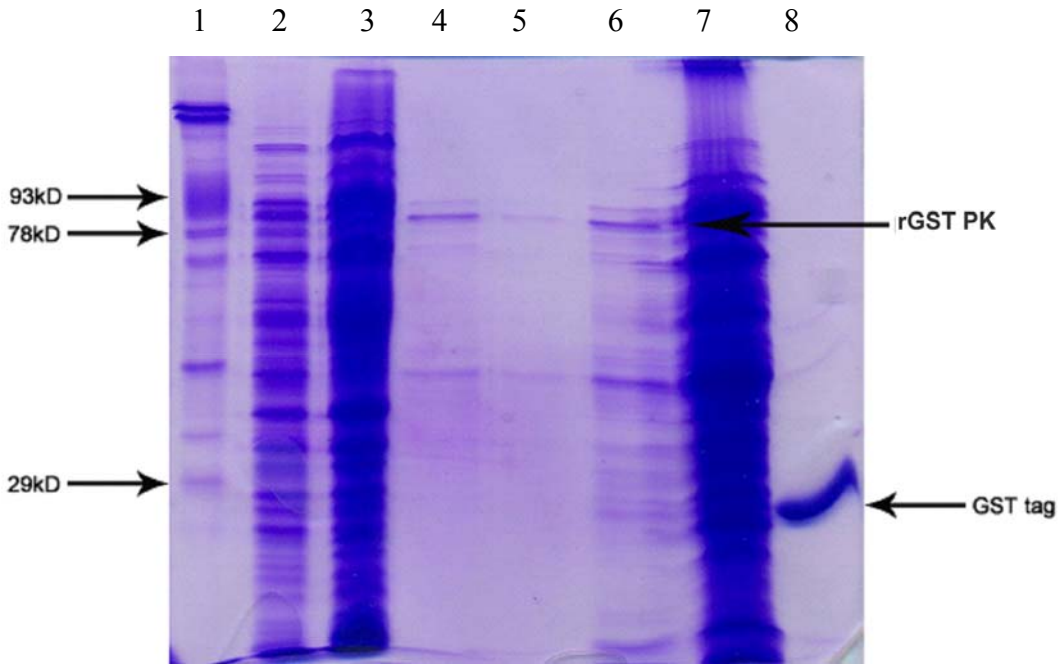
Figure 26. Amplified Rosetta vector constructs resolved on 1% agarose gel

All Rosetta 2 (DE3) colonies were positive for the PK insert (lanes 2-6). Lane 7 contained amplified PK from genomic P. falciparum DNA as a control.

3.5.4) Laemmli SDS-PAGE of rGST PK

E. coli cells were induced to express recombinant proteins that were analysed by SDS PAGE and immunoblotting. The rGST PK protein has a predicted molecular weight of 84.8kD. Soluble protein of the approximate molecular weight was detected in the eluted fractions as seen in lanes 4-6 of figure 27. PK 1st elution and PK 2nd elution contained GST-tagged protein sequestered from the beads during affinity purification.

The PK stripped fraction consisted of any proteins that remained attached to the beads after elution. The control containing only the GST tag resolved below the 29kD mark, as expected, with a molecular weight of 25.5kD.



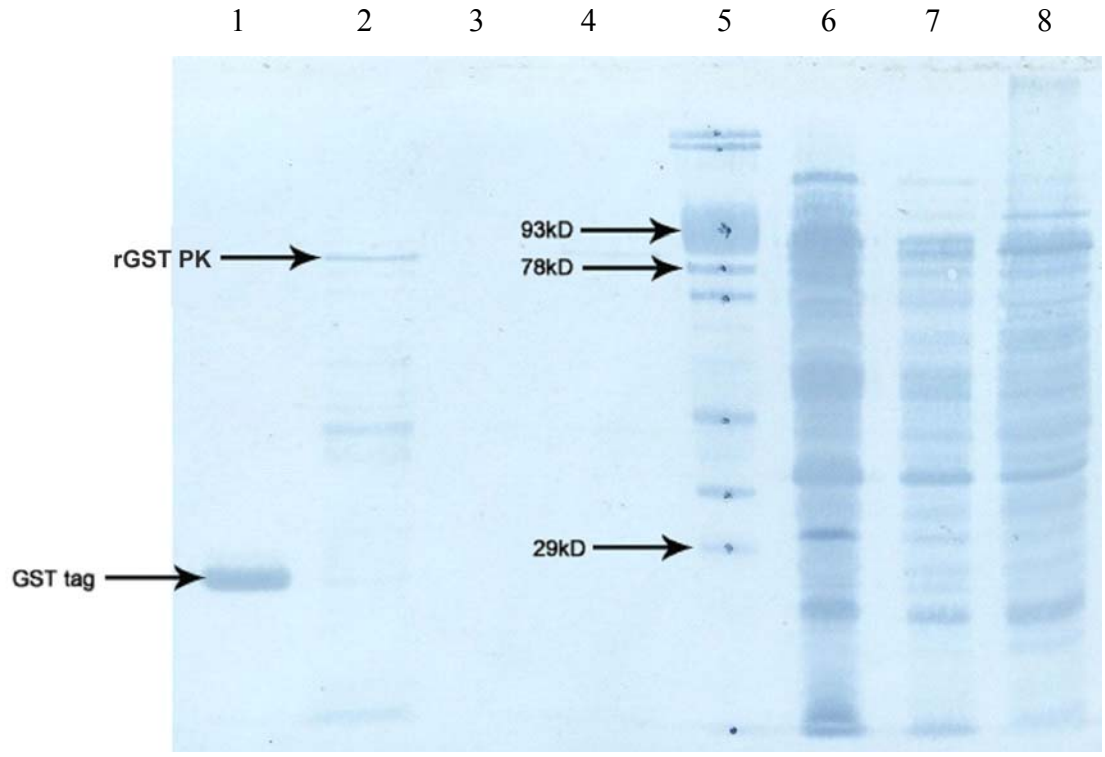
1 – RBC membrane marker
2 – PK total induced fraction
3 – PK insoluble fraction
4 – PK 1st elution
5 – PK 2nd elution
6 – PK stripped beads
7 – PK unbound fraction
8 – GST control

Figure 27. Purification of rGST PK protein

The protein fractions obtained during extraction and purification of rGST PK from Rosetta cells were resolved on a 12% SDS polyacrylamide gel. The purity of rGST PK in the first elution (lane 4) is 53.2 percent (section 3.5.7). Lane 8 contained GST protein expressed from E. coli cells containing non-recombinant pGEX 4T-2 plasmids. The 3 bands indicated in the RBC membrane marker (lane 1) are: Band 3 – 93kD; Protein 4.1 – 78kD; Stomatin/tropomyosin – 29kD.

3.5.5) Western blot of rGST PK

Figure 28 illustrates a nitrocellulose membrane stained after the transfer of proteins from the SDS acrylamide gel (*Figure 27*) and immunoblotting with an anti-GST primary antibody.



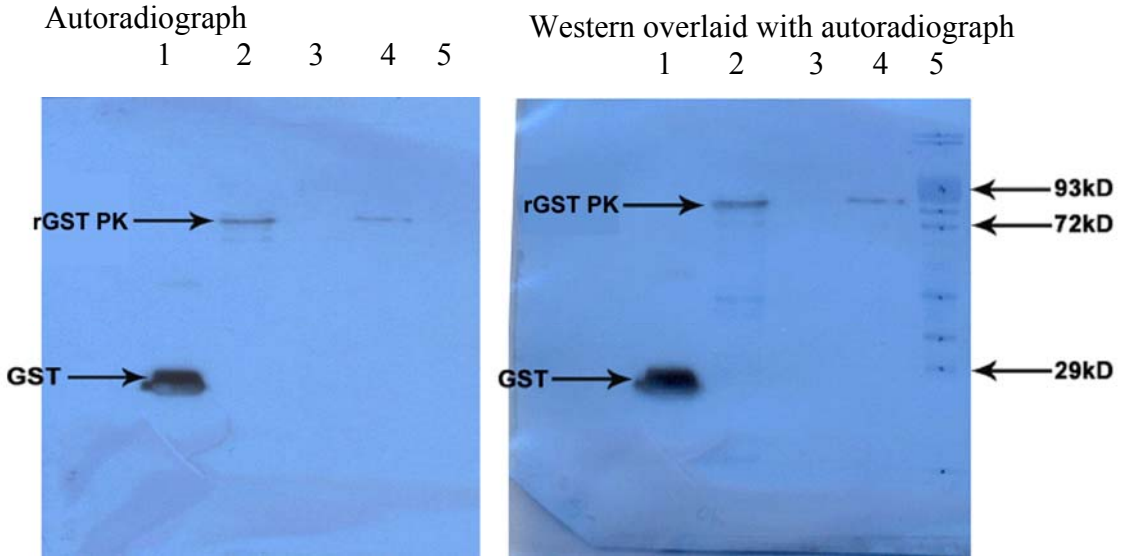
1 – GST control
2 – PK stripped beads
3 – PK 2nd elution
4 – PK 1st elution
5 – RBC membrane marker
6 – PK insoluble fraction
7 – PK total induced
8 – PK unbound fraction

Figure 28. rGST PK Western blot

The Western blot was stained in amido black to visualise proteins. Faint bands of approximately 84.8kD were obtained in the 1st elution and stripped fractions (lanes 2 and 4), indicating recombinant target protein in soluble form. Lane 8 contained GST protein expressed from E. coli cells containing non-recombinant pGEX 4T-2 plasmids. The RBC membrane marker (lane 5) consists of: Band 3 – 93kD; Protein 4.1 – 78kD; Stomatin/tropomyosin – 29kD. An anti-GST primary antibody was used for immunoblotting.

3.5.6) Immunoblot of rGST PK

An autoradiograph was obtained from the chemiluminescent immunoblot probed with an HRP-linked anti-GST primary antibody. This confirmed the presence of soluble protein in the correct molecular weight range for rGST PK in the stripped and 1st elution fractions as seen in lanes 2 and 4 (*Figure 29*). The presence of rGST PK in the stripped fraction indicated that the elutions were not completely effective and target protein was retained on the beads.



1 – GST control
2 – PK stripped beads
3 – PK 2nd elution
4 – PK 1st elution
5 – RBC membrane marker

Figure 29. Immunoblot of rGST PK

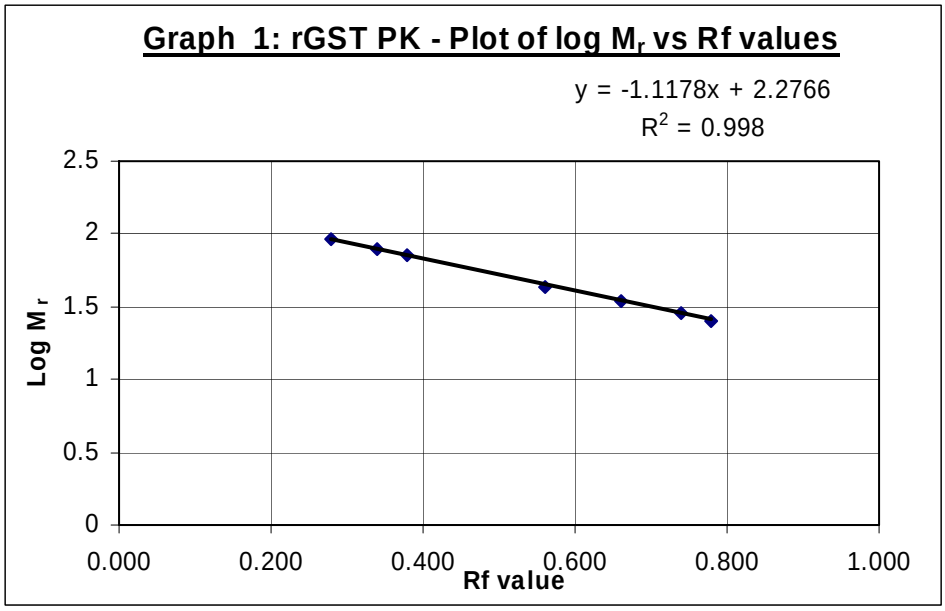
X-ray film was exposed to chemiluminescent immunoblots for 60 seconds to obtain autoradiographs. Lane 8 contained GST protein expressed from E. coli cells containing non-recombinant pGEX 4T-2 plasmids. The RBC membrane marker (lane 5) consists of: Band 3 – 93kD; Protein 4.1 – 78kD; Stomatin/tropomyosin – 29kD.

The molecular weight of rGST PK bands on the autoradiograph was determined via retardation factor (Rf) values of proteins making up the RBC membrane marker (*Table 11*).

TABLE 11: Data used to determine molecular weight of rGST PK

PROTEIN	M_r (kD)	Log M_r	MIGRATION DISTANCE (mm)	Rf VALUE
Band 3	93	1.9685	14	0.280
Protein 4.1	78	1.8921	17	0.340
Protein 4.2	72	1.8573	19	0.380
Actin	43	1.6335	28	0.560
G3PD	35	1.5441	33	0.660
Stomatin/tropomyosin	29	1.4624	37	0.740
GST	25.5	1.4065	39	0.780
PK (on autoradiograph)	**82.97**	1.9189	16	0.320
Dye front			50	

The molecular weight of rGST PK that was extrapolated from Graph 1 is 82.97kD. This is very close to the predicted value of 84.4kD. The difference between the two values could be due to the anomalous size migration of *P. falciparum* proteins on SDS polyacrylamide gels due to non-uniform binding of SDS, which has been reported previously. This occurs in low complexity regions of the parasite proteins which are highly charged (*Cooke, 2001*). The predicted molecular weight of 3D7 strain parasite proteins can also vary between different strains and isolates (http://www.expasy.org/tools/pi_tool.html). The R^2 value indicates that there is 99.8 percent correlation, signifying that the standard points accurately predict molecular weight under conditions of normal protein size migration.



3.5.7) Purity of rGST PK

The first elution fraction was scanned (*Figure 30*) from the SDS-PAGE gel of recombinant PK (*Figure 27*). There was no visible band of recombinant protein in the second elution fraction. The target protein is represented by peak 4.

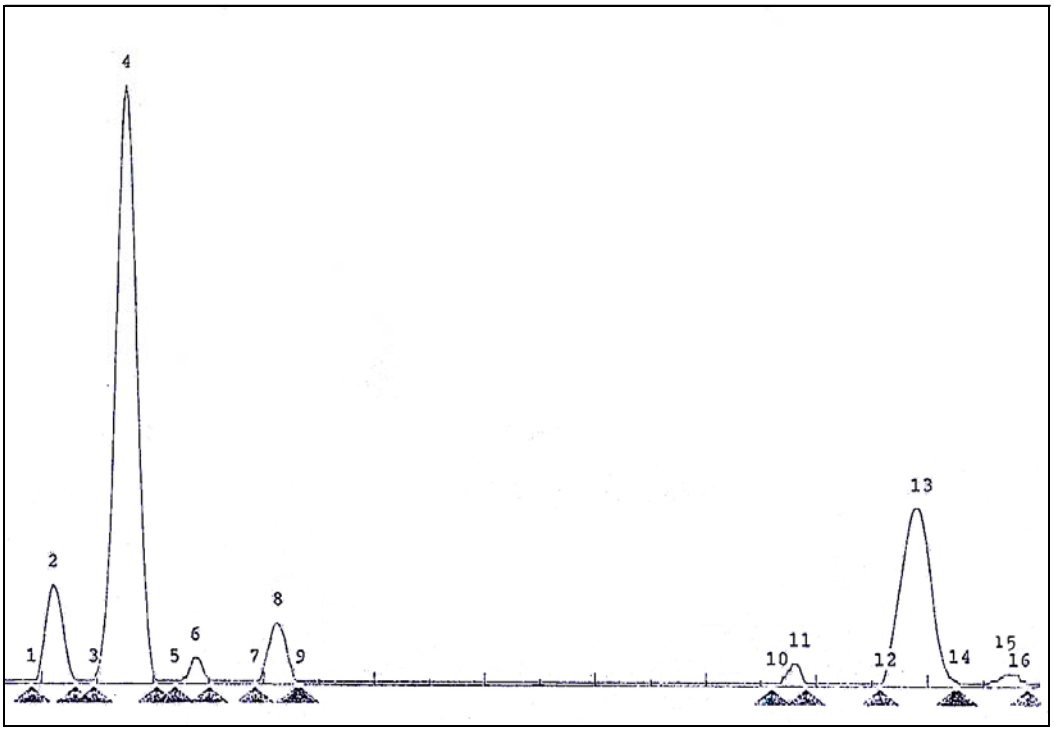


Figure 30. Densitometric scan of first elution of rGST PK

Peak 4 shows the amount of target protein; the other peaks consist of contaminating proteins.

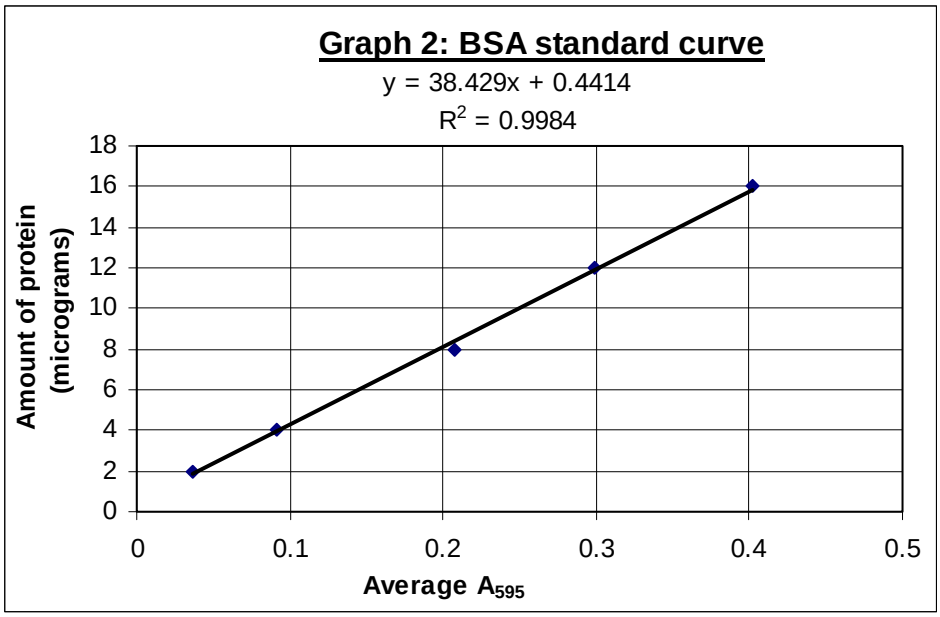
The seven peak areas are represented by numbers 2, 4, 6, 8, 11, 13 and 15 (*Table 12*). Taking these peaks into account, the rGST PK fraction had an approximate purity of 53.2 percent.

TABLE 12: Densitometric data for first elution of rGST PK

PEAK	AREA	PERCENTAGE AREA
2	319	8.2
4	2 066	53.2
6	60	1.5
8	201	5.2
11	262	6.7
13	912	23.4
15	69	1.8
TOTAL	3 889	100

3.5.8) Concentration of rGST PK

The Coomassie Assay Kit was used to determine the concentration of rGST PK. The BSA standard curve data obtained are shown in table 24 (*Appendix, section 5.6*) and graph 2. The R^2 value indicates that there is 99.8 percent correlation, which means that the standards closely predict the amount of protein present in the samples.



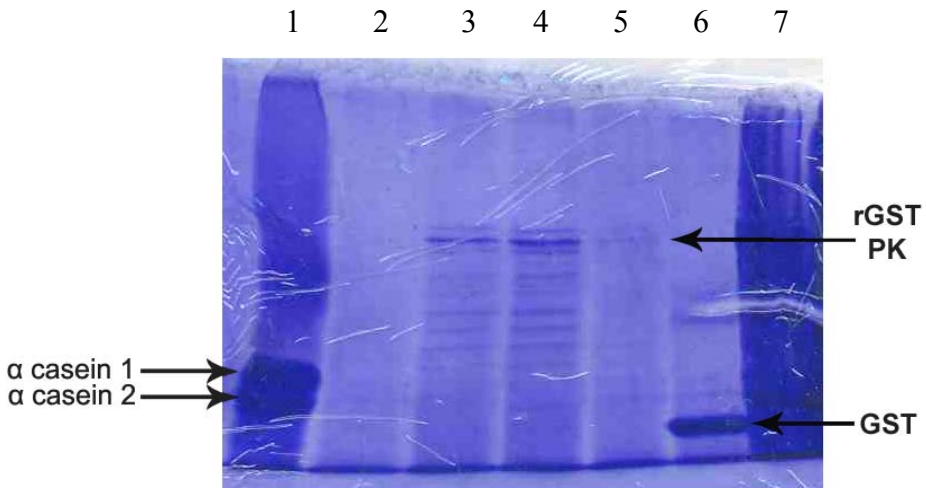
The extrapolated protein concentrations for rGST PK from three separate experiments is shown in table 13. The protein concentration of the first elution ranged from 0.35µg/µl to 0.4µg/µl in 30µl of elution buffer. With a purity of 53.2 percent, these amounts revert to 0.18µg/µl and 0.21µg/µl. The amount of pure rGST PK recovered from a 30ml culture volume, in 30µl of total elution buffer, is thus approximately 5.4µg to 6.3µg. This equates to 180-210µg of pure recombinant protein per litre of culture.

TABLE 13: Extrapolated protein concentration of rGST PK

rGST PK SAMPLE	VOLUME (µl)	A_{595}	PROTEIN (µg)	PROTEIN CONCENTRATION (µg/µl)
1 - 1st elution	10	0.080	3.5	0.35
2 - 1st elution	10	0.093	4.0	0.40
3 - 1st elution	10	0.085	3.7	0.37

3.5.9) Kinase assay

A standard kinase assay was carried out using 6µl of recombinant PK (*Table 13, sample 1*), resulting in approximately 1µg of pure recombinant enzyme in the first elution fraction (*Figure 31*). Lanes 2 to 7 contained radioactive assay mix along with the indicated PK samples, while lane 1 contained only bovine casein. It was inferred from research by Chishti et al (*1994*) that the PK would most probably function as a casein kinase, hence this exogenous protein was electrophoresed individually. Bovine casein consists of 3 protein subunits of varying molecular weights, namely: α casein 1 at 25kD; α casein 2 at 22kD and κ casein subunit at 19kD. The other two substrates, myelin basic protein and histone protein 1, have molecular weights of 18.4kD and 21.5kD, respectively.

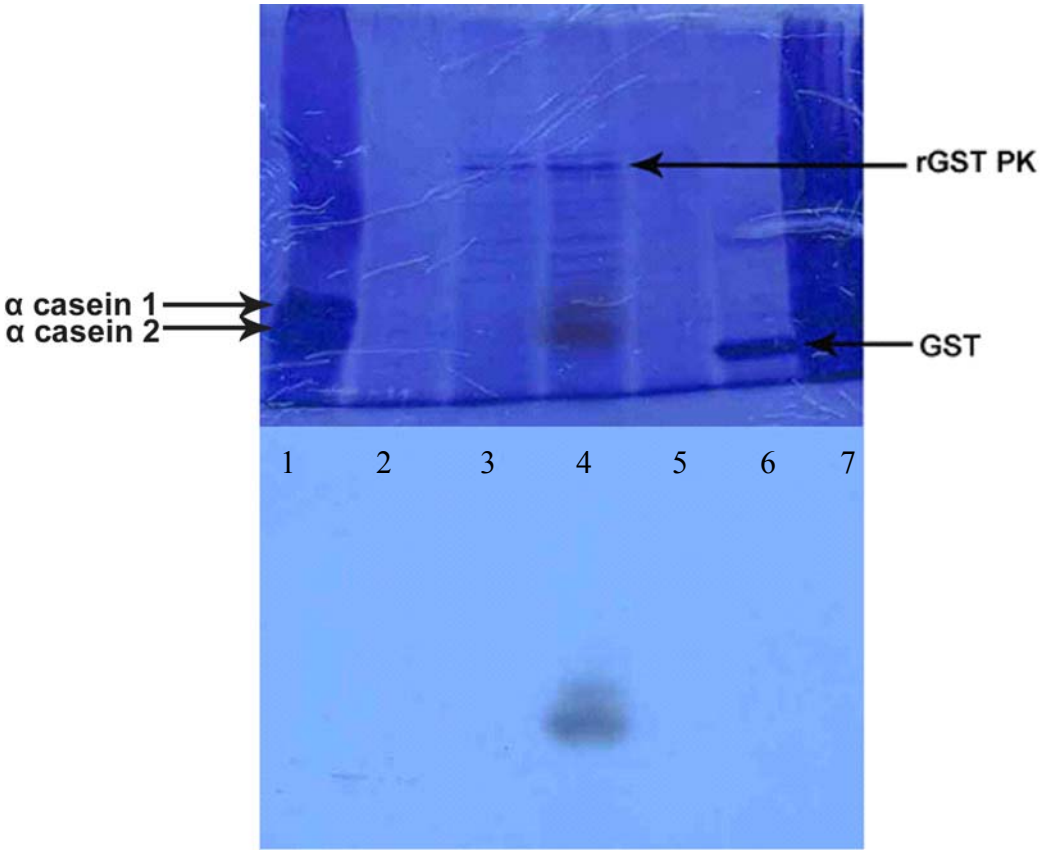


1 – Bovine casein
2 – PK total uninduced
3 – PK 2nd elution
4 – PK 1st elution
5 – PK stripped beads
6 – GST control
7 – PK total induced

Figure 31. Dried SDS acrylamide gel of kinase assay

40µg of bovine casein was loaded into lane 1. Lanes 2-7contained the radioactive kinase assay mix and recombinant enzyme, including 2µg bovine casein, 10µg histone protein 1 and 5µg myelin basic protein which were not visible on the gel due to their low concentration. rGST PK was soluble in the 1st and 2nd elution fractions in lanes 3 and 4.

An autoradiograph of the kinase assay gel showed that radioactivity was detected in lane 4, indicating that the purified recombinant enzyme was active (*Figure 32*). From the overlay, it was determined that the substrate that was phosphorylated was bovine casein, specifically the α subunits of the protein. The radioactive band on the autoradiograph seems to be a double band, with the upper section being fainter than the lower. This indicates that the α casein 1 subunit was phosphorylated to a lesser extent than the α casein 2 subunit. These results signify that the parasite PK functions as a casein kinase.



1 – Bovine casein
2 – PK total uninduced
3 – PK 2nd elution
4 – PK 1st elution
5 – PK stripped beads
6 – GST control
7 – PK total induced

Figure 32. Autoradiograph of kinase assay gel

The autoradiograph was obtained from overnight exposure to the dried radioactive kinase gel. The bands obtained in lane 4 lined up with the α1 and α2 subunits of bovine casein, indicating that this protein was the target of rGST PK phosphorylation.

3.5.10) Bioinformatic data for *P. falciparum* PK

3.5.10.1) Structure of *P. falciparum* PK

The Swiss Model Template Library was used to find three dimensional models of proteins to which the PK sequence had structural homology, as no 3D model of the PK was available on the PlasmoDB website. The Library is limited by the number of 3D models of crystal structures available in the database, so while other proteins may have better homology to the target, their 3D constructs may not yet have been created in the Swiss Model Template Library. The whole parasite PK sequence (2 485aa) had homology to template 1apmE (*Figure 33*). This template represents the crystal structure of the catalytic subunit of human cAMP-dependent protein kinase. This human enzyme is a serine/threonine protein kinase. The kinase domain shares sequence identity of only 26.4 percent with the template, indicating that errors in the sequence alignment algorithms may have occurred because the identity is less than 40 percent (*Appendix, section 5.11; figure 73*).

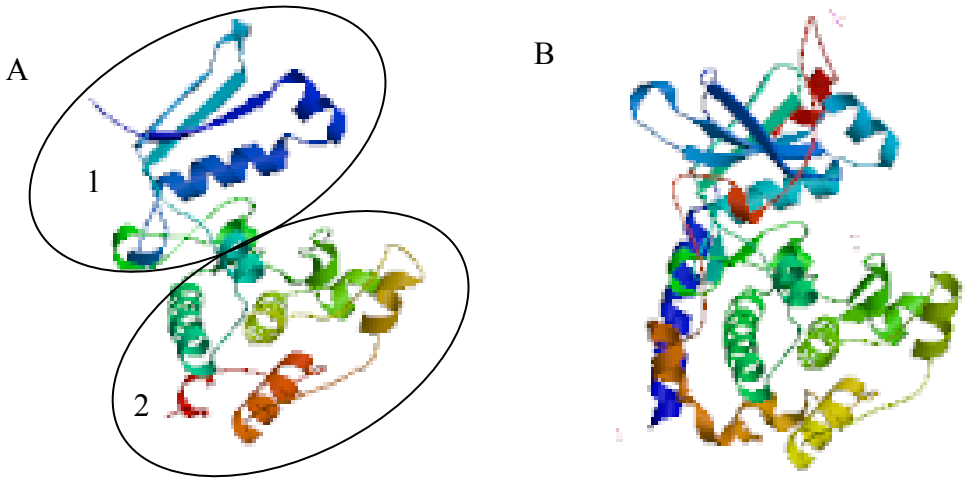


Figure 33. Putative 3D model of PK kinase domain

The putative model of the PK sequence (A) is based on template 1apmE (B), the catalytic subunit of human cAMP-dependent protein kinase. This template had most sequence identity to the PK sequence in the region of the kinase domain. The modeled residues are amino acids 2 117 to 2 353. The E-value is 2.80e-45. The areas designated 1 and 2 on model A indicate the bilobed structure of the kinase catalytic domain. The active site is found in the cleft between the two lobes (Hanks et al, 1998)

According to Hanks *et al* (1998) the 3D structure of eukaryotic PK catalytic domains is highly conserved, with the 300 amino acid residues constituting this area having a bilobal conformation. The active site is located in the cleft between the lobes, along with ATP and substrate binding sites (Hanks *et al*, 1998). This bilobed conformation can be seen in the putative model of the parasite PK sequence (*Figure 33*). The area of the PK that was modeled is indicated in figure 34.

A.

B.

EKDGYEEMNGGDKNEEMNGGDKNEEMNVGDKNGGINEEHKNEGINEEHKDELINKEHKNERINEEHKNERINEEHKN
EGINEEHKNEGINEEHKNERINEEHKNEGINKLTYHNMNKNNISNENNYNDDDSYDE**DNLVSLKIINLKYLSKK**
NSLKNILREVNFLKMCEHPNVVKYFESFFWPPCYLVIVCEYLSGGTLYDLYKNYGRISEDLLVYIL
DDVLNGLNYLHNECSSPLIHRDIKPTNIVLSKDGIAKIIDFGSCEELKNSDQSKELVGTIYYISPE
ILMRTNYDCSSDIWSLGITIYEIVLCTLPWKRNQSFENYIKTIINSSPKINITEGYSKHLCYFVEK
CLQKKPENRGNVKDLLNHKFLIKKRYIKKKPSSIYEIRDILKIYNGKGKTNIFRNFFKNLFFFNDKNKKKKPN
KMISSKSCDAEMFFEQLKRENFDFFEIKLKDDENSRSLNTFNINISKERDDISYSSLNLEKIKEHSLNMVASVVGTE
QSQK

Figure 34. Sequence homology of PK to 3D template

In A, the blue segment starting at amino acid 2 117 and ending at 2 353 indicates the section where the P. falciparum PK shares homology with template 1apmE. 26.4 percent of the blue segment shares sequence identity with template 1apmE. The green segment represents the whole PK protein of 2 485 amino acids.
In B, the amino acid sequence of the P. falciparum PK is represented. The red area indicates the kinase domain containing the enzyme active site which spans amino acids 2 079-2 354. The bold, larger red letters indicate the area of the parasite protein that shares homology with the 1apmE template, which falls within the kinase domain.

The Jnet secondary structure prediction algorithm (*Cuff and Barton, 2000*) predicted a conformation of mainly alpha helices and some beta sheets for the PK kinase domain (*Appendix, section 5.10, Figure 70*) which seems to be consistent with the 3D model in figure 33.

3.5.10.2) Homology of *P. falciparum* PK with other proteins

The entire *P. falciparum* PK protein sequence (2 485aa) was compared to the proteome of several organisms to determine sequence identity and positivity. Sequence identity refers to the two protein sequences having exactly the same amino acids in the same position; the sequence positivity value is the identity plus the number of amino acids of the same group residing in the same position. The first group that the sequence was compared to consisted of the following *Plasmodium* species: *P knowlesi, P. vivax, P. yoelii, P. chabaudi and P. berghei*. This search was performed to detect possible orthologues of the *P. falciparum* PK in the same genus. The results of the BLAST search are presented in table 14.

TABLE 14: Sequence identity of *P. falciparum* PK with other *Plasmodium* proteins

ORGANISM	GENE AND PROTEIN	SEQUENCE IDENTITY (percent)		SEQUENCE POSITIVITY (percent)		E-VALUE FOR KINASE DOMAIN
		Whole protein	Kinase domain	Whole protein	Kinase domain	
P. vivax	*PVX003590;* putative serine/threonine specific PK (6 378aa)	11	49	14	66	6.3e-141
P. knowlesi	*PKH041680;* putative PK (6 390aa)	10	57	13	74	3.0e-150
P. yoelii	*PY00029*; myosin light chain kinase (5 742aa)	3	30	5	47	1.6e-24
P. yoelii	*PY00403*; kinase Akt/PKB-related (2 178aa)	4	26	7	46	1.5e-24

The kinase domain of the *P. falciparum* PK, containing the catalytic site, had most sequence identity with the *Plasmodium sp* proteins in table 14 (*P.vivax PK BLAST, Appendix, section 5.9, Figure 65*). The putative PKs of *Plasmodium* species infecting humans – namely *P. vivax* and *P. knowlesi* – had most sequence identity to the *P. falciparum* PK. No paralogues or in-genera orthologues for the *P. falciparum* PK are listed on the PlasmoDB website. This means that the proteins from the other *Plasmodium sp*. mentioned in table 14 are not derived from the same ancestral gene, nor do they have exactly the same function as the *P. falciparum* PK.

The *T. gondii* proteome was searched using information from the ToxoDB website. This organism belongs to the same *Apicomplexa* phylum as *P. falciparum.* The results are shown in table 15.

TABLE 15: Sequence identity of *P. falciparum* PK with *T. gondii* proteins

PRODUCT	SEQUENCE IDENTITY (percent)		SEQUENCE POSITIVITY (percent)		E-VALUE FOR KINASE DOMAIN
	Whole protein	Kinase domain	Whole protein	Kinase domain	
42.m03467; serine/threonine-protein kinase-related (4 253aa)	7	42	11	61	5.1e-40
162.m00001; putative calcium-dependent protein kinase (583aa)	3	37	5	60	5.7e-24
641.m02549; serine/threonine-protein kinase (296aa)	3	29	5	47	2.2e-22
541.m00134; putative calmodulin-domain protein kinase (538aa)	3	32	6	51	6.4e-24

The region of the *P. falciparum* protein where most sequence identity with *T gondii* proteins was shared was the kinase domain containing the catalytic site (*42.m03467, Appendix, section 5.9, Figure 66*).

Finally, the human proteome was searched for molecules with homology to the *P. falciparum* PK protein sequence. This would determine the suitability of the parasite protein as a drug target; the fewer the hits and the lower the sequence identity between the human and parasite proteins, the more likely the utility as a drug target.

The whole *P. falciparum* PK protein has 3 percent identity and 5 percent positivity (E-value 2.0e-32) with *H. sapiens* serine/threonine protein kinase 24 found on chromosome 13, which was the best match (*Appendix, section 5.9, Figure 64*). However, most sequence homology was shared in the kinase domain where the *P. falciparum* PK had 37 percent identity and 55 percent positivity with the human enzyme. *H. sapiens* serine/threonine protein kinase 24 is 304 amino acids in length and is part of a group of non-specific serine/threonine protein kinases that do not have an activating compound, or whose specificity has not been analysed to date (*www.expasy.org*).

A BLASTP analysis on a non-redundant database encompassing all eukaryotes was carried out. The top thirty BLASTP hits (out of a total of 100) are shown in table 16 in descending order. The best hit occurred with a serine/ threonine-protein kinase of *Dictyostelium discoideum* (soil-living amoeba). Ninety percent of the sequences producing high-scoring hits were serine/ threonine-protein kinases from a number of organisms, as seen in table 16. The number of hits with MAPK kinases was at a much lower ten percent (yellow highlighted sections), although interestingly, two out of these three hits occurred within the plant proteome of *Arabidopsis thaliana* (Mustard plant).

