# THE ROLE OF GLYCEROL KINASE IN *PLASMODIUM FALCIPARUM*

**Kubendran Naidoo**

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

**Johannesburg, November 2012.**

# DECLARATION

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

Kubendran Naidoo

7th day of November 2012

# RESEARCH OUTPUT

**Publications:**

Naidoo, K. and Coetzer, T.L. (2012) Parasite growth is reduced in glycerol kinase knockout *Plasmodium falciparum* parasites during *in vitro* blood stage development. *Manuscript submitted*.

Durand, P.M., Naidoo, K. and Coetzer, T.L. (2008) Evolutionary Patterning: A novel approach to the identification of potential drug target sites in *Plasmodium falciparum* glycerol kinase. PLoS ONE 3(11): e3685.

**Conferences:**

Parasite growth is reduced in glycerol kinase knockout *Plasmodium falciparum* parasites during *in vitro* blood stage development.
Naidoo, K. and Coetzer, T.L.
Oral presentations at:

- South African Society for Biochemistry and Molecular Biology (SASBMB) and Federation of African Societies Biochemistry and Molecular Biology (FASBMB) Conference, Drakensburg, Kwazulu Natal, South Africa, January 2012.
- Molecular Bioscience Thrust Research Day 2011, University of the Witwatersrand, Johannesburg, South Africa, December 2011. (*Best oral presentation*).

Poster presentations at:

- Molecular Approaches to Malaria (MAM2012), Lorne, Australia, February 2012 (Presented by Prof T.L Coetzer).
- International Conference on OMICS Meets Disease and the 3rd Annual Meeting of Proteomics Society (India), Saha Institute of Nuclear Physics, Kolkata, India, December 2011.

An evolutionary and structural approach to the identification of potential drug target sites in *Plasmodium falciparum* Glycerol Kinase.

Naidoo, K., Durand, P.M and Coetzer, T.L.

Oral presentation at:

- Faculty of Health Sciences Research Day, University of the Witwatersrand, Johannesburg, South Africa, August 2008. (*Best oral presentation in the category of Molecular and Comparative Biosciences*).

*Plasmodium falciparum* Glycerol Kinase.

Naidoo, K. and Coetzer, T.L.

Poster presentation at:

- Molecular and Cell Biology Group Symposium, University of Pretoria, Pretoria, South Africa, October 2007. (*Best poster presentation*).

Expression and Purification of Recombinant *Plasmodium falciparum* Glycerol Kinase.

Naidoo, K. and Coetzer, T.L.

Oral presentation at:

- Molecular and Cell Biology Group Symposium, University of the Witwatersrand, Johannesburg, South Africa, October 2006.

Expression and Purification of Recombinant *Plasmodium falciparum* Glycerol Kinase.

Naidoo, K., Isokpehi, R.D. and Coetzer, T.L.

Poster presentations at:

- Faculty of Health Science Research Day, University of the Witwatersrand, Johannesburg, South Africa, August 2006.
- South African Society for Biochemistry and Molecular Biology XX$^{th}$, Conference, Pietermaritzburg, Kwazulu Natal, South Africa, July 2006.
- 45$^{th}$ Conference of the Federation of the South African Societies of Pathology, Rietvleidam, South Africa, July 2005.

Analysis of the Glycerol Kinase Gene in *Plasmodium falciparum*.

Naidoo, K., Isokpehi, R.D. and Coetzer, T.L.

Poster presentation at:

- South African Society for Biochemistry and Molecular Biology XIX$^{th}$ Conference, Stellenbosch, South Africa, January 2005.

# ABSTRACT

Malaria continues to be a devastating disease. *Plasmodium falciparum* is the most lethal human malaria parasite, responsible for the majority of the hundreds of millions of cases of malaria and approximately 665,000 thousand deaths in 2010. Understanding the biology of the parasite is vital in identifying potential drug targets necessary to develop novel treatments to combat the disease. During every 48-hour asexual intra-erythrocytic replication cycle, a single parasite can produce up to 32 progeny. This extensive proliferation implies that parasites require substantial amounts of lipid precursors. Glycerol kinase (GK) is a highly conserved enzyme that functions at the interface of lipid synthesis and carbohydrate metabolism. GK catalyzes the ATP-dependent phosphorylation of glycerol to glycerol-3-phosphate, a major phospholipid precursor. In this study, the full length 1,506bp *P. falciparum* glycerol kinase (*Pf*GK) gene was cloned and expressed as a glutathione S-transferase (GST) fusion protein in *E. coli*. The recombinant *Pf*GK (r*Pf*GK) enzyme was predominantly expressed as an insoluble aggregate, however, ~3µg soluble r*Pf*GK was purified from an 800ml induced culture. SDS-PAGE analysis showed that the protein migrated at ~73kDa and its enzyme activity was verified using an ADP-coupled spectrophotometric assay. The kinetic parameters for r*Pf*GK were $K_m$ = 15.7µM for glycerol and $K_m$ = 15.9µM for ATP. To evaluate the role of the enzyme in asexual blood-stage development, *Pf*GK was disrupted using double crossover homologous DNA recombination to generate a glycerol kinase knockout parasite line (3D7Δ*Pf*GK). Southern hybridization and *Pf*GK mRNA expression analysis verified that the gene had been disrupted. 3D7Δ*Pf*GK growth rates were evaluated using thiazole orange, a DNA staining dye, coupled to flow cytometry analysis for improved sensitivity. Highly synchronized ring stage parasites were monitored over one 48-hour developmental cycle and results showed that 3D7Δ*Pf*GK growth was significantly reduced to 56.5 ± 1.8% when compared to wild type parasites. This reduced proliferation of 3D7Δ*Pf*GK knockout parasites suggests that *Pf*GK is required for optimal proliferation during the blood stages but is not essential for viability and therefore, not a potential drug target. However, *Pf*GK mRNA

expression is markedly elevated in gametocytes and sporozoites. This suggests that *Pf*GK may play a significant role in the mosquito- and liver-stage parasites, with implications for a potential transmission-blocking target. Thus, using a novel bioinformatics method, Evolutionary Patterning, in combination with structural modelling, three potential drug target sites that were different to the human GK orthologue and less likely to develop resistance to compounds were identified. Further studies in the mosquito stages will provide insight into the role of *Pf*GK in the lifecycle of *P. falciparum* parasites.

# ACKNOWLEDGEMENTS

I have met many individuals through the years and have enjoyed their, sometimes brief, company. They are far too many to mention, but I would like to especially acknowledge:

- Professor Theresa Louise Coetzer. My mentor and a truly remarkable individual of whom I have had the privilege to study under and to work for. Thank you for everything, but mostly, your patience.

- Dr Pierre Marcel Durand. He pioneered the concept of evolutionary patterning and identified the potential drug target sites in the *Pf*GK protein used in this study. Thank you for the laughs.

- Dr Sonja Brigitte Lauterbach. Another special lady whose energy, presence and advice were greatly appreciated in the laboratory. Thank you for the time.

- Dewaldt Engelbrecht. Thank you for your assistance when operating the Flow Cytometer and during optimization of the growth assay protocol.

- Sadhaseevan Moodly, Dr Rajdeep Choudhury and Anthea Hean. Thank you for your friendship and support.

- Special thanks to:

    - Current and former members of the Plasmodium Molecular Research Unit, notably, Dr Daniel Meyersfeld, Dr Roberto Lanzillotti, Marcel Gil, Gillian Baker, Belinda Bezuidenhout, Sasha Roets, Warren Vieira, Melanie Wepener, Alisje Steyn and Dale Liebenberg for the company during the many hours we spent in the laboratory together.

- Dr Alexio Capovilla and Katherine Michler for the assistance and facilities during the initial phase of the insect cell-baculovirus based protein expression part of this study.

- Catherine Bell for sequencing some of the constructs.

- Professor Alan F. Cowman at the Walter and Eliza Hall Institute of Medical Research, Victoria, Australia for providing the pCC-1 and pCC-1-EBA knockout plasmid DNA.

- Professor Heinrich C. Hoppe at Rhodes University, Grahamstown, South Africa for providing the WR99210 drug.

- Beckman Coulter South Africa for their use and generous support of the FC500 flow cytometer.

- To the various funding bodies without which none of this work would have been possible:
  - National Health Laboratory Service (NHLS)
  - National Research Foundation (NRF)
  - University of the Witwatersrand

- Lastly, to my family, my mother, uncle Suren, Krishan, Kinesh, Revashnee, Shalini and Keyuri. Thank you for always being there and for the much needed support.

# ETHICS CLEARANCE

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

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# ABBREVIATIONS

| | |
|---|---|
| A | adenine |
| Å | angstrom |
| $A_{230nm}$ | absorbance at 230nm |
| $A_{260nm}$ | absorbance at 260nm |
| $A_{280nm}$ | absorbance at 280nm |
| $A_{340nm}$ | absorbance at 340nm |
| $A_{600nm}$ | absorbance at 600nm |
| AcNPV | *Autographa californica* nucleopolyhedrovirus |
| ACD | acid citrate dextrose |
| ACT | artemisinin-combination therapy |
| Acyl-CoA | acyl coenzyme A |
| AQP | aquaporin/aquaglyceroporin |
| ATP | adenosine triphosphate |
| | |
| bp | base pairs |
| BEVS | baculovirus expression vector system |
| BLAST | Basic Local Alignment Search Tool |
| BSA | bovine serum albumin |
| | |
| C | cytosine |
| CD/UPRT | cytosine deaminase/uracil phosphoribosyl transferase |
| CDP | cytidine diphosphate |
| CDP-DAG | cytidine diphosphate-diacylglycerol |
| CK | choline kinase |
| CEPT | choline/ethanolamine phosphotransferase |
| CQ | chloroquine |
| CTP | cytidine triphosphate |
| CCT | choline-phosphate cytidylyltransferase |
| | |
| Da | dalton |
| dATP | deoxyadenosine triphosphate |
| dCTP | deoxycytidine triphosphate |
| DAG | diacylglycerol |
| DMSO | dimethyl sulphoxide |
| dN | non-synonymous mutation |
| DNA | deoxyribonucleic acid |
| DNase | deoxyribonuclease |
| dNTP | deoxyribonucleoside triphosphate |
| dS | synonymous mutation |
| DTS | drug target site |
| DTT | 1,4-dithiothreitol |
| | |
| EC number | Enzyme Commission number |
| *Ec*GK | *E. coli* glycerol kinase |
| ECT | ethanolamine-phosphate cytidylyltransferase |

| | |
|---|---|
| EK | ethanolamine kinase |
| EP | evolutionary patterning |
| ER | endoplasmic reticulum |
| ET | evolutionary tracing |
| | |
| FA | fatty acid |
| | |
| G | guanine |
| G3P | glycerol-3-phosphate |
| GFP | green fluorescent protein |
| GK | glycerol kinase |
| GnHCl | guanidine-hydrochloride |
| GPAT | glycerol-3-phosphate acyltransferase |
| GST | glutathione-S-transferase |
| GPL | glycerophospholipid |
| | |
| hDHFR | human dihydrofolate reductase |
| HEPES | 4-(2-hydroxyethyl)-piperazine-1-ethanesulphonic acid |
| HT | hexose transporter |
| HMW | high molecular weight |
| | |
| IPTG | isopropyl-β-D-thiogalactopyranoside |
| iRBC | infected red blood cell |
| IS | insoluble |
| | |
| kb | kilo bases (1,000 base pairs) |
| kDa | kilo dalton |
| $K_m$ | Michaelis-Menten constant |
| | |
| LB | Luria broth |
| LMW | low molecular weight |
| LPA | lysophosphatidic acid |
| | |
| mA | milli ampere |
| MAFFT | multiple alignment using fast fourier transform |
| Mb | mega bases (1,000,000 base pairs) |
| MCS | multiple cloning site |
| ML | maximum likelihood |
| MOI | multiplicity of infection |
| MP | maximum parsimony |
| mRNA | messenger ribonucleic acid |
| MSA | multiple sequence alignment |
| MUSCLE | multiple sequence comparison by log expectation |
| | |
| $Na_2EDTA$ | disodium ethylenediaminetetraacetic acid |
| NPP | new permeation pathways |
| | |
| ORF | open reading frame |