TABLE 16: PK BLASTP analysis against eukaryotic proteomes

	ORGANISM	**SWISSPROT CODE**	**PROTEIN**	**P VALUE**
1.	Dictyostelium discoideum (soil-living amoeba)	PAKF_DICDI Q869T7	Serine/threonine-protein kinase	6.7e-33
2.	Squalus acanthias (Dogfish)	STK4_SQUAC Q802A6	Serine/threonine-protein kinase	8.5e-31
3.	Dictyostelium discoideum (soil-living amoeba)	SVKA_DICDI O61122	Serine/threonine-protein kinase	9.8e-31
4.	Homo sapiens	STK24_HUMAN Q9Y6E0	Serine/threonine-protein kinase 24	1.3e-30
5.	Dictyostelium discoideum (soil-living amoeba)	PAKB_DICDI Q869N2	Serine/threonine-protein kinase	1.3e-30

	ORGANISM	SWISSPROT CODE	PROTEIN	P VALUE
6.	Mus musculus	STK24_MOUSE Q99KH8	Serine/threonine-protein kinase 24	2.0e-30
7.	Dictyostelium discoideum (soil-living amoeba)	DST1_DICDI Q86IX1	Serine/threonine-protein kinase	3.9e-30
8.	Schizosaccharomyces pombe (Yeast)	PPK11_SCHPO O14047	Serine/threonine-protein kinase	1.1e-29
9.	Bos taurus	STK25_BOVIN Q3SWY6	Serine/threonine-protein kinase 25	1.1e-29
10.	Homo sapiens	STK25_HUMAN O00506	Serine/threonine-protein kinase 25	1.1e-29
11.	Xenopus laevis (Platana)	STK4_XENLA Q6PA14	Serine/threonine-protein kinase 4	2.3e-29
12.	Arabidopsis thaliana (Mustard plant)	M2K6_ARATH Q9FJV0	Mitogen-activated protein kinase kinase	2.3e-29
13.	Mus musculus	STK25_MOUSE Q9Z2W1	Serine/threonine-protein kinase 25	2.3e-29
14.	Cercopithecus aethiops (Vervet monkey)	STK4_CERAE A4K2Y1	Serine/threonine-protein kinase 4	5.4e-29
15.	Macaca mulatta (Rhesus monkey)	STK4_MACMU A4K2T0	Serine/threonine-protein kinase 4	5.4e-29
16.	Homo sapiens	STK4_HUMAN Q13043	Serine/threonine-protein kinase 4	5.4e-29
17.	Mus musculus	STK4_MOUSE Q9JI11	Serine/threonine-protein kinase 4	5.4e-29
18.	Papio anubis (Olive baboon)	STK4_PAPAN A4K2M3	Serine/threonine-protein kinase 4	5.4e-29
19.	Oryza sativa (cultivated rice)	M2K1_ORYSJ Q5QN75	Serine/threonine-protein kinase	6.9e-29
20.	Colobus guereza (Colobus monkey)	STK4_COLGU A4K2P5	Serine/threonine-protein kinase 4	7.1e-29
21.	Dictyostelium discoideum (soil-living amoeba)	DST2_DICDI Q55GC2	Mitogen-activated protein kinase kinase	7.9e-29
22.	Bos taurus	STK4_BOVIN Q5E9L6	Serine/threonine-protein kinase 4	9.3e-29
23.	Danio rerio (Zebrafish)	STK3_DANRE Q7ZUQ3	Serine/threonine-protein kinase 3	1.0e-28
24.	Lemur catta (Ring-tailed lemur)	STK4_LEMCA A4K2S1	Serine/threonine-protein kinase 3	1.3e-28
25.	Mus musculus	STK3_MOUSE Q9JI10	Serine/threonine-protein kinase 3	1.4e-28

	ORGANISM	SWISSPROT CODE	PROTEIN	P VALUE
26.	Rattus norvegicus	STK3_RAT O54748	Serine/threonine-protein kinase 4	1.6e-28
27.	Otolemur garnettii (Garnett's greater bushbaby)	STK4_OTOGA A4K2Q5	Serine/threonine-protein kinase 4	1.6e-28
28.	Arabidopsis thaliana (Mustard plant)	M2K1_ARATH Q94A06	Mitogen-activated protein kinase kinase	1.6e-28
29.	Homo sapiens	STK3_HUMAN Q13188	Serine/threonine-protein kinase 3	2.9e-28
30.	Xenopus laevis (Platana)	STK3_XENLA Q6IP06	Serine/threonine-protein kinase 4	6.2e-28

Table 17 shows a summary of all 100 hits obtained from the BLASTP process. Six out of the nine matches with MAPK kinases were with plant proteins, specifically *Arabidopsis thaliana* (Mustard plant). However, the vast majority of hits (84%) occurred with serine/ threonine protein kinases.

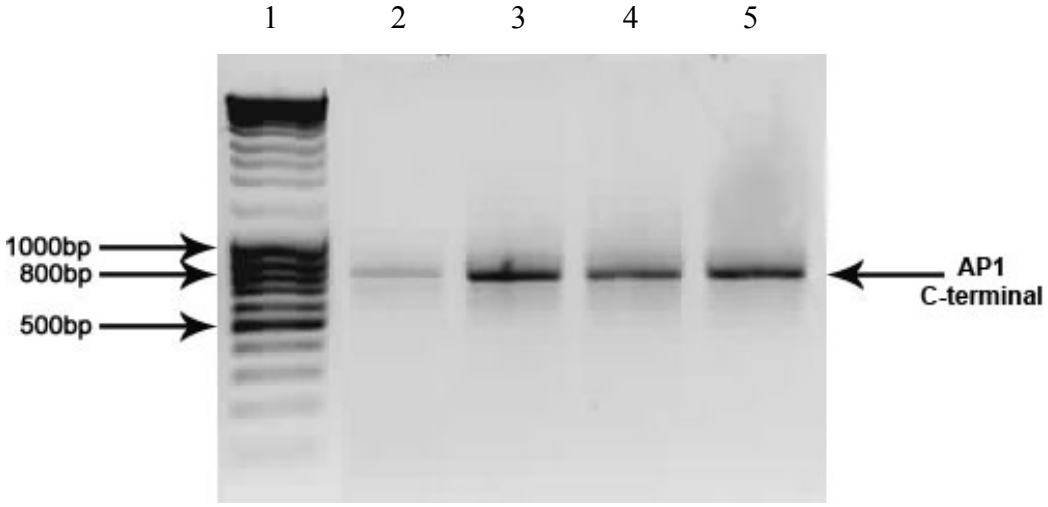
TABLE 17: Summary of BLASTP analysis

PROTEIN	NUMBER OF HITS
Serine/ threonine-protein kinase	84
Mitogen-activated protein kinase kinase (MAPK kinase)	9
Myosin IIIA	2
Cell division control protein 7	1
Dual specificity protein kinase	1
3-phosphoinositide-dependent protein kinase	1
Abscisic acid-inducible protein kinase	1
Protein kinase	1

3.6) Results for *P. falciparum* AP1 C-terminal

3.6.1) Verification of AP1 C-terminal vector construct

Plasmids ligated with theAP1 C-terminal insert were extracted from transformed DH5α colonies to ensure that the correct insert was present. Plasmids were digested with restriction endonucleases (*not shown*) and were amplified using insert specific primers. The PCR products were resolved on a 1% agarose gel (*Figure 35*). The AP1 C-terminal domain should be 723bp in size according to PlasmoDB, but migrated to between the 800bp and 900bp markers, as seen in figure 36. This aberrant migration occurred with all AP1 C-terminal-transformed colonies.



1 – MassRuler^{TM} DNA ladder mix
2 – Amplified *P. falciparum* DNA (control)
3 – Amplified vector construct from colony 1
4 – Amplified vector construct from colony 7
5 – Amplified vector construct from colony 10

Figure 35. Amplified AP1 C-terminal vector constructs from DH5α cells

The presence of the AP1 C-terminal insert was verified in plasmids extracted from DH5α colonies 1, 7 and 10 which were resolved on a 1% agarose gel.

3.6.2) AP1 C-terminal vector construct sequence data

The chromatogram for the AP1 C-terminal vector construct sequence showed well-delineated peaks with low background, indicating that the sequencing reaction was successful (*Figure 36*).

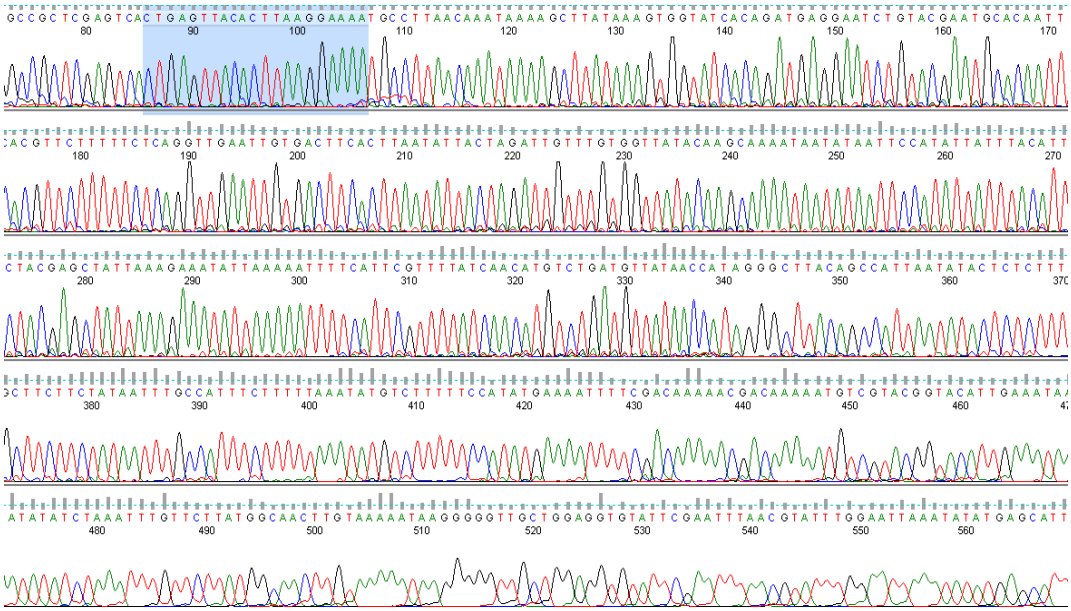
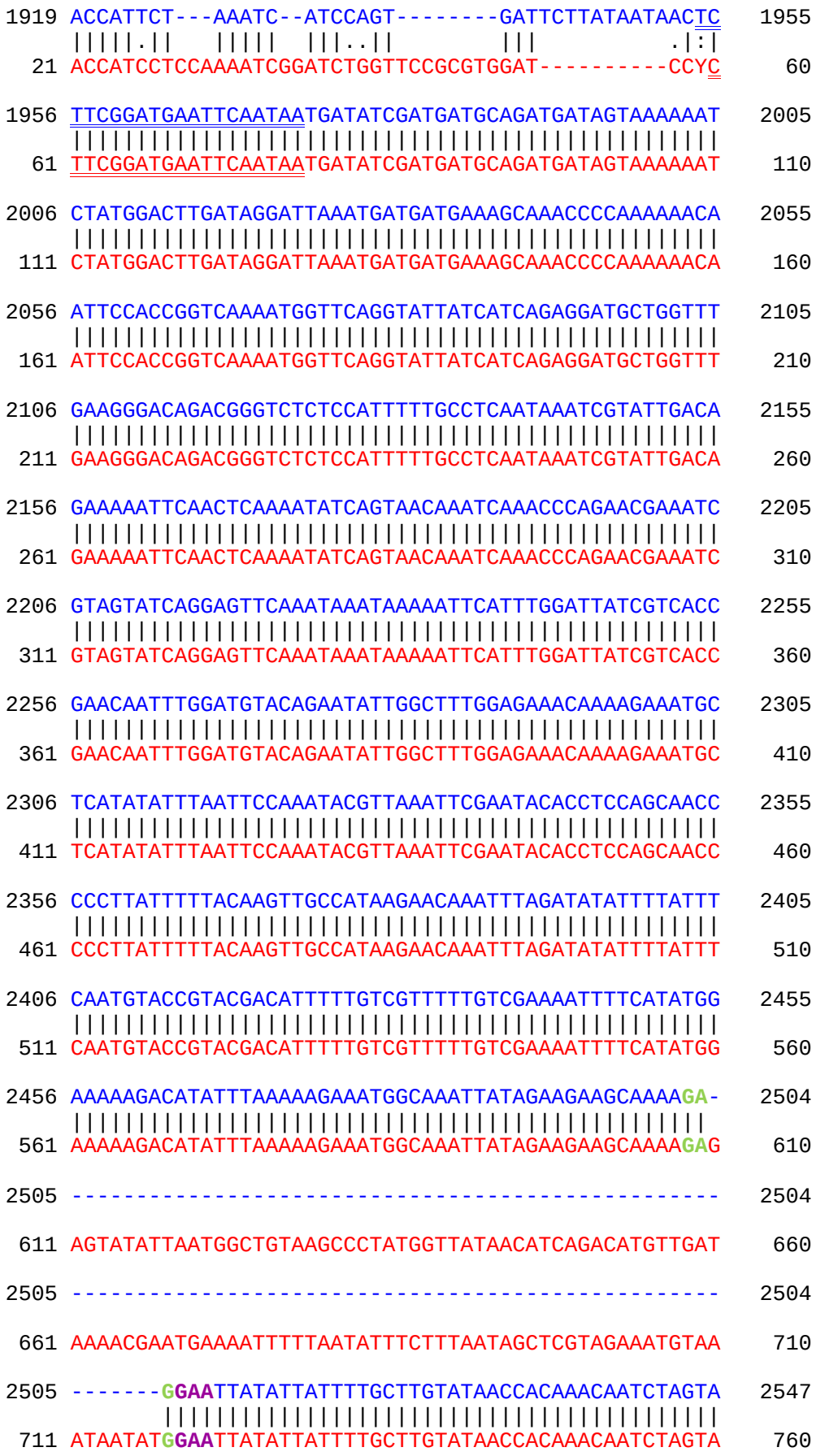


Figure 36. Section of AP1 C-terminal vector construct chromatogram

A vector construct from colony 1 was sequenced with a 5' pGEX 4T-2 primer. The P. falciparum sequence begins at base pair 86, highlighted in blue, with the reverse primer 5'- CTG AGT TAC ACT TAA GGA AAA -3'.

On aligning the subcloned region with the original sequence obtained from PlasmoDB it was found that the subcloned sequence contained an additional 108bp DNA insert (*Figure 37*). On examination of the intron-exon structure of the gene, it was determined that intron retaining mode of alternative splicing had occurred between exon 5 and 6. This indicates the existence of an alternative isoform of the *PFE1400c* gene that has not been documented on PlasmoDB to date. Apart from this, the two sequences aligned perfectly.


1919 ACCATTCT---AAATC--ATCCAGT--------GATTCTTATAATAACTC 1955
21 ACCATCCTCCAAAATCGGATCTGGTTCCGCGTGGAT----------CCYC 60
1956 TTCGGATGAATTCAATAATGATATCGATGATGCAGATGATAGTAAAAAAT 2005
61 TTCGGATGAATTCAATAATGATATCGATGATGCAGATGATAGTAAAAAAT 110
2006 CTATGGACTTGATAGGATTAAATGATGATGAAAGCAAACCCCAAAAAACA 2055
111 CTATGGACTTGATAGGATTAAATGATGATGAAAGCAAACCCCAAAAAACA 160
2056 ATTCCACCGGTCAAAATGGTTCAGGTATTATCATCAGAGGATGCTGGTTT 2105
161 ATTCCACCGGTCAAAATGGTTCAGGTATTATCATCAGAGGATGCTGGTTT 210
2106 GAAGGGACAGACGGGTCTCTCCATTTTTGCCTCAATAAATCGTATTGACA 2155
211 GAAGGGACAGACGGGTCTCTCCATTTTTGCCTCAATAAATCGTATTGACA 260
2156 GAAAAATTCAACTCAAAATATCAGTAACAAATCAAACCCAGAACGAAATC 2205
261 GAAAAATTCAACTCAAAATATCAGTAACAAATCAAACCCAGAACGAAATC 310
2206 GTAGTATCAGGAGTTCAAATAAATAAAAATTCATTTGGATTATCGTCACC 2255
311 GTAGTATCAGGAGTTCAAATAAATAAAAATTCATTTGGATTATCGTCACC 360
2256 GAACAATTTGGATGTACAGAATATTGGCTTTGGAGAAACAAAAGAAATGC 2305
361 GAACAATTTGGATGTACAGAATATTGGCTTTGGAGAAACAAAAGAAATGC 410
2306 TCATATATTTAATTCCAAATACGTTAAATTCGAATACACCTCCAGCAACC 2355
411 TCATATATTTAATTCCAAATACGTTAAATTCGAATACACCTCCAGCAACC 460
2356 CCCTTATTTTTACAAGTTGCCATAAGAACAAATTTAGATATATTTTATTT 2405
461 CCCTTATTTTTACAAGTTGCCATAAGAACAAATTTAGATATATTTTATTT 510
2406 CAATGTACCGTACGACATTTTTGTCGTTTTTGTCGAAAATTTTCATATGG 2455
511 CAATGTACCGTACGACATTTTTGTCGTTTTTGTCGAAAATTTTCATATGG 560
2456 AAAAAGACATATTTAAAAAGAAATGGCAAATTATAGAAGAAGCAAAAGA- 2504
561 AAAAAGACATATTTAAAAAGAAATGGCAAATTATAGAAGAAGCAAAAGAG 610
2505 ---------- 2504
611 AGTATATTAATGGCTGTAAGCCCTATGGTTATAACATCAGACATGTTGAT 660
2505 ---------- 2504
661 AAAACGAATGAAAATTTTTAATATTTCTTTAATAGCTCGTAGAAATGTAA 710
2505 -------GGAATTATATTATTTTGCTTGTATAACCACAAACAATCTAGTA 2547
711 ATAATATGGAATTATATTATTTTGCTTGTATAACCACAAACAATCTAGTA 760

```
2548 ATATTAAGTGAAGTCACAATTCAACCTGAGAAAAAGAACGTGAAATTGTG   2597
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 761 ATATTAAGTGAAGTCACAATTCAACCTGAGAAAAAGAACGTGAAATTGTG    810

2598 CATTCGTACAGATTCCTCATCTGTGATACCACTTTATAAGCTTTTATTTG   2647
     ||||||||||||||||||||||||||||||||||||||||||||||||||
 811 CATTCGTACAGATTCCTCATCTGTGATACCACTTTATAAGCTTTTATTTG    860

2648 TTAAGGCATTTTCCTTAAGTGTAACTCAG   2676
     ||||||||||||||||||||||||||||
 861 TTAAGGCATTTTCCTTAAGTGTAACTCAG    889
```

Figure 37. Alignment of AP1 C-terminal DNA sequences from 5' sequencing

The sequence obtained from subcloning (red) was aligned with the original sequence obtained from the PlasmoDB database (blue). The double underlined section indicates the forward primer position, while the single underlined section indicates the reverse primer. The 108bp insert indicates the inclusion of the intron between exons 5 and 6. The codon demarcating the end of exon 5 is coloured green (blue line 2 456). The codon distinguishing the start of codon 6 is coloured purple (blue line 2 505). Key to ambiguous bases: Y = C and T overlap. A dot between base pairs indicates a non-alignment, while a dash indicates a missing base.

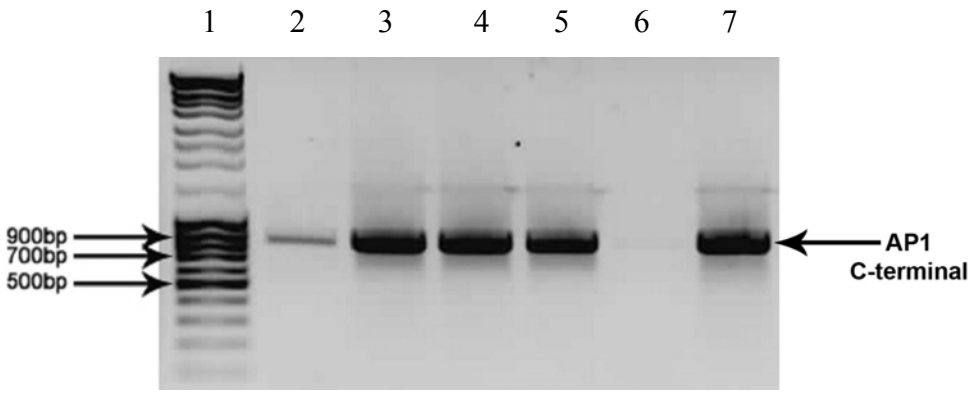
The 108bp insert codes for the following 36 amino acids:

SILMAVSPMVITSDMLIKRMKIFNISLIARRNVNNM.

This increases the length of rGST AP1 C-terminal to 496 amino acids and the molecular weight to 56.8 kD. The theoretical pI changes to 5.81.

3.6.3) Verification of AP1 C-terminal vector construct in Rosetta 2 cells

AP1 C-terminal vector constructs were verified by PCR of Rosetta 2 (DE3) colonies with *P. falciparum* gene specific primers (*Figure 38*). The insert migrated closer to the 800bp marker, corroborating the sequencing data which indicated an actual size of 831bp. This can be seen in lanes 3, 4, 5 and 7 (*Figure 38*), indicating that transformation was successful.



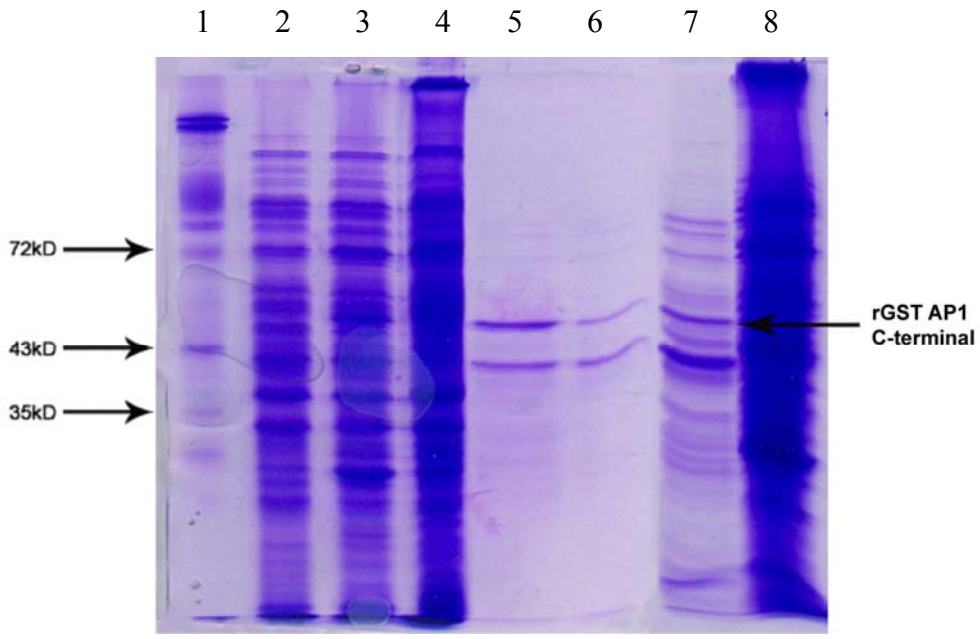
1 – Mass-Ruler™ DNA ladder mix
2 – Amplified P. falciparum DNA (control)
3-8 – Amplified vector constructs, colonies 1-6

Figure 38. Amplified Rosetta vector constructs resolved on 1% agarose gel

Four out of five Rosetta 2 (DE3) colonies were positive for the AP1 C-terminal insert (lanes 3-8).

3.6.4) Laemmli SDS-PAGE of rGST AP1 C-terminal

The estimated molecular weight of the rGST AP1 C-terminal was 53.2kD. However, with the extra 36 amino acids coded for by the 108bp insert, the molecular weight increases to 56.8kD. Soluble protein of the estimated molecular weight range was detected in the eluted fractions as seen in lanes 5-7 of figure 39.



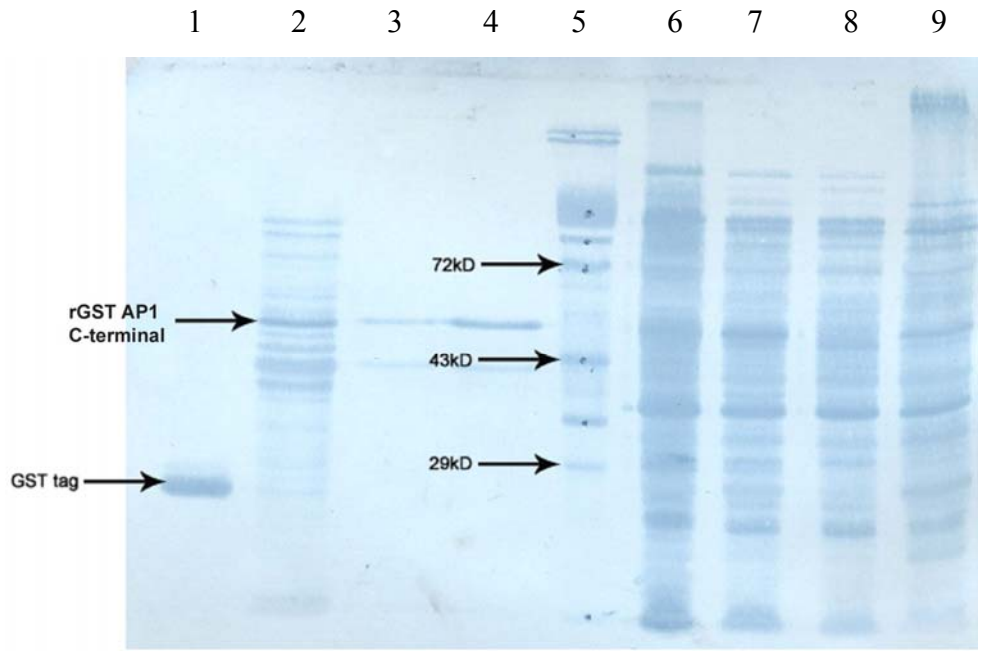
1 – RBC membrane marker
2 – AP1 C-terminal total uninduced
3 – AP1 C-terminal total induced
4 – AP1 C-terminal insoluble
5 – AP1 C-terminal 1st elution
6 – AP1 C-terminal 2nd elution
7 – AP1 C-terminal stripped beads
8 – AP1 C-terminal unbound fraction

Figure 39. Purification of rGST AP1 C-terminal protein

The protein fractions obtained during extraction and purification from Rosetta 2 (DE3) cells were resolved on a 12% SDS polyacrylamide gel. The RBC membrane marker (lane 1) consists of 3 indicated bands: Protein 4.2 – 72kD; Actin – 43kD; Glyceraldehyde 3-phosphate dehydrogenase (G3PD) – 35kD.

3.6.5) Western blot of rGST AP1 C-terminal

Proteins were transferred from the SDS acrylamide gel (*Figure 39*) onto a nitrocellulose membrane (*Figure 40*). The membrane was immunoblotted with an anti-GST primary antibody and subsequently stained in amido black. Soluble AP1 C-terminal protein of the estimated molecular weight range was confirmed in lanes 2-4.



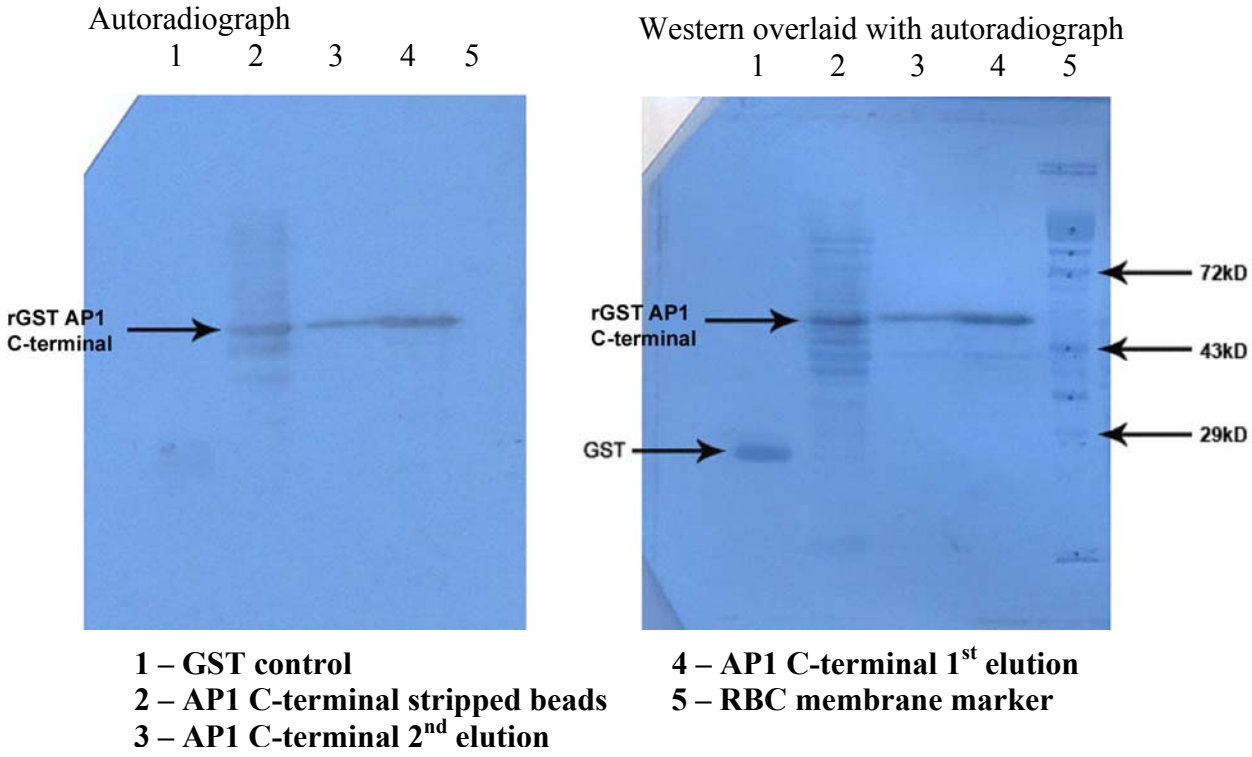
1 – GST control
2 – AP1 C-terminal stripped beads
3 – AP1 C-terminal 2nd elution
4 – AP1 C-terminal 1st elution
5 – RBC membrane marker
6 – AP1 C-terminal insoluble fraction
7 – AP1 C-terminal total induced
8 – AP1 C-terminal total uninduced
9 – AP1 C-terminal unbound fraction

Figure 40. rGST AP1 C-terminal Western blot

The Western blot was stained in amido black to visualise proteins. Bands of the expected size were obtained in the eluted and stripped fractions (lanes 2, 3 and 4), indicating that recombinant target protein was present in soluble form. Lane 1 contained GST expressed from E. coli cells containing non-recombinant pGEX 4T-2 plasmids. The RBC membrane marker (lane 5) has 3 bands labelled: Protein 4.2 – 72kD; Actin – 43kD; Stomatin/tropomyosin – 29kD. An anti-GST primary antibody was used for immunoblotting.

3.6.6) Immunoblot of rGST AP1 C-terminal

The presence of soluble protein in the estimated molecular weight range for the rGST AP1 C-terminal was confirmed by an autoradiograph (*Figure 41; lanes 2-4*). The chemiluminescent immunoblot – which had been probed with an HRP-linked anti-GST primary antibody – was exposed to x-ray film.



1 – GST control
2 – AP1 C-terminal stripped beads
3 – AP1 C-terminal 2nd elution
4 – AP1 C-terminal 1st elution
5 – RBC membrane marker

Figure 41. Immunoblot of rGST AP1 C-terminal

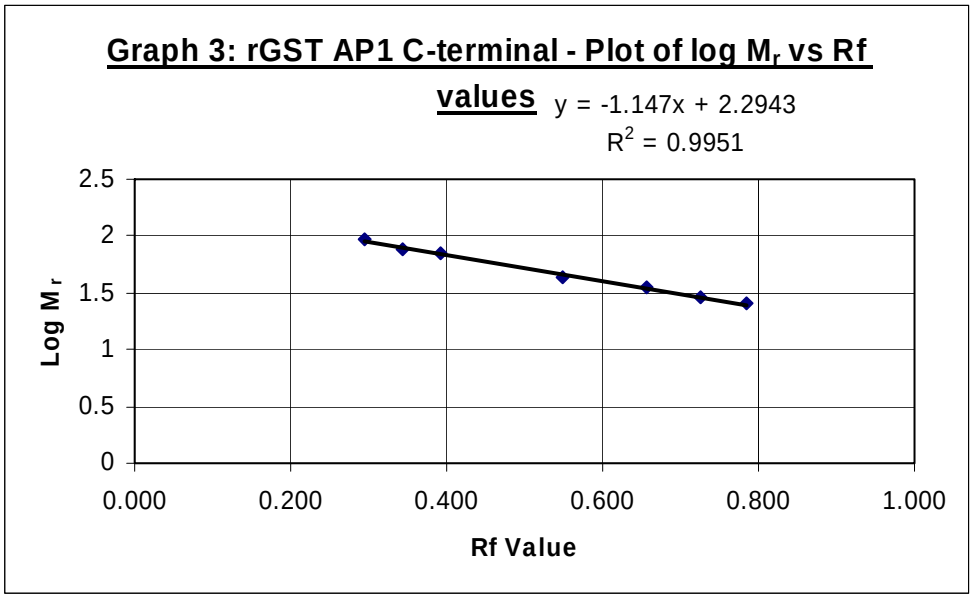
X-ray film was exposed to chemiluminescent immunoblots for 60 seconds to obtain autoradiographs. Lane 1 contained GST expressed from E. coli cells containing non-recombinant pGEX 4T-2 plasmids. The RBC membrane marker (lane 5) has 3 labelled bands: Protein 4.2 – 72kD; Actin – 43kD; Stomatin/tropo-myosin – 29kD.

The molecular weight of rGST AP1 C-terminal bands on the autoradiograph was determined via retardation factor (Rf) values of proteins making up the RBC membrane marker (*Table 18*).

TABLE 18: Data used to determine the M_r of rGST AP1 C-terminal

PROTEIN	M_r (kD)	Log M_r	MIGRATION DISTANCE (mm)	Rf VALUE
Band 3	93	1.9685	15	0.294
Protein 4.1	78	1.8921	17.5	0.343
Protein 4.2	72	1.8573	20	0.392
Actin	43	1.6335	28	0.549
G3PD	35	1.5441	33.5	0.657
7	29	1.4624	37	0.725
GST	25.5	1.4065	40	0.784
AP1 C-terminal (band on autoradiograph)	**53.96**	1.732	25	0.490
Dye front			51	

The molecular weight of the AP1 C-terminal as determined by the ExPASy ProtParam tool – including the extra 36 amino acids due to alternate splicing – was 56.8kD. The molecular weight extrapolated from Graph 3 is 53.96kD. The discrepancy in molecular weight values could be due to the anomalous size migration of *P. falciparum* proteins as described in section 3.5.6. The R^2 value indicates that the linear regression predicts the molecular weight with 99.51 percent accuracy.



3.6.7) Purity of rGST AP1 C-terminal

Densitometric scans of the first and second elution fractions were obtained from the SDS acrylamide gel of rGST AP1 C-terminal protein (*Figure 42*). The peaks detected for the first elution fraction are shown in figure 43. The target protein is represented by peak 2, while peak 5 consists of a contaminating protein.

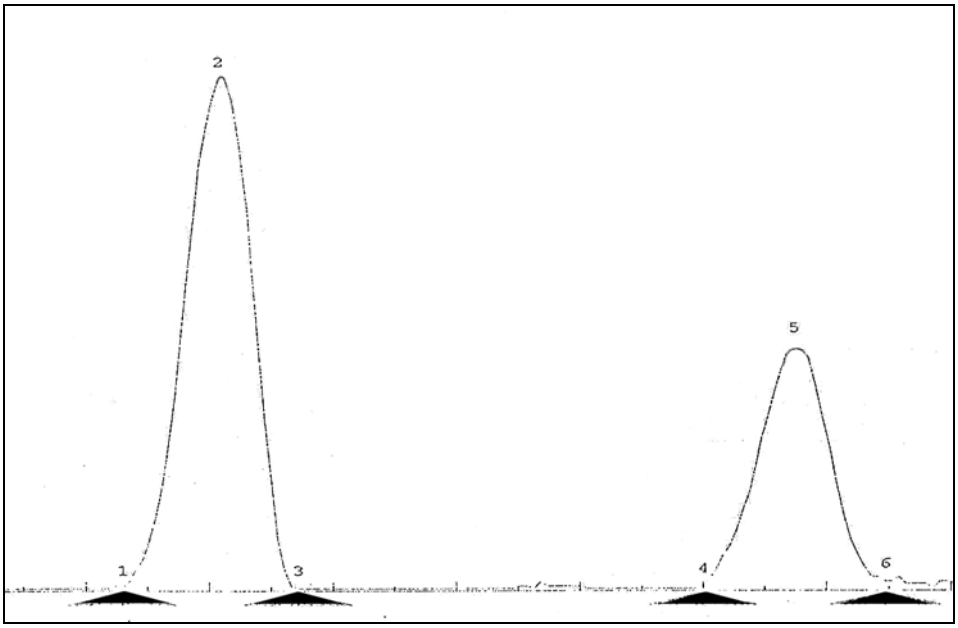


Figure 42. Densitometric scan of first elution of rGST AP1 C-terminal

Peak 2 consists of AP1 C-terminal protein with an estimated purity of 63.9 percent.

The first elution fraction had an approximate purity of 64 percent, as shown in table 19.

TABLE 19: Densitometric data for first elution of rGST AP1 C-terminal

PEAK	AREA	PERCENTAGE AREA
2	2 600	63.9
5	1 466	36.1
TOTAL	4 066	100

The peaks detected for the second elution fraction are shown in figure 43. The target protein is represented by peak 4, which represents about 29 percent of the total protein (*Table 20*).

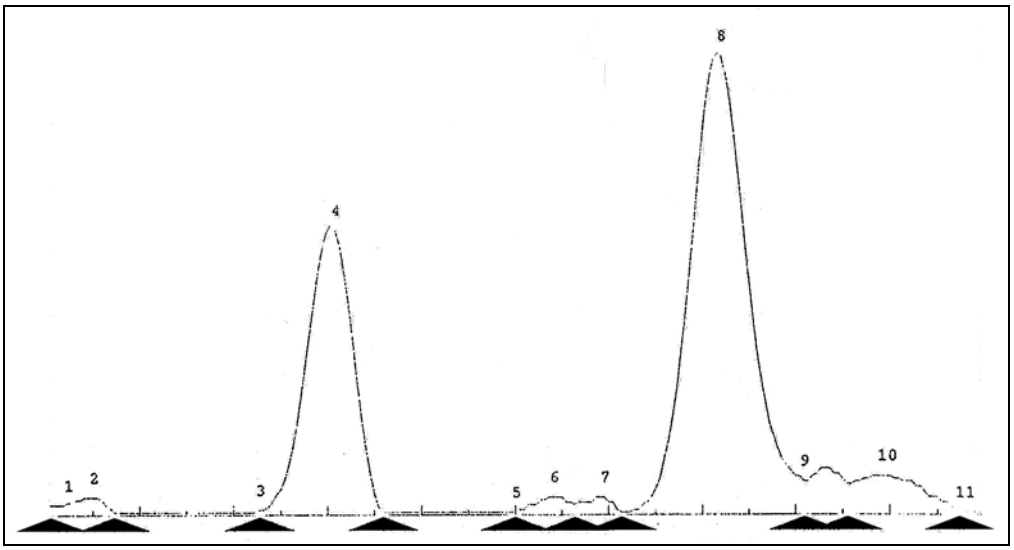


Figure 43. Densitometric scan of second elution of rGST AP1 C-terminal

The seven peak areas are represented by numbers 2, 4, 6, 7, 8, 9 and 10, and were determined to represent significant amounts of protein. The target protein is represented by peak 4.

TABLE 20: Densitometric data for second elution of rGST AP1 C-terminal

PEAK	AREA	PERCENTAGE AREA
2	89	1.8
4	1 441	28.7
6	103	2
7	58	1.2
8	2 869	57.1
9	169	3.4
10	296	5.8
TOTAL	5 025	100

The reduced purity of the second elution fraction is due to the increased levels of contaminating proteins, some of which may be *E. coli* proteins with GST-like sequences. These contaminants would compete with the target protein for binding to the glutathione-coated magnetic beads.

3.6.8) Bioinformatic data for *P. falciparum* AP1 C-terminal

3.6.8.1) Structure of *P. falciparum* AP1 C-terminal

The Swiss Model Template Library was used to find three dimensional models of proteins to which the AP-1 β subunit had structural homology, as no 3D model was available on the PlasmoDB website. The whole protein was compared to templates in the library and two hits resulted. Two separate templates were found that matched the C-terminal and the N-terminal of the whole protein, respectively. The AP1 C-terminal sequence had homology to template 1e42A (*Figure 44*). This template is human β2 appendage domain from clathrin adaptor AP-2. The target-template pair shares sequence identity of 25 percent, indicating that errors in the sequence alignment algorithms may have occurred because the target-template pair has less than 40 percent identity (*Appendix, section 5.11; figure 74*).

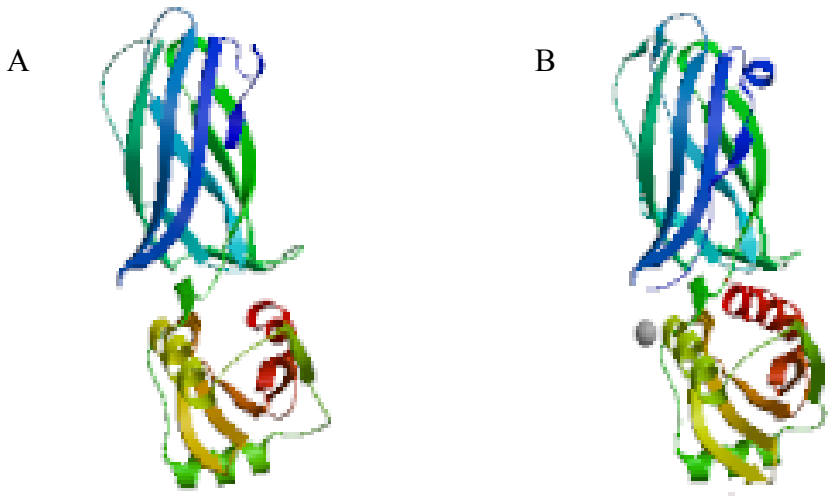


Figure 44. Predicted 3D model of AP1 C-terminal

The model (A) is based on template 1e42A (B), the human β2 appendage domain from clathrin adaptor AP-2, which had most sequence identity to the AP1 C-terminal sequence. The residues modeled span amino acids 694-911.

The predicted 3D model of the parasite AP1 C-terminal consists mostly of beta sheets, with three alpha helices (*Figure 44*). However, no beta sheets can be seen in the Heldwein model (*Figure 58).* The fact that the template did not accurately predict the structure of the *P. falciparum* AP1 C-terminal, due to the low sequence identity between the target-template pair, could account for this discrepancy. The area of the AP1 C-terminal that was modeled is indicated in figure 45.

A.

B.

NNSSDEFNNDIDDADDSKKSMDLIGLNDDESKPQKTIPPVKMVQ**VLSSEDAGLKGQTGLSIFASINRIDRKI QLKISVTNQTQNEIVVSGVQINKNSFGLSSPNNLDVQNIGFGETKEMLIYLIPNTLNSNTPPATPL FLQVAIRTNLDIFYFNVPYDIFVVFVENFHMEKDIFKKKWQIIEEAKEELYYFACITTNNLVILSE VTIQPEKKNVKLCIRTDSSSVIPLYK**LLFVKAFSLSVTQT

Figure 45. Homology of AP1 C-terminal to 3D template

In A, the blue and red segment starting at amino acid 694 and ending at 911 indicates the area where the AP1 C-terminal shares homology with template 1e42A. 25 percent of this segment shares sequence identity with template 1e42A. The blue bar is the AP1 N-terminal. The green segment is the whole AP1 protein of 925 amino acids.
In B, the AP1 C-terminal protein sequence is shown. The larger, bold type indicates the area with homology to template 1e42A.

The secondary structure of the AP1 C-terminal, as predicted using the Jnet secondary structure prediction algorithm (*Cuff and Barton, 2000*), showed that the principal conformation was that of beta sheets, with a small percentage of alpha helices (*Appendix, section 5.10, Figure 71*). This is reflected in the 3D model (*Figure 44*), which consists overwhelmingly of beta sheets.

3.6.8.2) Homology of *P. falciparum P. falciparum* AP-1 β subunit to other proteins

The whole putative AP-1 β subunit protein sequence (925aa) was compared to the proteomes of *P. knowlesi*, *P. vivax*, *P. chabaudi*, *P. berghei* and *P. yoelii*. Five proteins were found with high sequence identity to the *P. falciparum* sequence and these are listed as paralogues on the PlasmoDB website, indicating that the proteins are derived from the same ancestral gene. The results are presented in table 21.

TABLE 21: Sequence identity of *P. falciparum* AP-1 β subunit with other *Plasmodium* proteins

ORGANISM	PRODUCT	SEQUENCE IDENTITY	SEQUENCE POSITIVITY	E-VALUE
P. vivax	*PVX079845*; putative adapter-related protein complex 1 β 1 subunit	89 percent	95 percent	0
P. knowlesi	*PKH100410*; putative β adaptin protein	89 percent	95 percent	0
P. chabaudi	*PC000299.01.0*; putative β adaptin	83 percent	91 percent	1.6e-283
P. yoelii	*PY01282*; putative β adaptin-like protein	82 percent	91 percent	0
P. berghei	*PB000013.00.0*; putative β adaptin	81 percent	90 percent	0

The BLAST data for the *P. vivax* protein is shown in the appendix (*PVX079845, Appendix, section 5.10, Figure 67*).

No paralogues to the *P. falciparum* putative AP-1 β subunit were found in the *T. gondii* proteome, although several proteins showed sequence identity, as indicated in table 22. The protein that had the highest homology was *T. gondii* putative β adaptin protein, with an identity of 56 percent and positivity of 74 percent over the whole protein (*49.m00005, Appendix, section 5.9, Figure 68*).

The other *T. gondii* proteins had most sequence identity in the N-terminal region of the *P. falciparum* AP-1 β1 subunit.

TABLE 22: Sequence identity of *P. falciparum* AP-1 β subunit with *T.gondii* proteins

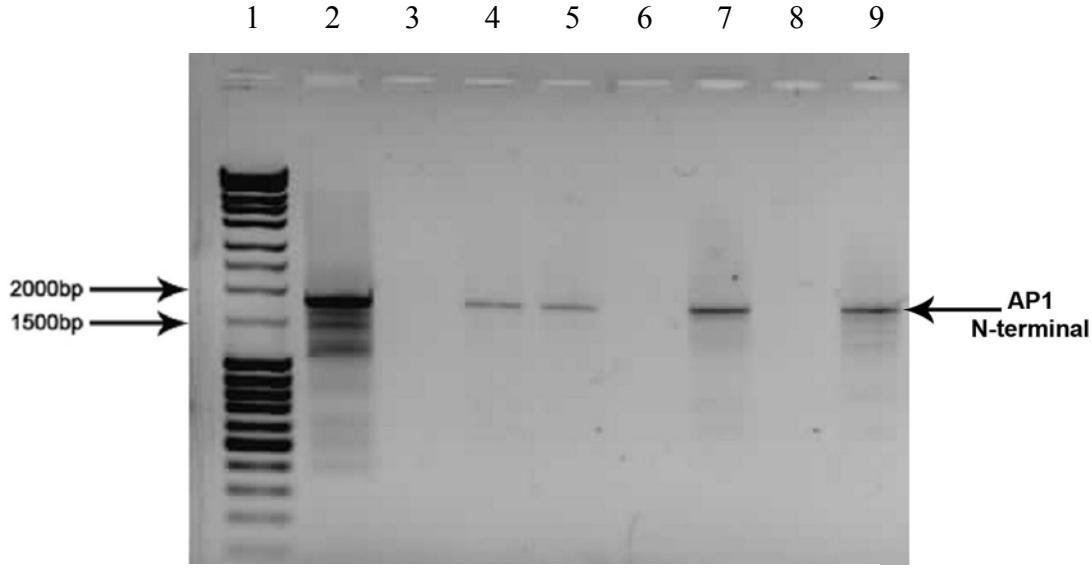
PRODUCT	SEQUENCE IDENTITY (percent)			SEQUENCE POSITIVITY (percent)			E-VALUE
	Whole protein	C-terminal	N-terminal	Whole protein	C-terminal	N-terminal	
49.m00005; putative β adaptin protein (925aa)	56	36	66	74	60	82	9.6e-274
57.m01782; putative β adaptin-like protein (915aa)	20	None	30	37	None	56	1.1e-82
80.m02192; putative adapter-related protein complex 3 β 2 subunit (1 187aa)	7	None	26	13	None	51	6.4e-34
57.m01829; putative β coat protein (1 104aa)	13	26	21	26	45	45	1.2e-16

The *P. falciparum* putative AP-1 β1 subunit protein sequence was compared to the human proteome. The orthologous *H. sapiens* AP-1 β1 subunit, found on chromosome 22, was the best hit with 46 percent identity and 62 percent positivity to the whole protein; 63 percent identity and 81 percent positivity with the parasite N-terminal protein sequence and a much lower 28 percent identity and 52 percent identity within the C-terminal region (E-value 2.0e-224) (*Appendix, section 5.9, figure 69*). The human protein is 949 amino acids in length compared to the 925 amino acid parasite protein sequence. The BLAST data indicate that the N-terminal of the AP-1 β subunit is more highly conserved than the C-terminal.

3.7) Results for *P. falciparum* AP1 N-terminal

3.7.1) Verification of AP1 N-terminal vector construct

The presence of PCR product (*lanes 4, 5, 7 and 9; Figure 46*) with the same length as the control – which contained amplified *P. falciparum* DNA – indicated that the AP1 N-terminal insert was present and that the transformation had been successful in DH5α colonies 2, 3, 5 and 7 (*Figure 46*). The expected size of the insert is 1 800bp, and the AP1 N-terminal band resolved between the 1 500bp and 2 000bp DNA markers as expected.



1 – MassRuler™ DNA ladder mix
2 – Amplified *P. falciparum* genomic DNA
3–9 Amplified vector constructs from colonies 1-

Figure 46. Digested AP1 N-terminal vector constructs from DH5α cells

The AP1 N-terminal insert was excised by digestion with the same restriction endonucleases used for the plasmid preparation (section 3.4.2) and was resolved on a 1% agarose gel.

3.7.2) AP1 N-terminal vector construct sequence data

Successful sequencing was carried out using the T7 terminator (*Appendix*) and promoter primers (*Figure 47*). The T7 promoter primer chromatogram for the AP1 N-terminal vector construct sequence demonstrated high amplitude peaks with minimal background, as seen in figure 47.

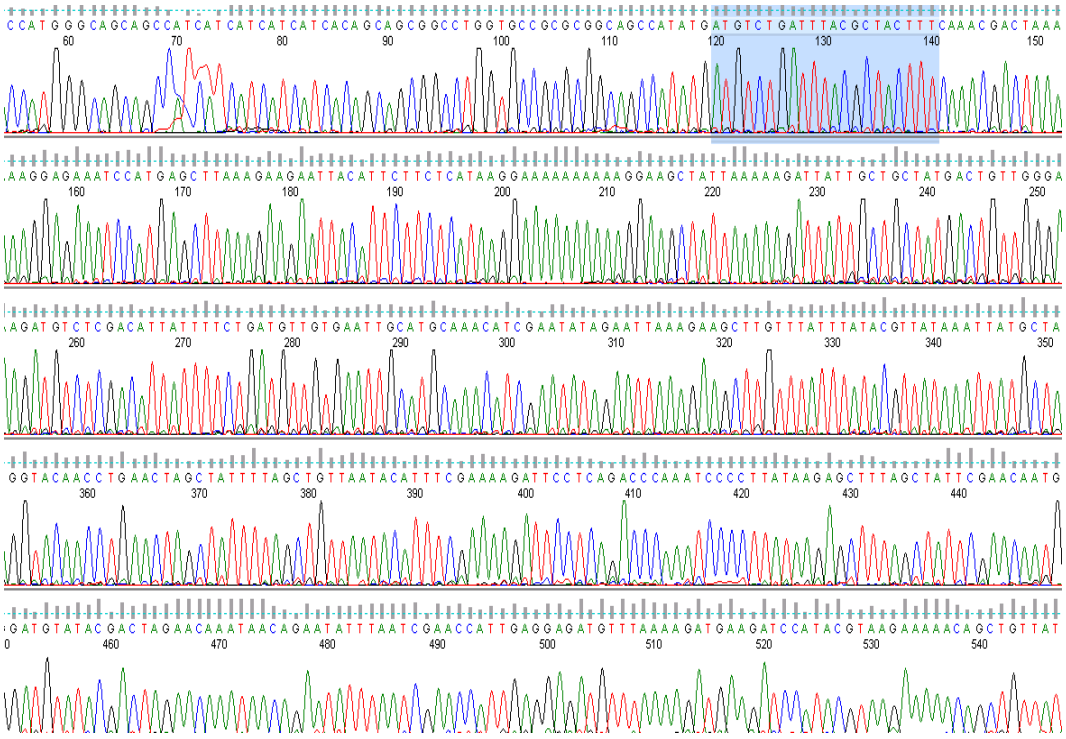
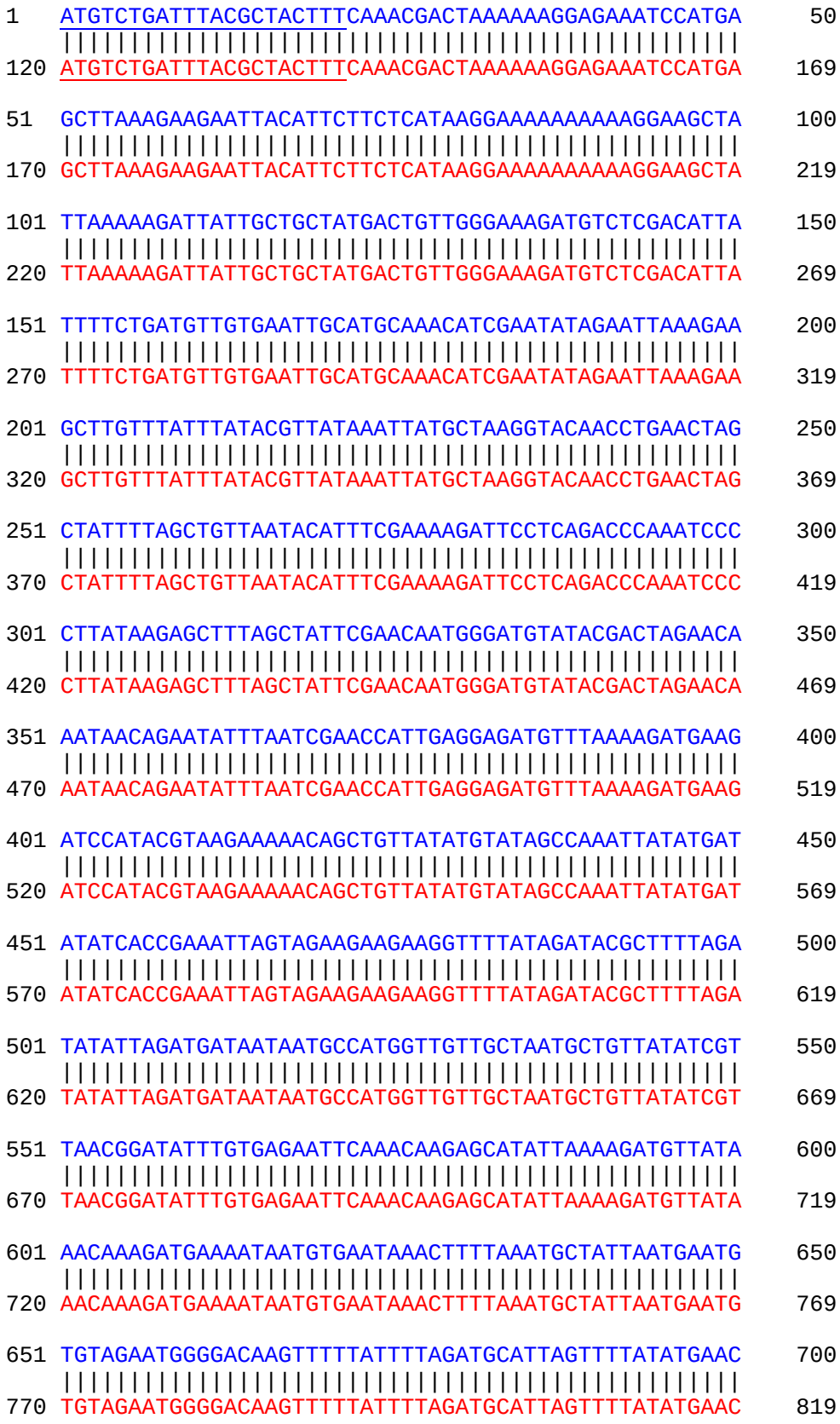


Figure 47. Section of AP1 N-terminal vector construct chromatogram

A vector construct from colony 5 was sequenced with a T7 promoter primer. The P. falciparum sequence begins at base pair 120, highlighted in blue, with the forward primer 5'- ATG TCT GAT TTA CGC TAC TTT -3'.

To confirm that the subcloned sequence was identical to the sequence obtained from the PlasmoDB website, the two sequences were aligned (*Figure 48*).


1 ATGTCTGATTTACGCTACTTTCAAACGACTAAAAAAGGAGAAATCCATGA 50
120 ATGTCTGATTTACGCTACTTTCAAACGACTAAAAAAGGAGAAATCCATGA 169
51 GCTTAAAGAAGAATTACATTCTTCTCATAAGGAAAAAAAAAAGGAAGCTA 100
170 GCTTAAAGAAGAATTACATTCTTCTCATAAGGAAAAAAAAAAGGAAGCTA 219
101 TTAAAAAGATTATTGCTGCTATGACTGTTGGGAAAGATGTCTCGACATTA 150
220 TTAAAAAGATTATTGCTGCTATGACTGTTGGGAAAGATGTCTCGACATTA 269
151 TTTTCTGATGTTGTGAATTGCATGCAAACATCGAATATAGAATTAAAGAA 200
270 TTTTCTGATGTTGTGAATTGCATGCAAACATCGAATATAGAATTAAAGAA 319
201 GCTTGTTTATTTATACGTTATAAATTATGCTAAGGTACAACCTGAACTAG 250
320 GCTTGTTTATTTATACGTTATAAATTATGCTAAGGTACAACCTGAACTAG 369
251 CTATTTTAGCTGTTAATACATTTCGAAAAGATTCCTCAGACCCAAATCCC 300
370 CTATTTTAGCTGTTAATACATTTCGAAAAGATTCCTCAGACCCAAATCCC 419
301 CTTATAAGAGCTTTAGCTATTCGAACAATGGGATGTATACGACTAGAACA 350
420 CTTATAAGAGCTTTAGCTATTCGAACAATGGGATGTATACGACTAGAACA 469
351 AATAACAGAATATTTAATCGAACCATTGAGGAGATGTTTAAAAGATGAAG 400
470 AATAACAGAATATTTAATCGAACCATTGAGGAGATGTTTAAAAGATGAAG 519
401 ATCCATACGTAAGAAAAACAGCTGTTATATGTATAGCCAAATTATATGAT 450
520 ATCCATACGTAAGAAAAACAGCTGTTATATGTATAGCCAAATTATATGAT 569
451 ATATCACCGAAATTAGTAGAAGAAGAAGGTTTTATAGATACGCTTTTAGA 500
570 ATATCACCGAAATTAGTAGAAGAAGAAGGTTTTATAGATACGCTTTTAGA 619
501 TATATTAGATGATAATAATGCCATGGTTGTTGCTAATGCTGTTATATCGT 550
620 TATATTAGATGATAATAATGCCATGGTTGTTGCTAATGCTGTTATATCGT 669
551 TAACGGATATTTGTGAGAATTCAAACAAGAGCATATTAAAAGATGTTATA 600
670 TAACGGATATTTGTGAGAATTCAAACAAGAGCATATTAAAAGATGTTATA 719
601 AACAAAGATGAAAATAATGTGAATAAACTTTTAAATGCTATTAATGAATG 650
720 AACAAAGATGAAAATAATGTGAATAAACTTTTAAATGCTATTAATGAATG 769
651 TGTAGAATGGGGACAAGTTTTTATTTTAGATGCATTAGTTTTATATGAAC 700
770 TGTAGAATGGGGACAAGTTTTTATTTTAGATGCATTAGTTTTATATGAAC 819

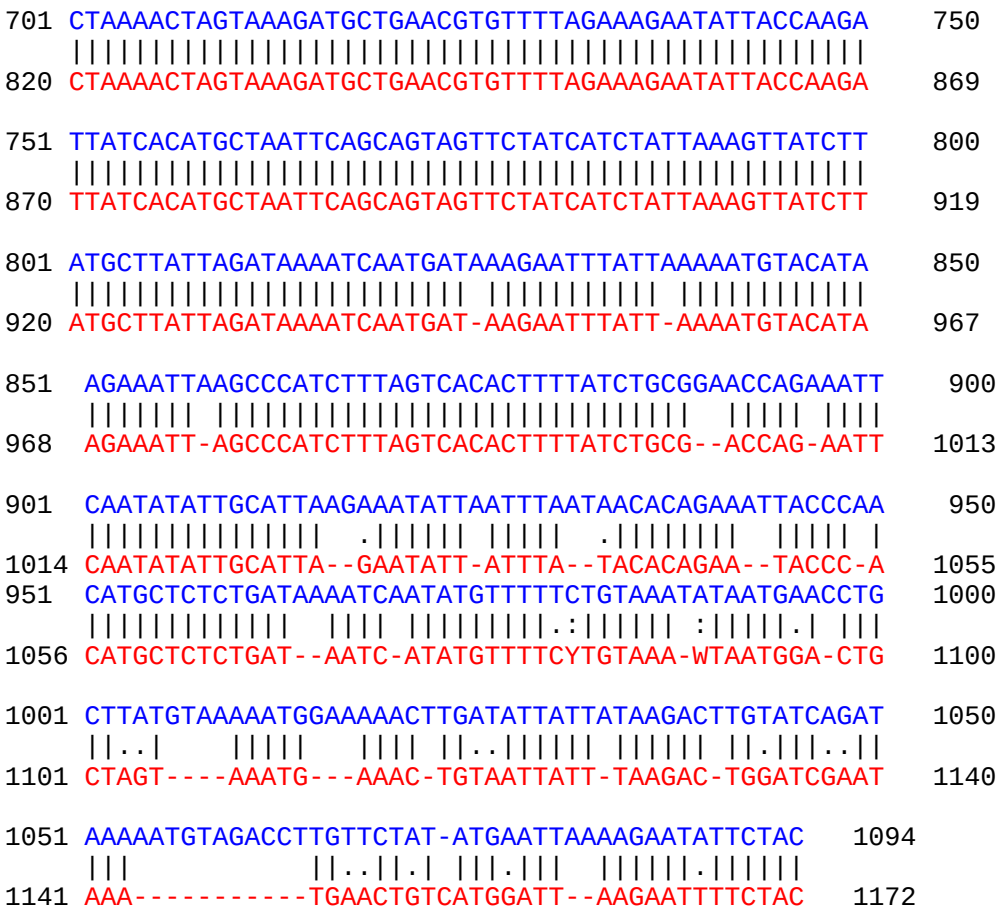


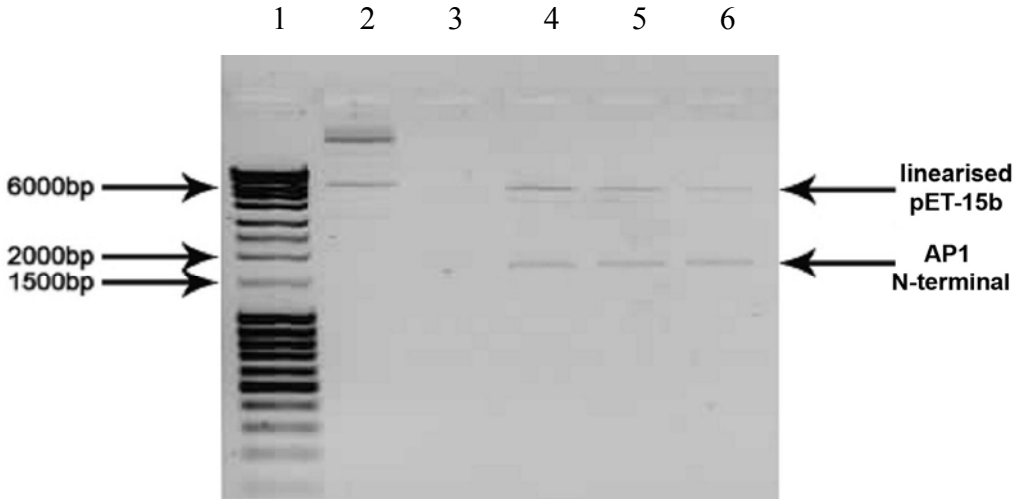
Figure 48. Alignment of AP1 N-terminal DNA sequences from T7 promoter primer sequencing

The sequence obtained from subcloning (red) was aligned with the original sequence obtained from the PlasmoDB database (blue). The underlined section indicates the P. falciparum forward primer position 5' - ATG TCT GAT TTA CGC TAC TTT -3'. Key to ambiguous bases: W = T and A overlap in sequence; Y = C and T overlap. A dot between base pairs indicates a non-alignment, while a dash indicates a missing base.

Discrepancies at the 3' end of the sequence were resolved by sequencing the reverse strand using the T7 terminator primer, which confirmed that the sequences aligned with 100 percent identity ((*Appendix, section 5.8 for T7 terminator primer sequence of AP1 N-terminal*).

3.7.3) Verification of AP1 N-terminal vector construct in Rosetta 2 cells

Successful transformations occurred in adaptin colonies 2, 3 and 4, as seen in lanes 4-6 of Figure 49, where two bands resolved on the gel after restriction endonuclease digestion of plasmids isolated from Rosetta 2 cells. The adaptin N-terminal domain is 1 800bp in length, and resolved to a position between the 1 500bp and 2 000bp standard markers.



1 – MassRuler™ DNA ladder mix
2 – pET-15b whole recombinant plasmid
3 – 6 Digested plasmids from colonies 1-4

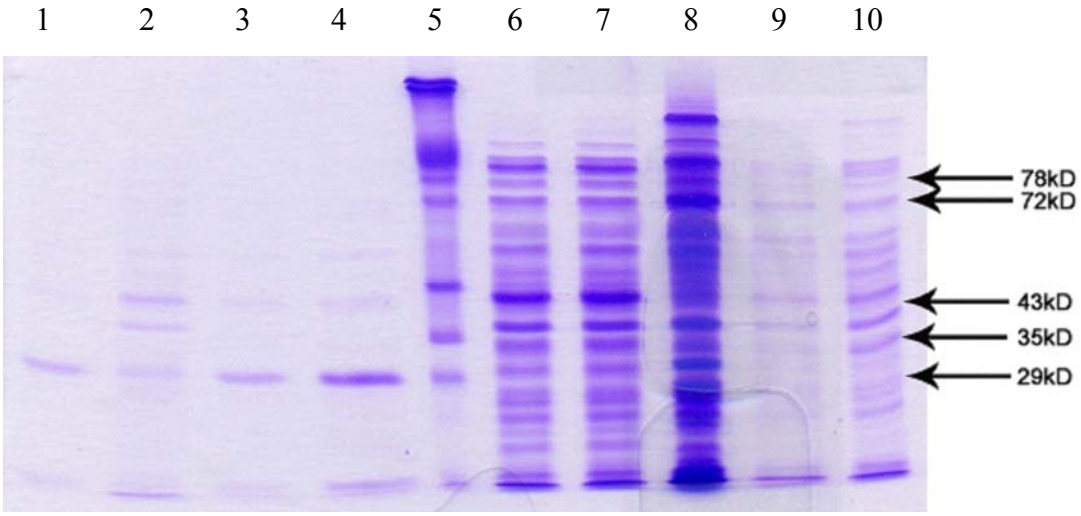
Figure 49. Digested Rosetta vector constructs resolved on 1% agarose gel

Rosetta 2 (DE3) colonies 2-4 were positive for the AP1 N-terminal insert (lanes 4-6). The restriction endonucleases used were the same as those used for digestion of plasmids (section 3.4.2).

3.7.4) Laemmli SDS-PAGE of rHis AP1 N-terminal

The rHis AP1 N-terminal protein has a predicted molecular weight of 72.2kD. Soluble protein of the correct molecular weight was not detected in the eluted fractions as seen in lanes 2-4 of figure 50. However, protein of the estimated correct molecular weight was observed in the unbound fraction (*lane 6*), soluble fraction (*lane 7*), insoluble fraction (*lane 8*) and the total induced protein fraction (*lane 10*).

It is possible that the protein band could consist of *E. coli* proteins and not the target parasite protein.



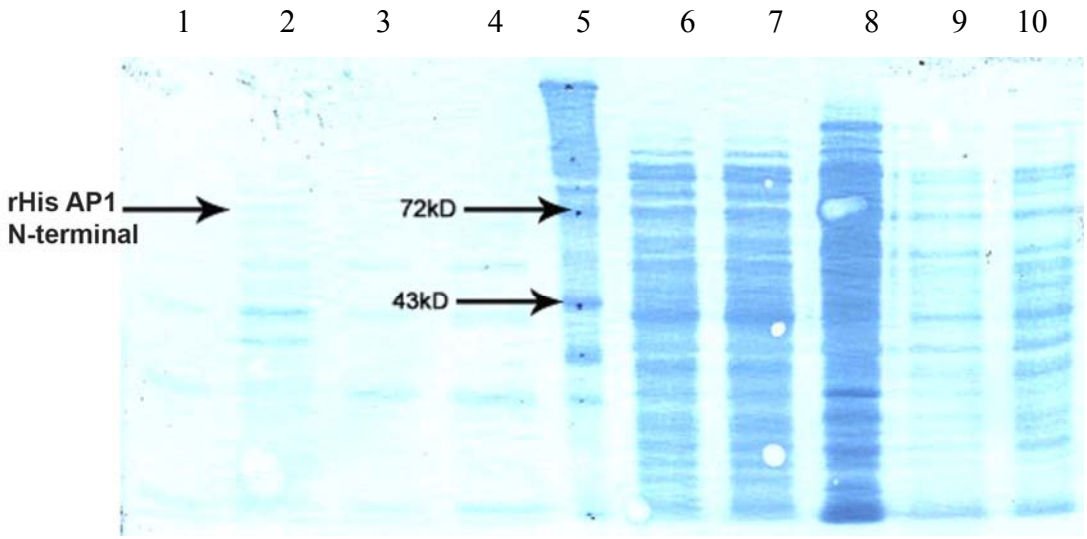
1 – His control
2 – AP1 N-terminal stripped beads
3 – AP1 N-terminal 2nd elution
4 – AP1 N-terminal 1st elution
5 – RBC membrane marker
6 – AP1 N-terminal unbound fraction
7 – AP1 N-terminal soluble fraction
8 – AP1 N-terminal insoluble fraction
9 – AP1 N-terminal total uninduced
10 – AP1 N-terminal total induced

Figure 50. Purification of rHis AP1 N-terminal protein

The protein fractions obtained during extraction and purification from Rosetta 2 cells were resolved on a 12% SDS polyacrylamide gel. The RBC membrane marker (lane 5) consists of 5 indicated bands: Protein 4.1 – 78kD; Protein 4.2 – 72kD; Actin – 43kD; Glyceraldehyde 3-phosphate dehydrogenase (G3PD) – 35kD; Stomatin/tropo-myosin – 29kD.

3.7.5) Western blot of rHis AP1 N-terminal

Proteins were transferred from the SDS acrylamide gel (*Figure 50*) onto a nitrocellulose membrane (*Figure 51*). The membrane was immunoblotted with an anti-His primary antibody and subsequently stained in amido black. A faint protein band within the correct molecular weight range for the AP1 N-terminal was seen in the stripped bead fraction (*Figure 51*, *lane 2*).