| | |
|---|---|
| P | precipitated/purified |
| PA | phosphatidic acid |
| PAGE | polyacrylamide gel electrophoresis |
| *Pb* | *Plasmodium berghei* |
| PBS | phosphate-buffered saline |
| PC | phosphatidylcholine |
| P-Cho | phosphocholine |
| PCR | polymerase chain reaction |
| PDB | Protein Data Bank |
| PE | phosphatidylethanolamine |
| PEMT | PE methyltransferase |
| P-Etn | phosphoethanolamine |
| *Pf* | *Plasmodium falciparum* |
| *Pf*GK | *Plasmodium falciparum* glycerol kinase |
| PI | phosphatidylinositol |
| PK | protein kinase |
| *Pk* | *Plasmodium knowlesi* |
| PL | phospholipid |
| PlasmoDB | *Plasmodium* Database |
| PMSF | phenylmethylsulfonyl fluoride |
| PMT | phosphoethanolamine methyltransferase |
| pRBC | parasite-infected red blood cell |
| PS | phosphatidylserine |
| PSD | phosphatidylserine decarboxylase |
| PV | parasitophorous vacuole |
| PVM | parasitophorous vacuole membrane |
| | |
| RBC | red blood cell |
| RIFINs | repetitive interspersed family |
| RNA | ribonucleic acid |
| RNase | ribonuclease |
| RPMI-1640 medium | Roswell Park Memorial Institute-1640 medium |
| | |
| S | soluble |
| SAM | S-adenosyl methionine |
| SD | serine decarboxylase |
| SDPM | serine decarboxylation-phosphoethanolamine methylation |
| SDS | sodium dodecyl sulphate |
| *Sf*9 | *Spodoptera frugiperda* |
| SM | sphingomyelin |
| | |
| T | thymine |
| TAG | triacylglycerol |
| TB | terrific broth |
| TBS | Tris buffered saline |
| TBST | Tris buffered saline with Tween®-20 |
| TCA | tricarboxylic cycle |
| TEA buffer | Tris EDTA acetate buffer |

| | |
|---|---|
| TE buffer | Tris EDTA buffer |
| Tris | Tris(hydroxymethyl)amino-methane |
| tRNA | transfer ribonucleic acid |
| UTR | untranslated region |
| $V_{max}$ | maximum velocity |
| WHO | World Health Organisation |
| WT | wild-type |

# Chapter 1: Introduction

## 1.1 Malaria

Malaria is the most deadly vector-borne disease in the world. Approximately half of the world's population (3.3 billion people) are at risk for the disease, most living in tropical countries in Africa, central and south America and central, south and southeast Asia (Figure 1). The World Health Organization (WHO) estimated that the number of malaria cases increased from 233 million in 2000 to 244 million in 2005, in line with population growth. In 2010 however, the number of cases decreased to 216 million and the number of estimated deaths was reduced from 985,000 in 2000 to 665,000 in 2010 due to malaria control programmes in the 106 malaria-endemic countries (WHO, 2011b). Most of the deaths occur in third-world and resource-poor countries, especially those in sub-Saharan Africa (~90% of cases) where malaria is especially severe in children under the age of five years. Malaria is caused by *Plasmodium* parasites and together with the expansion of other infectious diseases such as HIV/AIDS and tuberculosis in the last decade, co-infection has resulted in an overall increased risk of mortality and morbidity for each disease, having a significant negative impact on humankind (Hurtley *et al.*, 2010).

Several efforts have been made to ensure global coverage of malaria treatment and prevention programmes and initiatives introduced to curb the burden of the disease. Funding has increased significantly from US$200 million in 2004 to US$2 billion in 2011, and although this is still short of the US$5 billion per annum necessary to achieve universal coverage for all populations at risk, current funding has enabled the WHO and its partners to distribute insecticide-treated nets (ITNs) to 73% of the 800 million people at risk for malaria in sub-Saharan Africa and perform millions of residual indoor insecticide spraying for vector control, all geared towards the WHO Millennium Development Goal of preventing global deaths due to malaria by the end of 2015 (WHO, 2011b). Several successes have been achieved during the last four years, where a decrease in malaria burden was

observed in all WHO regions and malaria effectively eliminated in Morocco and Turkmenistan in 2010 (WHO, 2010b) and Armenia in 2011 (WHO, 2011b).

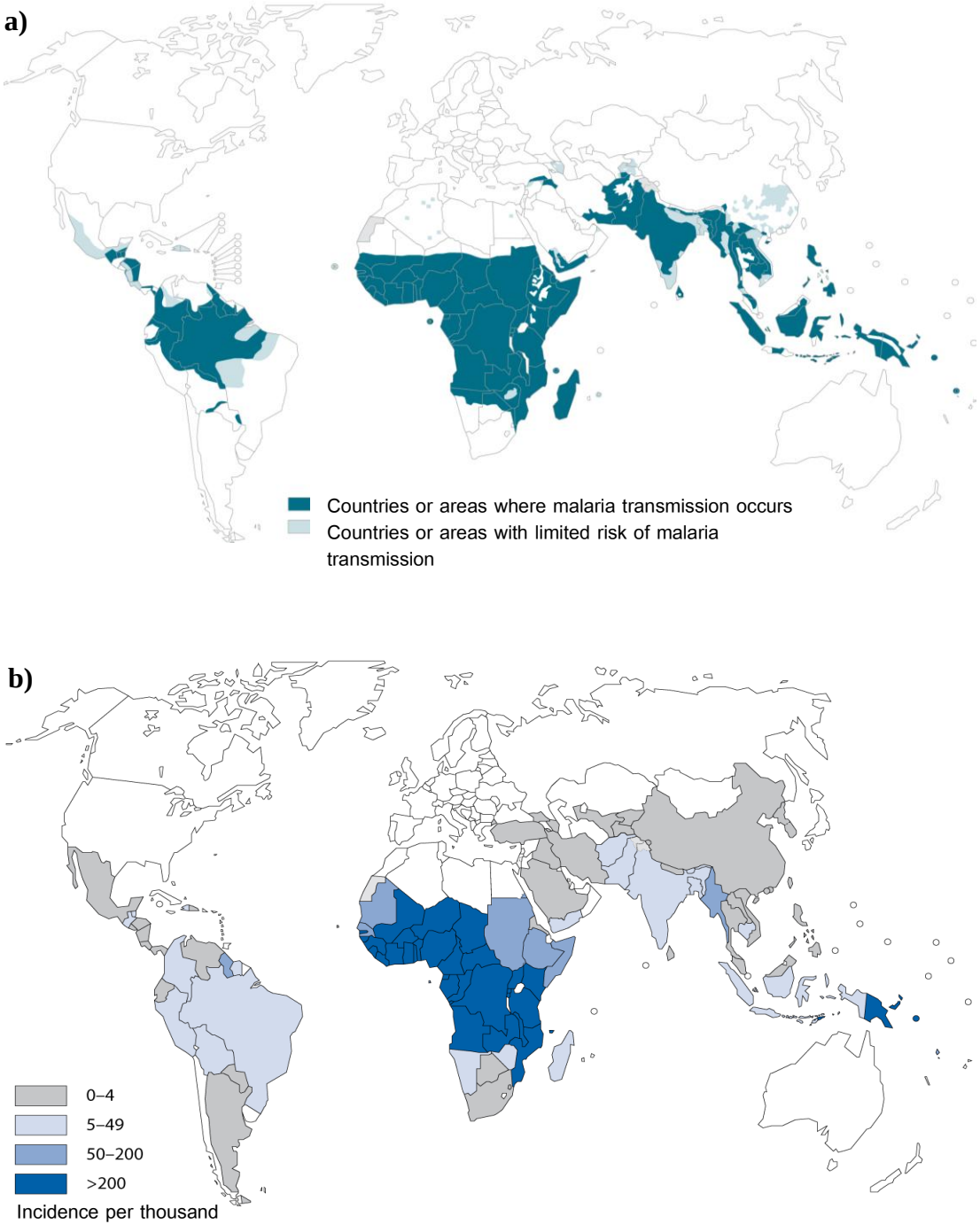



**Figure 1: The global distribution and incidence of malaria.**
**(a)** An illustration of areas with malaria transmission (blue) or with a limited risk of transmission (light blue) (WHO, 2011b); **(b)** the estimated incidence of malaria per 1,000 people (WHO, 2007).

There was evidence of an unexpected increase in the number of malaria cases in three countries (Rwanda, Sao Tome and Principe, and Zambia), in 2009. The specific reasons for resurgence in those areas are unclear; however, they do highlight the fragility of malaria control and the need for sustainable programmes. Indeed, all the gains achieved in controlling malaria may soon be in vain. Mosquito resistance to the single class of insecticides (pyrethroids) used in ITNs, has emerged in 41 countries since 2010 and the emergence of drug resistance to artemisinin-based monotherapies on the Thai-Cambodian border in 2009 (Dondorp *et al.*, 2009, Mok *et al.*, 2011) threatens the lifespan of artemisinin-based combinational therapies (ACTs), the current gold standard of treatment. There is currently no malaria vaccine (WHO, 2011b, WHO, 2010a). Phase III results of the current leading RTS,S/AS01 vaccine candidate, a fusion protein between a region of the circumsporozoite malaria antigen and a hepatitis B surface antigen, in combination with a potent adjuvant, will be released in 2014. However, the first set of three results, conducted at 11 sites across seven sub-Saharan African countries, showed that following a third dose, the vaccine reduced the incidence of clinical malaria by 55% in 6,000 children between the ages 5-17 months over a 12-month period (Agnandji *et al.*, 2011). Several other promising vaccine candidates in Phase I and II clinical trials are currently being explored (WHO, 2011a), however, these are a further 5-10 years behind that of the RTS,S/AS01 vaccine.

Therefore, the lack of an effective malaria vaccine, widespread parasite resistance to anti-malarial drugs, vector resistance to insecticides and the social and economic burden emphasize the need to identify novel drug targets and develop alternative treatment strategies to combat the disease.

### 1.1.1 The *Plasmodium* life cycle

*Plasmodium* is classified as: Superkingdom: Eukaryota with rank Alveolata; Phylum: Apicomplexa; Class: Aconoidasida; Order: Haemosporidia and Genus: *Plasmodium* (Phan *et al.*, 2003). Other animal pathogens, such as *Toxoplasma* and

*Theileria*, fall under the Apicomplexa grouping. Nearly 200 *Plasmodium* species are known to infect mammals, birds and reptiles (Rich and Ayala, 2006), five of which are responsible for disease in humans. The most prevalent and lethal is *Plasmodium falciparum*, responsible for ~90% of deaths due to malaria in humans. *P. vivax* (the most widespread of the malaria parasites), *P. malariae*, *P. ovale* and more recently *P. knowlesi* (Singh *et al.*, 2004) cause less complicated forms of the disease. The host specificity of *P. knowlesi* is not restricted to humans, as it also infects monkeys.

Malaria parasites exhibit a complex, multistage life cycle involving a vertebrate (human) and an invertebrate (*Anopheles* mosquito) host. During its lifecycle the parasite undergoes 10 morphological transitions in five different tissues, proliferates asexually in three of these tissues and propagates sexually following transfer from the human to the mosquito (Figure 2 and Figure 3).