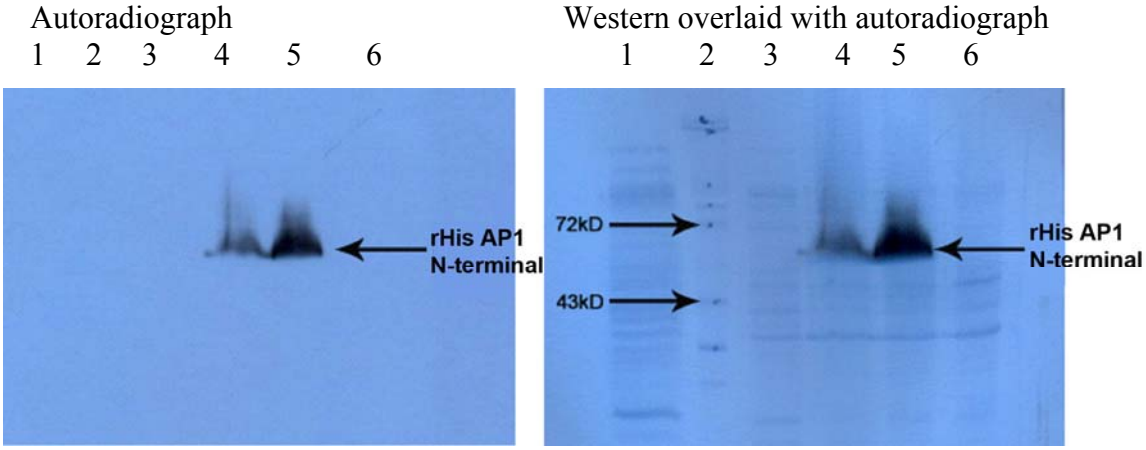
1 – His control
2 – AP1 N-terminal stripped beads
3 – AP1 N-terminal 2nd elution
4 – AP1 N-terminal 1st elution
5 – RBC membrane marker
6 – AP1 N-terminal unbound fraction
7 – AP1 N-terminal soluble fraction
8 – AP1 N-terminal insoluble fraction
9 – AP1 N-terminal total uninduced
10 – AP1 N-terminal total induced

Figure 51. rHis AP1 N-terminal Western blot

The Western blot was stained in amido black to visualise proteins. Bands of the expected size were faintly observed in the stripped fraction (lane 2), and in the unbound, soluble, insoluble and total fractions (lanes 6, 7, 8 and 10). The RBC membrane marker (lane 5) has 2 bands labelled: Protein 4.2 – 72kD; Actin – 43kD. An anti-His primary antibody was used for immunoblotting.

3.7.6) Autoradiograph of rHis AP1 N-terminal

The chemiluminescent immunoblot was exposed to x-ray film, resulting in an autoradiograph. No soluble protein in the correct molecular weight range for the rHis AP1 N-terminal was detected in the eluted fractions (*not shown*). However, the anti-His antibody bound to protein in the insoluble and total protein fractions (*Figure 52, lanes 4 and 5*) indicating that insoluble inclusion bodies had formed. The molecular weight of the bands obtained on the autoradiograph appears lower that 72kD.



1 – AP1 N-terminal unbound fraction
2 – RBC membrane marker
3 – AP1 N-terminal soluble fraction
4 – AP1 N-terminal insoluble fraction
5 – AP1 N-terminal total induced
6 – AP1 N-terminal total uninduced

Figure 52. Immunoblot of rHis AP1 N-terminal

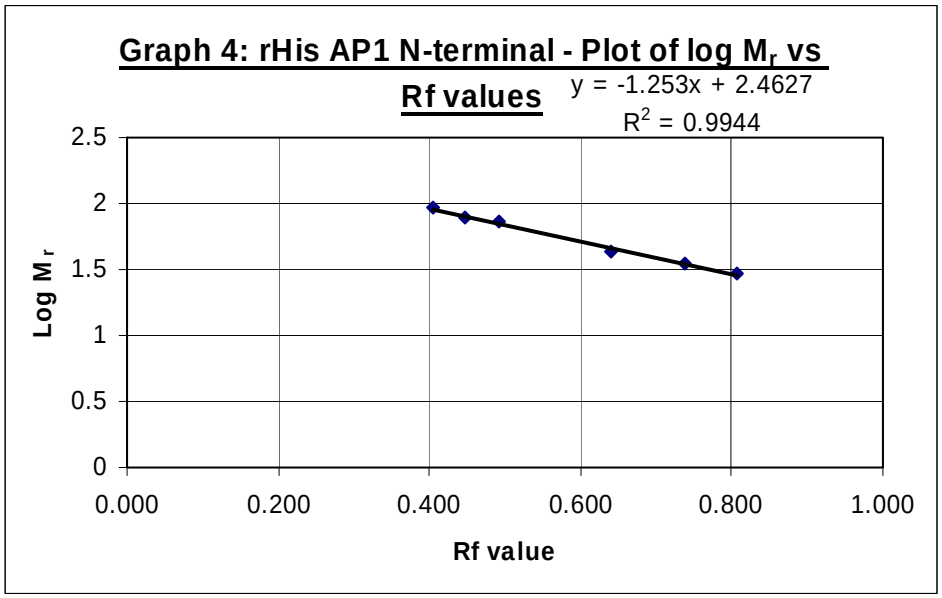
X-ray film was exposed to chemiluminescent immunoblots for 60 seconds to obtain autoradiographs. Protein was detected in lanes 4 and 5, but was below the predicted molecular weight of 72.2kD. The RBC membrane marker (lane 2) has 2 labelled bands: Protein 4.2 – 72kD; Actin – 43kD.

Using retardation factor (Rf) values of proteins making up the RBC membrane marker, the molecular weight of rHis AP1 N-terminal bands on the autoradiograph was determined (*Table 23*). The extrapolated value is shown in bold type.

TABLE 23: Data used to determine the M_r of rHis AP1 N-terminal

PROTEIN	M_r (kD)	Log M_r	MIGRATION DISTANCE (mm)	Rf VALUE
Band 3	93	1.9685	23	0.404
Protein 4.1	78	1.8921	25.5	0.447
Protein 4.2	72	1.8573	28	0.491
Actin	43	1.6335	36.5	0.640
G3PD	35	1.5441	42	0.737
7	29	1.4624	46	0.807
rHis AP1 N-terminal (band on autoradiograph)	**63.566**	1.8032	30	0.526
Dye front			57	

The molecular weight of adaptin that was extrapolated from Graph 4 is 63.6kD, which is much smaller that the estimated 72.2kD. With an R^2 value of 99.44 percent, the standard linear regression is accurate under conditions of normal protein size migration. However, due to the anomalous size migration of *P. falciparum* proteins, the value calculated for the parasite protein molecular weight is not exact, as described in section 3.5.6.



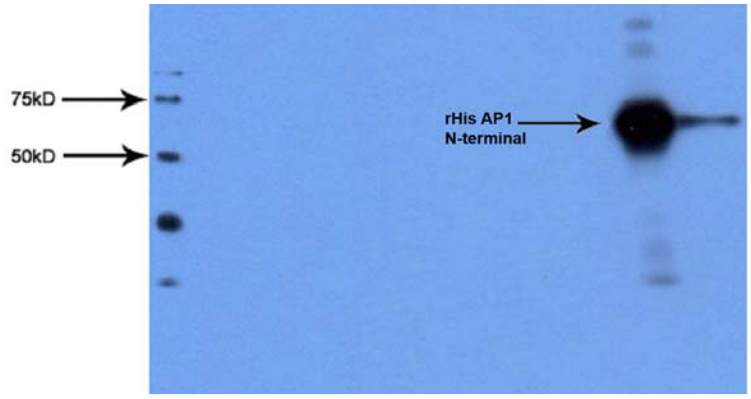
3.7.6.1) Western blot of rHis AP1 N-terminal extracted from inclusion bodies

Protein was extracted from inclusion bodies and refolded. The recovered protein was electrophoresed via SDS-PAGE (*not shown*) and then transferred onto a nitrocellulose membrane by means of Western blotting. The Western blot revealed protein bands in the estimated molecular weight range of 72.2kD in the soluble and insoluble fractions (*not shown*).

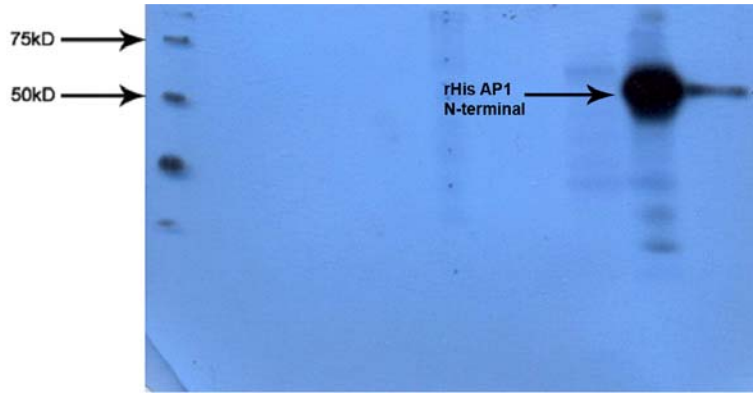
3.7.6.2) Autoradiograph of rHis AP1 N-terminal extracted from inclusion bodies

His-tagged proteins were detected in the insoluble fraction and the total induced fraction as seen in lanes 8 and 9 (*Figure 53*). These proteins appeared in the 50 – 75kD molecular weight range, according to the standard histidine ladder (*Lane 1*). This indicated that refolding of rHis AP1 N-terminal was not successful and the recombinant protein remained insoluble.

A.



B. 1 2 3 4 5 6 7 8 9



1 – 6x His ladder
2 – AP1 N-terminal stripped beads
3 – AP1 N-terminal 2nd elution
4 – AP1 N-terminal1st elution
5 – RBC membrane marker
6 – AP1 N-terminal unbound fraction
7 – AP1 N-terminal soluble fraction
8 – AP1 N-terminal insoluble fraction
9 – AP1 N-terminal total induced

Figure 53. Immunoblot of rHis AP1 N-terminal extracted from inclusion bodies

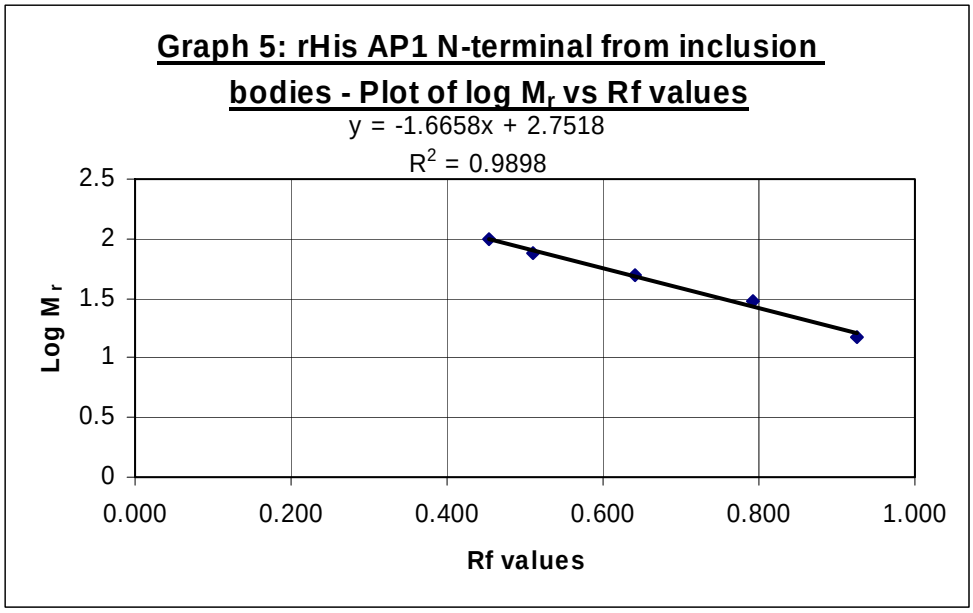
A is the autoradiograph. B is the Western blot overlaid with the autoradiograph. X-ray film was exposed to chemiluminescent immunoblots for 60 seconds to obtain autoradiographs. Protein of the estimated molecular weight for the AP1 N-terminal was detected in lanes 8 and 9. The 6X histidine ladder (lane 1) has 2 labelled bands: band 2 – 75kD; band 3 – 50D.

The retardation factor (Rf) values of proteins making up the RBC membrane marker were determined. The molecular weight of the rHis AP1 N-terminal bands were then determined from these values (*Table 24*). The extrapolated value is shown in bold type.

TABLE 24: Data used to determine the M_r of rHis AP1 N-terminal from insoluble fraction

6x HIS PROTEIN LADDER	M_r (kD)	Log M_r	MIGRATION DISTANCE (mm)	Rf VALUE
1	100	2	24	0.453
2	75	1.8750	27	0.509
3	50	1.6989	34	0.642
4	30	1.4771	42	0.792
5	15	1.1760	49	0.925
Adaptin inclusion bodies (x-ray film)	**64.401**	1.8088	30	0.566
Dye front			53	

The molecular weight of the rHis AP1 N-terminal that was extrapolated from Graph 5 is 64.4kD, which is much smaller that the predicted value of 72.2kD. This could be due to the same factors discussed in section 3.5.6. The R^2 value indicates that the standard linear regression predicts the molecular weight with 98.98 percent accuracy, under conditions of normal protein size migration.



3.7.7) Purity of rHis AP1 N-terminal

As recombinant adaptin was not expressed in a soluble form, densitometric analysis was not carried out for this protein.

3.7.8) Bioinformatic data for *P. falciparum* AP1 N-terminal

3.7.8.1) Structure of *P. falciparum* AP1 N-terminal

The Swiss Model Template Library was used to find three dimensional models of proteins to which the AP-1 β subunit had structural homology (*Section 3.6.8.1*). The AP1 N-terminal sequence had homology to template 2vglB (*Figure 54*). This template is human AP2 clathrin adaptor core. The target-template pair shares sequence identity of 63 percent, indicating that the template closely predicts the 3D structure of the parasite protein (*Appendix, section 5.11; figure 75*).This can be seen in figure 54, where the predicted model (A) has an almost identical structure to the template (B).

A B

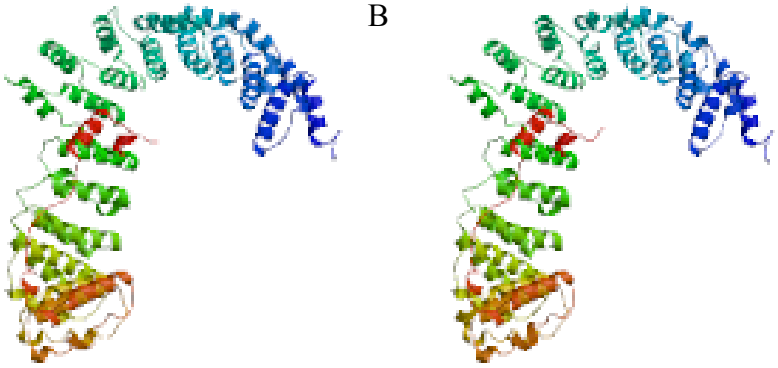


Figure 54. 3D model of AP1 N-terminal

Template 2vglB (B) is human AP-2 clathrin adaptor core, which had most sequence homology to the AP1 N-terminal sequence as modeled (A).

The putative model of the parasite protein *(figure 54, A*) consists mainly of alpha helices which is consistent with the crystal structure of AP-1 complexes as determined by Heldwein *et al (2004) (Figure 58).* The area of the AP1 N-terminal that shares sequence homology to the template is indicated in figure 55.

A.

B.

MSDLRYFQTTKK**GEIHELKEELHSSHKEKKKEAIKKIIAAMTVGKDVSTLFSDVVNCMQTSNIELKK**
LVYLYVINYAKVQPELAILAVNTFRKDSSDPNPLIRALAIRTMGCIRLEQITEYLIEPLRRCLKDE
DPYVRKTAVICIAKLYDISPKLVEEEGFIDTLLDILDDNNAMVVANAVISLTDICENSNKSILKDV
INKDENNVNKLLNAINECVEWGQVFILDALVLYEPKTSKDAERVLERILPRLSHANSAVVLSSIKV
ILCLLDKINDKEFIKNVHKKLSPSLVTLLSAEPEIQYIALRNINLITQKLPNMLSDKINMFFCKYN
EPAYVKMEKLDIIIRLVSDKNVDLVLYELKEYSTEVDVEFVKKSVRAIGSCAIKLPQSSEKCINIL
LDLIDTKINYVIQECIVVIKDIFRKYPNKYESIITILCENLESLDESNAKASLIWIIGEYVERIDN
ADELIDSFLENFSDEPYNVQLQILTASVKLFLKCSKNTKDIITKVLKLSTEESDNPDLRDRAYIYW
RLLSKNIDVAKKIVLADKPPIQEENKITDTKVLNKLIKNISMLSSVYHKLPETFISKKNSYSLNSDNN
NDHMQDDHYDDDDYDKDNHVL

Figure 55. Sequence homology of AP1 N-terminal to 3D template

In A, the blue and red segment starting at amino acid 12 and ending at 584 indicates the area where the AP1 N-terminal shares homology with template 2vglB. 63 percent of this segment has sequence identity to the template. The blue segment is the AP1 C-terminal. The green segment represents the whole AP-1 β subunit of 925 amino acids. In B, the protein sequence of the AP1 N-terminal is represented. The larger, bold black sequence represents the area sharing sequence homology with template 2vglB.

Using the Jnet secondary structure prediction algorithm (*Cuff and Barton, 2000*), the protein secondary structure of the AP1 N-terminal was predicted to consist mostly of alpha helices, with only one section of beta sheets (*Appendix, section 5.10, Figure 72*). These data are supported by the 3D model (*Figure 55*) which is predominantly made up of alpha helical conformations.

3.7.8.2) Homology of *P. falciparum* AP1 N-terminal to other proteins

The data for the AP1 N-terminal are detailed in section 3.6.8.2, where the entire AP-1 β1 subunit was compared to the proteomes of *P. knowlesi*, *P. vivax*, *P. chabaudi*, *P. berghei* and *P. yoelii*, as well as *T. gondii* and *H sapiens.*

CHAPTER 4

DISCUSSION

4.1) *PFB0150c* – A *P. falciparum* PK

The catalytic core of all PKs is conserved across eukaryotes, viruses and some bacteria and consists of about 260 amino acids (*Hanks & Hunter, 1995*). In higher eukaryotes, this catalytic domain contains an ATP-binding region at the N-terminal extremity which is made up of a glycine-rich region in the locale of a lysine residue. Conserved aspartic acid residues are situated in the carboxy terminus lobe and are significant for the catalytic activity of the enzyme. Higher eukaryotic PKs have twelve subdomains. The amino terminus lobe consists of subdomains I, II, III, IV and V. The carboxy terminus lobe is made up of subdomains VIA, VIB, VII, VIII, IX, X and XI (*Figure 56*) (*Hanks, 2003*).

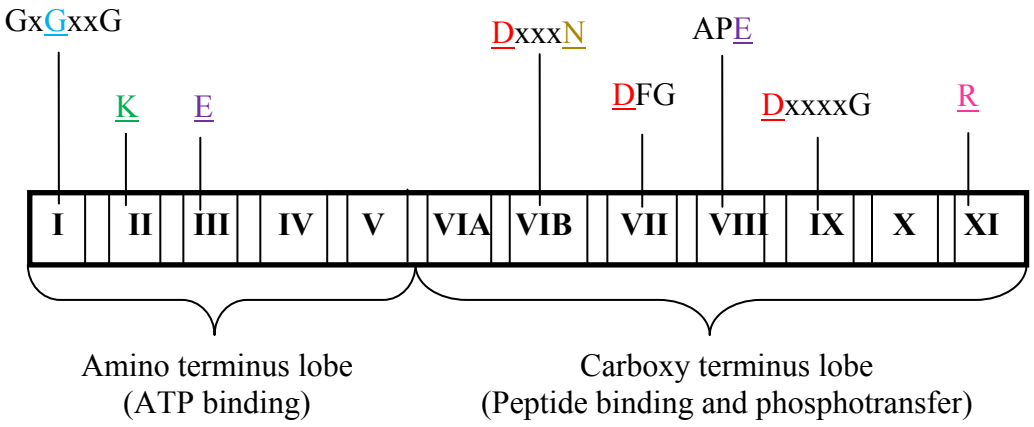


Figure 56A. Catalytic core of PKs
The higher eukaryote PK catalytic domain. Invariant amino acids are underlined.

In the case of *PFB0150c*, the catalytic core is 276 amino acids in length (*Figure 57*). The parasite kinase domain contains all the invariant residues that are present in eukaryotic PKs *(Figure 56B)*.

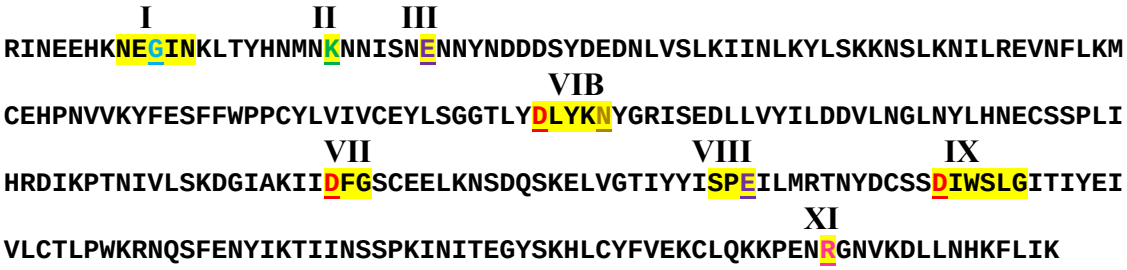


Figure 56B. Catalytic core of P. falciparum PK

The invariant residues found in the catalytic core of eukaryotic PKs are present in the catalytic core of the P. falciparum PK, spanning amino acids 2 079 – 2 354.

3D templates available in the Swiss Model Template Library, determined from x-ray crystallography of proteins, did not closely predict the structure of the entire parasite PK, with the catalytic subunit of human cAMP-dependent protein kinase sharing the most sequence identity of only 26.4 percent in the kinase domain. This indicates that the predicted 3D model is not a reliable representation of the quaternary PK structure because the template does not share sequence identity of more than 40 percent (*Arnold et al, 2006; Schwede et al, 2003; Guex and Peitsch, 1997)*. However, the area of homology between the target-template pair occurred exclusively in the kinase domain, supporting the proposal of PK catalytic core conservancy across eukaryotes (*Hanks & Hunter, 1995*).

BLAST data indicated that the *P. falciparum* PK shares identity with other eukaryotic PKs in the region of the catalytic domain, while the protein as a whole has very low homology. This could be due to different sizes of the PKs among the different organisms – the P. falciparum enzyme is 2 485aa, the *Plasmodium sp*. PKs range from 2 178 to 6 390aa, the human enzyme is 304aa and the *T. gondii* PKs are between 296 and 4 253aa in length. The *P. vivax* and *P. knowlesi* PKs had the most homology with the *P. falciparum* PK, more so than the *P. yoelii* PKs.

This is probably due to the fact that that the former three species infect humans, while the latter is a rodent malaria. Interestingly, the PKs with the best matches from the *H. sapiens* and *T. gondii* proteomes shared similar homology with the *P. falciparum* PK. Moreover, the number of conserved catalytic core residues in the human PK (*Appendix, section 5.9, Figure 64*) matched those in the *P. falciparum* PK better (*T. gondii PK homology, appendix, section 5.9, Figure 66*). This is unusual as one would expect more similarity between parasite proteins from the same phylum, rather than between the human and *P. falciparum* enzymes.

4.1.1) ***PFB0150c*** protein and mRNA expression

Mass spectrometry-based evidence indicates that the protein is expressed in notable quantities during the sporozoite, merozoite and gametocyte stages of the parasite life cycle, with no detectable expression during the trophozoite stage *(Florens et al, 2002)*. Abundance of the PK is greatest during the gametocyte stage with 90 percent (percentage of the protein sequence covered by identified peptides) of whole protein being detected; this is followed by 70 percent abundance during the merozoite stage and 40 percent during the sporozoite stage *(Florens et al, 2002)*. Photolithographic oligo array data for mRNA expression by Le Roch *et al* (2003) is represented in figure 57. From these data it can be deduced that the expression of the *PFB0150c* gene is greatest during the gametocyte stage, followed by the sporozoite, late schizont and then merozoite stage, respectively. Differences in the amount of protein and the expression of the gene may be attributed to the fact that not all mRNA transcripts are translated immediately into protein. Therefore, mRNA expression does not always mirror the amount of protein produced.

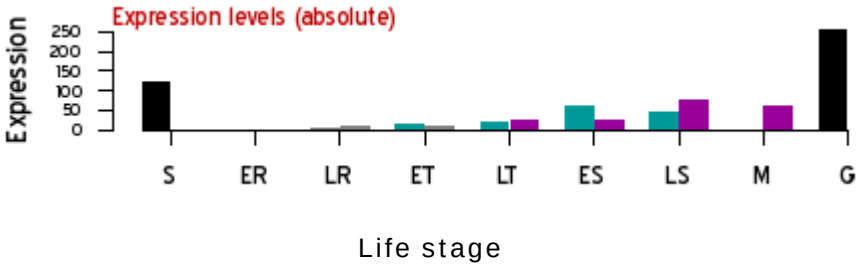


Figure 57. Intraerythrocytic P. falciparum 3D7 photolithographic oligo array data for PFB0150c mRNA expression (Le Roch et al, 2003)

x-axis: *Plasmodium developmental stages* synchronised by* ***sorbitol*** *and* ***temperature****. Data for gametocyte sample corresponds to synchronisation by sorbitol only, and for sporozoite sample represents average of two replicates.*
y-axis: *Absolute mRNA expression levels*
**** Developmental stages:*** *ER = Early Rings; LR = Late Rings; ET = Early Trophozoites; LT = Late Trophozoites; ES = Early Schizonts; LS = Late Schizonts; M = Merozoites; S = Sporozoites; G = Gametocytes*

4.1.2) Recombinant *PFB0150c* expression in *E. coli*

Expression of the rGST PK insert was successful despite the predicted insolubility of 63.2 percent, which was the highest prediction for all targeted domains. Presumably its pI of 5.88 predisposed the sequence to expression, as it fell within the ideal parameters of 3.45 and 6.8, where most *P. falciparum* proteins are expressed in soluble form (*Mehlin et al, 2006)*. The yield of purified rGST PK obtained, namely 180-210µg per litre of culture, was quite low. At least 2mg of protein is needed in order to carry out crystallisation (*Vedadi et al, 2007*), which would allow for further elucidation of its structure and biochemical function.

4.1.3) Classification of *PFB0150c* PK

In 2005, Anamika *et al* carried out *in silico* studies and identified a 2 985 residue long putative kinase-like sequence (PKLS) annotated *PFB0150c.* The catalytic domain of this PKLS, one of 90 PKs with a conserved catalytic aspartic acid, was found to share 30 percent identity with MAP kinase kinase (MAPK kinase) and p21-activated kinases. MAPK kinases are serine/threonine/tyrosine-specific protein kinases that are only activated once phosphorylated by a MAPK kinase kinase (MAP3K). The function of MAPK kinases is to activate MAPKs by phosphorylating threonine and tyrosine residues within a specific consensus sequence on their target molecule (*Cobb, 1996*). They postulated that the product of *PFB0150c* was a MAPK kinase involved in a cascade regulating various transcription factors. However, bioinformatic data collected in the current study indicate that *PFB0150c* codes for a serine/threonine PK. MAPK kinases did not feature among the top ten BLAST hits and 84 percent of the top 100 hits were with serine/threonine PKs *(Table 16)*. However, it is interesting to note that out of the top 30 hits, three were plant MAPK kinases which corroborates the postulated close evolutionary link between plant and parasite kinases (*Anamika et al, 2005*). The parasite PK also functioned without prior phosphorylation by a MAP3K, which is an essential characteristic of MAPK kinases. It is possible that *PFB0150c* is an atypical MAPK kinase variant, but the data from the current study strongly support the hypothesis that the parasite enzyme is a casein kinase. For example, the results of the kinase assay showed that the α casein 1 and 2 subunits of exogenous bovine casein were phosphorylated by the recombinant enzyme. However, the parasite PK does not have all the characteristics of higher eukaryotic isoforms of casein kinase 1 (CK1). While both the parasite PK and eukaryotic CK1's contain the domains common to all serine/ threonine PKs, the CK1's additionally have a Ser-Ile-Asn (SIN) motif in domain VIII *(Gross & Anderson, 1998)*. In the parasite this motif has the sequence Ser-Pro-Glu (SPE) *(Figure 57)*. Additionally, all CK1 isoforms contain a near-consensus SV40 T-antigen putative nuclear localisation sequence in the kinase catalytic domain which is made up of the following amino acids: TKKQKY *(Gross & Anderson, 1998)*. This motif is not present in the parasite PK. Determining whether the parasite enzyme is inactivated by MAPK kinase inhibitors or casein kinase inhibitors would aid with further characterisation.

The hypothesis that *PFB0150c* is a PK that has protein 4.1 and/ or spectrin as its substrates would be futher substantiated by the presence of a PEXEL/ VTS export motif to facilitate trafficking of the parasite enzyme into the host RBC. This motif has the consensus sequence R/KxLxE/Q, where x is any amino acid, and it is present in the *PFB0150c* protein sequence (*Appendix, figure 76*).

4.1.4) Probable function of *PFB0150c* PK

A host-parasite PPI was found to exist between protein 4.1, as well as spectrin, and a PK encoded by the *PFB0150c* gene through the use of a *P. falciparum* phage display library (*Lauterbach et al, 2003*). The binding sequence spans amino acids 786-815 of the parasite enzyme, which is outside the kinase domain (2 079 – 2 354) (*Lauterbach et al, 2003*). This implies that once the substrate has been bound, the folding conformation of the PK brings the phosphorylation target in proximity to the catalytic domain. Recent protein expression data using mass spectrometry (*(Florens et al, 2002)*) show that the enzyme is present with greatest abundance in gametocytes, followed by merozoites and then sporozoites. No enzyme expression was detected in the trophozoite stage. The kinase assay performed in the current study indicated that the *PFB0150c* product has casein kinase activity. The enzyme could therefore be involved in the phosphorylation of protein 4.1 and/or spectrin – which weakens the host RBC membrane – and thereby aid rapid merozoite invasion of host RBCs. Research by Chishti *et al* (*1994*) determined that an enzyme that phosphorylated protein 4.1 in parasitised erythrocytes was a casein kinase of parasite origin, by showing that phosphorylation of the RBC membrane protein was prevented by casein kinase I and II inhibitors. Phosphorylated protein 4.1 was detected in pRBC at the trophozoite/schizont stage, but not at the ring stage. This phosphorylation was postulated to increase the flexibility of the host membrane and facilitate intraerythrocytic growth and exit of the parasite. Phosphorylation of protein 4.1 during invasion by merozoites was not documented by Chishti *et al* (1994). However, merozoite invasion is extremely rapid, due to the need to evade the host immune system, and it would thus be almost impossible to accurately obtain RBC membranes at this specific stage in order to capture these data. Therefore, further investigations are necessary to fully characterise the role of this enzyme, but it is tempting to speculate that the PK analysed in the current study is the enzyme first discovered by Chishti *et al* (*1994*).

It is possible that the PK has other substrates *in vivo*, as it is also found in abundance in the gametocyte and sporozoite stages where its roles are unknown; but one can speculate that in sporozoites it could function in invasion of hepatocytes in the human host, while in gametocytes the PK could play a role in growth and survival of the parasite within the mosquito.

The fact that the kinase domain makes up such a small part of the PK – 276 aa of a total 2 485 – indicates that the enzyme may be multifunctional, although no additional roles have currently been attributed to it (*www.plasmodb.org, 2008*).

4.1.5) Drug-target potential of *PFB0150c*

Malarial PKs have long been investigated for their important roles in the replication and differentiation of the parasite (*Doerig et al, 1995; Wiser, 1995; Kappes et al, 1995; Kappes et al, 1999; Harmse, 2006*) and for the extensive changes they cause in the phosphorylation profile of the red blood cell membrane proteins to facilitate the invasion and development of *P. falciparum* (*Rangachari et al, 1986; Jones & Edmundson, 1990; Chisti et al, 1994*). *P. falciparum* PKs have recently been highlighted as targets for potential curative and transmission-blocking drugs (*Syin et al, 2001*; *Doerig and Meijer, 2007*).

Traditionally, one of the key criteria that a drug target should fulfill is that its protein sequence must be significantly different to host proteins. This prevents any non-selective inhibition of the host molecules. However, it has transpired that even in the case of host-parasite orthologues sharing sequence identity of 40 to 60 percent, selectivity for the parasite protein can be achieved (*Doerig and Meijer, 2007*). An example of specific targeting of a drug to the protein of one species, even in the presence of an orthologue from another species, was demonstrated by Gray *et al* (*1998*). In the presence of human cyclin dependant kinase (CDK) 1and its orthologue in yeast, the CDK inhibitor purvalanol B demonstrated a 200-fold difference in IC_{50} values between the two (*Gray et al, 1998*).

Interestingly, *P. falciparum* casein kinase (CK) 1 is also a major target of purvalanol B, which has antiproliferative effects on the parasite (*Harmse et al, 2001; Knockaert et al, 2000*). Further studies will show whether the same species-selective inhibition occurs in the presence of human CK1 (*Doerig and Meijer, 2007*).

The data from this study indicated that the PK product of *PFB0150c* may represent a feasible drug target, since only the small stretch of amino acids making up the kinase domain (276aa) is relatively well conserved across eukaryotes. The rest of the protein could be targeted, where the binding of inhibitors could cause steric hindrance that interferes with the catalytic domain. Alternatively, the sequence that binds to protein 4.1 and spectrin, which does not fall within the kinase domain spanning amino acids 2 079-2 354, could be blocked so that the enzyme is prevented from docking with its phosphorylation target. Its probable function in invasion and growth of the erythrocytic stages of the parasite means that inhibiting this enzyme could prevent this stage from occurring and would thereby result in parasite death. Experiments with casein kinase inhibitors will determine whether there is parasite protein specificity of the drug, even in the presence of human orthologues. Gene knockout and/or knockdown experiments would also be required to validate this PK as a drug target.

4.2) *PFE1400c* – A putative AP-1 β subunit

The function of the AP-1 β subunit – which forms a vital part of the AP-1 complex – has been well characterised in mammals (*Ohno, 2006)*, where it induces the assembly of clathrin triskelions to form a latticed coat that supports the formation of vesicles for transporting proteins within cells (*Brodsky et al, 2001).* The putative function of the AP-1 β subunit in malaria transport pathways is deemed to be the same as that for mammals, although it has not been proven (*Cooke et al, 2004*).

Heldwein *et al* (2004) are the only group to have determined the crystal structure of AP-1 complexes, specifically in *Rattus norvegicus* and *Mus musculus*, but due to the highly conserved nature of these proteins, the structure is probably similar in most eukaryotes. As a case in point, the AP-1 complex β subunit of *R. norvegicus* used by Heldwein *et al* (*2004*) has 61 percent identity and 78 percent positivity to the *P. falciparum* putative AP-1 complex β subunit. All AP-1 complexes consist of four subunits – two large adaptin subunits designated γ-1 and β-1, a medium subunit termed μ-1 and a small subunit called σ-1 (*Heldwein et al, 2004*) (*Figure 58*).

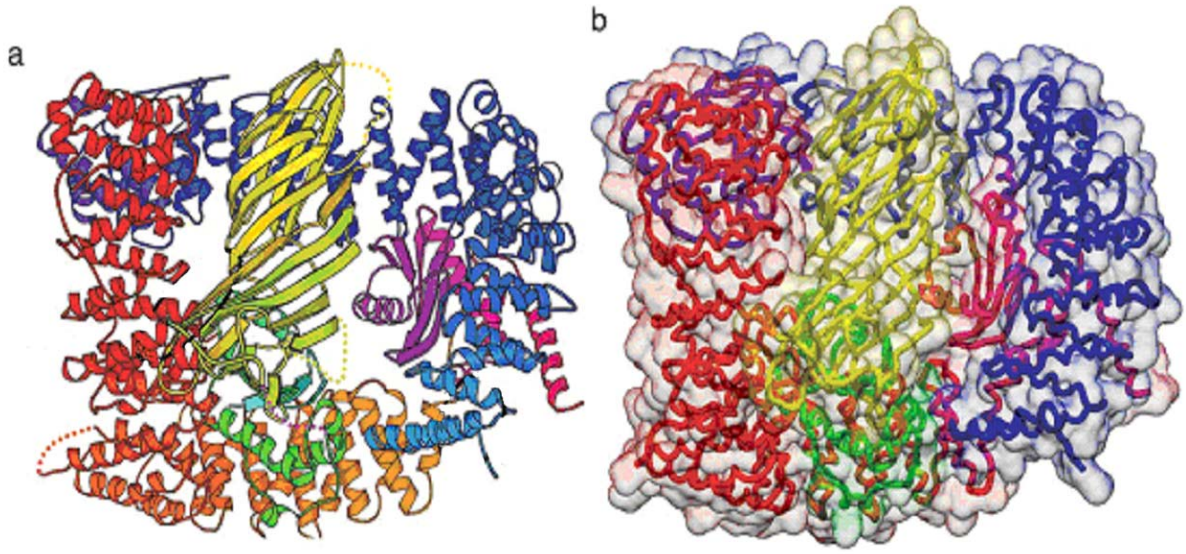


Figure 58. Structure of the AP-1 core. (Heldwein et al, 2004)

(a) Ribbon representation. (b) Ribbon with molecular surface representation
Colour coding: red - β-1 subunit; blue - γ-1 subunit; yellow/orange - μ-1C subunit; green - μ-1N subunit; magenta – σ-1 subunit.