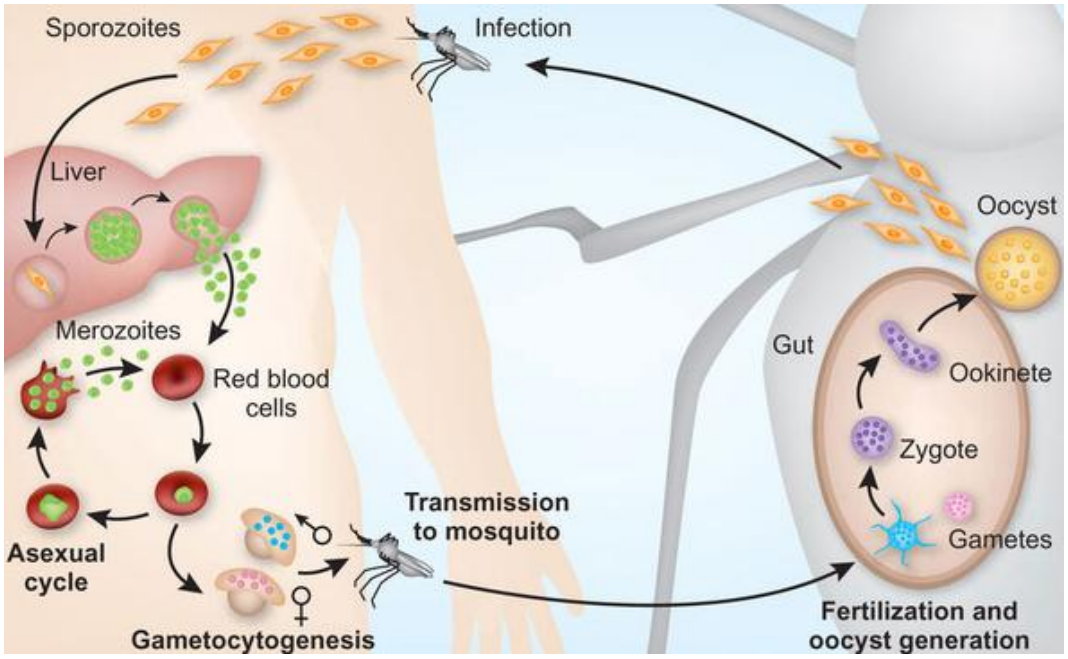



**Figure 2: The life cycle of *Plasmodium*.**
The development of *Plasmodium* parasites in the human (**left**) and mosquito (**right**). Three distinct stages with different parasite forms exist during development: the mosquito-stage, and exo-erythrocytic liver stage and an asexual intra-erythrocytic stage (Pasvol, 2010).

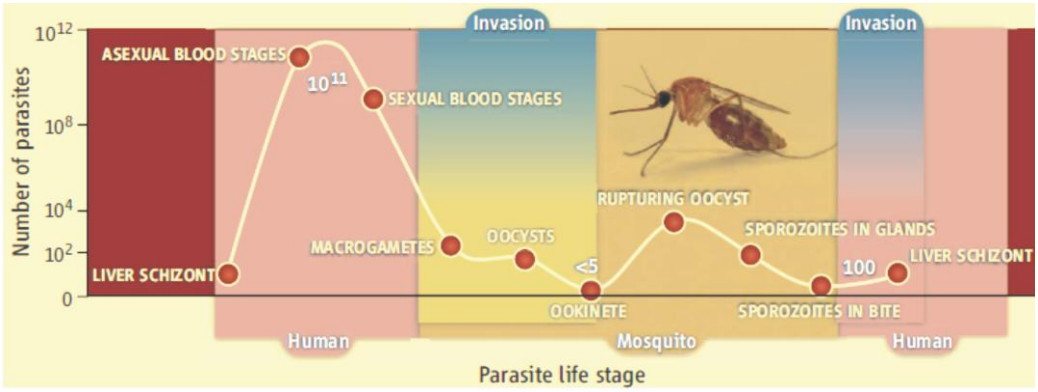



**Figure 3: Population of *Plasmodium* parasites at each developmental stage.**
The number of malaria parasites during growth and proliferation within each host. Parasites can range from less than five ookinetes in the mosquito to trillions in the human bloodstream (Vogel, 2010).

When a female mosquito takes a blood meal, sporozoites from its salivary glands are transmitted to the subcutaneous tissue of the human host. The sporozoites travel rapidly to the liver and invade hepatocytes where they develop into an exo-erythrocytic stage called hepatic or tissue schizonts. After 6-10 days, schizonts undergo schizogony, multiplying via mitosis until they rupture the infected hepatocyte and release thousands of merozoites into the bloodstream. Merozoites invade erythrocytes and after a single 48-hour asexual developmental and replication cycle (72 hours in the case of *P. malariae* and 24 hours for *P. knowlesi*), a single merozoite produces 16-32 daughter merozoites. Progeny merozoites subsequently invade other erythrocytes and it is this stage of the parasite life cycle that is responsible for the clinical features of the disease, including fever, anaemia and cerebral malaria. A small proportion of merozoites do not multiply following invasion, but instead differentiate into sexual parasite forms termed gametocytes. When gametocytes are ingested by a mosquito during a subsequent blood meal, environmental conditions such as (i) the presence of organic chemical signals, including xanthurenic acid (Billker *et al.*, 1998), (ii) an approximately 5ºC decrease in temperature (iii) an increase in pH from 7.2 to 8.0 (Billker *et al.*, 1997), (iv) a decrease in glucose concentration and (v) an increase in $Na^+$ as well as a $HCO_3^-/Cl^-$ ion exchange shuttle, initiate gametogenesis (Kawamoto *et al.*, 1992, Alano and Billker, 2005). Both the male microgametocyte and female macrogametocyte differentiate into gametes and

egress from the enveloping human erythrocyte within 10 minutes post-activation. The male microgamete exflagellates and undergoes three rounds of mitosis to form eight haploid microgametes. Fertilization occurs within the next hour resulting in the formation of a zygote. This brief diploid stage allows for sexual recombination of genetic material, including those genes responsible for drug resistance. During the next 24 hours, the zygote differentiates into a motile ookinete which, over 36-48 hours, traverses across the mosquito midgut epithelium and lodges beneath the basal lamina facing the mosquito body cavity (haemocoel). During the next 10-15 days, ookinetes differentiate into mature oocysts, each producing thousands of sporozoites. Upon oocyst rupture, oocyst-derived sporozoites are released into the haemocoel where they migrate to and invade mosquito salivary glands. Sporozoites are injected back into a human host during another blood meal, completing the *Plasmodium* life cycle (Moreira *et al.*, 2004).

### 1.1.2 Drug-resistant malaria

*Plasmodium's* resistance to antimalarials is a major obstacle in the fight against malaria. Several factors including (i) high mutation rates of the parasite, (ii) low fitness cost associated with resistance mutations (iii) strength of drug selection, which accelerates the spread of the drug-resistant parasites whilst decreasing the prevalence of the competing sensitive wild-type strain, and (iv) non-treatment compliance have led to the emergence of drug-resistant parasites (Petersen *et al.*, 2011). Since the 1940s, fast-acting and inexpensive antimalarials such as chloroquine (CQ) successfully controlled the disease. Worldwide drug use and selective pressure though, led to the emergence and spread of drug-resistant parasites. History, therefore, shows that the introduction of drug treatment is usually followed by parasite resistance to the drug, with some cases more rapid than others (Figure 4). *P. falciparum* resistance to pyrimethamine-sulfadoxine or to atovaquone, for example, developed in the same year that the drugs were introduced. Quinine, one of the oldest antimalarial compounds, has been used to treat malaria since the 17$^{th}$ century. Resistance to quinine developed sporadically,

first in Brazil in 1908 (Clyde, 1972) and in southeast Asia in 1965 (Hyde, 2005). The short half-life (8-10 hours) of quinine probably limited the wide spread of resistance to the drug (Petersen *et al.*, 2011). Therefore, quinine can still be used and is reserved, in combination with antibiotics, as a second-line treatment in severe cases of malaria (WHO, 2010a).

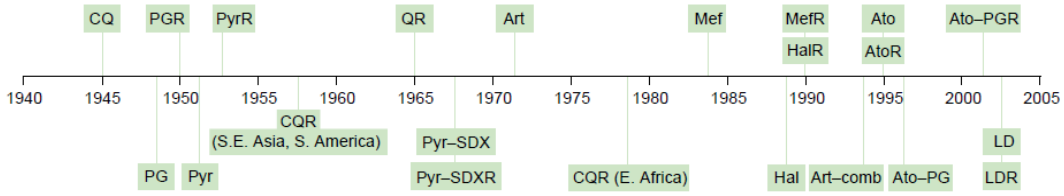



**Figure 4: Timeline of anti-malarial drug introduction and resistance.**
Approximate dates (1945-2005) from the introduction of anti-malarial drug compounds until observed treatment failure. Resistance to artemisinin has emerged in Western Cambodia in 2009 (Dondorp *et al.*, 2009). **Abbreviations:** chloroquine (CQ), proguanil (PG), pyrimethamine (Pyr), sulfadoxine (SDX), quinine (Q), artemisinin (Art), mefloquine (Mef), halofantrine (Hal), atovaquone (Ato), LapDap (LD) and resistance (R) (Hyde, 2005).

CQ (a synthetic derivative of quinine) together with vector control measures was the cornerstone of the malaria eradication campaign in the 1950-60s. Its efficacy, affordability and safety made CQ the gold standard of treatment for several years. CQ is involved in the haem detoxification pathway (Orjih *et al.*, 1994, Mita *et al.*, 2009), whereby parasites digest the red cell haemoglobin and polymerize the toxic haem moiety into inert biocrystals known as haemozoin (Egan, 2008). CQ-resistant strains, however, emerged 10-15 years later in southeast Asia and south America and by 1978, had spread to Africa (Mita *et al.*, 2009). CQ is the first-line treatment for *P. vivax* and despite widespread resistance, can still be used to treat *P. falciparum* in countries where the drug is still effective (WHO, 2010a). Interestingly, the withdrawal of CQ in some countries has led to the re-emergence of the wild-type CQ-sensitive strains. Malawi's treatment policy changed from CQ to ACTs in 1993. Thirteen years later, CQ treatment was observed to have a 99% curative rate in children with uncomplicated malaria (Laufer *et al.*, 2006). The mechanism of emergence of the CQ-sensitive phenotype is unclear, but it is

thought to restore the fitness of the parasite in the absence of drug selection pressure.

The antifolate drugs interfere with nucleotide synthesis and include the sulfa drugs (sulfadoxine and dapsone) which inhibit the dihydropteroate synthase enzyme; and pyrimethamine and proguanil which inhibit the dihydrofolate reductase (*Pf*DHFR) moiety of the bifunctional dihydrofolate reductase/thymidylate synthase (*Pf*DHFR/TS) enzyme (Yuvaniyama *et al.*, 2003). The sulfadoxine-pyrimethamine drug combination was introduced following the emergence of CQ-resistant parasites. However, mutations in both target enzymes rendered the drug ineffective. Sulfadoxine-pyrimethamine is now used in intermittent preventative treatment (WHO, 2010b).

The WHO has recommended artemisinin-based combination therapies as the first-line treatment of uncomplicated malaria since 2001. However, resistance to artesunate monotherapy has now emerged on the Thai-Cambodian border (Dondorp *et al.*, 2009) and because the three artemisinin derivatives (artesunate, artemether and dihydroartemisinin) are chemically closely related, resistance to one compound is likely to lead to parasite resistance to all.

Although increased funding has renewed interest for academics and pharmaceutical companies in the malaria field, generally, the drug discovery and development pipeline is a long, complex and costly process that can take up to 15 years and cost in excess of US$1.8 billion (Paul *et al.*, 2010). Despite advances in molecular and computational biology and expansion of chemically diverse compound libraries, a success rate of 5-6 ‘first-in-class’ new molecular entities (i.e. small-molecule drugs containing an active ingredient with a new or unique mechanism of action that elicits a specific response upon the drug-target interaction, and that has not been previously approved) per annum, across the pharmaceutical industry and for all diseases, is poor (Paul *et al.*, 2010). Indeed, recent data from a survey of nine large pharmaceutical companies showed that

only two compounds were approved by the FDA in 2010, a poor return from an investment of approximately US$60 billion (Bunnage, 2011).

Several thousand promising antimalarial compounds have recently been identified (Gamo *et al.,* 2011, Guiguemde *et al.,* 2011) and a few other lead compounds are in the pipeline (Fidock, 2011). However none have reached Phase I clinical trials and it will be several years before novel antimalarials are introduced into the market. Whether they will be effective against resistant strains (cross resistance) or whether new resistant strains will emerge, and how rapidly, remains to be seen. While the development of a fully protective malaria vaccine remains elusive, drug therapy is currently key to controlling malaria and if the current generation of ACTs fail, no solid contingency plans are in place.