This model indicates that the β subunit consists almost exclusively of alpha helices. This aligns with the putative 3D model obtained for the *P. falciparum* AP1 N-terminal (*Figure 54*) and with its predicted secondary structure (*Appendix, section 5.10, figure 72*). However, no beta sheets can be seen in the Heldwein model (*Figure 58*), which is contrary to the putative 3D model (*Figure 44*) and secondary structure predictions (*Appendix, section 5.10, figure 71*) for the AP1 C-terminal of the β subunit. However, the 3D model did not accurately predict the structure of the *P. falciparum* AP1 C-terminal due to the low sequence identity between the target-template pair, so a discrepancy is expected. Unfortunately, the Heldwein 3D model is not available in the Swiss Model Template Library, but one would assume that it would come up as the best match since it has higher sequence homology with the parasite transport protein than what the human AP-2 proteins have (*sections 3.6.8.1 and 3.7.8.1*).

The *P. falciparum* AP1 C-terminal and the AP1 N-terminal domain that were subcloned in the current project constitute two important areas of the β subunit (*Figure 58, red subunit*). In eukaryotes, the AP1 C-terminal domain is found at the C-terminal end of the β subunit and recruits accessory proteins that modulate clathrin binding and polymerisation. The AP1 N-terminal domain is proposed to interact with a uniform area of the clathrin-coated vesicles. Together, the β subunit as a whole triggers the assembly of clathrin triskelions into an extensive latticed-coat which stabilises vesicles for the transport of cargo from the TGN to endosomes (*Wang et al, 1995*).

The AP-1 heterotetramer is further grouped into a core and an appendage section. The core consists of the trunk (the N-terminals of the two large subunits, γ-1 and β-1) along with the μ-1 and σ-1 subunits. The appendages are the C-terminal regions of the two large subunits, which are attached to the core by a hinge segment (*Heldwein et al, 2004*) (*Figure 59*).

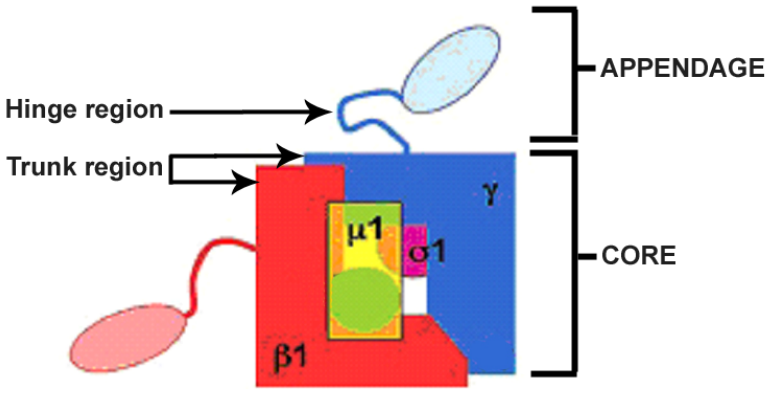


Figure 59. Proposed AP-1 subunit organisation *(Adapted from Heldwein et al, 2004)*
The proposed AP-1 subunit organisation is based on the structure of AP-2. AP-1 is a heterotetramer composed of 4 separate chains or subunits that assemble to form the functional AP complex. The trunk region consists of the N-terminals of the large γ-1 and β-1 subunits. The appendage region consists of the C-terminals of the large γ-1 and β-1 subunits. The core consists of the trunk region and the μ-1 and σ-1 subunits. The appendage region is joined to the core by a hinge region.

The trunk is composed almost entirely of 14 consecutive α-zigzags – known as HEAT elements – that are twisted to form a boomerang-like structure (*Figure 58a, red domain*). This is consistent with the secondary structure prediction for the N-terminal (*Appendix, section 5.10, figure 72*). This region acts as a point of contact for the small σ-1 and medium μ-1 subunits, thus stabilising the assembly of the four chains. According to Heldewein *et al* (*2004*), the heterotetramer will only be marginally stable if the trunk region is not able to interact with the other subunits, indicating that the N-terminal regions of the γ-1and β-1 subunits are vital for AP-1 complex stability.

Paralogues of the *P. falciparum* β subunit were found in five other *Plasmodium* species, with *P. vivax* and *P. knowlesi* putative β subunits sharing the highest sequence homology. The *T. gondii* and *H. sapiens* β subunits also had high sequence homology to the *P. falciparum* protein, and interestingly shared similar levels of identity and positivity. This indicates extreme conservation of this protein among these eukaryotes. Overall, the *P. falciparum* AP1 N-terminal region had most homology with all proteins that were detected using the BLAST. This shows that the larger AP1 N-terminal region is highly conserved in eukaryotes, while the small C-terminal is less so. This difference in levels of conservation could be attributed to the function of each of the terminals.

The N-terminal interacts with a uniform area of clathrin-coated vesicles, which is probably common in most eukaryotes as clathrin is a ubiquitous structural protein. The C-terminal, on the other hand, is proposed to interact with various accessory proteins that modulate AP complex function and these accessory proteins could be quite varied in type and structure in different species.

4.2.1) *PFE1400c* protein and mRNA expression

According to mass spectrometry-based evidence, the expression of *PFE1400c* protein is high during all life-cycle stages *(Florens et al, 2002)*. This is supported by the photo-lithographic oligo array data for *PFE1400c* mRNA expression (*Figure 60*) determined by Le Roch *et al* (2003). Marginally higher levels were detected during the early schizont (ES) stage, and lower expression occurred during the gametocyte (G) stage, but generally the expression levels remain similar and relatively high when compared to that of the PK (*Section 4.1.1*). This is expected, as trafficking of proteins occurs throughout the parasite life cycle and one trafficking molecule is physically needed to transport protein molecules.

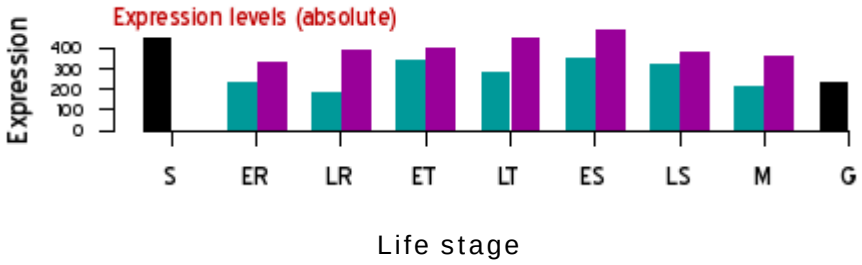


Figure 60. Intraerythrocytic P. falciparum 3D7 photolithographic oligo array data for PFE1400c mRNA expression (Le Roch et al, 2003)

x-axis: *Plasmodium developmental stages* synchronized by* ***sorbitol*** *and* ***temperature****. Data for gametocyte sample corresponds to synchronization by sorbitol only, and for sporozoite sample represents average of two replicates.*
y-axis: *Absolute mRNA expression levels*
Developmental stages: *ER = Early Rings; LR = Late Rings; ET = Early Trophozoites; LT = Late Trophozoites; ES = Early Schizonts; LS = Late Schizonts; M = Merozoites; S = Sporozoites; G = Gametocytes*

4.2.2) Recombinant *PFE1400c* expression in *E. coli*

The AP1 C-terminal was expressed in soluble form despite the predicted insolubility of 56.9 percent. The expression of 36 additional amino acids – due to the alternative splicing of the 108bp intron – increased the molecular weight from 53.2kD to 56.8kD, and the pI from 5.4 to 5.81. The change in the pI would probably not have affected the expression of the sequence as the new value still falls within the range of 3.45 to 6.8 which is associated with optimal expression of *P. falciparum* proteins (*Mehlin et al, 2006*). The inclusion of the 108bp making up the intron between exon 5 and 6 of gene *PFE1400c* is due to alternative splicing, which is a novel finding for this gene as PlasmoDB makes no mention of parasite AP-1 β-subunit isoforms. Alternative splicing is known to occur in *P. falciparum* pre-mRNA transcripts (*Knapp et al, 1991; Singh et al, 2004*). Consensus splice sites for *P. falciparum* have currently not been determined (*PlasmoDB, 2008*), and the consensus sequences for RNA splicing in higher eukaryotes (*Alberts et al, 1994*) were not found in the 108bp insert (*Figure 61*).

AGAGAGTATATTAATGGCTGTAAGCCCTATGGTTATAACATCAGACATGTTGATAAAACGAATGAAAATTTT
TAATATTTCTTTAATAGCTCGTAGAAATGTAAATAATATG

Figure 61. Exon/intron boundary sequences of PFE1400c intron 5

The 108bp insert (red) comprising the intron between exons 5 and 6 in PFE1400c may have been retained in the final mRNA due to the mechanism of alternative splicing. In higher eukaryotes, the consensus sequence for the 5' splice site, known as the donor site, is AAG or CAG on the 3' end of the first exon and an invariant GT dinucleotide on the intron. These consensus sequences were not found in PFE1400c exon 5 (blue nucleotides) or the 5' end of the intron (underlined). In higher eukaryotes the consensus sequence for the 3' splice site, known as the acceptor site, begins with an invariant AG dinucleotide at the intron 3' end, which was not found in the above insert (underlined) followed by an A or a G (green) at the 5' end of the second exon, as seen in the above insert.

The exon/intron boundaries in the rest of the *P. falciparum* AP1 C-terminal are shown in table 25.

TABLE 25: Donor and acceptor sites in non-alternatively spliced exon/intron boundaries of *P. falciparum* AP1 C-terminal

DONOR SITE		ACCEPTOR SITE	
Exon 1, 3' end ACA	Intron 1, 5' end GT	Intron 1, 3' end AG	Exon 2, 5' end G
Exon 2, 3' end TTC	Intron 2, 5' end GT	Intron 2, 3' end AG	Exon 3, 5' end A
Exon 3, 3' end AAG	Intron 3, 5' end GT	Intron 3, 3' end AG	Exon 4, 5' end A
Exon 4, 3' end CAA	Intron 4, 5' end GT	Intron 4, 3' end AG	Exon 5, 5' end G
Exon 6, 3' end GTG	Intron 6, 5' end GT	Intron 6, 3' end AG	Exon 7, 5' end A

The exon donor sites are all different and do not conform to the eukaryotic consensus sequences. However, all the other sites contain the eukaryotic invariant residues or the A/G consensus nucleotides. The fact that intron 5 (*Figure 61*) does not contain the invariant donor and acceptor nucleotides indicates why the spliceosome did not recognise it and why the 108bp insert was retained in the mRNA.

There are a number of speculations as to why alternative splicing occurs in eukaryotes (*Kriventseva et al, 2003*). It has been proposed that a greater level of efficiency is attained because genetic information can be stored more economically. Multiple proteins are encoded by a single DNA sequence that could have only coded for a single protein. The mechanism of alternative splicing enables the evolution of new proteins without altering the original DNA sequence, which not only means faster evolution of proteins, but also prevents loss of integrity of established proteins (*Kriventseva et al, 2003*).

The importance of this mechanism for *P. falciparum* is evident when the complex life cycle of the parasite is considered. The parasite is exposed to the immune systems of two organisms, the mosquito and the human, making the fast evolution of proteins in order to avoid detection an essential survival strategy. The parasite also needs to adapt quickly to new environmental factors –such as drugs, and the difference in mosquito and human cells – in order to survive.

The AP1 N-terminal was predicted as the most soluble (49.6 percent) of the three domains analysed in this study, but despite this the protein was expressed in insoluble inclusion bodies. Interestingly, even though the pI of 6.3 was within the soluble expression range of 3.45 and 6.8 (*Mehlin et al, 2006*), the AP1 N-terminal protein was insoluble. It is possible that the size of the protein played a role in its insoluble status; the recombinant AP1 N-terminal protein contained the longest of all the expressed *P. falciparum* domains at 618 amino acids, compared to the AP1 C-terminal with 244 amino acids and the PK with 503 amino acids. However, the latter two had GST tags of an additional 220 amino acids, while the AP1 N-terminal had only a 6 amino acid His tag, along with a vector sequence of 9 amino acids. This resulted in the recombinant AP1 N-terminal not being the largest protein in terms of molecular weight. However, while larger protein size has been correlated with decreased expression – with only 20 percent of proteins >60kD being expressed – no relationship was found between gene/protein size and solubility (*Mehlin et al, 2006*). The results obtained in the current study are therefore not consistent with the general conclusion of Mehlin *et al* (*2006*) as the AP1 N-terminal was expressed despite its size (72.2kD), but it was insoluble.

This highlights the fact that recombinant expression of each *P. falciparum* protein cannot be accurately predicted and remains a matter of trial and error. Numerous other factors could have influenced the solubility of the AP1 N-terminal. The presence of the GST tag could have influenced protein solubility, as the two proteins with this tag were soluble. However, estimations of solubility for the AP1 N-terminal were higher with the His tag when compared to the GST tag. This brings the accuracy of the recombinant protein solubility prediction tool into question, and indicates that it can only be used as a rough guide. Alternatively, more complex folding may have been required for the AP1 N-terminal protein secondary structure, and the prokaryotic environment of *E. coli* does not contain the chaperones necessary for this process. The presence of eukaryotic protein chaperones could better facilitate the folding and subsequent soluble expression of the parasite protein (*Sorenson and Mortenson, 2005)*.

4.2.3) Drug-target potential of *PFE1400c*

Molecules involved in protein trafficking within the parasite and pRBC offer novel targets for therapeutic intervention. By targeting these transport molecules, entire groups of cargo proteins could be prevented from reaching their destination and carrying out their function (*Phelps et al, 2003*). Because of the specificity of transport, therapies that interfere with the attachment of cargo to APs would have minimal cross-reactivity (*Phelps et al, 2003*). Based on the function of the AP1 C-terminal domain, it can be inferred that the knockout of this region would result in a cessation of clathrin binding and subsequent polymerisation to form a coat. CCVs would be incomplete, lacking the outer protein coat, and transport of associated cargo proteins would halt. If inhibition of this protein could be achieved in *P. falciparum*, the parasite would perish as vital proteins would be left stranded within the TGN, unable to reach their target destination and carry out their effector function.

The AP1 N-terminal domain that was subcloned and expressed in this project forms a significant part of the trunk region that stabilises the interaction of the subunits making up the AP-1 complex; therefore, if this region could be knocked out or blocked in some way, the AP-1 molecule would have decreased stability and may not be able to carry out its function in formation of clathrin-coated vesicles. This potential halt in the transport of proteins from the TGN to other endosomes would stop the development of erythrocytic–stage *P. falciparum*. For example, the areas of the Golgi membrane containing RAMA would be prevented from budding and forming vesicles (*Topolska et al, 2003*). If this transport is inhibited, the biogenesis of the rhoptries may also be inhibited. A parasite lacking rhoptries would be unable to invade subsequent erythrocytes and would be sequestered by the host immune system.

Studies involving double knockout mice lacking the μ-1subunit gene and γ-1 subunit gene have shown that the animals die early in their embryonic development, as the inability of the molecules to form a stable tetramer causes the degradation of the remaining subunits. The large γ-1 subunit was found to be particularly important, as no AP-1 subunits were found in the knockout mice at all, whereas the μ-1 knockout mice showed non-functional trimeric compounds consisting of the remaining subunits (*Hinners & Tooze, 2003)*. It is inferred from these data that the other large subunit, namely the β subunit, is as critical to the AP-1 complex as the γ-1 subunit is. This provides confirmation that knocking-out or inactivating one of the vital large subunits could prevent the trafficking of key molecules within the parasite and lead to the eradication of infective erythrocytic-stage malaria parasites, making the AP-1 complex an attractive drug target.

Targeting *P. falciparum* invasion proteins is difficult because they are highly redundant. By inhibiting the protein trafficking complex, transport of invasion proteins to the micronemes and rhoptries would be blocked and they would be unable to reach their final destination and carry out their function. Thus invasion of host RBC by these defective merozoites would be prevented. A drawback is the high sequence identity of the human AP-1 β1 subunit (46 percent) with the parasite protein sequence, which could make it difficult to target the latter without harming the host. However, the C-terminal has much lower identity (28 percent) and could be exclusively targeted to prevent cross-reactivity.

4.3) Future studies

A number of experiments could be done to further characterise the structure and function of the *P. falciparum* proteins examined in the current study. Gene knockdown and/or knockout studies could be carried out to elucidate the function of the proteins and determine whether they are necessary for parasite survival. DNA microarrays could be used on knockout/knockdown parasites to see which genes are up- or downregulated and thereby infer which functional pathways they are involved in. Microarrays could also confirm the expression profile of the genes during each of the developmental stages, while protein arrays could indicate additional PPIs. Localisation studies using immunofluorescence and confocal microscopy may indicate the sites where the proteins reside within the parasite or pRBC, and co-localisation studies – for example with protein 4.1, spectrin and the parasite PK; or with clathrin and the AP1 N- and C-terminals – could indicate which molecules the proteins interact with. The technique of two-hybrid screening could be utilised to determine which proteins associate with the AP1 N- and C-terminal domains. The AP1 sequences could be used as the bait proteins to identify binding partners from a parasite lysate of prey proteins (*Fields & Song, 1989*). Kinetic studies would specify the optimum pH and temperature for PK activity, while doing a kinase assay with purified protein 4.1 and spectrin would determine whether they are phosphorylation targets for the PK *in vitro*.

Measuring the effect of the casein kinase inhibitors used in the Chishti *et al* study (1994) on the PK would provide some evidence on whether the enzyme is the same as that described by these authors, as hypothesised in the present study. To determine whether the enzyme functions as a MAPK kinase, inhibition by PD 98059 and U0126 - which prevent activation of classical MAPK kinases - could be carried out and measured (*Dudley et al, 1995; Favata et al, 1998*). Additionally, including further negative controls in the kinase assay would ensure that the observed enzyme activity was not carried out by a co-purified *E. coli* protein. A negative control could be generated via site-directed mutagenesis to eradicate the enzyme activity.

To further elucidate the 3D and secondary structure of the molecules, protein crystallisation could be performed. For some of these studies – especially the protein array, protein crystallisation and kinetic experiments – increased yields and solubility of the proteins are needed.

A number of methods could be utilised to improve the expression of soluble recombinant *P. falciparum* proteins in future studies (*Birkholtz et al, 2008*). A method for preventing the formation of insoluble inclusion bodies – which are loosely aggregated folding intermediates of the recombinant protein – is the co-expression of protein chaperones (*Sorenson and Mortenson, 2005*). While the expression of these plasmid-encoded chaperones may not guarantee improved folding and solubility of the recombinant protein, they may minimise the level of protein aggregation, making it easier for native protein to be prepared from inclusion bodies (*Sorenson and Mortenson, 2005*). This process may prove effective for the AP1 N-terminal, as solubilisation and refolding of the recombinant protein – using guanidine hydrochloride and dialysis against a native buffer – was not successful. This result may have been due to the high initial concentration of guanidine hydrochloride.

According to Singh and Panda (*2005*), the use of extremely high initial concentrations of chaotropic agents – like urea and guanidine hydrochloride – increases the rate of protein aggregation. These agents cause areas of hydrophobic amino acids to be exposed on the protein through the production of random coil structures. The intermolecular interactions between the hydrophobic regions on protein molecules compete with intramolecular bonds and cause aggregation of the molecules. To prevent the exposure of hydrophobic patches and subsequent protein aggregation, lower initial concentrations of chaotropic agents can be coupled with a pH shock during refolding (*Singh & Panda, 2005*).

Singh and Panda (*2005*) found that the use of 2M urea, with a refolding buffer pH that was more alkaline than the pI of the protein, destabilised unwanted intermolecular interactions. In the case of the AP1 N-terminal, the pH of the refolding buffer was 8, whereas the pI of the recombinant protein is 6.3, which seems to be a viable combination. However, an initial concentration of 6M guanidine hydrochloride was used for refolding which could have prevented the recovery of bioactive recombinant protein.

Another factor that could have affected the expression of the recombinant *PFE1400c* AP1 N-terminal is codon bias (*Baneyx, 1999*). Due to the different codon usage of the parasite to that of the *E. coli* host, recombinant protein expression could have been stalled. Rare codons cause mRNA and plasmid instability via ribosome stalling which results in slow, error-prone translation and can completely inhibit protein synthesis and host cell growth (*Zahn, 1996*). An additional setback to expression is that the levels of tRNAs to their associated rare codons remain very low in *E. coli* cells (*Baneyx, 1999*).

There are two strategies for preventing the unwanted effects of codon bias. The first is to use a plasmid that encodes rare tRNAs that are lacking in the usual *E. coli* tRNA pool. This approach was used in the current study, but did not facilitate the expression of the AP1 N-terminal. This finding is supported by several studies that have shown inconsistent results when employing this method (*Hannig & Makrides, 1998*).

The other strategy involves codon optimisation (*Zhou et al, 2004*). This process involves the production of a synthetic gene by changing the native codons to those most commonly used by *E. coli* cells, without changing the amino acid composition of the translated product. In a study on recombinant malaria candidate vaccine proteins by Zhou *et al* (*2004*), the use of a codon optimised gene was found to increase the expression level threefold, as well as promoting better host cell growth. While this method has been successful for the expression of several *P. falciparum* proteins – namely dihydrofolate reductase, merozoite surface protein-1 and serine repeat antigen – the production of the synthetic gene tends to be a costly process and is not always successful (*Zhou et al, 2004*).

Alternatively, the process of codon harmonisation can be utilised to increase the heterologous *P. falciparum* protein expression in *E. coli* (*Angov et al, 2008*). The method involves determining the codon usage frequencies of both *P. falciparum* and *E.coli* using a bioinformatic algorithm and then substituting the native codons with synonymous ones having analogous usage frequencies in the host (*Angov et al, 2008*). Genes that were recoded in this manner had expression levels 4 to 1 000 times greater than the original genes (*Angov et al, 2008*).

The use of more advanced, eukaryotic expression systems would also increase the probability of successful recombinant *P. falciparum* protein expression. The use of yeast cells, or a baculovirus system utilising insect cells, could result in improved solubility due to the greater similarity in the microenvironment between the eukaryotic organisms (*Possee, 1997*). Different tags, like maltose binding protein (MBP) (*New England Biolabs, www.neb.com)* or NusA (*Novagen, www.emdbiosciences.com*), could also be incorporated and may result in improved solubility of the recombinant proteins.

4.4) Conclusion

To conclude, the three proteins analysed in this study present varying promise as drug targets. The PK has high sequence similarity with a human enzyme in the catalytic kinase domain, but areas in the proximity of the active site could be targeted to interfere with its function, for example, by steric hindrance or by blocking the binding of ATP. The human kinase is also only 304aa in length compared to the large 2 485aa parasite PK, so it is hypothesised that the domains not present in the former could be inhibited in the latter with no toxic side effects for the host. However, further elucidation of the function of these domains is needed in order to substantiate this claim.

In the case of the *P. falciparum* AP-1 β subunit, the N-terminal is very well conserved in eukaryotic organisms and high homology to its human counterpart renders it unsuitable as a drug target. However, the smaller C-terminal shows promise, as it shares minimal identity with the human trafficking protein.

For future studies with these *P. falciparum* proteins, the expression of soluble recombinant AP1 N-terminal and increased yield of AP1 C-terminal and PK could be most economically achieved through the use of a eukaryotic expression system. If inclusion bodies are still present in the case of the AP1 N-terminal, mild solubilisation during the refolding stage could aid in recovery of active protein. Codon harmonisation and optimisation are more involved and more expensive processes, which could be utilised at a later stage if the eukaryotic expression system fails. The increased recovery of soluble protein will facilitate future functional studies and attempts at crystallisation.

CHAPTER 5
APPENDIX

All sterile solutions were filtered using a Millipore syringe-driven filter unit (0.22µm), unless stated otherwise.

5.1) Malaria culturing techniques

Sterile incomplete medium

1x GIBCO™ RPMI Medium 1640 (*Invitrogen, USA*) with L-Glutamine and 25mM HEPES
4g glucose
0.044g hypoxanthine
0.05g gentamycin
1L autoclaved Milli-Q water
Store in sterile Schott bottles at 4°C

Sterile complete medium

10% heat-inactivated AB plasma
5g $NaHCO_3$
Make up to 100ml with incomplete medium
Store at 4°C and use within 5 days

Sterile 5% $NaHCO_3$

2.5g $NaHCO_3$
50ml autoclaved Milli-Q water
Store at 4°C

Sterile PBS (pH 7.2 – 7.4)

8g NaCl
0.2g KCl
1.78g $Na_2HPO_4.2H_2O$
0.2g KH_2PO_4
Make up to 1L using Milli-Q water
Store at room temperature

Sterile 12% NaCl

12g NaCl
100ml Milli-Q water
Store at 4°C

Sterile 1.6% NaCl

1.6g NaCl
100ml Milli-Q water
Store at 4°C

Sterile 0.9% NaCl / 0.2% Glucose

0.9g NaCl
0.2g glucose
100ml Milli-Q water
Store at 4°C

PBS for freezing

7.2g NaCl
14.8g $Na_2HPO_4.2H_2O$
4.4g KH_2PO_4
Make up to 1L with Milli-Q water
Store at 4°C

Sterile 60 % Glycerol solution

40ml PBS for freezing
60ml glycerol
Decant into sterile tubes and store at -20°C until needed

5.2) Reagents for DNA and RNA experiments

PBS (pH 7.2 - 7.4)

8g NaCl
0.2g KCl
1.78g $Na_2HPO_4.2H_2O$
0.2g KH_2PO_4
Make up to 1L using Milli-Q water

5% Saponin in PBS

0.05g saponin (*USB Corporation, USA*)
1ml PBS

Lysis Buffer

Make up lysis buffer from stock solutions of 1M Tris-HCl and 0.5M EDTA:

1M Tris-HCl (pH 8.0)

12.11g Tris
Add 90ml Milli-Q water
pH to 8.0 using HCl
Make up to 100ml with Milli-Q water

0.5M EDTA (pH 8.0)

18.61g EDTA
Add 90ml Milli-Q water
pH to 8.0 using NaOH
Make up to 100ml with Milli-Q water

Making up 100ml lysis buffer from stock solutions:

4ml of 1M Tris-HCl (final concentration 40mM)
16ml of 0.5M EDTA (final concentration 80mM)
2g SDS (final concentration 2%)
Add 90ml Milli-Q water and check pH 8.0
Adjust pH if necessary
Make up to 100ml with Milli-Q water
NB: Add 0.1mg/ml Proteinase-K just before use

3M Sodium Acetate (pH 5.2)

24.61g sodium acetate
Add 90ml Milli-Q water
pH to 5.2 using HCl
Make up to 100ml with Milli-Q water

TE Buffer

0.12g Tris
0.037g EDTA
Add 90ml Milli-Q water
pH to 8.0 using HCl
Make up to 100ml with Milli-Q water
OR

1ml of 1M Tris-HCl (pH 8.0) (final concentration 10mM)
200µl of 0.5M EDTA (final concentration 1mM)
Add 90ml Milli-Q water and check pH 8.0
Adjust pH with HCl if necessary
Make up to 100ml with Milli-Q water

0.8% Agarose gel for DNA

0.32g agarose
40ml 1X TAE buffer

Dissolve in microwave
Cool to about 60°C
Add 3ul of 10ug/µl ethidium bromide
Mix and pour
Allow gel to set for 20-30 minutes at room temperature
For RNA prepare 1% Agarose gel with 0.4g of agarose

50x TAE buffer

24.2g Tris
10ml 0.5M EDTA
5.71ml acetic acid
Make up to 100ml with Milli-Q water

1x TAE buffer

20ml 50x TAE buffer
Make up to 1L with Milli-Q water

5.3) Reagents for sub-cloning experiments

Luria broth
10g Bacto™ Tryptone (*Becton Dickinson Biosciences, USA*)
5g yeast extract (*Oxoid, UK*)
10g NaCl
10ml 1M Tris-HCl (pH 7.5)
Make up to 1L with Milli-Q water
Autoclave solution and cool

Sterile 100mg/ml ampicillin

1g ampicillin
10ml Milli-Q water
Store in 500ul aliquots at -20°C and use within 6 months

Sterile 50mg/ml chloramphenicol

0.5g chloramphenicol
10ml Milli-Q water
Store in 500ul aliquots at -20°C and use within 6 months

Agar plates

15g agar
1L Luria broth
Autoclave the solution and allow it to cool
For DH5α cell cultures (*Invitrogen, USA)*, add 100mg/ml ampicillin to a final

concentration of 100ug/ml.
For Rossetta 2 (DE3) cell cultures (*Novagen, USA),* add 100mg/ml ampicillin to a final concentration of 100ug/ml and 50mg/ml chloramphenicol to a final concentration of 50ug/ml.

Bacterial cell lysis buffer

0.9g glucose
0.37g EDTA
2.5ml 1M Tris-HCl (pH 8.0)
Make up to 100ml with autoclaved Milli-Q water
Store at 4°C

0.4M NaOH

1.6g NaOH
100ml autoclaved Milli-Q water
Store at room temperature

10% SDS

1g SDS
10ml autoclaved Milli-Q water
Store at room temperature

Fresh 0.2M NaOH/2% SDS solution

5ml 0.4M NaOH
200ul 10% SDS
4.8ml Milli-Q water

5M Potassium acetate

49.075g potassium acetate
100ml autoclaved Milli-Q water
Store at 4°C

5M Acetic acid

29ml 17.4M acetic acid
71ml autoclaved Milli-Q water
Store at 4°C

Fresh Potassium acetate (pH 4.8)

2ml 5M potassium acetate
4ml 5M acetic acid
Mix by inversion

TE Buffer

See section 5.2

5.4) Reagents for protein experiments

5.4.1) Reagents for recombinant protein purification

MagneGST™ binding/wash buffer (pH 8.5)

0.075g $Na_2HPO_4.2H_2O$ (4.2mM)
0.035g KH_2PO_4 (2mM)
2.92g NaCl (500mM)
0.0745g KCl (10mM)
Add 100ml Milli-Q water and check the pH

1M Glutathione stock (pH 7.0-8.0)

3.07g glutathione
Add 9ml Milli-Q water and check the pH
Adjust pH with HCl
Make up to 10ml with Milli-Q water

MagneGST™ elution buffer

1ml 1M Glutathione stock (pH 7.0-8.0) (final concentration 100mM)
500ul 1M Tris-HCl (pH 8.0) (final concentration 50mM)
8.5ml Milli-Q water
Store at -20°C in 500µl aliquots

0.5M sodium phosphate (pH 8.0)

7.8g $NaH_2PO_4.2H_2O$
100ml Milli-Q water

HIS-Select™ equilibration/wash buffer

10ml 0.5M sodium phosphate (final concentration 50mM)
1.75g NaCl (0.3M)
0.068g imidazole (10mM)
90ml Milli-Q water

HIS-Select™ elution buffer

10ml 0.5M sodium phosphate (final concentration 50mM)
1.75g NaCl (0.3M)
1.7g imidazole (250mM)

90ml Milli-Q water

5.4.2) Reagents for SDS PAGE analysis of recombinant protein

5x Suspension solution

0.303g Tris (50mM)
0.093g EDTA (5mM)
2.5g SDS (5%)
12.5g sucrose (25%)
Add 40ml Milli-Q water.
Adjust pH to 8.0 with HCl
Make up volume to 50ml with Milli-Q water
Store at room temperature

Bromophenol blue solution

0.25g sucrose
0.05g Bromophenol blue
Make up to 10ml with Milli-Q water and store at 4°C

30% Acrylamide w/v

30g acrylamide (*Promega, USA*)
Make up to 100ml with Milli-Q water
Store at 4°C in a dark bottle

1% Bisacrylamide w/v

1g bisacrylamide (*Promega, USA*)
Make up to 100ml with Milli-Q water
Store at 4°C in a dark bottle

4x Stacking gel buffer (pH 6.8)

12.12g Tris (0.5M)
Add 180ml Milli-Q water and pH with HCl
Make up to 200ml with Milli-Q water
Store at 4°C

4x Separating gel buffer (pH 8.8)

18.17g Tris (1.5M)
Make up to 90ml with Milli-Q water and pH with HCl
Adjust volume to 100ml with Milli-Q water
Store at 4°C

10% SDS

1g SDS
10ml Milli-Q water

10% APS (Make fresh each time)

0.1g APS
1ml Milli-Q water

Electrophoresis buffer

3.03g Tris (0.125M)
14.4g glycine (0.096mM)
1g SDS (0.1%)
Make up to 1L with Milli-Q water

12% Separating gel for SDS-PAGE

6ml 30% acrylamide
1.6ml 1% bisacrylamide
3.75ml 4x running gel buffer
80µl 10% SDS
3.5ml Milli-Q water
100µl 10% APS
7.5µl TEMED

4% Stacking gel for SDS-PAGE

1.3ml 30% acrylamide
1ml 1% bisacrylamide
2.5ml 4x stacking gel buffer
20µl 10% SDS
4.9ml Milli-Q water
200µl 10% APS
7.5µl TEMED

0.05% Coomassie blue stain

1g Coomassie blue R-250 (*Merck, Germany*)
500ml Isopropanol
200ml Acetic acid
Make up to 2L with Milli-Q water

1st Destain

10% Acetic acid
10% Methanol
Add 200ml of acetic acid and 200ml of methanol to 1.6L of Milli-Q water

2nd Destain

10% Acetic acid
Add 200ml of acetic acid to 1.8L of Milli-Q water

5.4.3) Reagents for immunoblot analysis of recombinant protein

Transblot buffer

6.06g Tris (25mM)
28.8g glycine (192mM)
Dissolve the reagents in 1.5L Milli-Q water
Add 400ml methanol (20%) and adjust the final volume to 2L with Milli-Q water

Ponceau S (Nitrocellulose membrane stain)

700µl acetic acid
9.3ml Milli-Q water
0.1g Ponceau S (*Sigma-Aldrich Inc, USA*)

TBS buffer (pH 7.5)

6.06g Tris (25mM)
9g NaCl (150mM)
Add 900ml Milli-Q water
Adjust to pH 7.5 with HCl
Adjust to 1L with Milli-Q water

0.5% Tween-TBS

0.5ml Tween-20 (*Sigma-Aldrich Inc, USA*)
1L TBS buffer (pH 7.5)

3% BSA – Anti-GST HRP conjugated blocking buffer

0.75g BSA
25ml TBS buffer (pH 7.5)

Anti-GST HRP-conjugated antibody

(1:100 000)
0.5g BSA in 50ml TBS buffer (pH 7.5)
0.5µl anti-GST HRP-conjugated antibody (*Amersham Biosciences, UK*)

Anti-His HRP-conjugated blocking buffer (*Qiagen, UK*)

0.1g Blocking Reagent
20ml 1x Blocking Reagent Buffer (0.5% w/v)

Heat mixture to 70°C until dissolved
Add 20µl Tween-20 (0.1% v/v)
Allow to cool to room temperature before use

Anti-His HRP-conjugated antibody

(1:2 000)
12.5µl anti-His HRP-conjugated antibody (*Qiagen, UK*)
25ml anti-His HRP-conjugated blocking buffer

Amido black stain

1% amido black (*AppliChem, Germany*)
10% methanol
2% acetic acid
88ml Milli-Q water
Store at room temperature

Amido black destain

50% methanol
7% acetic acid
43ml Milli-Q water
Store at room temperature

5.4.4) Reagents for inclusion body experiments

6M Guanidine HCl buffer

14.32g guanidine HCl (6M)
1.25ml 1M Tris pH 8.0 (final concentration 50mM)
0.15g NaCl (100mM)
0.093g EDTA (10mM)
0.039g DTT (10mM)
Make up to 25ml with Milli-Q water

Refolding buffer stock

100ml 1M Tris-HCl pH 8.0 (final concentration 200mM)
0.93g EDTA (10mM)
43.55g L-Arginine (1M) (*Merck, Germany*)
Make up to 500ml with Milli-Q water
Store at room temperature

To make complete refolding buffer, add the following reagents fresh each time:

0.00044g PMSF (0.1mM)
0.0154g reduced glutathione (2mM) (*Sigma-Aldrich Inc, USA*)

0.0031g oxidised glutathione (0.2mM) (*Sigma-Aldrich Inc, USA*)
Make up to 25ml with refolding buffer base

Note: PMSF is soluble in aqueous solution to a concentration of 2mM; for concentrations >2mM, dissolve in an organic solvent (*Price & Stevens, 2002*).

10mM DTT

0.00154g DTT (*Sigma-Aldrich Inc, USA*)
Make up to 1ml with Milli-Q water
Make up fresh each time

5.5) Reagents for PK assay

Kinase reaction mix

6µl 0.1M Tris-HCl pH 7.4 (final concentration 20mM)
6µl 0.1M $MgCl_2$ (final concentration 20mM)
0.6µl 0.1M $MnCl_2$ (final concentration 2mM)
3µl 0.1M β-glycerophosphate (final concentration 10mM)
3µl 0.1M NaF (final concentration 10mM) (*Calbiochem, Germany*)
3µl 0.1mM ATP (final concentration 10µM) (*Novagen, USA*)
0.5µl 250µCi γ-[^{32}P] ATP (final concentration 2.5µCi) (*PerkinElmer, USA*)
0.5µl 4µg/µl casein (*Merck, Germany*)
0.5µl 10µg/µl myelin basic protein (*Sigma-Aldrich Inc, USA*)
1µl 10µg/µl histone protein 1 (*Merck, Germany*)
Add 5.9µl of recombinant protein to result in a total volume of 30µl

0.1M Tris-HCl stock (pH 7.4)

0.1211g Tris
9ml Milli-Q water
pH to 7.4 with HCl
Make up to 10ml with Milli-Q water
Store at 4°C

0.1M $MgCl_2.6H_2O$ stock

0.2033g $MgCl_2.6H_2O$
10ml Milli-Q water
Store at 4°C

0.1M $MnCl_2.2H_2O$ stock

0.1618g $MnCl_2.2H_2O$
10ml Milli-Q water
Store at 4°C

0.1M NaF stock

0.042g NaF
10ml Milli-Q water
Store at 4°C

0.1M β-glycerophosphate stock

0.3061g β-glycerophosphate (*Calbiochem, Germany*)
10ml Milli-Q water
Store at 4°C

Gel drying solution

4% glycerol
20% ethanol
76ml Milli-Q water

5.6) Data for graphs

TABLE 26: BSA standard curve data from Coomassie assay

SAMPLE	VOLUME (µl)	AMOUNT (µg)	A_{595}	AVERAGE A_{595}
BSA Standard 1	1	2	0.033	0.0365
Duplicate 1	1	2	0.04	
BSA Standard 2	2	4	0.085	0.0915
Duplicate 2	2	4	0.098	
BSA Standard 3	4	8	0.205	0.207
Duplicate 3	4	8	0.209	
BSA Standard 4	6	12	0.295	0.299
Duplicate 4	6	12	0.303	
BSA Standard 5	8	16	0.386	0.4015
Duplicate 5	8	16	0.417	

5.7) Vector maps

pGEX-4T-2 (27-4581-01)

Thrombin

Leu Val Pro Arg↓Gly Ser Pro Gly Ile Pro Gly Ser Thr Arg Ala Ala Ala Ser
CTG GTT CCG CGT GGA TCC CCA GGA ATT CCC GGG TCG ACT CGA GCG GCC GCA TCG TGA

BamH I · EcoR I · Sma I · Sal I · Xho I · Not I · Stop codon

pGEX-4T-3 (27-4583-01)

Thrombin

Leu Val Pro Arg↓Gly Ser Pro Asn Ser Arg Val Asp Ser Ser Gly Arg Ile Val Thr Asp
CTG GTT CCG CGT GGA TCC CCG AAT TCC CGG GTC GAC TCG AGC GGC CGC ATC GTG ACT GAC TGA

BamH I · EcoR I · Sma I · Sal I · Xho I · Not I · Stop codons

pGEX-3X (27-4803-01)

Factor Xa

Ile Glu Gly Arg↓Gly Ile Pro Gly Asn Ser Ser
ATC GAA GGT CGT GGG ATC CCC GGG AAT TCA TCG TGA CTG ACT GAC

BamH I · Sma I · EcoR I · Stop codons

pGEX-5X-1 (27-4584-01)

Factor Xa

Ile Glu Gly Arg↓Gly Ile Pro Glu Phe Pro Gly Arg Leu Glu Arg Pro His Arg Asp
ATC GAA GGT CGT GGG ATC CCC GAA TTC CCG GGT CGA CTC GAG CGG CCG CAT CGT GAC TGA

BamH I · EcoR I · Sma I · Sal I · Xho I · Not I · Stop codons

pGEX-5X-2 (27-4585-01)

Factor Xa

Ile Glu Gly Arg↓Gly Ile Pro Gly Ile Pro Gly Ser Thr Arg Ala Ala Ala Ser
ATC GAA GGT CGT GGG ATC CCC GGA ATT CCC GGG TCG ACT CGA GCG GCC GCA TCG TGA

BamH I · EcoR I · Sma I · Sal I · Xho I · Not I · Stop codon

pGEX-5X-3 (27-4586-01)

Factor Xa

Ile Glu Gly Arg↓Gly Ile Pro Arg Asn Ser Arg Val Asp Ser Ser Gly Arg Ile Val Thr Asp
ATC GAA GGT CGT GGG ATC CCC AGG AAT TCC CGG GTC GAC TCG AGC GGC CGC ATC GTG ACT GAC TGA

BamH I · EcoR I · Sma I · Sal I · Xho I · Not I · Stop codons

pGEX-6P-1 (27-4597-01)

PreScission™ Protease

Leu Glu Val Leu Phe Gln↓Gly Pro Leu Gly Ser Pro Glu Phe Pro Gly Arg Leu Glu Arg Pro His
CTG GAA GTT CTG TTC CAG GGG CCC CTG GGA TCC CCG GAA TTC CCG GGT CGA CTC GAG CGG CCG CAT

BamH I · EcoR I · Sma I · Sal I · Xho I · Not I

pGEX-6P-2 (27-4598-01)

PreScission™ Protease

Leu Glu Val Leu Phe Gln↓Gly Pro Leu Gly Ser Pro Gly Ile Pro Gly Ser Thr Arg Ala Ala Ala Ser
CTG GAA GTT CTG TTC CAG GGG CCC CTG GGA TCC CCA GGA ATT CCC GGG TCG ACT CGA GCG GCC GCA TCG

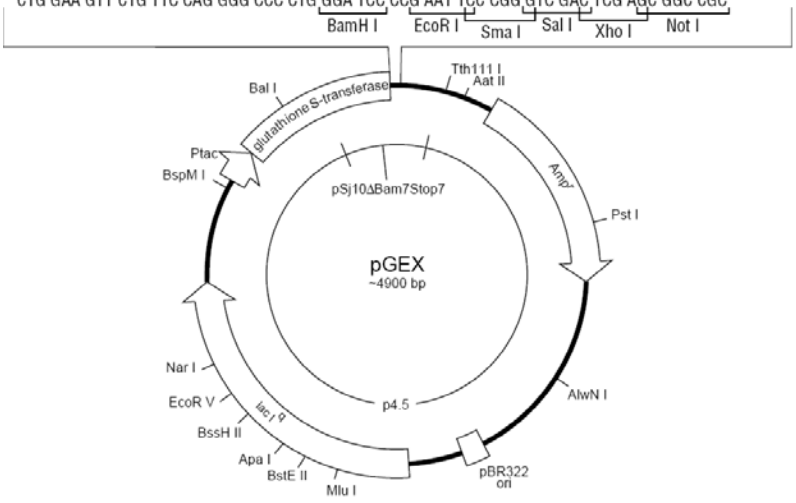
BamH I · EcoR I · Sma I · Sal I · Xho I · Not I

pGEX-6P-3 (27-4599-01)

PreScission™ Protease

Leu Glu Val Leu Phe Gln↓Gly Pro Leu Gly Ser Pro Asn Ser Arg Val Asp Ser Ser Gly Arg
CTG GAA GTT CTG TTC CAG GGG CCC CTG GGA TCC CCG AAT TCC CGG GTC GAC TCG AGC GGC CGC

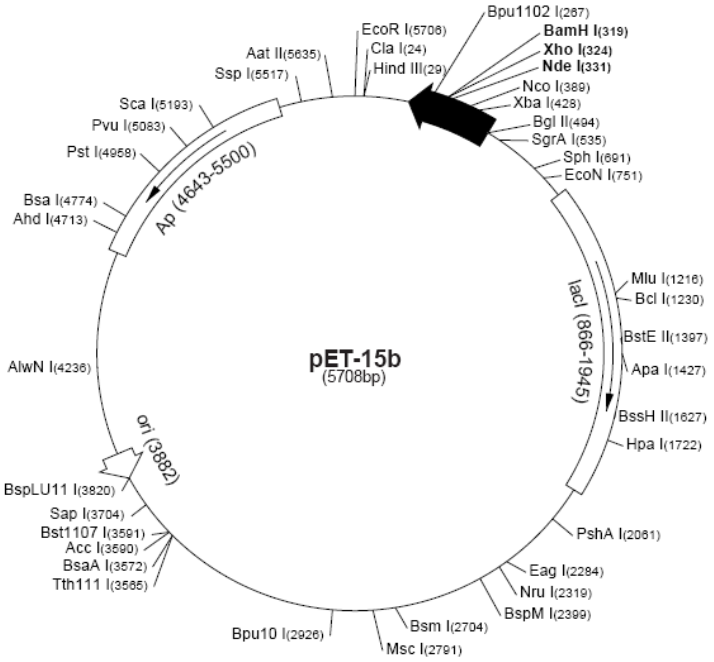
BamH I · EcoR I · Sma I · Sal I · Xho I · Not I

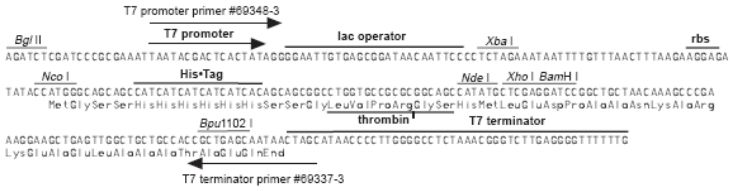


pET-15b Vector

The pET-15b vector (Cat. No. 69661-3) carries an N-terminal His• Tag® sequence followed by a thrombin site and three cloning sites. Unique sites are shown on the circle map. Note that the sequence is numbered by the pBR322 convention, so the T7 expression region is reversed on the circular map. The cloning/expression region of the coding strand transcribed by T7 RNA polymerase is shown below.

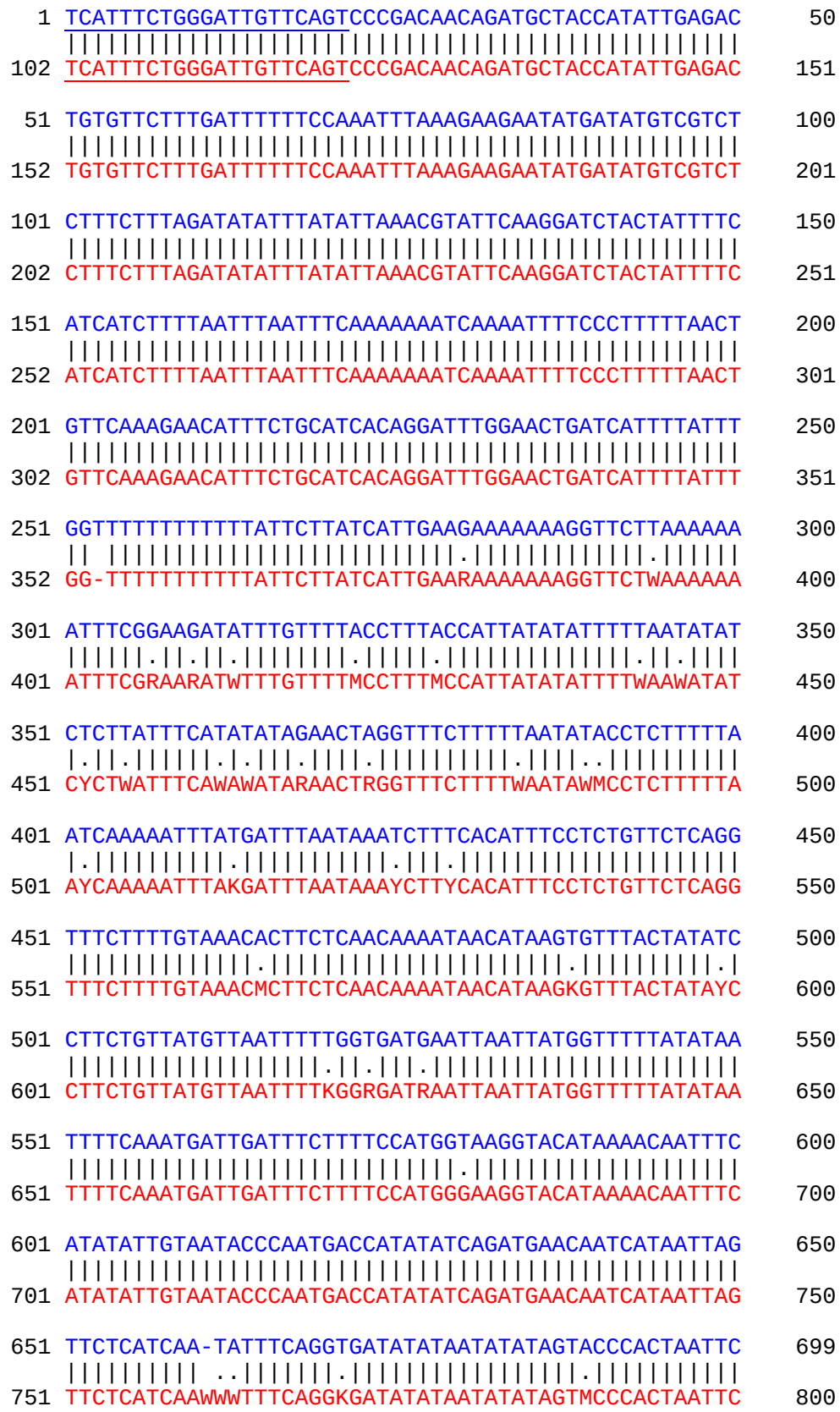
pET-15b sequence landmarks	
T7 promoter	453-469
T7 transcription start	452
His• Tag coding sequence	362-380
Multiple cloning sites (*Nde* I - *Bam*H I)	319-335
T7 terminator	213-259
lacI coding sequence	(866-1945)
pBR322 origin	3882
bla coding sequence	4643-5500





pET-15b cloning/expression region

5.8) Sequencing data



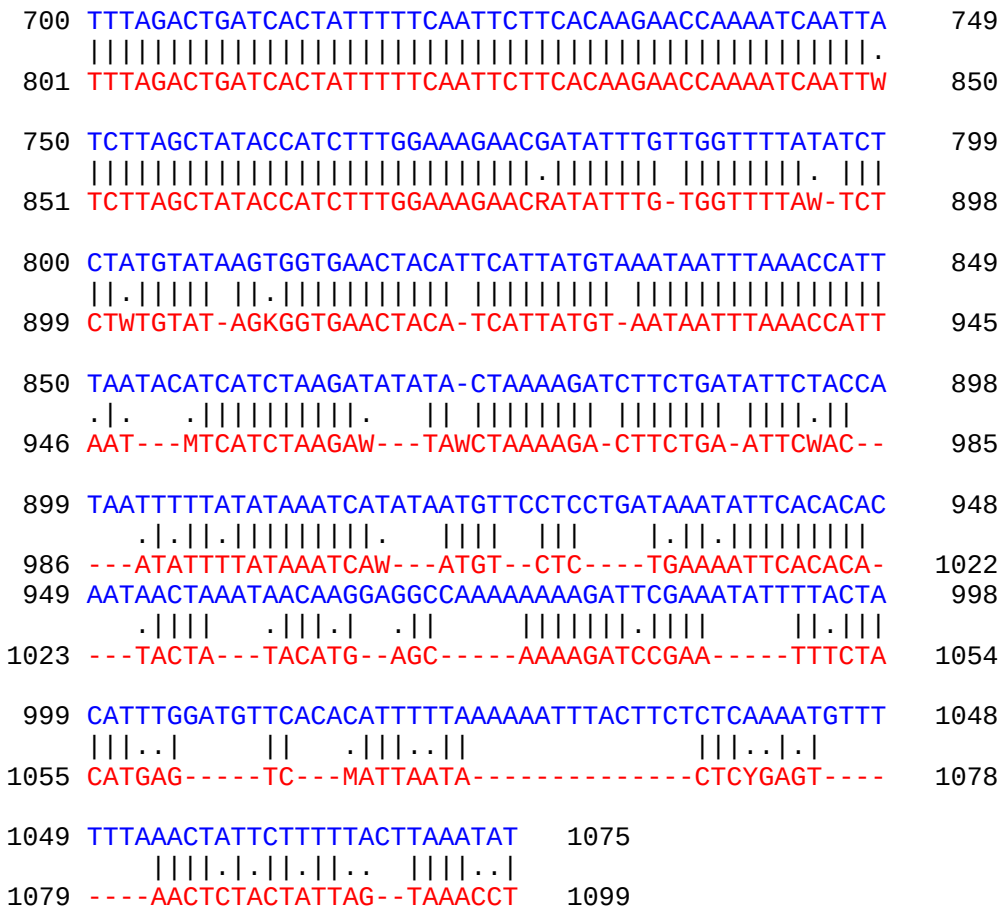
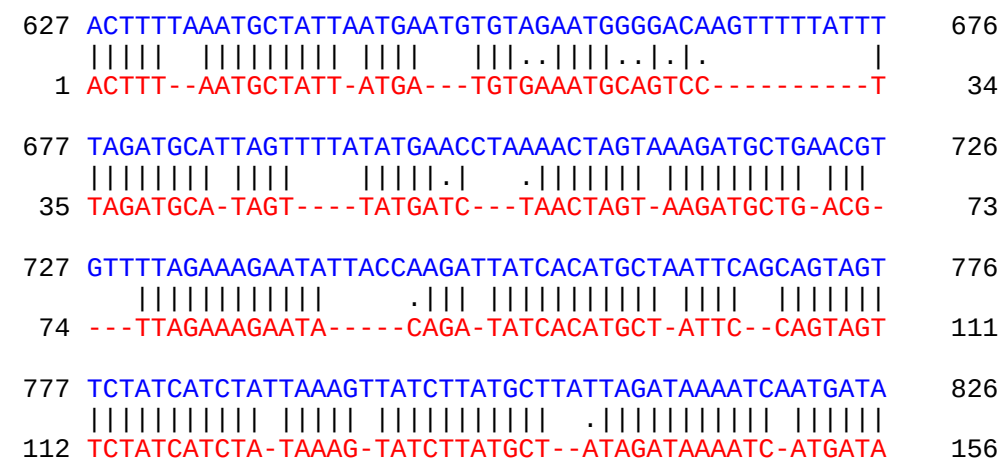
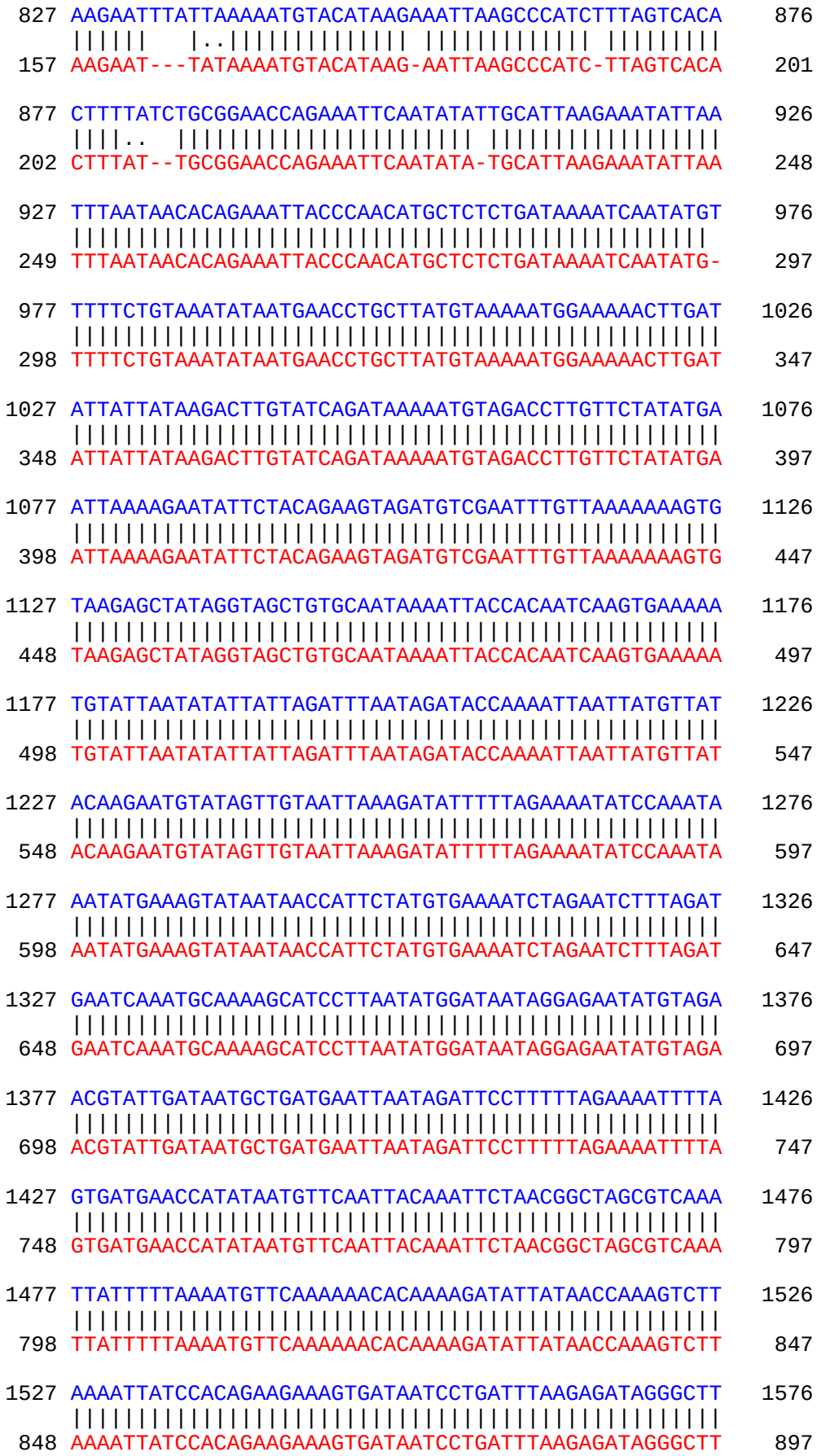


Figure 62. Alignment of PFB0150c DNA sequences from 3' sequencing

The sequence obtained from subcloning (red) was aligned with the original sequence obtained from the PlasmoDB database (blue). The underlined sections indicate the reverse primer position 5'-TCA TTT CTG GGA TTG TTC AGT -3'. Key to ambiguous bases: W = T and A overlap in sequence; M = multiple overlap of peaks in sequence; R = A and G overlap; Y = C and T overlap. A dot between base pairs indicates a non-alignment, while a dash indicates a missing base.




827 AAGAATTTATTAAAAATGTACATAAGAAATTAAGCCCATCTTTAGTCACA 876
157 AAGAAT---TATAAAATGTACATAAG-AATTAAGCCCATC-TTAGTCACA 201
877 CTTTTATCTGCGGAACCAGAAATTCAATATATTGCATTAAGAAATATTAA 926
202 CTTTAT--TGCGGAACCAGAAATTCAATATA-TGCATTAAGAAATATTAA 248
927 TTTAATAACACAGAAATTACCCAACATGCTCTCTGATAAAATCAATATGT 976
249 TTTAATAACACAGAAATTACCCAACATGCTCTCTGATAAAATCAATATG- 297
977 TTTTCTGTAAATATAATGAACCTGCTTATGTAAAAATGGAAAAACTTGAT 1026
298 TTTTCTGTAAATATAATGAACCTGCTTATGTAAAAATGGAAAAACTTGAT 347
1027 ATTATTATAAGACTTGTATCAGATAAAAATGTAGACCTTGTTCTATATGA 1076
348 ATTATTATAAGACTTGTATCAGATAAAAATGTAGACCTTGTTCTATATGA 397
1077 ATTAAAAGAATATTCTACAGAAGTAGATGTCGAATTTGTTAAAAAAAGTG 1126
398 ATTAAAAGAATATTCTACAGAAGTAGATGTCGAATTTGTTAAAAAAAGTG 447
1127 TAAGAGCTATAGGTAGCTGTGCAATAAAATTACCACAATCAAGTGAAAAA 1176
448 TAAGAGCTATAGGTAGCTGTGCAATAAAATTACCACAATCAAGTGAAAAA 497
1177 TGTATTAATATATTATTAGATTTAATAGATACCAAAATTAATTATGTTAT 1226
498 TGTATTAATATATTATTAGATTTAATAGATACCAAAATTAATTATGTTAT 547
1227 ACAAGAATGTATAGTTGTAATTAAAGATATTTTTAGAAAATATCCAAATA 1276
548 ACAAGAATGTATAGTTGTAATTAAAGATATTTTTAGAAAATATCCAAATA 597
1277 AATATGAAAGTATAATAACCATTCTATGTGAAAATCTAGAATCTTTAGAT 1326
598 AATATGAAAGTATAATAACCATTCTATGTGAAAATCTAGAATCTTTAGAT 647
1327 GAATCAAATGCAAAAGCATCCTTAATATGGATAATAGGAGAATATGTAGA 1376
648 GAATCAAATGCAAAAGCATCCTTAATATGGATAATAGGAGAATATGTAGA 697
1377 ACGTATTGATAATGCTGATGAATTAATAGATTCCTTTTTAGAAAATTTTA 1426
698 ACGTATTGATAATGCTGATGAATTAATAGATTCCTTTTTAGAAAATTTTA 747
1427 GTGATGAACCATATAATGTTCAATTACAAATTCTAACGGCTAGCGTCAAA 1476
748 GTGATGAACCATATAATGTTCAATTACAAATTCTAACGGCTAGCGTCAAA 797
1477 TTATTTTTAAAATGTTCAAAAAACACAAAAGATATTATAACCAAAGTCTT 1526
798 TTATTTTTAAAATGTTCAAAAAACACAAAAGATATTATAACCAAAGTCTT 847
1527 AAAATTATCCACAGAAGAAAGTGATAATCCTGATTTAAGAGATAGGGCTT 1576
848 AAAATTATCCACAGAAGAAAGTGATAATCCTGATTTAAGAGATAGGGCTT 897

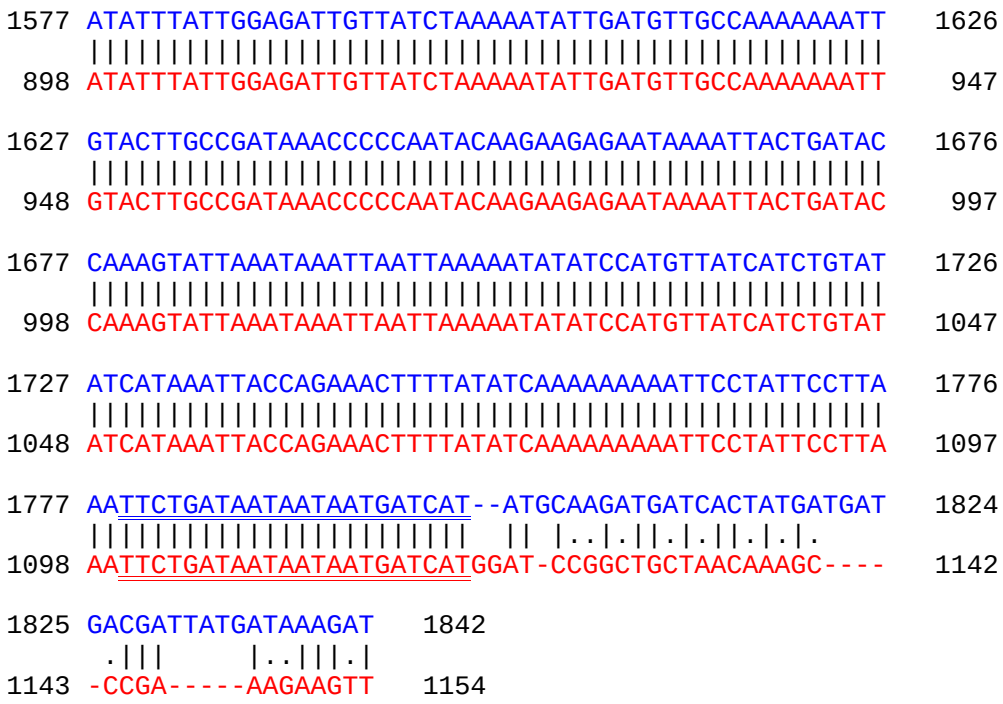


Figure 63. Alignment of AP1 N-terminal domain DNA sequences from T7 terminal primer sequencing

The sequence obtained from subcloning (red) was aligned with the original sequence obtained from the PlasmoDB database (blue). The underlined sections indicate the reverse primer position 5'-TTC TGA TAA TAA TAA TGA TCA T-3'. Key to ambiguous bases: W = T and A overlap in sequence; M = multiple overlap of peaks in sequence; R = A and G overlap; Y = C and T overlap. A dot between base pairs indicates a non-alignment, while a dash indicates a missing base.

5.9) BLAST data

Note: For a given score, the E-value is the number of hits in a database search that we expect to see by chance with this score or better. The lower the E-value, the more significant the score.

```
3ckx_A mol: protein length: 304 Serine/threonine-protein kinase 24
Organism: Homo sapiens Length 304

Score = 355 (130.0 bits), Expect = 2.0e-32, P = 2.0e-32
Identities = 87/2485 (3%), Positives = 127/2485 (5%)
Identities = 87/230 (37%), Positives = 127/230 (55%)

Hs:    59 LSQCDSPYVTKYYGSYL-KDTKLWIIMEYLGGGSALDLLEP-GPLDETQIATILREILKG 116
          L  C+ P V KY+ S+      L I+ EYL GG+  DL +  G + E  +  IL ++L G
Pf:  2146 LKMCEHPNVVKYFESFFWPPCYLVIVCEYLSGGTLYDLYKNYGRISEDLLVYILDDVLNG 2205

Hs:   117 LDYLHSEKK---IHRDIKAANVLLSEHGEVKLADFGVAGQLTDTQIKRNXFVGTPFWMAP 173
          L+YLH+E     IHRDIK  N++LS+ G  K+ DFG   +L ++   +   VGT ++++P
Pf:  2206 LNYLHNECSSPLIHRDIKPTNIVLSKDGIAKIIDFGSCEELKNSDQSKE-LVGTIYYISP 2265

Hs:   174 EVIKQSAYDSKADIWSLGITAIELARGEPPHSELHPMKVLFLIPKNNPPTL---EGNYSK 230
          E++ ++ YD  +DIWSLGIT  E+     P       +       N+ P +   EG YSK
Pf:  2266 EILMRTNYDCSSDIWSLGITIYEIVLCTLPWKRNQSFENYIKTIINSSPKINITEG-YSK 2325

Hs:   231 PLKEFVEACLNKEPSFRPTAKELLKHKFILRNA---KKTSYLTELIDRYK 277
           L  FVE CL K+P  R   K+LL HKF+++     KK S + E+ D  K
Pf:  2326 HLCYFVEKCLQKKPENRGNVKDLLNHKFLIKKRYIKKKPSSIYEIRDILK 2375
```

Figure 64. BLAST data for P. falciparum PK with Homo sapiens protein

The parasite PK insert shares most homology with human serine/threonine protein kinase 24 within the kinase domain which spans amino acids 2 079-2 354. The sequence homology begins at parasite PK amino acid 2 146 and ends at 2 375. Conserved residues found in the catalytic core of higher eukaryotes are highlighted. The invariant residues are underlined. A dash indicates no amino acid match. A plus sign indicates sequence positivity, while a matching letter shows areas of sequence identity. Pf indicates the parasite protein sequence and Hs the amino acid sequence of the human protein kinase.

```
>gb|PVX_003590 | organism=Plasmodium_vivax_SaI-1 |
product=serine/threonine-specific protein kinase, putative |
location=CM000445:748722-755099(+) | length=2125
        Length = 2125

 Score = 1291 (459.5 bits), Expect = 6.3e-141, Sum P(5) = 6.3e-141
Identities = 265/2485 (11%), Positives = 354/2485 (14%)
 Identities = 265/533 (49%), Positives = 354/533 (66%)

Query:  1932 ENLNLENKKKGYIDETN-VNENYESDNE-YDSDEDDTESDNDDEQNKENERGDEKDGYEE 1989
              E  NL+ +K      +N + EN+E   E  +   DDT           N  G  K   EE
Sbjct:  1532 EEHNLKRQKSDSTRSSNEMLENFERIAERINFILDDTVEFFKKNLYVHNGYGSVKVSKEE 1591

Query:  1990 MNGGDKNE--EMNGGDKNEEMNVGDKNGGINEEHKNEGINEEHKDELINKE--HKNERIN 2045
                    +N   E +   K  ++     +G +    + E   ++ ++E    E  H  ++ N
Sbjct:  1592 TGWLGENTLGEWSHVYKINKVVCKGAHGVVFSAWRGEDSAKQCEEEAERGEAGHSTDQGN 1651

Query:  2046 EEHKNERINEEHKNEGINEEH-KNEGINEEHKNERINEEHKNEGINKLTYH-NMNKNNIS 2103
                E   E         G  EE+ K +  +    N + N +  + G ++     N+ +N+
Sbjct:  1652 AEGSGEG-GGGGSGGGSGEENGKGKDQDNAQGNAQGNAQGNDRGNDQENVEENVEENDQG 1710
```

```
Query:  2104 NENNYNDDDSYDEDNLVSLKIINLKYLSKKNSLKNILREVNFLKMCEHPNVVKYFESFFW 2163
             N+ +  D +SYDE+ LV+LKI+NL+YLSKKNSL+ I++EV+FL++C+HPN+VKY ESFFW
Sbjct:  1711 NDRD-GDQESYDEEKLVTLKIVNLRYLSKKNSLRRIMKEVHFLQICDHPNIVKYHESFFW 1769

Query:  2164 PPCYLVIVCEYLSGGTLYDLYKNYGRISEDLLVYILDDVLNGLNYLHNECSSPLIHRDIK 2223
             PPCYLVIVCE+LSGGTL+DLYK  GRI+ED+LV+ILDDVL  L YLHNEC+S L+HRDIK
Sbjct:  1770 PPCYLVIVCEFLSGGTLFDLYKKCGRITEDVLVHILDDVLKALQYLHNECTSCLVHRDIK 1829

Query:  2224 PTNIVLSKDGIAKIIDFGSCEELKNSDQSKELVGTIYYISPEILMRTNYDCSSDIWSLGI 2283
             PTNIV SK G+AKI+DFGSCE +++  +  E+VGT+YYISPEIL R  YDCS+DIWSLGI
Sbjct:  1830 PTNIVFSKSGVAKIVDFGSCERVEDL-KMHEVVGTLYYISPEILKREKYDCSADIWSLGI 1888

Query:  2284 TIYEIVLCTLPWKRNQSFENYIKTIINSSPKINITEGYSKHLCYFVEKCLQKKPENRGNV 2343
             TIYE+V+C LPWK  +  E  IK I+ SSPKIN+  G++K  C+FVE CLQ  P  R N
Sbjct:  1889 TIYEVVMCALPWKGKKHIEESIKQIVGSSPKINLCSGFTKQFCFFVESCLQNDPGKRANA 1948

Query:  2344 KDLLNHKFLIKKRYIKKKPSSIYEIRDILKIYNGKGKTNIFRNFFKNLFFFNDKNKKKKP 2403
               LL HKFL KKR +++KPSSI+EIRDILK+ NGKGK NIFRNFFKNLFF NDKNK+++
Sbjct:  1949 AHLLGHKFLTKKRLLRRKPSSIFEIRDILKVNNGKGKNNIFRNFFKNLFFLNDKNKRRR- 2007

Query:  2404 NKMISSKSCDAEMFFEQLKRENFDFFEIKLKDDENSRSLNTFNINISKERDDI 2456
             NK + SKSC+ EMF+ +LKRENFDFFEI+L+D  +SRSL   ++   K  +++
Sbjct:  2008 NKALGSKSCEPEMFYRKLKRENFDFFEIRLRDG-SSRSLGHLHVGEGKREEEV 2059
```