## 1.2 *P. falciparum* genome and proteome

The *P. falciparum* nuclear genome was published in 2002 (Gardner *et al.*, 2002) and consists of 14 chromosomes ranging in size from 0.643kb for chromosome one to 3.29Mb for chromosome 14. The 22,853,764bp (~23Mb) nuclear genome is comparable in size to other *Plasmodium* genomes and encodes for 5,268 predicted open reading frames (ORFs) or genes, at an average gene density of one per 4,338bp. The A+T content of the genome is 80.6%, the highest sequenced to date, and that rises to ~90% in introns and intergenic regions. The mean gene length is 2,283bp, significantly larger than the average 1,300-1,600bp in other organisms. There is no clear explanation of the increased gene length. More than 60% (3,208) of ORFs identified by the malaria genome sequencing consortium lack similarity to genes from any other known organism. Ascribing putative roles to these ‘hypothetical proteins’ in the absence of homology remains a challenging task, however, their unique nature is important in understanding parasite biology and identifying *Plasmodium*-specific pathways and proteins that may represent targets for both vaccine or antimalarial chemotherapeutic strategies. Introns are present in 53.9% of genes and there is an average of 2.39 exons per gene, with a mean length of 939bp and 76.3% A+T content. No transposable elements or

retrotransposons have been identified which suggest that mobile genetic elements have not had any profound impact on *P. falciparum's* genome.

*P. falciparum* also contains two non-nuclear genomes: (i) a 6kb linear mitochondrial genome encoding three proteins (cytochrome *b* and subunits I and II of cytochrome *c*) and ribosomal RNA and (ii) a 35kb circular plastid-like or apicoplast genome encoding 30 proteins that are involved in various metabolic processes including fatty acid (FA) synthesis, haem and isoprenoid biosynthesis (Gardner *et al.*, 2002). The plastid is homologous to chloroplasts of plants and algae and thought to originate via a secondary endosymbiotic event (Lim and McFadden, 2010, Moore *et al.*, 2008). Approximately 551 nuclear-encoded proteins are post-translationally targeted to the apicoplast. Most of the proteins encoded by the plastid genome are of prokaryotic origin, making them attractive drug targets. Indeed, inhibitors against plastid-associated isoprenoid biosynthesis, DNA replication and translation have been shown to kill parasites and clear infection in mice, demonstrating the essential nature of the plastid (Fichera and Roos, 1997, Jomaa *et al.*, 1999). Sixty genes, including those which code for proteins involved in the tricarboxylic cycle (TCA) and oxidative damage protection were identified as being of probable mitochondrial origin (Gardner *et al.*, 2002).

Of the 5,268 predicted proteins, 733 were identified as enzymes, of which only 435 (or ~8% of proteins) were assigned Enzyme Commission (EC) numbers. These are significantly fewer than the 25-33% genes in bacterial and archael genomes that can be mapped to the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway diagrams or the 17% of *S. cerevisiae* proteins that have been assigned EC numbers. However, the data suggest that either (i) *P. falciparum* devotes a smaller proportion of its genome to enzymes, (ii) enzymes could not be identified through homology-based methods or (iii) a few genes escaped detection in smaller regions/gaps of the genome that remain unresolved (Gardner *et al.*, 2002).

## 1.3 Carbon metabolism in *P. falciparum*

Genomic, transcriptomic and proteomic analysis during the last decade has assisted greatly towards understanding energy and carbon metabolism in *Plasmodium* parasites. During intra-erythrocytic development, *Plasmodium* parasites are secluded within a parasitophorous vacuole (PV) of which its membrane (PVM) is derived from both the host red blood cell (RBC) and the parasite. Though this compartmentalization shelters the parasite from host defence mechanisms, the parasite has to develop strategies to access host nutrients and expel waste products. *P. falciparum* acquires metabolites through active modification of the infected RBC (iRBC) membrane and the expression of a variety of transporters on the surface of these cells. The export of parasite proteins to the RBC surface and the induction of new permeation pathways (NPP) dramatically change the properties of the plasma membrane. Activation of NPP results in the import of metabolites essential for parasite growth (ions, sugars and peptides) as well as the export of waste products, such as lactate from glycolysis (Kirk, 2001). More than 2.5% of the *P. falciparum* genome is dedicated to genes coding for channels and transporters, however, only a small proportion have been characterized (Martin *et al.*, 2009).

Most of the central carbon metabolism work has focused on the intra-erythrocytic stages of *P. falciparum* parasite development. Collectively, studies suggest that the core components of metabolism are (i) glycolysis, (ii) the pentose phosphate pathway, (iii) glycosylation and aminosugars, (iv) aspects of TCA metabolism and (v) lipid biosynthesis (Figure 5) (Olszewski and Llinas, 2010). Glycerol metabolism, boxed in red in Figure 5, is the central theme of this thesis and is discussed in section 1.5.

### 1.3.1 Glycolysis

*In vitro* studies showed that *Plasmodium* parasites are voracious consumers of glucose. Uninfected, mature red blood cells consume glucose at a rate of

5µmol/24hrs/$10^9$ RBCs, with a 90% glucose to lactate conversion. However, upon *Plasmodium* invasion, glucose consumption is increased up to 100-fold, particularly in the metabolically active trophozoite and schizonts stages (Roth, 1990).

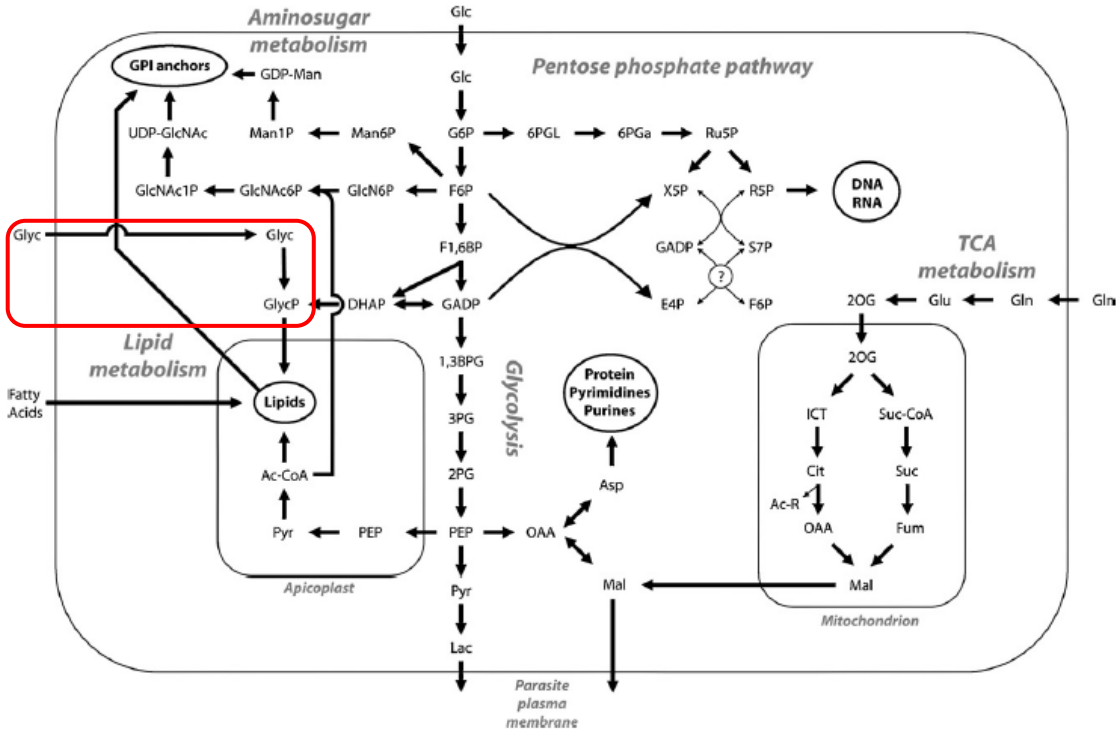



**Figure 5: Carbon metabolism in *P. falciparum*.**
Arrows represent the proposed direction of net metabolite flux. Text in circles represents major biomass components. The circled question mark indicates uncertainty about the existence of the enzyme transaldolase. Cofactors such as ATP and NADH are not shown for clarity. **Abbreviations:** Glc, glucose; G6P, glucose-6-phosphate; F6P, fructose-6-phosphate; F1,6BP, fructose-1,6-bisphosphate; DHAP, dihydroxyacetone phosphate; GADP, glyceraldehyde-3-phosphate; 1,3BPG, 1,3-bisphosphoglycerate; 3PG, 3-phosphoglycerate; 2PG, 2-phosphoglycerate; PEP, phosphoenolpyruvate; Pyr, pyruvate; Lac, lactate; Ac-CoA, acetyl-coenzyme A; Ac-R, acetate or acetyl-CoA; GlycP, glycerol-3-phosphate; Glyc, glycerol; Man6P, mannose-6-phosphate; Man1P, mannose-1-phosphate; GDP-Man, GDP-mannose; GlcN6P, glucosamine-6-phosphate; GlcNAc6P, N-acetyl-glucosamine-6-phosphate; GlcNAc1P, N-acetyl-glucosamine-1-phosphate; UDP-GlcNAC, UDP-N-acetyl-glucosamine; 6PGL, 6-phosphoglucono-δ-lactone; 6PGa, 6-phosphogluconate; Ru5P, ribulose-5-phosphate; R5P, ribose-5-phosphate; X5P, xylulose-5-phosphate; S7P, sedoheptulose-7-phosphate; E4P, erythrose-4-phosphate; Asp, aspartate; Gln, glutamine; Glu, glutamate; 2OG, 2-oxoglutarate; ICT, isocitrate; Cit, citrate; OAA, oxaloacetate; Mal, malate; Suc-CoA, succinyl-CoA; Suc, succinate; Fum, fumarate; GPI, glycophosphatidylinositol (Olszewski and Llinas, 2010).

The glucose to lactate conversion in parasitized RBCs (pRBC), however, is reduced to 60-70%. This reflects the commitment of glucose to nucleic acid, lipid and protein biomass. Only a small proportion of glucose is completely oxidized into carbon dioxide. This correlates with the lack of a fully functional TCA under the anaerobic intra-erythrocytic growth conditions (Olszewski and Llinas, 2010). *P. falciparum* acquires glucose (and fructose) via an essential hexose transporter (*Pf*HT) and rodent models have shown that the *P. berghei* HT is expressed throughout parasite development (Slavic *et al.*, 2010). All the enzymes necessary for a fully functional glycolytic pathway have been identified in *P. falciparum* (Gardner *et al.*, 2002) and are expressed during the blood-stages (Bozdech *et al.*, 2003).

*Plasmodium* parasites can utilize alternative carbon sources; *P. falciparum* can survive on fructose *in vitro*, albeit at a reduced proliferation rate (Woodrow *et al.*, 2000). *P. falciparum* parasites, however, lack fructose bisphosphatase, implying that gluconeogenesis is absent. Enzymes involved in the synthesis of trehalose, glycogen or other carbohydrate stores were also not identified in the *P. falciparum* genome (Gardner *et al.*, 2002). Ultimately, *Plasmodium* parasites depend on the fermentation of a constant and abundant supply of glucose, irrespective of alternative energy sources.

### 1.3.2 Pentose phosphate pathway

The pentose phosphate pathway (PPP) is an essential conserved pathway composed of two interconnected branches (Figure 5): (i) the oxidative arm, where glucose-6-phosphate is oxidized to ribose-5-phosphate (R5P) to generate the nucleic acid ribose backbone and NADPH and (ii) the non-oxidative arm, where R5P is recycled back into glycolysis when NADPH is required to maintain redox balance within the cell in response to oxidative stress (Bozdech and Ginsburg, 2005). Genes for all enzymes in the conventional PPP have been identified, with the exception of transaldolase (Gardner *et al.*, 2002).

### 1.3.3 Glycosylation and aminosugars

The malaria parasite expresses a variety of glycoconjugated proteins, several of which are present in the intra-erythrocytic stage and are involved in parasite invasion and pathogenesis of the disease. For example, the merozoite surface proteins MSP-1 and MSP-2 comprise ~67% of the total membrane-associated surface coat. The glycosylation machinery present in the *Plasmodium* genome appears to be restricted to the biosynthesis of glycophosphatidylinositols (GPIs), a class of glycolipids that anchors proteins to the plasma membrane. The GPI biosynthetic pathway occurs in the endoplasmic reticulum (ER) and utilizes a variety of sources including hexose sugars, inositol, glycerol-3-phosphate, ethanolamine and lipids (Figure 5) (Olszewski and Llinas, 2010, von Itzstein *et al.*, 2008).

### 1.3.4 Tricarboxylic acid metabolism

In most eukaryotes, the mitochondrion is a versatile central hub of carbon metabolism in which carbon derived from glycolysis is fully oxidized in a unidirectional pathway to produce energy by oxidative phosphorylation. The pyruvate generated from glycolysis enters the mitochondria where it is converted to acetyl coenzyme A (Acetyl-CoA) and carbon dioxide. Acetyl-CoA enters the TCA cycle, where following a series of enzyme-mediated reactions, citrate is converted to oxaloacetate. The movement of protons and electrons across the inner mitochondrial membrane and the electron transport chain drives the multi-protein complex $F_0F_1$-ATP synthase to produce ATP. This process produces 36 molecules of ATP from one molecule of glucose.

A complete set of genes encoding TCA enzymes has been identified in the *P. falciparum* nuclear genome (Gardner *et al.*, 2002) (Figure 5) and apart from malate dehydrogenase, they also possess mitochondrial signal sequences. However, the minimal oxygen consumption, rapid fermentation of glucose and the presence of a single mitochondrion with minimal cristae imply that the

contribution of oxidative phosphorylation during intra-erythrocytic parasite development is uncertain. In addition, *Plasmodium* has dispensed with several functions that utilize ATP from the TCA cycle, such as the *de novo* amino acid biosynthetic pathway and instead, acquires amino acid precursors through scavenging from the host serum or haemoglobin catabolism (Liu *et al.*, 2006).

Proteomic profiling suggests that the TCA cycle is more prominent in the mosquito-stages rather than the blood-stages (Lasonder *et al.*, 2008). However, enzymes involved in TCA metabolism are upregulated under starvation conditions during *in vivo* intra-erythrocytic development (Daily *et al.*, 2007). Labelling studies have shown that *P. falciparum* utilizes glutamate and glutamine in an unusual branched TCA cycle to generate ATP and acetyl-CoA (Figure 5) (Olszewski *et al.*, 2010). It appears that the TCA cycle in *P. falciparum* parasites consists of an oxidative pathway that generates succinyl-CoA essential for haem biosynthesis and a reductive pathway to generate citrate. Interestingly, the Acetyl-CoA generated in the TCA is targeted to the nucleus for the acetylation of histone proteins; whereas glucose-derived acetyl-CoA in the apicoplast is used in the synthesis of amino sugars in the ER. Thus, the parasite appears to generate functionally distinct acetyl-CoA from alternate carbon sources (Olszewski *et al.*, 2010, Ginsburg, 2011).

### 1.3.5 Phospholipid synthesis

Throughout their life cycle, *Plasmodium* parasites repeatedly undergo rapid cell growth and division, necessitating lipid acquisition and membrane biogenesis. Lipids are among the most critical components of the parasite and there is substantial evidence supporting their role in (i) synthesis of membranes that serve as permeability barriers between the parasite and its host, (ii) membrane networks in the cytoplasm of pRBCs and (iii) lipid-derived signaling molecules that regulate parasite development during growth and proliferation. *Plasmodium* lipid content, lipid acquisition as well as the enzymes and pathways involved in lipid biogenesis have been studied in several *Plasmodium* species, however, mostly at

the intra-erythrocytic stages of development of infected human (*P. falciparum*), primate (*P. knowlesi*) and rodent (*P. berghei* and *P. yoelli*) RBCs. Although there are differences in phospholipid (PL) metabolic pathways between rodent and non-rodent malaria parasites (Dechamps *et al.*, 2010a), such as the serine decarboxylase-phosphoethanolamine methyltransferase (SDPM) pathway discussed later, it is expected that the broad outline of PL metabolism should be conserved among the *Plasmodium* species due to similarities in their life cycle.

PLs, in particular, glycerophospholipids (GPLs), are the main constituents of parasite membranes during blood-stage development, where extensive proliferation is accompanied by a marked increase in phosphatidylcholine (PC), phosphatidylethanolamine (PE), phosphatidylserine (PS), phosphatidylinositol (PI), fatty acids, as well as the diacylglycerol (DAG) and triacylglycerol (TAG) neutral lipid content. The increase in DAG and TAG is significant when expressed as a percentage, however, it is low in relation to the GPLs. Little is known about lipid acquisition during non-erythrocytic stages, however, extensive proliferation during exo-erythrocytic development and during sporozoite formation within oocysts imply a need for massive amounts of lipids at these stages (Vial and Mamoun, 2005, Dechamps *et al.*, 2010b).

In uninfected red cells, PLs and cholesterol are the major components of red cell membranes, co-existing at a molar ratio of approximately 1.0 (Maguire and Sherman, 1990, Vial and Mamoun, 2005). PC, PE, sphingomyelin (SM) and PS are the main PLs. Other phospholipids, such as PI, phosphatidic acid (PA), cardiolipin (CL) and lyso-PLs account for less than 5% of the total PLs, and neutral lipids and fatty acids are scarcely detectable (Maguire and Sherman, 1990, Vial and Mamoun, 2005). Upon *Plasmodium* infection, a 5-6-fold increase (Vial and Mamoun, 2005) in erythrocyte PL content, specifically PC, PE and PI, is observed (Figure 6). There is no substantial modification to the host membrane with respect to lipid content and class distribution (Vial and Mamoun, 2005); although observations of a loss in SM content has been reported (Maguire and Sherman, 1990) and the two FA molecules attached to the GPL backbone are

modified in both chain length and degree of saturation (Krishnegowda and Gowda, 2003).

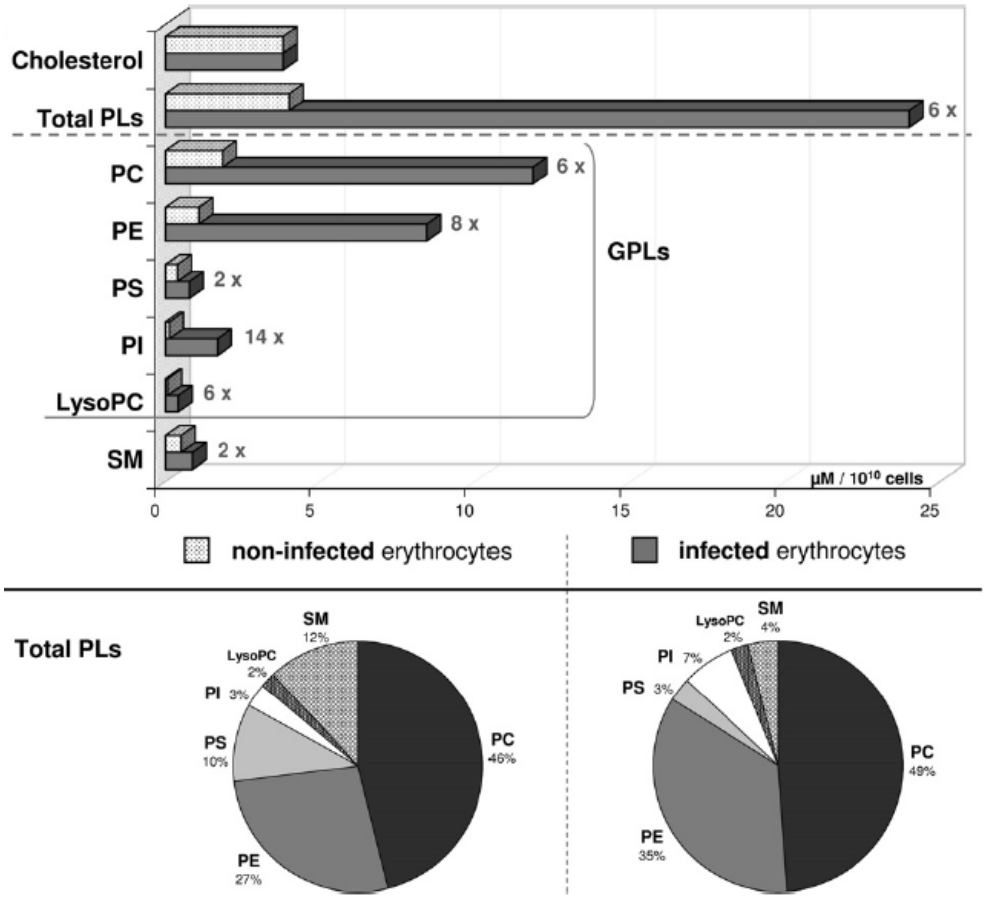



**Figure 6: Cholesterol and phospholipid composition in non-infected and *P. falciparum*-infected erythrocytes.**
**(a)** The fold-increase in phospholipid (PL) content of pRBCs and **(b)** the molar percentage in phospholipid content of non-infected and pRBCs are indicated. **Abbreviations:** GPLs, glycerophospholipids; LysoPC, lysophosphatidylcholine; PC, phosphatidylcholine; PE, phosphoethanolamine; PI, phosphatidylinositol and SM, sphingomyelin (Vial and Ancelin, 1992, Dechamps *et al.*, 2010b).

Metabolic labelling and gas chromatography studies showed that *P. falciparum* parasites (i) efficiently incorporated C12-C18, C16:1, C18:1 and C18:2 FAs into the PL backbone, but could not incorporate very long-chain unsaturated FAs such as arachidonic acid (C20:4), (ii) preferentially incorporated saturated and unsaturated FAs at the *sn*-1 and *sn*-2 positions of the PL backbone, respectively (Krishnegowda and Gowda, 2003) and (iii) have some capacity to desaturate and

elongate serum-derived FAs, although the parasite's FAs are similar to those in the serum (Mi-Ichi *et al.*, 2006). The FA composition is similar to those in parasite PLs, suggesting that *Plasmodium* is capable of fine modulation of the host membrane and indicates dynamic PL traffic between the host and parasite membrane (Vial *et al.*, 2003). In addition, while PS is present on the inner leaflet of the plasma membrane in RBCs, pRBCs display PS on the outer leaflet, which contributes to the cytoadherence of the infected cell (Eda and Sherman, 2002). Parasite membranes contain predominantly PC (40-50%), PE (35-45%) and PI (4-11%), with SM and PS accounting for <5%. The PE content is unusually high for a eukaryotic organism (Vial and Mamoun, 2005). In marked contrast to erythrocytes, there is hardly any cholesterol in parasite membranes due to the parasite's inability to synthesize sterols (Holz, 1977).

The major PLs are synthesized via several pathways, summarized in Figure 7: (i) an ancient CDP-diacylglycerol (CDP-DAG)-dependent pathway which utilizes PA precursors to form PI and PS, with the latter being decarboxylated to PE (Vial and Mamoun, 2005), (ii) eukaryotic *de novo* cytidine diphosphate (CDP)-choline and CDP-ethanolamine pathways (Kennedy pathways), which incorporate exogenous choline and ethanolamine into lipid components (Vial and Mamoun, 2005) and (iii) an additional plant-like SDPM pathway in *P. falciparum* parasites that synthesizes PC from serine and ethanolamine (Pessi *et al.*, 2004).

## *Formation of phosphatidic acid*

In most organisms, glycerolipid metabolism begins with the acylation at the *sn*-1 position of glycerol-3-phosphate (G3P) to produce lysophosphatidic acid (LPA). The formation of LPA represents the commitment to GPL biosynthesis and is catalyzed by glycerol-3-phosphate acyltransferase (GPAT). In *P. falciparum*, GPAT is located in the ER and is developmentally regulated, with activity highest during trophozoite development, the stage of active PL synthesis (Santiago *et al.*, 2004).