Figure 65. BLAST data for P. falciparum PK with P. vivax protein

The P. falciparum PK shares homology with P. vivax putative serine/threonine specific protein kinase within the kinase domain which spans amino acids 2 079-2 354 of the entire protein sequence. The sequence homology begins at P. falciparum PK amino acid 2 119 and ends at 2 456. Conserved residues found in the catalytic core of higher eukaryotes are highlighted. The invariant residues are underlined. A dash indicates no amino acid match. A plus sign indicates sequence positivity, while a matching letter shows areas of sequence identity. Query indicates the P. falciparum protein sequence and Sbjct the amino acid sequence of the P. vivax protein kinase.

```
Toxoplasma_gondii|X|42.m03467|Annotation|Toxoplasma_gondii_TIGR|(protein
coding) serine/threonine-protein kinase-related / mediator complex
subunit SOH1-related
        Length = 4253

Score = 286 (105.7 bits), Expect = 5.1e-40, Sum P(5) = 5.1e-40
Identities = 181/2485 (7%), Positives = 276/2485 (11%)
Identities = 128/305 (42%), Positives = 185/305 (61%)

Query:  2263 SPEILMRTNYDCSSDIWSLGITIYEIVLCTLPWKRNQSFENYIKTIINSSP-KINITEGY 2321
             SP       YD   D+W+LGIT YE+ + +LPW R    E+ + TI+  SP +IN+ EGY
Sbjct:  1870 SPGFGASEGYDFRVDLWALGITAYEVAVGSLPWPRRMRLEDLLGTILEGSPPRINLNEGY 1929

Query:  2322 SKHLCYFVEKCLQKKPENRGNVKDLLNHKFLIKK-RYIKKKPSSIYEIRDILKIYNGKGK 2380
              K  C+FVE+CL+K  E RG   +LL H F  K     + +P+S  E+R+ +++ NGK K
Sbjct:  1930 EKTFCFFVERCLRKNSEERGTATELLQHPFFKKFCASSRSRPASHQELREAIEVANGKRK 1989

Query:  2381 TNIFRNFFKNLFFFNDK---NKKKK 2402
              N+  N  K + FF      N +KK
Sbjct:  1990 ANVLSNLLKFIPFFRKPQPTNGRKK 2014
-------------------------------------------------------------------
Query:  2111 DDSYDEDNLVSLKIINLKYL----------SKKNSLKNILREVNFLKMCEHPNVVKYFES 2160
             ++  +ED  V+LKII+L             ++ N L  ++REV+ LK C H  +VK +E+
Sbjct:  1633 EEDVEEDE-VALKIIDLDGALQLQGATDKEARVNYLNQVMREVDVLKQCNHEGIVKMYEA 1691
```

```
Query:  2161 FFWPPCYLVIVCEYLSGGTLYDLYKNYGRISEDLLVYILDDVLNGLNYLHN 2211
             F WPPCYLV   E L GG+L DLY + G + E L+  +L DVL  L YLHN
Sbjct:  1692 FQWPPCYLVFSMELLPGGSLRDLYASAGPLPEPLIAALLQDVLKALKYLHN 1742
```

```
Query:  1897 SCIYKINKIVRKGAHGVVFSAWRSENV--------DFFNHSFFENLNLENKKKGYIDETN 1948
             S +Y + ++V KGAHG VF   R              F  +F + ++   +KK   +ET
Sbjct:  1554 SAVYVVEELVGKGAHGAVFKCTRYRKAARPEEKTGGSFLGAFGKKVSEAREKKA--EETR 1611

Query:  1949 VN-ENYESDNEY-DSDEDDTESDNDDEQNKENERGDEKDGYEEMNGG-DKNEEMN 2000
                E  E   +  D DE D  ++ D E+++   +  + DG  ++ G  DK   +N
Sbjct:  1612 PGVERPEGARQVSDGDERDG-AEEDVEEDEVALKIIDLDGALQLQGATDKEARVN 1665
```

```
Query:   686 PKENNIIYTSGKS 698
             P+EN ++ +SGKS
Sbjct:    21 PEENALLGSSGKS 33
```

Figure 66. BLAST data for P. falciparum PK with T. gondii proteins

The P. falciparum PK shares most homology with the T. gondii serine/threonine protein kinase-related complex within the kinase domain which spans amino acids 2 079-2 354 of the entire protein sequence. The sequence homology occurs at many different areas throughout the T. gondii protein, with large gaps of no homology in between. Conserved residues found in the catalytic core of higher eukaryotes are highlighted. The invariant residues are underlined. A dash indicates no amino acid match. A plus sign indicates sequence positivity, while a matching letter shows areas of sequence identity. Query indicates the P. falciparum protein sequence and Sbjct the T. gondii protein kinase.

```
gb|PVX_079845 | organism=Plasmodium_vivax_SaI-1 | product=adapter-
related protein complex 1 beta 1 subunit, putative |
location=CM000451:144521-148411(+) | length=930
        Length = 930

 Score = 4205 (1485.3 bits), Expect = 0., P = 0.
 Identities = 837/925 (89%), Positives = 893/925 (95%)

Query:     1 MSDLRYFQTTKKGEIHELKEELHSSHKEKKKEAIKKIIAAMTVGKDVSTLFSDVVNCMQT 60
             MSDLRYFQTTKKGEIHELKEELHSSHKEKKKEAIKKIIAAMTVGKDVSTLFSDVVNCMQT
Sbjct:     1 MSDLRYFQTTKKGEIHELKEELHSSHKEKKKEAIKKIIAAMTVGKDVSTLFSDVVNCMQT 60

Query:    61 SNIELKKLVYLYVINYAKVQPELAILAVNTFRKDSSDPNPLIRALAIRTMGCIRLEQITE 120
             SNIELKKLVYLYVINYAKVQPELAILAVNTFRKDSSDPNPLIRALAIRTMGCIRLEQITE
Sbjct:    61 SNIELKKLVYLYVINYAKVQPELAILAVNTFRKDSSDPNPLIRALAIRTMGCIRLEQITE 120

Query:   121 YLIEPLRRCLKDEDPYVRKTAVICIAKLYDISPKLVEEEGFIDTLLDILDDNNAMVVANA 180
             YLIEPLRRCLKDEDPYVRKTAVICIAKLYDISPKLVEEEGFI+TLL+ILDDNNAMVVANA
Sbjct:   121 YLIEPLRRCLKDEDPYVRKTAVICIAKLYDISPKLVEEEGFIETLLNILDDNNAMVVANA 180

Query:   181 VISLTDICENSNKSILKDVINKDENNVNKLLNAINECVEWGQVFILDALVLYEPKTSKDA 240
             +ISLTDICENSNKSILKDVINKDENNVNKLLNAINECVEWGQVFILDALVLYEPKTSKDA
Sbjct:   181 IISLTDICENSNKSILKDVINKDENNVNKLLNAINECVEWGQVFILDALVLYEPKTSKDA 240

Query:   241 ERVLERILPRLSHANSAVVLSSIKVILCLLDKINDKEFIKNVHKKLSPSLVTLLSAEPEI 300
             ERVLERILPRLSHANSAVVLSSIKVIL LLDKINDKEFIKNVHKKLSPSLVTLLSAEPEI
Sbjct:   241 ERVLERILPRLSHANSAVVLSSIKVILSLLDKINDKEFIKNVHKKLSPSLVTLLSAEPEI 300

Query:   301 QYIALRNINLITQKLPNMLSDKINMFFCKYNEPAYVKMEKLDIIIRLVSDKNVDLVLYEL 360
             QYIALRNINLITQKLP+ML+DKINMFFCKYNEPAYVKMEKLDIIIRLVSDKNVDLVLYEL
Sbjct:   301 QYIALRNINLITQKLPHMLADKINMFFCKYNEPAYVKMEKLDIIIRLVSDKNVDLVLYEL 360
```

```
Query:   361 KEYSTEVDVEFVKKSVRAIGSCAIKLPQSSEKCINILLDLIDTKINYVIQECIVVIKDIF 420
             KEYSTEVDVEFVKKSVRAIGSCAIKLPQSSEKCINILLDLIDTKINYVIQECIVVIKDIF
Sbjct:   361 KEYSTEVDVEFVKKSVRAIGSCAIKLPQSSEKCINILLDLIDTKINYVIQECIVVIKDIF 420

Query:   421 RKYPNKYESIITILCENLESLDESNAKASLIWIIGEYVERIDNADELIDSFLENFSDEPY 480
             RKYPNKYESIITILCENLESLDESNAKASLIWIIGEYVERI+NADELIDSFLENF+DEPY
Sbjct:   421 RKYPNKYESIITILCENLESLDESNAKASLIWIIGEYVERIENADELIDSFLENFTDEPY 480

Query:   481 NVQLQILTASVKLFLKCSKNTKDIITKVLKLSTEESDNPDLRDRAYIYWRLLSKNIDVAK 540
             NVQLQILTASVKLFLKCSKNTKDIITKVLKLSTEESDNPDLRDRA+IYWRLLSKNI++AK
Sbjct:   481 NVQLQILTASVKLFLKCSKNTKDIITKVLKLSTEESDNPDLRDRAFIYWRLLSKNIEIAK 540

Query:   541 KIVLADKPPIQEENKITDTKVLNKLIKNISMLSSVYHKLPETFISKKNSYSLNSDNNNDH 600
             KIVLA+KPPIQE+NKITDTKVLNKLIKNISMLSSVYHKLPETFISKK +Y  ++D ++D
Sbjct:   541 KIVLAEKPPIQEDNKITDTKVLNKLIKNISMLSSVYHKLPETFISKKANYVFHNDKDDD- 599

Query:   601 MQDDHYDD-DDYDKDNHVLKIKKQMDKQKYDSYSSDNKKSNHSKSSSDSYNNSSDEFNND 659
              +D+  D+ DDY+    V K KKQM++QKYDSYSS+++KSN S+SSS+S NNSSD+ N+D
Sbjct:   600 -EDNRIDNMDDYN----VSKFKKQMERQKYDSYSSESRKSNRSRSSSESSNNSSDDANDD 654

Query:   660 IDD---ADDSKKSMDLIGLNDDES---KPQKTIPPVKMVQVLSSEDAGLKGQTGLSIFAS 713
              DD   ADDSKKSMDLIGLNDDE+   KP+++ PPVK+VQVLS ED GLKGQTGLSI +S
Sbjct:   655 KDDDEDADDSKKSMDLIGLNDDEASNAKPRRSAPPVKLVQVLSPEDTGLKGQTGLSILSS 714

Query:   714 INRIDRKIQLKISVTNQTQNEIVVSGVQINKNSFGLSSPNNLDVQNIGFGETKEMLIYLI 773
             INRI+ KIQLKI+VTNQT N +V+SGVQINKNSFGLSSPNNLD+QN+ FGETKE+LI L+
Sbjct:   715 INRIEGKIQLKIAVTNQTPNPVVISGVQINKNSFGLSSPNNLDIQNVSFGETKEILILLV 774

Query:   774 PNTLNSNTPPATPLFLQVAIRTNLDIFYFNVPYDIFVVFVENFHMEKDIFKKKWQIIEEA 833
             PN LNSNTPP+TPLFLQVAIRT++DIFYFNVPYDIF+VFVENF+MEKDIFKKKWQ+IE++
Sbjct:   775 PNLLNSNTPPSTPLFLQVAIRTSIDIFYFNVPYDIFIVFVENFNMEKDIFKKKWQLIEDS 834

Query:   834 KESILMAVSPMVITSDMLIKRMKIFNISLIARRN----ELYYFACITTNNLVILSEVTIQ 889
             KESILMA SPMVITSD+LIKRMKIFNISLIARRN    ELYYFAC+TTNNLVILSEV IQ
Sbjct:   835 KESILMASSPMVITSDILIKRMKIFNISLIARRNVNNMELYYFACLTTNNLVILSEVAIQ 894

Query:   890 PEKKNVKLCIRTDSSSVIPLYKLLFVKAFSLSVTQT 925
             PEKK VKLC+RTDS+SVIPLYKLLFVKAFSLSVTQT
Sbjct:   895 PEKKVVKLCVRTDSTSVIPLYKLLFVKAFSLSVTQT 930
```

Figure 67. BLAST data for P. falciparum AP-1 β1 subunit with P. vivax protein

The P. falciparum AP-1 β1 subunit shares homology with P. vivax putative adaptor related complex 1 β1subunit across the entire protein sequence of 925 amino acids. The sequence homology begins at P. falciparum amino acid 1 and ends at 925. The N-terminal is coloured red and spans amino acids 1 to 618, while the C-terminal is shown in blue and spans amino acids 650 to 925. A dash indicates no amino acid match. A plus sign indicates sequence positivity, while a matching letter shows areas of sequence identity. Query indicates the P. falciparum protein sequence and Sbjct the amino acid sequence of the P. vivax protein.

>Toxoplasma_gondii|VI|49.m00005|Annotation|Toxoplasma_gondii_TIGR|(prote
in coding) beta adaptin protein, putative
Length = 925

Score = 2616 (925.9 bits), Expect = 9.6e-274, P = 9.6e-274
Identities = 519/925 (56%), Positives = 689/925 (74%)

```
Query:     1 MSDLRYFQTTKKGEIHELKEELHSSHKEKKKEAIKKIIAAMTVGKDVSTLFSDVVNCMQT 60
             M+D  YFQ  K+GE+HELKEELHSS+KEKKKEA+KK+IAAMTVGKDVS+LF DVVNCMQT
Sbjct:     1 MTDGNYFQPAKRGELHELKEELHSSNKEKKKEAVKKVIAAMTVGKDVSSLFPDVVNCMQT 60

Query:    61 SNIELKKLVYLYVINYAKVQPELAILAVNTFRKDSSDPNPLIRALAIRTMGCIRLEQITE 120
             +N+ELKKLVYLYVINYAK QPELAILA+NTFRKDS DPNPLIRALA+RTMGCIRLE+ITE
Sbjct:    61 TNMELKKLVYLYVINYAKAQPELAILAINTFRKDSLDPNPLIRALAVRTMGCIRLEEITE 120

Query:   121 YLIEPLRRCLKDEDPYVRKTAVICIAKLYDISPKLVEEEGFIDTLLDILDDNNAMVVANA 180
             YL+EPLRR  KD DPYVRKTA IC+AKL+ I P +V EEGFI+ L  +L D+N +VVANA
Sbjct:   121 YLVEPLRRSCKDPDPYVRKTAAICVAKLFSIRPDMVGEEGFIEELTTMLSDSNPVVVANA 180

Query:   181 VISLTDICENSNKSILKDVINKDENNVNKLLNAINECVEWGQVFILDALVLYEPKTSKDA 240
             V +L++I ENS ++ +K+++N  E+NVNKLL A+NEC EWGQVFILDAL  +EP+T + A
Sbjct:   181 VAALSEISENSGRNYMKNILNAKESNVNKLLAALNECTEWGQVFILDALAQFEPETPRAA 240

Query:   241 ERVLERILPRLSHANSAVVLSSIKVILCLLDKINDKEFIKNVHKKLSPSLVTLLSAEPEI 300
             E VL+R+  RLSHANSAVVLS+IKV++ LLDK+ + + ++ VH+KL P LVTLLSAEPEI
Sbjct:   241 ESVLDRVTARLSHANSAVVLSAIKVVMKLLDKVTNPDVVRAVHRKLCPPLVTLLSAEPEI 300

Query:   301 QYIALRNINLITQKLPNMLSDKINMFFCKYNEPAYVKMEKLDIIIRLVSDKNVDLVLYEL 360
             QY+ALRNI LI QK P++L+ ++ MFFCKYN+P YVK+EKLDI++RLVS+KNVD VL EL
Sbjct:   301 QYVALRNIELIVQKRPSILASEVKMFFCKYNDPVYVKIEKLDILVRLVSEKNVDQVLSEL 360

Query:   361 KEYSTEVDVEFVKKSVRAIGSCAIKLPQSSEKCINILLDLIDTKINYVIQECIVVIKDIF 420
             KEY+TEVDV+FV+K+VR IG CAIKL  ++E+C+ +LLDLI TK+NYV+QE IV IKDIF
Sbjct:   361 KEYATEVDVDFVRKAVRCIGRCAIKLDCAAERCVAVLLDLIQTKVNYVVQEAIVAIKDIF 420

Query:   421 RKYPNKYESIITILCENLESLDESNAKASLIWIIGEYVERIDNADELIDSFLENFSDEPY 480
             RKYPN+YES+I+ LCENLE+LDE  AKAS++WI+GEYV+RIDNADEL+++FLE F DEP
Sbjct:   421 RKYPNQYESMISTLCENLETLDEPAAKASMVWIVGEYVDRIDNADELLETFLETFHDEPS 480

Query:   481 NVQLQILTASVKLFLKCSKNTKDIITKVLKLSTEESDNPDLRDRAYIYWRLLSKNIDVAK 540
              VQLQ+LTA+VKLFLK   +T+D++TKVLK++TEE+ NPDLRDRAYIYWR+L++N + AK
Sbjct:   481 IVQLQLLTATVKLFLKKPAHTQDLVTKVLKMATEETYNPDLRDRAYIYWRMLARNPEAAK 540

Query:   541 KIVLADKPPIQEENKITDTKVLNKLIKNISMLSSVYHKLPETFISKKNSYSLNSDNNNDH 600
             K+V A KPPI E+    D   L++LI NIS+LSSVYHK PETF+++    S
Sbjct:   541 KVVFAPKPPINEDADALDYNTLDRLIGNISLLSSVYHKAPETFVARAMPPSAALPK---- 596

Query:   601 MQDDHYDDDDYDKDNHVLKIKKQMDKQKYDSYSSDNKKSNHSKSSSDSYNNSSDEFNNDI 660
              +      D    D  V + K+ M KQ Y S   + + S  S   SDS ++   +
Sbjct:   597 -EVGSCSSDGESTDARVEQAKQSMQKQHYSSDEKEEESSPTSSEDSDSDSDGPTDLLGLS 655

Query:   661 DDADDSKKSMDLIGLND--DESKPQKTIPPVKMVQVLSSEDAGLKGQTGLSIFASINRID 718
             D+A   K+S      +D  D S P +    V    VL+++  G +G+TGL + A++ R
Sbjct:   656 DEATPRKRSSKESS-DDLFDLSSPPEDPRGVGKTLVLAADRPGNQGRTGLQVSAALTRAH 714

Query:   719 RKIQLKISVTNQTQNEIVVSGVQINKNSFGLSSPNNLDVQNIGFGETKEMLIYLIPNTLN 778
              +IQL +++ N++   +    +Q N+NSFGL+   NL V ++  G++ E  + ++P  L
Sbjct:   715 GRIQLHLTLANKSSMTLNGWAIQFNRNSFGLAPAANLQVADLLSGQSAETTVPVVPGQLM 774

Query:   779 SNTPPATPLFLQVAIRTNLDIFYFNVPYDIFVVFVENFHMEKDIFKKKWQIIEEAKESIL 838
             SN  P  PL LQVA++TNLDIF F VP+D+ VV  EN   +KD+F+++WQ I EA++S L
Sbjct:   775 SNAAPEQPLSLQVAVKTNLDIFCFTVPFDLSVVLQENSSADKDVFRQRWQNIGEARQSSL 834
```

```
Query:   839 MAVSPMVITSDMLIKRMKIFNISLIARRN----ELYYFACITTNNLVILSEVTIQPEKKN 894
             MA +P   +   + K+M+  NISL+A+R+    +  YF+  TTNNLV+L+EV++Q
Sbjct:   835 MASAPSSQSPQAVTKQMQAANISLVAQRSADTFDALYFSATTTNNLVVLAEVSLQRNGNA 894

Query:   895 VKLCIRTDSSSVIPLYKLLFVKAFSLS 921
             VKL  R+++++++PL+      A  L+
Sbjct:   895 VKLVTRSEAAALLPLFTSTVCAALRLT 921
```

Figure 68. BLAST data for P. falciparum AP-1 β1 subunit with T. gondii protein

The P. falciparum AP-1 β1 subunit shares homology with T. gondii putative β adaptin protein across the entire protein sequence of 925 amino acids. The N-terminal is coloured red and spans amino acids 1 to 618, while the C-terminal is shown in blue and spans amino acids 650 to 921. A dash indicates no amino acid match. A plus sign indicates sequence positivity, while a matching letter shows areas of sequence identity. Query indicates the P. falciparum protein sequence and Sbjct the amino acid sequence of the T. gondii protein.

```
Q10567| organism=Human | product= AP-1 complex subunit beta-1
length=949

Score = 1971 (698.9 bits), Expect = 2.0e-224, Sum P(2) = 2.0e-224
Identities = 428/925 (46%), Positives = 577/925 (62%)
Identities = 374/585 (63%), Positives = 478/585 (81%)

Query:     1 MTDSKYFTTTKKGEIFELKAELNSDKKEKKKEAVKKVIASMTVGKDVSALFPDVVNCMQT 60
             M+D +YF TTKKGEI ELK EL+S  KEKKKEA+KK+IA+MTVGKDVS LF DVVNCMQT
Sbjct:     1 MSDLRYFQTTKKGEIHELKEELHSSHKEKKKEAIKKIIAAMTVGKDVSTLFSDVVNCMQT 60

Query:    61 DNLELKKLVYLYLMNYAKSQPDMAIMAVNTFVKDCEDPNPLIRALAVRTMGCIRVDKITE 120
              N+ELKKLVYLY++NYAK QP++AI+AVNTF KD  DPNPLIRALA+RTMGCIR+++ITE
Sbjct:    61 SNIELKKLVYLYVINYAKVQPELAILAVNTFRKDSSDPNPLIRALAIRTMGCIRLEQITE 120

Query:   121 YLCEPLRKCLKDEDPYVRKTAAVCVAKLHDINAQLVEDQGFLDTLKDLISDSNPMVVANA 180
             YL EPLR+CLKDEDPYVRKTA +C+AKL+DI+ +LVE++GF+DTL D++ D+N MVVANA
Sbjct:   121 YLIEPLRRCLKDEDPYVRKTAVICIAKLYDISPKLVEEEGFIDTLLDILDDNNAMVVANA 180

Query:   181 VAALSEIAESHPSSNLLDL---NPQSINKLLTALNECTEWGQIFILDCLANYMPKDDREA 237
             V +L++I E+   S L D+   +  ++NKLL A+NEC EWGQ+FILD L  Y PK  ++A
Sbjct:   181 VISLTDICENSNKSILKDVINKDENNVNKLLNAINECVEWGQVFILDALVLYEPKTSKDA 240

Query:   238 QSICERVTPRLSHANSAVVLSAVKVLMKFMEMLSKDLDYYGTLLKKLAPPLVTLLSAEPE 297
             + + ER+ PRLSHANSAVVLS++KV++  ++ ++ D ++   + KKL+P LVTLLSAEPE
Sbjct:   241 ERVLERILPRLSHANSAVVLSSIKVILCLLDKIN-DKEFIKNVHKKLSPSLVTLLSAEPE 299

Query:   298 LQYVALRNINLIVQKRPEILKHEMKVFFVKYNDPIYVKLEKLDIMIRLASQANIAQVLAE 357
             +QY+ALRNINLI QK P +L  ++ +FF KYN+P YVK+EKLDI+IRL S  N+  VL E
Sbjct:   300 IQYIALRNINLITQKLPNMLSDKINMFFCKYNEPAYVKMEKLDIIIRLVSDKNVDLVLYE 359

Query:   358 LKEYATEVDVDFVRKAVRAIGRCAIKVEQSAERCVSTLLDLIQTKVNYVVQEAIVVIKDI 417
             LKEY+TEVDV+FV+K+VRAIG CAIK+ QS+E+C++ LLDLI TK+NYV+QE IVVIKDI
Sbjct:   360 LKEYSTEVDVEFVKKSVRAIGSCAIKLPQSSEKCINILLDLIDTKINYVIQECIVVIKDI 419

Query:   418 FRKYPNKYESVIATLCENLDSLDEPEARAAMIWIVGEYAERIDNADELLESFLEGFHDES 477
             FRKYPNKYES+I  LCENL+SLDE  A+A++IWI+GEY ERIDNADEL++SFLE F DE
Sbjct:   420 FRKYPNKYESIITILCENLESLDESNAKASLIWIIGEYVERIDNADELIDSFLENFSDEP 479

Query:   478 TQVQLQLLTAIVKLFLKKPTETQELVQQVLSLATQDSDNPDLRDRGYIYWRLLSTDPVAA 537
               VQLQ+LTA VKLFLK    T++++ +VL L+T++SDNPDLRDR YIYWRLLS +   A
Sbjct:   480 YNVQLQILTASVKLFLKCSKNTKDIITKVLKLSTEESDNPDLRDRAYIYWRLLSKNIDVA 539
```

```
Query:   538 KEVVLAEKPLISEETDLIEPTLLDELICYIGTLASVYHKPPSAFV 582
             K++VLA+KP I EE  + +  +L++LI  I  L+SVYHK P  F+
Sbjct:   540 KKIVLADKPPIQEENKITDTKVLNKLIKNISMLSSVYHKLPETFI 584

 Score = 201 (75.8 bits), Expect = 2.0e-224, Sum P(2) = 2.0e-224
 Identities = 54/190 (28%), Positives = 99/190 (52%)

Query:   731 GLEISGTFTRQVGSISMDLQLTNKALQVMTDFAIQFNRNSFGLAPAAPLQVHAPLSPNQT 790
             GL I  +  R    I + + +TN+    +    +Q N+NSFGL+    L V   +   +T
Sbjct:   707 GLSIFASINRIDRKIQLKISVTNQTQNEIVVSGVQINKNSFGLSSPNNLDVQN-IGFGET 765

Query:   791 VEISLPL--STVGS-VMKMEPLNNLQVAVKNNIDVFYFSTLYPLHILFVEDGKMDRQMFL 847
              E+ + L  +T+ S      PL  LQVA++ N+D+FYF+  Y + ++FVE+  M++ +F
Sbjct:   766 KEMLIYLIPNTLNSNTPPATPLF-LQVAIRTNLDIFYFNVPYDIFVVFVENFHMEKDIFK 824

Query:   848 ATWKDIPNENEAQFQ-IRDCPLNAEAASSKLQSSNIFTVAKRNVEGQDMLYQSLKLTNGI 906
               W+ I    E+    +    + ++    +++  NI  +A+RN    ++ Y +   TN +
Sbjct:   825 KKWQIIEEAKESILMAVSPMVITSDMLIKRMKIFNISLIARRN----ELYYFACITTNNL 880

Query:   907 WVLAELRIQP 916
              +L+E+ IQP
Sbjct:   881 VILSEVTIQP 890
```

Figure 69. BLAST data for P. falciparum AP-1 β1 subunit with H. sapiens protein

The P. falciparum AP-1 β1 subunit shares 46 percent identity and 62 percent positivity with H. sapiens AP-1 complex subunit β1. The sequence homology begins at P. falciparum amino acid 1 and ends at 584 which is part of the AP1 N-terminal (red). This area shares 63 percent identity and 81 percent positivity with the H. sapiens protein. Part of the AP1 C-terminal, from amino acid 707 to 890 (blue), shares only 28 percent identity and 52 percent positivity with the H. sapiens protein. A dash indicates no amino acid match. A plus sign indicates sequence positivity, while a matching letter shows areas of sequence identity. Sbjct indicates the P. falciparum protein sequence and Query the amino acid sequence of the H. sapiens protein.

5.10) Secondary structure data

```
1---------11--------21--------31--------41--------51--------61--------71--------81--------91--------101-------111
EKDGYEEMNGGDKNEEMNGGDKNEEMNVGDKNGGINEEHKNEGINEEHKDELINKEHKNERINEEHKNERINEEHKNEGINEEHKNEGINEEHKNERINEEHKNEGINKLTYH
----------------------------------------------------HHHHHHHHHHHHHHHHHHH-----------------------------------EE

-------121-------131-------141-------151-------161-------171-------181-------191-------201-------211-------221
NMNKNNISNENNYNDDDSYDEDNLVSLKIINLKYLSKKNSLKNILREVNFLKMCEHPNVVKYFESFFWPPCYLVIVCEYLSGGTLYDLYKNYGRISEDLLVYILDDVLNG
EEEEEE----EEEEEEEE-----EEEEEEE-------HHHHHHHHHHHHHHHH------EEEEEEE-----EEEEEEEE-----HHHHHH------HHHHHHHHHHHHHH

-------231-------241-------251-------261-------271-------281-------291-------301-------311-------321-------331
LNYLHNECSSPLIHRDIKPTNIVLSKDGIAKIIDFGSCEELKNSDQSKELVGTIYYISPEILMRTNYDCSSDIWSLGITIYEIVLCTLPWKRNQSFENYIKTIINSSPKI
HHHHHH----------------------EEEEE-----------------------------------HHHHHHHHHHHHHHHH---------HHHHHHHHH-------

-------341-------351-------361-------371-------381-------391-------401-------411-------421-------431-------441
NITEGYSKHLCYFVEKCLQKKPENRGNVKDLLNHKFLIKKRYIKKKPSSIYEIRDILKIYNGKGKTNIFRNFFKNLFFFNDKNKKKKPNKMISSKSCDAEMFFEQLKREN
-------HHHHHHHHHHHH--------HHHHHHH-------------------------------------------------------------------------

-------451-------461-------471-------481-------491-------501
FDFFEIKLKDDENSRSLNTFNINISKERDDISYSSLNLEKIKEHSLNMVASVVGTEQSQK
----EEE-----------------------------------------------------
```

Figure 70. Secondary structure prediction of PK

The P. falciparum PK sequence from amino acid 1 982-2 485 (the sequence expressed in E. coli) has a predicted secondary structure consisting mainly of alpha helices (H) and beta sheets (E). The areas with a '-' indicate that no secondary structure conformations were predicted by Jnet.

```
1---------11--------21--------31--------41--------51--------61--------71--------81--------91--------101-------111
NNSSDEFNNDIDDADDSKKSMDLIGLNDDESKPQKTIPPVKMVQVLSSEDAGLKGQTGLSIFASINRIDRKIQLKISVTNQTQNEIVVSGVQINKNSFGLSSPNNLDVQNIGF
------------HH----------------------------EEEEE------------EEEEEEEE----EEEEEEEEE--------EEEEEE---EEEEEE-----------

-------121-------131-------141-------151-------161-------171-------181-------191-------201-------211-------221
GETKEMLIYLIPNTLNSNTPPATPLFLQVAIRTNLDIFYFNVPYDIFVVFVENFHMEKDIFKKKWQIIEEAKEELYYFACITTNNLVILSEVTIQPEKKNVKLCIRTDSS
---EEEEEEE-----------------EEEE-----EEEEEE---EEEEE-------HHHHHHHHHH------EEEEEE-------EEEEEEEEE----EEEEEEE----

-------231-------241-
SVIPLYKLLFVKAFSLSVTQT
HHHHHHHHHHHHH--HE----
```

Figure 71. Secondary structure prediction of AP1 C-terminal

The P. falciparum AP1 C-terminal spanning amino acids 650-925 has a predicted secondary structure consisting mainly of beta sheets (E) with a few alpha helices (H), according to Jnet. The areas with a '-' indicate that no secondary structure conformations were predicted.

```
1---------11--------21--------31--------41--------51--------61--------71--------81--------91--------101-------111
MSDLRYFQTTKKGEIHELKEELHSSHKEKKKEAIKKIIAAMTVGKDVSTLFSDVVNCMQTSNIELKKLVYLYVINYAKVQPELAILAVNTFRKDSSDPNPLIRALAIRTMGCI
-----------HHHHHHHHHHH----HHHHHHHHHHHHHHHH------HHHHHHHHH------HHHHHHHHHHHH------HHHHHHHHHHHHH-----HHHHHHHHHHHH--

-------121-------131-------141-------151-------161-------171-------181-------191-------201-------211-------221
RLEQITEYLIEPLRRCLKDEDPYVRKTAVICIAKLYDISPKLVEEEGFIDTLLDILDDNNAMVVANAVISLTDICENSNKSILKDVINKDENNVNKLLNAINECVEWGQV
-----HHHHHHHHHHH----HHHHHHHHHHHHHHHH----------HHHHHHHHH------HHHHHHHHHHHHH--------------HHHHHHHHHH------HHHH

-------231-------241-------251-------261-------271-------281-------291-------301-------311-------321-------331
FILDALVLYEPKTSKDAERVLERILPRLSHANSAVVLSSIKVILCLLDKINDKEFIKNVHKKLSPSLVTLLSAEPEIQYIALRNINLITQKLPNMLSDKINMFFCKYNEP
HHHHHHHH------HHHHHHHHHHHHH-----HHHHHHHHHHHH------HHHHHHHHHHHHHHHHH-----HHHHHHHHHHHHHHH-HHHHH---EEEEE-----

-------341-------351-------361-------371-------381-------391-------401-------411-------421-------431-------441
AYVKMEKLDIIIRLVSDKNVDLVLYELKEYSTEVDVEFVKKSVRAIGSCAIKLPQSSEKCINILLDLIDTKINYVIQECIVVIKDIFRKYPNKYESIITILCENLESLDE
HHHHHHHHHHHHH-----HHHHHHHHHHHH----HHHHHHHHHHHHHHHH-----HHHHHHHHHHHH------HHHHHHHHHHHH---HHHHHHHHHHHHHH----

-------451-------461-------471-------481-------491-------501-------511-------521-------531-------541-------551
SNAKASLIWIIGEYVERIDNADELIDSFLENFSDEPYNVQLQILTASVKLFLKCSKNTKDIITKVLKLSTEESDNPDLRDRAYIYWRLLSKNIDVAKKIVLADKPPIQEE
-HHHHHHHHHHH--------HHHHHHHHHH-----HHHHHHHHHHHHHHHHH-----HHHHHHHHHHHH------HHHHHHHHHH---HHHHHHHH----------

-------561-------571-------581-------591-------601-------611-----
NKITDTKVLNKLIKNISMLSSVYHKLPETFISKKNSYSLNSDNNNDHAANKARKEAELAAATAEQ
-----HHHHHHHHHH-------------------------------------------------
```

Figure 72. Secondary structure prediction of AP1 N-terminal

The P. falciparum AP1 N-terminal spanning amino acids1-618 has a predicted secondary structure consisting mainly of alpha helices (H) with one section of beta sheets (E), according to Jnet. The areas with a '-' indicate that no secondary structure conformations were predicted.