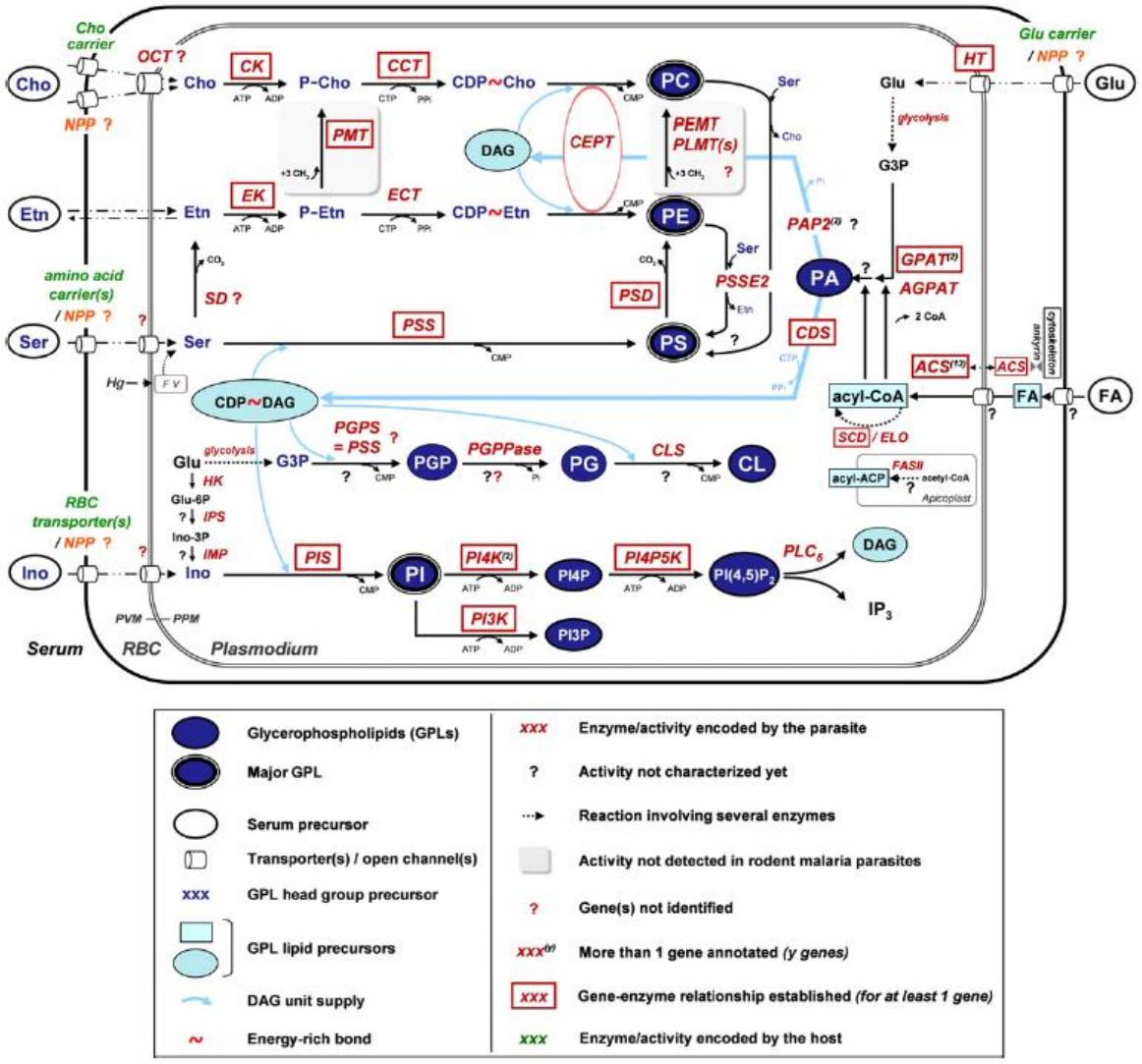



**Figure 7: Overview of glycerophospholipid biosynthesis in *Plasmodium* (Dechamps *et al.*, 2010b).**

GPL pathways are described in the text. Metabolite uptake via the parasitophorous vacuole membrane (PVM) and/or the parasite plasma membrane (PPM) are not distinguished. Colours, shapes and symbols representing metabolites, genes and proteins are defined in the key. **Abbreviations:** FV, food vacuole; Hg, haemoglobin. *Metabolites*: CDP, cytidine diphosphate; CMP, cytidine monophosphate; CTP, cytidine triphosphate; Cho, choline; CL, cardiolipin; CoA, coenzyme A; DAG, diacylglycerol; Etn, ethanolamine; FA, fatty acid; Glu, glucose; Glu-6P, glucose-6-phosphate; G3P, glycerol 3-phosphate; Ino, myo-inositol; Ino-3P, myo-inositol-3-phosphate; IP3, inositol 1,4,5-triphosphate; P-Cho, phosphocholine; P-Etn, phosphoethanolamine; PA, phosphatidic acid; PC, phosphatidylcholine; PE, phosphatidylethanolamine; PG, phosphatidylglycerol; PGP, phosphatidylglycerophosphate; PI, phosphatidylinositol; PI3P, PI 3-phosphate; PI4P, PI 4-phosphate; PI(4,5)P2, PI 4,5-bisphosphate; Pi, inorganic phosphate; PPi, pyrophosphate; PS, phosphatidylserine; Ser, serine. *Transporters*: HT, hexose transporter; OCT, organic-cation transporter (genes not identified). Enzymes involves in PL and neutral lipid synthesis are summarized in Table 1 (Dechamps *et al.*, 2010b).

**Table 1: Enzymes involved in *P. falciparum* GPL synthesis (Dechamps *et al.*, 2010b).**

| Enzyme abbreviation | Enzyme Name | *P. falciparum* gene(s) | PlasmoDB ID[1] |
|---|---|---|---|
| ACS | Acyl-CoA synthetase | PfACS1a, PfACS1b, PfACS2, PfACS3, PfACS4, PfACS5, PfACS6, PfACS7, PfACS8, PfACS9, PfACS10, PfACS11,PfACS12 | PF14 0761, PF14 0751, PFC0050c, PFL2570w, MAL13P1.485, PF07 0129, PFD0085c, PFL0035c, PFB0695c, PFB0685c, PFE1250w, PFL1880w, PFF0945c |
| AGPAT | 1-Acyl-G3P acyltransferase | PfPlsC | PF14_0421 |
| CCT | Choline-phosphate cytidylyltransferase | PfCCT | MAL13P1.86 |
| CDS | CDP-DAG synthase | PfCDS | PF14_0097 |
| CEPT | Choline/ethanolamine-phosphotransferase | PfCEPT | PFF1375c |
| CK | Choline kinase | PfCK | PF14_0020 |
| CLS | CL synthase | PfCLS | PFF0465c |
| ECT | Ethanolamine-phosphate cytidylyltransferase | PfECT | PF13_0253 |
| EK | Ethanolamine kinase | PfEK | PF11_0257 |
| ELO | Long-chain FA elongase | | PF0290w, PF11_0344, PFI0980w |
| GPAT | G3P acyltransferase | PfGAT, PfPIsB | PFL0620c, PF13_0010 |
| HK | Hexokinase | PfHK | PFF1155w |
| IMP | Inositol monophosphatase | PfIMP | PF07_0024 |
| IPS | Inisitol-3-phosphate synthase | PfIPS | PFE0585c |
| PAP | PA phosphatase | PfPAP2, PfPAP2a | PFF1210w, MAL8P1.202 |
| PEMT | PE *N*-methyltransferase | | NI-BA |
| PLMT | PL *N*-methyltransferase | | NI-BA |
| PGPPase | PGP phosphatase | | NI |
| PGPS | PGP synthase | | NI (PfPSS?) |
| PIS | PI synthase | PfPIS | MAL13P1.82 |
| PI3K | PI3-kinase | PfPI3K | PFE0795w |
| PI4K | PI4-kinase | | PFE0485w |
| PI4P5K | PI4-phosphate 5-kinase | PfPIP5K/NCS | PFA0515w |
| PMT | Phosphoethanolamine *N*-methyltransferase | PfPMT | MAL13P1.214 |
| PSD | PS decarboxylase | PfPSD | PFI1370c |
| PSS | PS synthase CDP-DAG-dependent | PfPSS | MAL8P1.58 |
| PSSE | PS synthase via exchange | PfPSSE | |
| SCD | Stearoyl-CoA Δ9-desaturase | PfSCD | PFE0555w |
| SD | Serine decarboxylase | | NI-BA |

[1] PlasmoDB gene identity
NI = not identified
BA = Biochemical activity reported
Regions shaded in grey = established gene-enzyme relationship

A second acylation step, catalyzed by 1-acyl-G3P- transferase (AGPAT) at the *sn*-2 position of the glycerol backbone, produces PA, an important structural and regulatory metabolite that serves as a substrate for (i) phosphatidic acid phosphatase (PAP) to produce DAG and TAGs or (ii) for CDP-DAG synthase to produce CDP-DAG.

## *The CDP-DAG dependent pathways*

In prokaryotes, CDP-DAG is the major precursor for PG, CL, PS (weakly represented) and PE, i.e. the entire GPL pool. In eukaryotes, PA is partitioned between CDP-DAG and DAG. The CDP-DAG dependent pathway conducts synthesis of PG, CL, PI as well as PS when required. DAG is a precursor for the *de novo* synthesis of PC and PE from CDP-choline and CDP-ethanolamine respectively (the Kennedy pathways). PS decarboxylation to form PE occurs in all eukaryotes, whereas methylation of PE into PC is absent in some eukaryotes (Kent, 1995, Dowhan, 1997).

*Plasmodium* parasites synthesize PLs similar to eukaryotes, although they have retained the prokaryote system of synthesizing PS from CDP-DAG (Vial and Ancelin, 1998) (Figure 7). In *P. falciparum* and *P. knowlesi*, serine is incorporated into PS, PE and to a lesser extent, PC. In rodent parasites, however, serine is not incorporated into PC (Dechamps *et al.*, 2010a), which highlights a difference in lipid metabolism between the malaria parasites. Although PS is present at relatively low levels, synthesis is increased 2-fold upon *P. falciparum* infection (Figure 6) and it is a key metabolite in the synthesis of PE (Elabbadi *et al.*, 1997).

PI is also synthesized via the CDP-DAG dependent pathway and is increased 14-fold, the highest relative PL increase, following parasite infection (Figure 6). PI synthase (PIS) activity, for the synthesis of PI from CDP-DAG and inositol, has been observed in the mature blood stage *P. falciparum* parasites (Elabbadi *et al.*, 1994), at the time of active membrane biogenesis. The PI inositol head group can be phosphorylated by specific PI3- and PI4-kinases to produce phosphoinositides

(Figure 7) that play key roles in a several cellular processes such as trafficking of haemoglobin to the food vacuole of *P. falciparum* parasites (Vaid *et al.*, 2010) and the signal events involved in the activation of *P. berghei* gametocytes (Raabe *et al.*, 2011). PG and CL are barely detected in pRBCs.

In *P. falciparum*, DAG enters the Kennedy pathways for synthesis of PC and PE, and also serves as a substrate for the synthesis of TAGs via DAG acyltransferases (Vielemeyer *et al.*, 2004). TAGs are usually contained in lipid droplets within the cell. They are present throughout the intra-erythrocytic stage of parasite development, though enriched during trophozoite formation. Their specific role is unclear; however, TAGs are thought to provide acyl groups for lipid synthesis when resources are limited (Palacpac *et al.*, 2004).

## *The CDP-Choline and CDP-Ethanolamine pathways*

The *de novo* Kennedy pathways involve two distinct kinases, CTP-cytidylyltransferases and a common choline/ethanolamine phosphotransferase (CEPT) to produce PC and PE, the PL species that constitute 75-95% of the parasite membrane (Figure 7). These enzymes are expressed throughout mature blood-stage parasite development (Alberge *et al.*, 2010, Aurrecoechea *et al.*, 2009, Dechamps *et al.*, 2010b), during active membrane biogenesis.

Choline enters the red blood cell via the erythrocytic choline carrier and the NPP of the parasite via an efficient organic-cation transporter (OCT) (Lehane *et al.*, 2004). The parasite's choline kinase (CK) phosphorylates choline to produce phosphocholine (P-Cho), the first commitment step in the CDP-cho pathway. The P-Cho intermediate can also be obtained from sphingomyelin via phospholipase C and in some *Plasmodium* parasites, from the direct methylation of phosphoethanolamine (P-Etn) via phosphoethanolamine methyltransferase (PMT) (Dechamps *et al.*, 2010a, Pessi *et al.*, 2004). P-Cho is coupled to CTP by a choline-phosphate cytidylyltransferase (CCT) and further converted to PC by CEPT in the presence of DAG.

A similar *de novo* pathway exists for the synthesis of PE. Ethanolamine enters the parasite via passive diffusion, is phosphorylated by ethanolamine kinase (EK) to produce P-Etn, coupled to CTP by ethanolamine-phosphate cytidylyltransferase (ECT) and converted to PE by CEPT in the presence of DAG. Metabolic studies using radiolabelled ethanolamine showed that PE can also be methylated to form PC in *P. falciparum* and *P. knowlesi* parasites (Elabbadi *et al.*, 1997, Dechamps *et al.*, 2010a). The reaction is possibly catalyzed via PE methyltransferase (PEMT), although *Plasmodium* PEMT genes have not yet been identified. In some *Plasmodium* parasites, ethanolamine can also be generated from the SDPM pathway.

## *The SDPM pathway*

Serine and ethanolamine radiolabelling studies showed that *P. knowlesi* and *P. falciparum* possess an alternative pathway for the synthesis of PC from non-choline precursors (Elabbadi *et al.*, 1997, Pessi *et al.*, 2004). Ethanolamine can be generated from serine, either transported from the host or via haemoglobin degradation. Serine is decarboxylated via serine decarboxylase (SD) in *P. falciparum* and *P. knowlesi* to form ethanolamine which can then enter the Kennedy pathways (Elabbadi *et al.*, 1997) (Figure 7). The SD gene is widely represented in plants and algae, however SD genes have not yet been identified (Dechamps *et al.*, 2010a). This is possibly due to low sequence homologies of *Plasmodium* SDs to known orthologues. *P. falciparum*, *P. vivax* and *P. knowlesi* parasites can convert P-Etn to P-Cho via PMT to ultimatedly produce PC (Dechamps *et al.*, 2010a). The SDPM pathway is present in plants, which lack phosphatidylserine decarboxylases (PSD) to form PE from PS. This pathway is absent in mammals (Rontein *et al.*, 2001).

## 1.4 Targeting *Plasmodium* phospholipid synthesis

### 1.4.1 The Kennedy pathways

Enzymes involved in the Kennedy pathways (*Pb*CK, *Pb*CCT, *Pb*ECT and *Pb*CEPT) are refractory to genetic disruption in *P. berghei* parasites (Dechamps *et al.*, 2010c). This implies that the pathway may be essential in intra-erythrocytic development of the rodent parasite. However, their significance in *P. falciparum* parasites needs to be explored, where differences such as the connection between the CDP-Cho and CDP-Etn pathways via PMT activity exist. *In vitro* analysis showed that *Pf*CK and *Pf*EK enzymes were selectively inhibited by the choline analogue, hemicholinium-3, and the ethanolamine analogue, 2-amino-1-butanol, respectively, however at non-physiological concentrations with $IC_{50s}$ in the millimolar range (Alberge *et al.*, 2010).

### 1.4.2 The SDPM pathway

PMT orthologues have been identified in *P. falciparum*, *P. vivax* and *P. knowlesi* parasites and the absence of PMT activity in mammalian cells has made it an attractive drug target candidate (Mamoun *et al.*, 2010). Wild type and transgenic *P. falciparum* analysis showed that *Pf*PMT was negatively regulated by choline in a dose-dependent manner, and by P-Cho. The finding that PMT was inhibited by its product suggested that PMT may be inhibited by P-Cho analogs; indeed recombinant *Pf*PMT activity was inhibited by 50% in the presence of 50μM hexadecylphosphocholine and *P. falciparum* parasite proliferation was reduced at an $IC_{50} \approx 80\mu M$ *in vitro* (Pessi *et al.*, 2004). *Pf*PMT knockout parasites displayed altered DNA replication, reduced multiplication rates, delayed parasite growth and increased cell death. In addition, *Pf*PMT knockout parasites abolished the incorporation of ethanolamine into PC, which suggests that *Pf*PMT may also have PEMT activity since PE was not methylated into PC. The gene though, is not essential for development due to the availability of choline. However, a 10-fold

increase in the physiological concentration of choline did not complement parasite growth in *Pf*PMT knockout parasites, suggesting that the PMT's function is not redundant (Witola *et al.*, 2008).

### 1.4.3 PC synthesis

Compounds that mimic the structure of membrane precursors such as choline, ethanolamine, serine and FAs are able to inhibit parasite proliferation with $IC_{50}$ values in the low micromolar range. Biochemical analysis showed that these analogs are transported to the pRBC and are incorporated into parasite membranes; however, there they dramatically alter the lipid composition and physio-chemical properties of the membrane, leading to parasite death. The most promising of these anti-malarial compounds are those that block/interfere with choline transport to the parasite (Vial and Ancelin, 1992).

Rationally designed ammonium compounds, analogous to choline, have demonstrated potent anti-malarial activity with $IC_{50}$ values in the low nanomolar range (Ancelin *et al.*, 2003). The bis-quaternary ammonium compound, G25 [1.16-hexadecamethylenebis(*N*-methylpyrrolidinium) dibromide], was effective against both sensitive and drug resistant *P. falciparum* parasites with $IC_{50}$ values between 1-5.3nM and was 1000-fold less toxic in mammalian cells. G25 was able to clear infection in *P. falciparum*- and *P. cynomolgi*-infected *Aotus* monkeys with no recrudescence after 60 days. G25 also exhibited a cross-species specificity and was able to clear *P. vivax*-infected rhesus monkeys with $IC_{50}$ values and 50% effective doses comparable to that of *P. falciparum* (Wengelnik *et al.*, 2002). An interesting feature of the choline analogue was its ability to accumulate several hundred-fold, and into the millimolar range, into pRBCs and not uninfected RBCs. This accumulation also occurred in the parasite's food vacuole, where the bis-quaternary ammonium compounds interacted with haem, interfering with haemozoin formation and enhancing their anti-malarial effect (Biagini *et al.*, 2003). This probably contributes to the compound's potency, specificity and low toxicity. Biochemical analysis showed that the mode of action of the choline

analogs is at the level of choline uptake in the *de novo* PC biosynthetic pathway (Biagini *et al.*, 2004) and is specific to the trophozoite stage of intra-erythrocytic development.

Choline uptake alone, however, cannot account for the anti-malarial activity of the choline analogs since choline is not essential for parasite development. Transcriptional analysis of parasites treated with the T4 choline analogue showed cell cycle arrest and general induction of genes involved in gametocytogenesis, with no apparent changes in genes involved in PC biosynthesis, while proteomic analysis displayed a decrease in *Pf*CEPT (Le Roch *et al.*, 2008). The possibility of choline analogs having multiple targets, however, is an advantage. Currently, one bis-thiazolium salt, T3/SAR97276 is under Phase II clinical trials for parenteral cure of severe malaria (Alberge *et al.*, 2010).

### 1.4.4 Fatty acid synthesis

FAs are the building units of GPLs and their distribution at the *sn*-1 and *sn*-2 positions of the G3P backbone determines the physical properties and function of the lipid bilayer. Radiolabelling studies have shown that the polar head groups of PLs, as well as FAs, are taken up from the host serum (Krishnegowda and Gowda, 2003, Vial and Ancelin, 1992). Parasites though, do have the capacity for *de novo* type II Fatty Acid synthesis (FAS II) (Gardner *et al.*, 2002). FA chain extension is catalyzed by four key enzymes: FabB/F (β-ketoacyl-ACP-synthase I/II), FabG (β-ketoacyl-ACP-reductase), FabI (enoyl-ACP-reductase) and FabZ (β-hydroxyacyl-ACP-dehydrase), each attached to an acyl carrier protein (ACP). The enzymes are of bacterial origin and the pathway is different to the FAS I pathway in humans, making these *Plasmodium* enzymes attractive drug target candidates.

The upregulation of FAS II transcripts and enzymes in the liver-stage of *P. yoelli* parasite development suggested that FAS II may be important at that stage of proliferation (Tarun *et al.*, 2008). Indeed, targeted disruption of the FabB/F and FabZ enzymes in *P. yoelli* parasites demonstrated their essential nature in liver-

stage proliferation. Intra-erythrocytic, mosquito-stage development and the ability to invade hepatocytes in FabB/F and FabZ knockout parasites were unaffected, however, knockout parasites failed to form exo-erythrocytic parasites (Vaughan *et al.*, 2009). Similar results were seen in FabI knockout *P. berghei* parasites (Yu *et al.*, 2008). FAS II enzymes are highly conserved among *Plasmodium* parasites and FabI knockout *P. falciparum* parasites showed no reduction in *in vitro* blood-stage development, possibly further supporting its critical role in liver-stage development (Vaughan *et al.*, 2009). *In vivo* analysis, however, suggested that FA synthesis may be required in intra-erythrocytic stages, during starvation (Daily *et al.*, 2007). Several *P. falciparum* FAS II inhibitors have been reported, with $IC_{50}$ values in the low micromolar range (Lee *et al.*, 2009, Lu *et al.*, 2005). However, the search for inhibitors is typically focused on the blood-stages and with the inhibitor selection process, where potent inhibitors are further evaluated for activity against cultured *P. falciparum* parasites, a pool of compounds capable of inhibiting parasites *in vivo*, during liver-stage development, may have been lost.

### 1.4.5 Sphingomyelin synthesis

SM is the most abundant of the sphingolipids (SLs), another class of membrane components, in mammals and is widely distributed in eukaryotes. Almost all metabolites involved in SL biosynthesis or degradation have important roles in cell growth, differentiation, apoptosis and cell signaling (Hannun and Obeid, 2008). The pathway is poorly understood in *Plasmodium*, however it is thought that the breakdown of host SM into ceramide may be used to (i) modulate cell cycle progression of intra-erythrocytic parasites or (ii) supply parasite sphingomyelinase (SMase) a substrate for synthesis of parasite-derived SM (Hanada *et al.*, 2000). Evidence for the latter includes an up to 47% degradation of the host SM (Maguire and Sherman, 1990) and the upregulation of *Plasmodium* SMase transcripts during trophozoite and erythrocytic schizonts stages (Lauer *et al.*, 2001). Inhibitor-based studies showed that *Plasmodium* parasites lack a *de novo* ceramide biosynthesis pathway (Lauer *et al.*, 2001, Gerold and Schwarz, 2001). However, *P. falciparum* trophozoite maturation into schizonts was

impaired and proliferation blocked when treated with the mammalian SMase inhibitor, scyphostatin ($IC_{50} \approx 3\mu M$). PPMP and PDMP sphingolipid analogues also blocked parasite proliferation ($IC_{50} \approx 6\mu M$), suggesting that drugs targeting *Plasmodium* SM biosynthesis may be explored as possible lead compounds (Hanada *et al.*, 2002).

## 1.5 Glycerol metabolism

Glycerol is known to play fundamental roles in several physiological processes. Glucose-deprived bacteria, for example, have a measured growth rate in glycerol of approximately 74% that achieved on glucose (Holms, 1996) and *S. cerevisiae* can use glycerol as the sole carbon source under aerobic conditions (Nevoigt and Stahl, 1997).

All PLs have a glycerol backbone and GPL biosynthesis is initiated when two FA molecules are transferred from a donor to G3P, which functions at the interface of lipid and carbohydrate metabolism (Lin, 1977). In *Plasmodium*, G3P is thought to be produced primarily from the glycolytic intermediate, dihydroxyacetone phosphate (DHAP) (Dechamps *et al.*, 2010b), since glucose consumption is dramatically increased in pRBCs. Glycolysis then is thought to supply the parasite's energy as well as PL needs. However, metabolic labelling studies in *P. knowlesi* showed that glycerol is also incorporated into PLs (Rock, 1971).