5.11) Target-template pair alignments

```
TARGET    2117          DN LVSLKIINLK YLSKKNSLKN ILREVNFLKM CEHPNVVKYF
1apmE     60    vkhkes--gn hyamkildkq kvvklkqieh tlnekrilqa vnfpflvkle

TARGET    2159  ESFFWPPCYL VIVCEYLSGG TLYDLYKNYG RISEDLLVYI LDDVLNGLNY
1apmE     108   fsfkdn-snl ymvmeyvagg emfshlrrig rfaepharfy aaqivltfey

TARGET    2209  LHNECSSPLI HRDIKPTNIV LSKDGIAKII DFGSCEELKN SDQSKELVGT
1apmE     157   lhs---ldli yrdlkpenll idqqgyiqvt dfgfakrvkg --rtwtlcgt

TARGET    2259  IYYISPEILM RTNYDCSSDI WSLGITIYEI VLCTLPWKRN QSFENYIKTI
1apmE     202   peylapeiil skgynkavdw walgvliyem aagyppffad qpiq-iyeki

TARGET    2309  INSSPKINIT EGYSKHLCYF VEKCLQKKPE NRGN-----V KDLLNHKFLI
1apmE     251   vsgkv--rfp shfssdlkdl lrnllqvdlt krfgnlkngv ndiknhkwfa
```

Figure 73. Target-template pair alignment for PK kinase domain 3D model

The target template pair shared homology in the kinase domain of the PK, from amino acids 2 117-2 353. Target – PK protein sequence; 1apmE - catalytic subunit of human cAMP-dependent protein kinase protein sequence.

```
TARGET    694             VLSSEDAGLK GQTGLSIFAS INRIDRKIQL KISVTNQTQN
1e42A     705   ggyvapka-- vwlpa---vk ak-gleisgt fthrqghiym emnftnkalq

TARGET    734   EIVVSGVQIN KNSFGLSSPN NLDVQNI-GF GETKEMLIYL IPNTLNSNTP
1e42A     749   hmtdfaiqfn knsfgvipst plaihtplmp nqsidvslpl ntlgpvmkme

TARGET    783   PATPLFLQVA IRTNLDIFYF NVPYDIFVVF VENFHMEKDI FKKKWQIIEE
1e42A     799   p--lnnlqva vknnidvfyf scliplnvlf vedgkmerqv flatwkdipn

TARGET    833   AKESILMAVS PMVITSDMLI KRMKIFNISL IARRN----E LYYFACITTN
1e42A     847   ene-lqfqik echlnadtvs sklqnnnvyt iakrnvegqd mlyqslkltn

TARGET    879   NLVILSEVTI QPEKKNVKLC IRTDSSSVIP LYK ------ --
1e42A     896   giwilaelri qpgnpnytls lkcrapevsq yiyqvydsil kn
```

Figure 74. Target-template pair alignment for AP1 C-terminal 3D model

The target template pair shared homology from amino acids 694-911. Target – PK protein sequence; 1e42A – human β2 appendage domain from clathrin adaptor AP-2.

```
TARGET    12      KGEIHELK EELHSSHKEK KKEAIKKIIA AMTVGKDVST LFSDVVNCMQ
2vglB     12      kgeifelk aelnnekkek rkeavkkvia amtvgkdvss lfpdvvncmq

TARGET    60    TSNIELKKLV YLYVINYAKV QPELAILAVN TFRKDSSDPN PLIRALAIRT
2vglB     60    tdnlelkklv ylylmnyaks qpdmaimavn sfvkdcedpn pliralavrt

TARGET    110   MGCIRLEQIT EYLIEPLRRC LKDEDPYVRK TAVICIAKLY DISPKLVEEE
2vglB     110   mgcirvdkit eylceplrkc lkdedpyvrk taavcvaklh dinaqmvedq

TARGET    160   GFIDTLLDIL DDNNAMVVAN AVISLTDICE NSNKSILKDV INKDENNVNK
2vglB     160   gfldslrdli adsnpmvvan avaalseise shpnsnlldl ---npqnink
```

```
TARGET    210   LLNAINECVE WGQVFILDAL VLYEPKTSKD AERVLERILP RLSHANSAVV
2vglB     207   lltalnecte wgqifildcl snynpkddre aqsicervtp rlshansavv

TARGET    260   LSSIKVILCL LDKI-NDKEF IKNVHKKLSP SLVTLLSAEP EIQYIALRNI
2vglB     257   lsavkvlmkf lellpkdsdy ynmllkklap plvtllsgep evqyvalrni

TARGET    309   NLITQKLPNM LSDKINMFFC KYNEPAYVKM EKLDIIIRLV SDKNVDLVLY
2vglB     307   nlivqkrpei lkqeikvffv kyndpiyvkl ekldimirla sqaniaqvla

TARGET    359   ELKEYSTEVD VEFVKKSVRA IGSCAIKLPQ SSEKCINILL DLIDTKINYV
2vglB     357   elkeyatevd vdfvrkavra igrcaikveq saercvstll dliqtkvnyv

TARGET    409   IQECIVVIKD IFRKYPNKYE SIITILCENL ESLDESNAKA SLIWIIGEYV
2vglB     407   vqeaivvird ifrkypnkye siiatlcenl dsldepdara amiwivgeya

TARGET    459   ERIDNADELI DSFLENFSDE PYNVQLQILT ASVKLFLKCS KNTKDIITKV
2vglB     457   eridnadell esflegfhde stqvqltllt aivklflkkp setqelvqqv

TARGET    509   LKLSTEESDN PDLRDRAYIY WRLLSKNIDV AKKIVLADKP PIQEENKITD
2vglB     507   lslatqdsdn pdlrdrgyiy wrllstdpvt akevvlsekp liseetdlie

TARGET    559   TKVLNKLIKN ISMLSSVYHK LPETFI
2vglB     557   ptlldelich igslasvyhk ppnafv-
```

Figure 75. Target-template pair alignment for AP1 N-terminal 3D model

The target template pair shared homology from amino acids 12-584. Target – PK protein sequence; 2vglB – human AP-2 clathrin adaptor core

```
MKTTKENDNNNIVHYVDWINQIFKKNSLQCDLYFLDDNKEKDVSKKRKAQLKDEYDNISR
SKENINNSKKIKNELSIKDNMHDYIYDDRIYNNDKEKNVIKSDNKNVIKSDNKNVIKSDN
KNVIKSDNKNVIKSDNKNVIKSDNKNVIKSDYKSDDRNACDIYKSNKKNVPDNCHIYDDN
SSVENLDGKNKLNNIRNIHNDNSSSCDISDIKSEDEYIEPYEKKNEENINEYKNKKNIAN
ENIKEGKSSIYNDEHNYNSLLYNSCNGEISKINKISSHNNIDNNMDNYNTFANVNNFIIY
SSDDEDNISNYYNGKDVLNDEIMFPIKFNFEKLKKNIYVIEHIDKIYYDTFLNKNPSEKS
VFMNDESTGYLKNDVNDKCVVDNINVINPSSVNTLSNISNIRNEKIENNNKNEKLIKSYP
TQSKNVMSTFSFWNIEKETFITKPLYAQNLRKKQFSLLDESEEMIRNYSSNQYSIKFVPR
HLLYVMSQVASRSFFDPLYRKQLFFLELENRSGYKKRKKKKWNNKSTGQDKFTNKDIISE
EKEEGLDIECGHNILGDVQYDGTYNINEQVKKNSLFYFKCKEEINLKDGNIILDDKNRKV
DDINITGDDKNIKVDDKNIKVDDKNITGEDKNITGEDKNITGDDKNIIFDVDEILIHQHN
TSNSNIYINCNDNNNDIRNSSNVQHYYNDKIKENINKQNKKYVLINDYINNKYILSKNKT
CKINKGKKLIKKKKVNNISRRRNHILYKCRNKLYNGNVFSDDIIKSEVNVCNSLTVLHKN
YNINMDNYLDDNIHTNNSNIYDINYTNENVINSTCRYYPIGNNNTLSKDEVTKSSSKINS
LSYFDDIINVNKNDIPILHDKENINIISNKESCHKDEKEEEKYIMYNSNLVEEKKQKKMI
WNSLNVLPIDILLKNGHDEINKEICKKKKKSFFSQNDIKSKMLYNNKSYSKSEKVLYTNN
KNSNTFIPIFFLNKVGDKFKNSENIYDMYNNKKNVYIHDKKIYTNMYSNKLKQKHYYSTS
NINLLYNNIGKVLDNGLHLSNNMYCRLNSNPPYKSISLINNNVFFYKKRKSNSNNNNNNN
NISSSSSSSSKKNHVIINKKISSYNIHYKERKDSFKENFLFFKEKILPSKKDTCVFNERQ
KDLFEKSNEHIKCVSSFNNTSDDISSHSSVNKKEPFFALKNNSIRHIPKENNIIYTSGKS
FNHVQDKEKTVLLKKKKEINDKNTFSSCLINHNITTYTLQNGVNKNLNMLGIRDSIYKID
EKNNMLKECYNGNNDSNNKKKKKKKKLSFSCDIINDNITPYESDKEKNNSNNIKSMDIFN
YVKRKSNLYNNLSSNRDSTVDMHNKYNSEEYINIQRTNKIYELSNKRIRNYKLYSMDEIF
KVSLKEKKYIDNISNNMERVTYKNEMINEKISKMDDILYPCDKNKSLNMSCPVIIENNIS
REENEKNSSVILNKKKNENMFNCVGRLHCHMGKMNNQDNIYDQGNIKKNEEEITKHDEYI
SREEKNKYNSKCIRNFDDYKYEQVLSYHTLDEDKKKNDMNNLIDMNNEAIIETVNGVINN
IILDRKDNNSRKDMEKEMEKEMEKKMEKEMEKVMEKEMEKVMEKEVEKELKNEMNNRMNN
RMNNEMKNEINIYKNNEIYVDNDKELEIVNEEKKLIYPFNYESDVHKNMNMSININNCKD
DYNNILKEYVDNSCLAQKEENIFRPLFNLNKKDKVWKRFNIKNNIKTIIHNEEMKRIYQT
```

INKNVFPIYNFNRYENFLINHLTYNFPKNDLFKLSYKVSMNNIRNLYIANKHINNNYDYM
NKLYNQNIYTLKYQVANIDNDHHICKKGGGLDYINMNISKECKNRKDKTYLNKIFHYKKK
KDARFFINDEIGSNDYMYDIKKKYSNDENNYKLNEKMNISMSNDEDMIPTLNSEHGNNFP
SCQPNLLEKKSTYIDLNLYDSNSMDDFTEEKYNFVNNENDLFNTKRWKFNFSKGKNLFNN
KFFNVSNEDGVFSFFKNMNLFRELNKSNNSLKLESVKNSNNNCSNNKGDDNIGNMENMNT
TNVTIASDEHISTKGDIHDESFSRDDNDCILLKIEGRSKKYSDITLYNEDKSNLENDNET
INEYENVCSNIDVNEWEDKVNGTCNSVGDKETEKNNEKNNEKNNEKNNEKNNEKNNEKNN
EKNNEKNNEENNEGNNEENNEENNEENNEENNDIEKNDIKDNNSGQVKENIIVMNNTNNM
DVDNDDNNNNYNNVSTDEGIDIIKNIKSEMNDYIYNDNIMIKINNKSIDLMNIKNQKNEP
FLNYTNEKDIHMKSNSSYNVNDKMNLFNNNEKTEKNNTSLNDLLYKRKEELDDEKISEYK
DTNLTNNTFEHIAKRINLILNDTIEFFQKHTYLHNGYGNVQVCKKNKRKLEKKKLKKWSC
IYKINKIVRKGAHGVVFSAWRSENVDFFNHSFFENLNLENKKKGYIDETNVNENYESDNE
YDSDEDDTESDNDDEQNKENERGDEKDGYEEMNGGDKNEEMNGGDKNEEMNVGDKNGGIN
EEHKNEGINEEHKDELINKEHKNERINEEHKNERINEEHKNEGINEEHKNEGINEEHKNE
RINEEHKNEGINKLTYHNMNKNNISNENNYNDDDSYDEDNLVSLKIINLKYLSKKNSLKN
ILREVNFLKMCEHPNVVKYFESFFWPPCYLVIVCEYLSGGTLYDLYKNYGRISEDLLVYI
LDDVLNGLNYLHNECSSPLIHRDIKPTNIVLSKDGIAKIIDFGSCEELKNSDQSKELVGT
IYYISPEILMRTNYDCSSDIWSLGITIYEIVLCTLPWKRNQSFENYIKTIINSSPKINIT
EGYSKHLCYFVEKCLQKKPENRGNVKDLLNHKFLIKKRYIKKKPSSIYEIRDILKIYNGK
GKTNIFRNFFKNLFFFNDKNKKKKPNKMISSKSCDAEMFFEQLKRENFDFFEIKLKDDEN
SRSLNTFNINISKERDDISYSSLNLEKIKEHSLNMVASVVGTEQSQK

Figure 76. Protein sequence of PFB0150c

The PEXEL/ VTS motif consensus sequence (R/KxLxE/Q) is present in the PFB0150c protein sequence as shown in red. This indicates that the parasite protein is exported into the host erythrocyte. The bold underlined section indicates the catalytic kinase domain.

REFERENCES

1) Afonso A, Hunt P, Cheesman S, Alves AC, Cunha CV, do Rosario V, Cravo P (2006) **"Malaria parasites can develop stable resistance to artemisinin but lack mutations in candidate genes *atp6* (encoding the sarcoplasmic and endoplasmic reticulum Ca2_ATPase), *tctp*, *mdr1*, and *cg10"*** *Antimicrobial Agents and Chemotherapy;* Vol. 50; 480-489

2) Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson JD (1994) Chapter 8: The cell nucleus: RNA synthesis and RNA processing, In "**The Molecular Biology of the Cell**" *Garland;* 3rd edition; 300-384

3) Anamika N, Srinivasan N, Krupa A (2005) "**A genomic perspective of protein kinases in *Plasmodium falciparum***" *PROTEINS: Structure, Function and Bioinformatics;* Vol. 58; 180-189

4) Angov E, Hillier CJ, Kincaid RL, Lyon JA (2008) "**Heterologous protein expression is enhanced by harmonising the codon usage frequencies of the target gene with those of the expression host**" *PLoS ONE;* Vol. 3; 2 189-2 199

5) Arnold K, Bordoli L, Kopp J, Schwede T (2006) "**The SWISS-MODEL Workspace: A web-based environment for protein structure homology modeling**" *Bioinformatics*; Vol. 22; 195-201

6) Baca AM and Hol WGJ (2000) "**Overcoming codon bias: A method for high-level overexpression of *Plasmodium* and other AT-rich parasite genes in *Escherichia coli***" *International Journal for Parasitology;* Vol. 30; 113-118

7) Bain J, Mclauchlan H, Elliott M, Cohen P (2003) "**The specificities of protein kinase inhibitors: an update**" *Journal of Biochemistry*; Vol. 371, 199-204

8) Bain J, Plater L, Elliott M, Shpiro NC, Hastie J, Mclauchlan H, Klevernic IJ, Arthur SC, Alessi DR, Cohen P (2007) "**The selectivity of protein kinase inhibitors: a further update**" *Journal of Biochemistry*; Vol. 408, 297–315

9) Baneyx F (1999) "**Recombinant protein expression in Escherichia coli**" *Current Opinion in Biotechnology*; Vol. 10; 411-421

10) Bannister LH, Hopkins JM, Fowler RE, Krishna S, Mitchell GH (2000) "**A brief illustrated guide to the ultrastructure of *Plasmodium falciparum* asexual blood stages"** *Parasitology Today;* Vol. 16, 427-433

11) Behari R and Haldar K (1994) "***Plasmodium falciparum:* Protein localisation along a novel, lipid-rich tubovesicular membrane network in infected erythrocytes**" *Experimental Parasitology*; Vol. 79; 250-259

12) Bimboim HC & Doly J (1979) "**A rapid alkaline extraction procedure for screening recombinant plasmid DNA**" *Nucleic Acids Research*; Vol. 7; 1513-1523

13) Birkholtz L-M, Blatch G, Coetzer TL, Hoppe HC, Human E, Morris EJ, Ngete Z, Oldfield L, Roth R, Shonhai A, Stephens L, Louw A (2008) **"Heterologous expression of plasmodial proteins for structural studies and functional annotation"** *Malaria Journal;* Vol. 197; http://www.malariajournal.com/content /pdf/1475-2875-7-197.pdf

14) Brodsky FM, Chen CY, Knuehl C, Towler MC, Wakeham DE (2001) "**Biological basket weaving: Formation and function of clathrin-coated vesicles"** *Annual Review Cell Developmental Biology;* Vol. 17; 517-568

15) Chishti AH, Maalouf GJ, Marfatia S, Palek J, Wang W, Fisher D, Liu SC (1994) "**Phosphorylation of protein 4.1 in *Plasmodium falciparum*-infected human red blood cells"** *Blood;* Vol. 83; 3339-3345

16) Cobb MH (1996) **"MAP Kinase signalling pathways"** *Promega Notes Magazine***;** Vol. 59; 37

17) Cooke BM, Lingelbach K, Bannister LH, Tilley L (2004) "**Protein trafficking in *Plasmodium falciparum*-infected red blood cells**" *Trends in Parasitology;* Vol. 20; 581-589

18) Cuff JA and Barton GJ (2000) "**Application of enhanced multiple sequence alignment profiles to improve protein secondary structure prediction**" *PROTEINS: Structure, Function and Genetics;* Vol. 40; 502-511

19) Daily JP, Le Roch KG, Sarr O, Fang X, Zhou Y, Ndir O, Mboup S, Sultan A, Winzeler EA, Wirth DF (2004) "***In vivo* transcriptional profiling of *Plasmodium falciparum***" *Malaria Journal;* Vol. 30; 223-231

20) Davies SP, Reddy H, Caivano M, Cohen P (2000) "Specificity and mechanism of **action of some commonly used protein kinase inhibitors**" *Journal of Biochemistry;* Vol. 351, 95-105

21) Diaz CA, Allocco J, Powles MA, Yeung L, Donald RGK, Anderson JW, Liberator PA (2006) "**Characterisation of *Plasmodium falciparum* cGMP-dependent protein kinase (*Pf*PKG): Antiparasitic activity of a PKG inhibitor**" *Molecular & Biochemical Parasitology;* Vol. 146; 78-88

22) Doerig CD, Doerig C, Horrocks P, Coyle J, Carlton J, Sultan A, Arnot D, Carter R (1995) "**Pfcrk-1, a developmentally regulated cdc2-related protein kinase of *Plasmodium falciparum***" *Molecular and Biochemical Parasitology*; Vol. 70; 167-174

23) Doerig C and Meijer L (*2007*) "**Antimalarial drug discovery: targeting protein kinases**" *Expert Opinion Therapeutic Targets;* Vol. 11; 279-290

24) Druker BJ (2002) "**Inhibition of the Bcr-Abl tyrosine kinase as a therapeutic strategy for CML**" *Oncogene;* Vol. 21; 8541–8546

25) Duden R (2003) "**ER-to-Golgi transport: COP I and COP II function**" *Molecular Membrane Biology;* Vol. 20; 197-207

26) Dudley DT, Pang L, Decker SJ, Bridges AJ, Saltiel AR (1995) "**A synthetic inhibitor of the mitogen-activated protein kinase cascade**" *Proceedings of the National Academy of Science*; Vol. 92; 7686-7689

27) Dyer MD, Murali TM, Sobral BW (2007) "**Computational prediction of host-pathogen protein interactions**" *Bioinformatics;* Vol. 23; 159-166

28) Eckstein-Ludwig U, Webb RJ, Van Goethem ID, East JM, Lee AG, Kimura M, O'Neill PM, Bray PG, Ward SA, Krishna S (2003) "**Artemisinins target the SERCA of *Plasmodium falciparum*"** *Nature*; Vol. 424; 957-961

29) Ehrt S & Schnappinger D (2003) "**Isolation of plasmids from *E. coli* by alkaline lysis**" *Methods in Molecular Biology;* Vol. 235; 75-78

30) Favata MF, Horiuchi KY, Manos EJ, Daulerio AJ, Stradley DA, Feeser WS, Van Dyk DE, Pitts WJ, Earl RA, Hobbs F (1998) "**Identification ofa novel inhibitor of mitogen-activated protein kinase kinase**" *Journal of Biological Chemistry*; Vol. 273; 18623-18632

31) Fields S & Song O (1989). "**A novel genetic system to detect protein-protein interactions**" *Nature;* **Vol. 340;** 245–246

32) Flick K, Chen Q (2004) "***var* genes, PfEMP1 and the human host**" *Molecular and Biochemical Parasitology;* Vol. 134; 3-9

33) Florens L, Washburn MP, Raine JD, Anthony RM, Grainger M, Haynes JD, Moch JK, Muster N, Sacci JB, Tabb DL, Witney AA, Wolters D, Wu Y, Gardner MJ, Holder AA, Sinden RE, Yates JR, Carucci DJ (2002) "**A proteomic view of the *Plasmodium falciparum* life cycle"** *Nature;* Vol. 419; 520-526

34) Foley M and Tilley L (1998) "**Protein trafficking in malaria-infected erythrocytes**" *International Journal for Parasitology;* Vol. 28; 1 671-1 680

35) Gallagher PG (2006) Chapter 44: Disorders of the red blood cell membrane: Hereditary spherocytosis, elliptocytosis and related disorders, In: Lichtman MA, Beutler E, Kipps TJ, Seligsohn U, Kaushansky K, Prchal JT, editors; "**Williams Haematology**" *McGraw-Hill Medical;* Seventh edition; 571-579

36) Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson JE, Bowman S, Paulsen IT, James K, Eisen JA, Rutherford K, Salzberg SL, Craig A, Kyes S, Chan M-S, Nene V, Shallom SJ, Suh B, Peterson J, Angiuoli S, Pertea M, Allen J, Sellengurt J, Haft D, Mather MW, Vaidya AB, Martin DMA, Fairlamb AH, Fraunholz MJ, Roos DS, Ralph SA, MacFadden GI, Cummings LM, Mani Subramanian G, Mungall C, Venter JC, Carucci DJ, Hoffman SL, Nubold C, Davis RW, Fraser CM, Barrell B (2002) "**Genome sequence of the human malaria parasite *Plasmodium falciparum***" *Nature;* Vol. 419; 498-511

37) Glushakova S, Yin D, Li T, Zimmerberg J (2005) "**Membrane transformation during malaria parasite release from human red blood cells**" *Current Biology;* Vol. 15, 1645–1650

38) Gray NS, Wodicka L, Thunnissen AM (1998) "**Exploiting chemical libraries, structure and genomics in the search for kinase inhibitors**" *Science;* Vol. 281; 533-538

39) Gross SD & Anderson RA (1998) "**Casein kinase 1: Spatial organisation and positioning of a multifunctional protein kinase family**" *Cell Signal;* Vol. 10; 699-711

40) Guex N and Peitsch M. C. (1997) "**SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling**" *Electrophoresis;* Vol. 18: 2714-2723

41) Haeggström M and Schlichtherle M (2004) Thawing of glycerolyte-frozen parasites with NaCl, In: Ljungström I, Perlmann H, Scherf A, Schlichtherle M, Wahlgren M, editors; 4th edition "**Methods in Malaria Research**" *Malaria Research and Reference Reagent Resource Center;* 1-265

42) Hanks SK, Quinn AM, Hunter T. (1998) "**The protein kinase family: conserved features and deduced phylogeny of the catalytic domains**." *Science*; Vol. 241; 42–52

43) Hanks SK & Hunter T (1995) "**Protein kinase 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification**" *The Journal of the Federation of American Societies for Experimental Biology;* Vol. 9; 576-596

44) Hanks SK (2003) "**Genomic analysis of the eukaryotic protein kinase superfamily: a perspective**" Genome Biology; Vol. 14; 1-7

45) Hannig G and Makrides SC (1998) "**Strategies for optimising heterologous protein expression in *Escherichia coli***" *Trends in Biotechnology;* Vol. 16; 54-60

46) Harmse L (2006) "**Characterisation of two *Plasmodium falciparum* cell cycle related kinases and the effect of kinase inhibitors on the parasite**" *PhD thesis; University of the Witwatersrand, Johannesburg;* 1-220

47) Harmse L, van Zyl R, Gray N (2001) "**Structure-activity relationships and inhibitory effects of various purine derivatives on the *in vitro* growth of *Plasmodium falciparum***" *Biochemical Pharmacology;* Vol. 62; 341-348

48) Haynes, DJ (1993) "**Erythrocytes and Malaria**" *Current Opinion in Haematology;* Vol. 8; 79-89

49) Healer J, Crawford S, Ralph S, McFadden G, Cowman A.F (2002) **"Independent Translocation of Two Micronemal Proteins in Developing *Plasmodium falciparum* Merozoites"** *Infection and Immunity;* Vol. 70; 5 751-5 758

50) Heldwein EE, Macia E, Wang J, Yin HL, Kirchhausen, Harrison SC (2004) "**Crystal structure of the clathrin adaptor protein 1 core**" *Proceedings of the National Academy of Sciences;* Vol.101; 14 108-14 113

51) Hiller NL, Bhattacharjee S, van Ooij C, Liolios K, Harrison T, Lopez-Estraño C, Haldar K (2004) "**A host-targeting signal in virulence proteins reveals a secretome in malarial infection**" *Science;* Vol. 306; 1934–1937

52) Hinners I & Tooze SA (2003) "**Changing directions: clathrin-mediated transport between the Golgi and endosomes**" *Journal of Cell Science;* Vol. 116; 763-771

53) Hoffman SL, Subramanian GM, Collins FH, Venter JC (2002) "***Plasmodium,* human and *Anopheles* genomics and malaria**" *Nature;* Vol. 415; 702-709

54) Hoffmann-Rohrer UW & Kruchen B, editors (2006) Chapter 1: Working with DNA, In: **Roche Applied Science Lab FAQS,** 2nd edition; Roche Diagnostics

55) Hoffmann-Rohrer UW & Kruchen B, editors (2006) Chapter 2: Working with RNA, In: **Roche Applied Science Lab FAQS,** 2nd edition; Roche Diagnostics

56) Hubbard MJ and Cohen P. (1993) "**On target with a new mechanism for the regulation of protein phosphorylation**" *Trends in Biochemical Science;* Vol. 18; 172-177

57) Jambou R, Legrand E, Niang M, Khim N, Lim P, Volney B, Ekala MT, Bouchier C, Esterre P, Fandeur T, Mercereau-Puijalon O (2005) "**Resistance of *Plasmodium falciparum* field isolates to in-vitro artemether and point mutations of the SERCA-type PfATPase6"** *Lancet;* Vol. 366; 1960-1963

58) Jones GL & Edmundson HM (1990) "**Protein phosphorylation during the asexual life cycle of the human malarial parasite *Plasmodium falciparum***" *Biochimica et Biophysica Acta*; Vol. 1053; 118-124

59) Kappes B, Doerig CD, Graeser R (1999) **"An overview of *Plasmodium* protein kinases"** *Parasitology Today; Vol. 15; 449-453*

60) Kappes B, Yang J, Suetterlin BW, Rathgeb-Szabo K, Lindt MJ, Franklin RM (1995) "***A Plasmodium falciparum* protein kinase with two unusually large kinase inserts**" *Molecular and Biochemical Parasitology*; Vol. 72; 163-1772

61) Kiefhaber T, Rudolph R, Kohler HH, Buchner J (1991) "**Protein aggregation *in vitro* and *in vivo*: A quantitative model of the kinetic competition between folding and aggregation**" *Biotechnology;* Vol. 9; 825-829

62) Knapp B, Nau U, Hundt E, Kiipper HA (1991) "**Demonstration of alternative splicing of a pre-mRNA expressed in the blood stage form of *Plasmodium falciparum"*** *The Journal of Biological Chemistry;* Vol. 266; 7 148-7 154

63) Knockaert M, Gray N, Damiens E (2000) "**Intracellular targets of cyclin-dependant kinase inhibitors: identification by affinity chromatography using immobilised inhibitors**" *Chemical Biology;* Vol. 7; 411-422

64) Kopp J and Schwede T (2004) "**The SWISS-MODEL Repository of annotated three-dimensional protein structure homology models**" *Nucleic Acids Research*; Vol. 32; D230-D234.

65) Kriventseva EV, Koch I, Apweiler R, Vingron M, Bork P, Gelfand MS, Sunyaev S (2003) "**Increase of functional diversity by alternative splicing**" *Trends in Genetics;* Vol. 19; 125-128

66) Laemmli UK (1970) "**Cleavage of structural proteins during the assembly of the head of bacteriophage T4**" *Nature;* Vol.227; 680-685

67) Lambros C and Vanderberg JP (1979) "**Synchronisation of *Plasmodium falciparum* erythrocytic stages in culture**" *The Journal of Parasitology;* Vol. 65; 418-420

68) Lauterbach SB, Lanzillotti R, Coetzer TL (2003) "**Construction and use of *Plasmodium falciparum* phage display libraries to identify host parasite interactions"** *Malaria Journal*; www.malariajournal.com/content/2/1/47

69) Le Roch KG, Zhou Y, Blair PL, Grainger M, Moch JK, Haynes JD, De la Vega P, Holder AA., Batalov S, Carucci DJ, Winzeler EA (2003) "**Discovery of gene function by expression profiling of the malaria parasite life cycle"** *Science;* Vol. 301; 1503-1508

70) Letunic I, Copley RR, Pils B, Pinkert S, Schultz J, Bork P (2006) "**SMART 5: domains in the context of genomes and networks"** *Nucleic Acids Research*; Vol. 34, 257-260

71) Ling Goh L, Loke P, Singh M, Suan Sim T (2003) "**Soluble expression of a functionally active *Plasmodium falciparum* falcipain-2 fused to maltose-binding protein in *Escherichia coli***" *Protein Expression & Purification;* Vol. 32; 194-201

72) Luna EJ & Hitt AL (1992) "**Cytoskeleton-plasma membrane interactions**" S*cience;* Vol. 258; 955-964

73) Maier AG, Duraisingh MT, Reeder JC, Patel SS, Kazura JW, Zimmerman PA, Cowman AF (2003) "***Plasmodium falciparum* erythrocyte invasion through glycophorin C and selection for Gerbich negativity in human populations**" *Nature Medicine;* Vol. 9; 87-92

74) Maier AG, Rug M, O'Neill MT, Beeson JG, Marti M, Reeder J, Cowman AF (2007) "**Skeleton-binding protein 1 functions at the parasitophorous vacuole membrane to traffic PfEMP1 to the *Plasmodium falciparum*–infected erythrocyte surface**" *Blood;* Vol. 109; 1 289 - 1 297

75) Marti M, Good RT, Rug M, Knuepfer E, Cowman AF (2004) "**Targeting malaria virulence and remodeling proteins to the host erythrocyte**" *Science;* Vol. 306; 1930–1934

76) Mayer DCG, Mu J, Feng X, Su X, Miller LH (2002) "**Polymorphism in a *Plasmodium falciparum* erythrocyte-binding ligand changes its receptor specificity**" *The Journal of Experimental Medicine;* Vol. 196; 1523-1528

77) Mehlin C, Boni E, Buckner FS, Engel L, Feist T, Gelb MH, Haji L, Kim D, Liu C, Mueller N, Myler PJ, Reddy JT, Sampson JN, Subramanian E, Van Voorhuis WC, Worthey E, Zucker F, Hol WGJ (2006) "**Heterologous expression of proteins from *Plasmodium falciparum*: Results from 1000 genes"** *Molecular & Biochemical Parasitology;* Vol. 148; 144-160

78) Moore D & Dowhan D (2002) Preparation and analysis of DNA, In: Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K, editors; "**Current Protocols in Molecular Biology**" *John Wiley & Sons, Inc;* Vol. 1; 2.1.1-2.1.10

79) Ngô HM, Yang M, Paprotka K, Pypaert M, Hoppe H, Joiner KA (2003) "**AP-1 in *Toxoplasma gondii* mediates biogenesis of the rhoptry secretory organelle from a post-Golgi compartment"** *The Journal of Biological Chemistry;* Vol. 278; 5 343-5 352

80) Nunes MC, Goldring D, Doerig C, Scherf1 A (2007) "**A novel protein kinase family in *Plasmodium falciparum* is differentially transcribed and secreted to various cellular compartments of the host cell**" *Molecular Microbiology;* Vol. 63; 391-403

81) Ohno H (2006) "**Clathrin-associated adaptor protein complexes**" *Journal of Cell Science;* Vol. 119; 3 719-3 721

82) Owen DJ, Vallis Y, Pearse BMF, McMahon HT, Evans PR (2000) "**The structure and function of the β2-adaptin appendage domain**" *The EMBO Journal;* Vol. 19; 4 216-4 227

83) Phelps MA, Foraker AB, Swaan PW (2003) "**Cytoskeletal motors and cargo in membrane trafficking: opportunities for high specificity in drug intervention**" *Drug Discovery Today;* Vol. 8; 495-502

84) Possee RD (1997) "**Baculoviruses as expression vectors**" *Current Opinion in Biotechnology;* Vol. 8; 569-572

85) Price NC & Stevens L (2002) Chapter 9: Techniques for enzyme extraction, In: Eisenthal R & Danson MJ, editors; 2nd edition "**Enzyme assays: A practical approach**" *Oxford University Press;* 209-224

86) Ramya TNC, Karmodiya K, Surolia A, Surolia N (2007) "**15-Deoxyspergualin primarily targets the trafficking of apicoplast proteins in *P. falciparum***" *The Journal of Biological Chemistry;* Vol. 282; 6 388-697

87) Rangachari K, Dluzewski A, Wilson RJM, Gratzer WB (1986) "**Control of malarial invasion by phosphorylation of the host cell membrane cytoskeleton**" *Nature;* Vol. 324; 364-365

88) Roberts LS, Janovy J (1996) Chapter 9: Phylum *Apicomplexa* – Malaria Organisms and Piroplasms, In: 5th Edition "**Foundations of Parasitology**" *WCB Publishers;* 137-162

89) Rodnina MV, Beringer M, Wintermeyer W (2006) "**How ribosomes make peptide bonds**" *Trends in Biochemical Sciences;* Vol. 32; 20-26

90) Sasse J & Gallagher SR (2003) Detection of proteins on blot transfer membranes, In: Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K, editors; "**Current Protocols in Molecular Biology**" *John Wiley & Sons, Inc;* Vol. 2; 10.7.1-10.7.6

91) Schaefer E, Hsaio B, Guimond S (1996) "**Detection and quantitation of protein tyrosine kinases** " *Promega Notes Magazine;* Vol. 59, 110-118

92) Schlichtherle M and Wahlgren M (2004) Small-scale genomic DNA isolation from *Plasmodium falciparum*, In: Ljungström I, Perlmann H, Scherf A, Schlichtherle M, Wahlgren M, editors; 4th edition "**Methods in Malaria Research**" *Malaria Research and Reference Reagent Resource Center;* 1-265

93) Schneider AG & Mercereau-Puijalon O (2005) "**A new *Apicomplexa*-specific protein kinase family: multiple members in *Plasmodium falciparum*, all with an export signal**" *BMC Genomics;* Vol. 6; 1471-1486

94) Schwede T, Kopp J, Guex N, Peitsch MC (2003) "**SWISS-MODEL: an automated protein homology-modeling server**." *Nucleic Acids Research;* Vol. 31; 3381-3385

95) Sefton BM & Shenolikar S (1996) Overview of protein phosphorylation: Labeling studies, In: Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K, editors; "**Current Protocols in Molecular Biology**" *John Wiley & Sons, Inc;* Vol. 4; 18.1.1-18.1.4

96) Simonian MH & Smith JA (2003) Spectrophotometric and colorimetric determination of protein concentration, In: Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K, editors; "**Current Protocols in Molecular Biology**" *John Wiley & Sons, Inc;* Vol. 2; 10.1A.1-10.1A.5

97) Singh N, Preiser P, Renia L, Balu B, Barnwell J (2004) "**Conservation and developmental control of alternative splicing in maebl among malaria parasites**" *Journal of Molecular Biology;* Vol. 343; 589-599

98) Singh SM & Panda AK (2005) "**Solubilisation and refolding of bacterial inclusion body proteins**" *Journal of Bioscience and Bioengineering;* Vol. 99; 303-310

99) Sorensen HP & Mortensen KK (2005) "**Advanced genetic strategies for recombinant protein expression in *Escherichia coli***" *Journal of Biotechnology;* Vol. 115; 113-128

100) Suetterlin BW, Kappes B, Franklin RM (1991) "**Localisation and stage specific phosphorylation of *Plasmodium falciparum* phosphoproteins during the intraerythrocytic cycle**" *Molecular and Biochemical Parasitology;* Vol. 46; 113-122

101) Syin C, Parzy D, Traincard F, Boccaccio I, Joshi MB, Lin DT, Yang X-M, Assemat K, Doerig C, Langsley G (2001) "**The H89 cAMP-dependent protein kinase inhibitor blocks *Plasmodium falciparum* development in infected erythrocytes**" *European Journal of Biochemistry;* Vol. 268; 4 842-4 849

102) Templeton TJ, Deitsch KW (2005) "**Targeting malaria parasite proteins to the erythrocyte**" *Trends in Parasitology;* Vol. 21; 399-402

103) Topolska AE, Lidgett A, Truman D, Fujioka H, Coppel RL (2003) "**Characterisation of a membrane-associated rhoptry protein of *Plasmodium falciparum***" *The Journal of Biological Chemistry;* Vol. 279; 4 648-4 656

104) Topolska AE, Wang L, Black CG, Coppel RL (2004) Chapter 13: Merozoite cell biology, In: Waters AP, Janse CJ, editors "**Malaria Parasites: Genomes and Molecular Biology**" *Caister Academic Press*; 365-444

105) Trager W and Jensen JB (1976) "**Human malaria parasites in continuous culture**" *Science;* Vol. 193; 673–675

106) Vedadi M, Lewa J, Artz J, Amania M, Zhao Y, Donga A, Wasney GA, Gao M, Hills T, Brokx S, Qiu W, Sharma S, Diassiti A, Alama Z, Melone M, Mulichak A, Wernimont A, Bray J, Loppnau P, Plotnikova O, Newberry K, Sundararajan E, Houston S, Walker J, Tempel W, Bochkarev A, Kozieradzki I, Edwards A, Arrowsmith C, Kain K, Roose D, Hui R (2007) "**Genome-scale protein expression and structural biology of *Plasmodium falciparum* and related Apicomplexan organisms**" *Molecular and Biochemical Parasitology*; Vol. 151; 100-110

107) Von Heijne G (1985) "**Signal sequences: The limits of variation**" *Journal of Molecular Biology;* Vol. 184; 99-105

108) Wang L-H, Sudhof TC, Anderson RGW (1995) "**The appendage domain of alpha-adaptin is a high affinity binding site for dynamin"** *Journal of Biological Chemistry;* Vol. 270; 10 079-10 083

109) Wang Y & Roach (1993) "**Purification and assay of mammalian protein (serine/threonine) kinases in protein phosphorylation: A practical approach**" *Oxford University Press;* 121-144

110) Ward P, Equinet L, Packer J, Doerig C (2004) "**Protein kinases of the human malaria parasite *Plasmodium falciparum*: the kinome of a divergent eukaryote**" *BMC Genomics;* Vol. 5; 75-79

111) Wiser MF (1995) "**Proteolysis of a 34 kDa phosphoprotein coincident with a decrease in protein kinase activity during the erythrocytic schizont stage of the malaria parasite**" *Journal of Eukaryotic Microbiology;* Vol. 42; 659-664

112) World Health Organisation (2007) **Roll Back Malaria program**; *http://www.rbm.who.int*

113) Yeo H-J, Larvor M-P, Ancelin M-L, Vial HJ (1997) "***Plasmodium falciparum* CTP phosphocholine cytidylyltransferase expressed in Escherichia coli: purification, characterization and lipid regulation**" *Journal of Biochemistry;* Vol. 324; 903-910

114) Zahn K (1996) "**Overexpression of an mRNA dependent on rare codons inhibits protein synthesis and cell growth**" *Journal of Bacteriology;* Vol. 178; 2 926–2 933.

115) Zeugin JA and Hartley JA (1985) "**Ethanol precipitation of DNA**" *Focus;* Vol. 7; 2-3

116) Zhou Z, Schnake P, Xiao L, Lal A (2004) "**Enhanced expression of a recombinant malaria candidate vaccine in *Escherichia coli* by codon optimisation**" *Protein Expression and Purification;* Vol. 34; 87-94