Glycerol kinase (GK, ATP: phosphotransferase, EC 2.7.1.30) is a highly conserved enzyme that catalyzes the ATP-dependant phosphorylation of glycerol to G3P (Zwaig *et al.*, 1970). A *P. falciparum* GK (*Pf*GK) orthologue has been annotated in the *Plasmodium* database (PlasmoDB gene ID PF13_0269) and its proposed role in glycerolipid synthesis and glycolysis is summarized in Figure 8.

### 1.5.1 *P. falciparum* glycerol kinase

DNA sequence analysis available in PlasmoDB shows that the *Pf*GK gene is a single 1,506bp exon (Appendix A1) located between nucleotide positions 2,051,547-2,053,052 on chromosome 13 (Gene ID: PF13_0269).

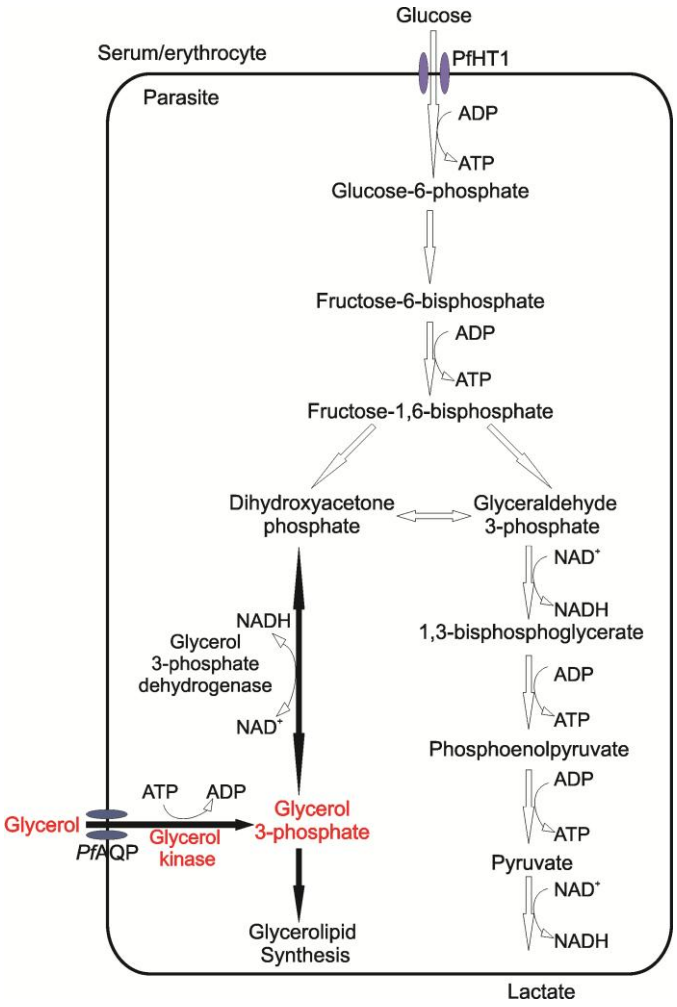



**Figure 8: The proposed role of *P. falciparum* glycerol kinase (*Pf*GK).**
Glycerol enters the parasite via the *P. falciparum* aquaglyceroporin facilitator (*Pf*AQP). Glycerol kinase (*Pf*GK) then catalyzes an ATP-dependant phosphorylation of glycerol to glycerol-3-phosphate, a precursor for glycerolipid synthesis and glycolysis. Image modified from Hansen *et al* (2002).

The A+T composition of the *Pf*GK gene is 70%, lower than both the average 80.6% A+T content of the *P. falciparum* genome and 76.3% of the average exon (Gardner *et al.*, 2002).

Protein family domain analysis via the Pfam database (http://pfam.sanger.ac.uk/) shows that the 501 amino acid *Pf*GK protein consists of two structurally related N- (amino acids 3-253) and C- (amino acids 256-461) terminal domains belonging to the FGGY family of carbohydrate kinases (Figure 9).

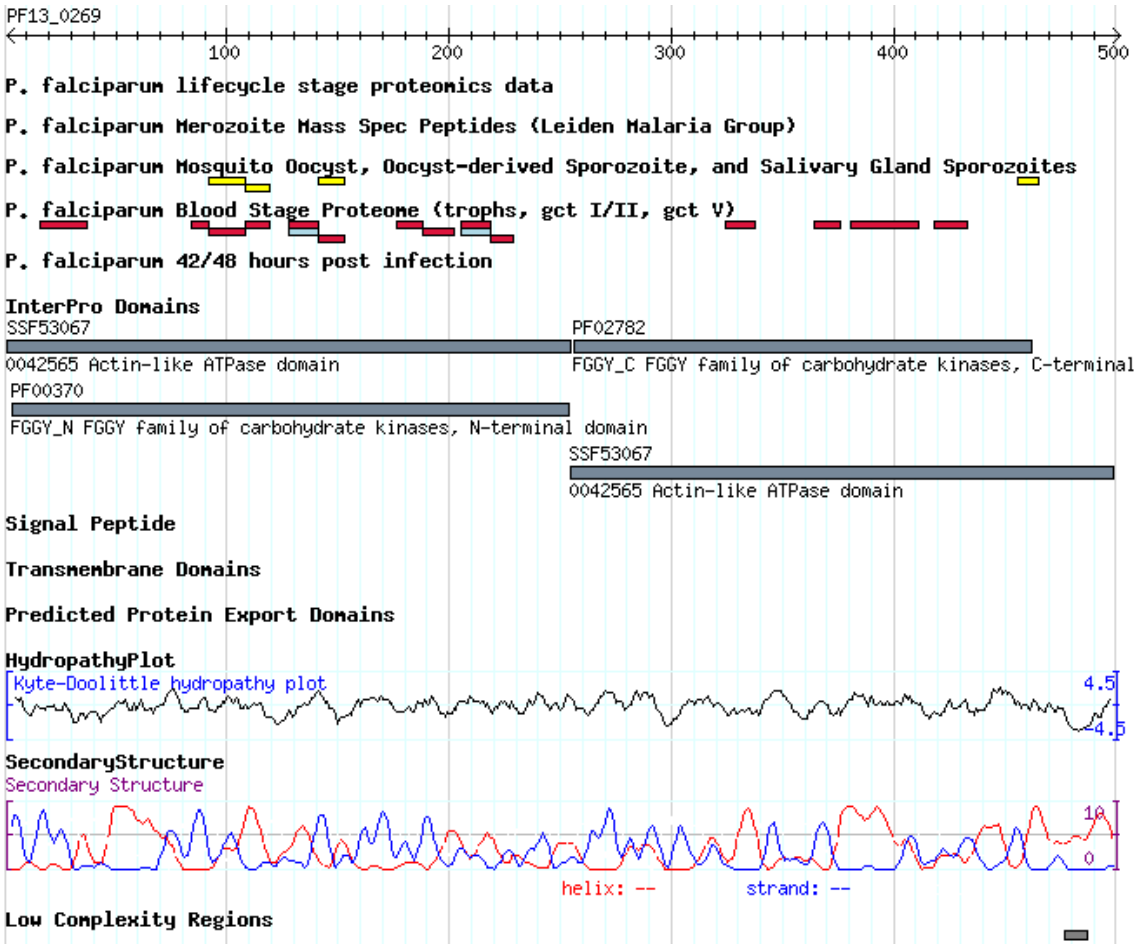



**Figure 9: Primary structural analysis of the *Pf*GK gene.**
Analysis of the *Pf*GK primary structure predicts that the *Pf*GK protein has structurally related Actin-like ATPase and FGGY domains at both the N- and C-terminus and a 10 residue low complexity region between amino acids 477-486. Eighteen mass-spectrometry peptides were identified, two in early gametocytes (light blue), 14 in late gametocytes (red) and four in mosquito salivary gland sporozoites (yellow). The *Pf*GK PlasmoDB gene ID (PF13_0269) and the number of amino acids, in increments of 100, are indicated (Aurrecoechea *et al.*, 2009).

The FGGY domain adopts a ribonuclease H-like fold. *Pf*GK is also member of the ATPase superfamily and shares a common ßßßαßαßα tertiary fold motif with actin, hexokinase and heat shock protein-70 (Hsp70) (Kabsch and Holmes, 1995, Hurley,

1996). Although these proteins are dissimilar in amino acid sequence and function, the ATP-binding site and catalytic mechanism are highly conserved within this actin-like ATPase domain (Figure 10).

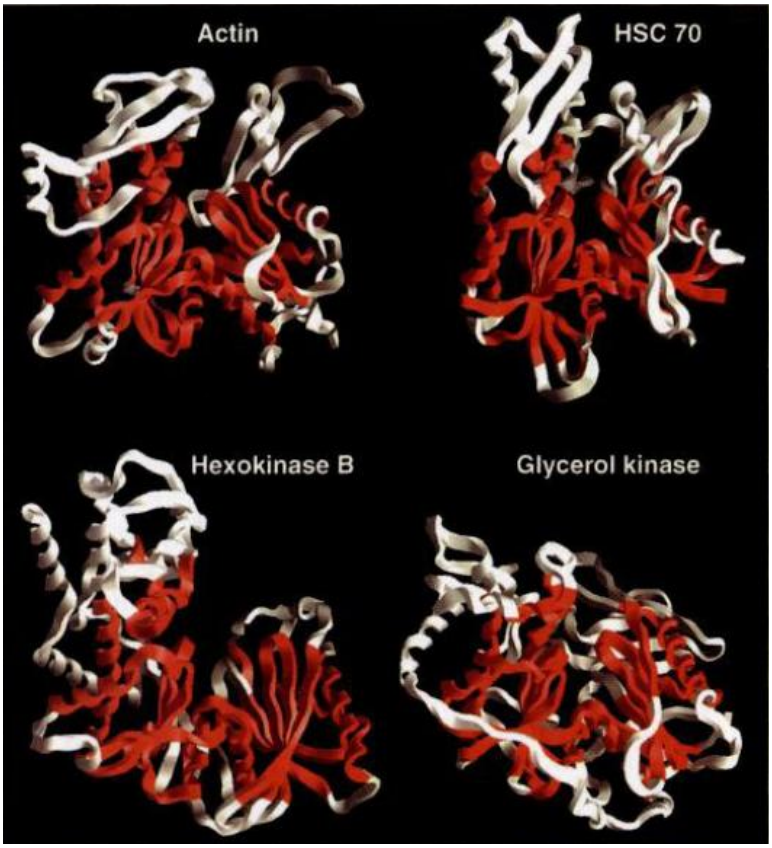



**Figure 10: The actin ATPase domain.**
The architecture of the actin ATPase core domain (red) in unrelated proteins (actin, heat-shock protein-70, hexokinase and glycerol kinase) that interacts with ATP and $Mg^{2+}$ or $Ca^{2+}$ (Kabsch and Holmes, 1995).

Mass-spectrometry analysis (Figure 9) identified two peptides in early gametocytes, 14 peptides in late gametocytes (Silvestrini *et al.*, 2010, Aurrecoechea *et al.*, 2009) and four peptides in the mosquito salivary gland sporozoites (Lasonder *et al.*, 2008, Aurrecoechea *et al.*, 2009).

A low complexity region (LCR) is located between amino acids 477-486 (KMDDKKRKKK; Figure 9). LCRs have no definite structure and their function is unclear. LCRs are (i) thought to be highly evolving regions that expanded from

DNA replication slippage and unequal crossover recombination events (DePristo *et al.*, 2006), (ii) characteristically composed of a limited set of amino acids, often repeats of one or more amino acids, notably lysine and asparagine (Pizzi and Frontali, 2001), (iii) largely depend on genomic A+T content (Xue and Forsdyke, 2003, DePristo *et al.*, 2006), (iv) exist as surface-exposed and predominantly hydrophilic disordered coils (Wootton, 1994, Aravind *et al.*, 2003) and (v) may not interfere with the functioning of the molecule (DePristo *et al.*, 2006).

No protein export or transmembrane domains were identified.

### 1.5.2 Structure of *E. coli* glycerol kinase

The *E. coli* glycerol kinase (*Ec*GK) enzyme has been well characterized. Under physiological conditions, *Ec*GK exists in an equilibrium of functional dimers and inactive tetramers (Ormo *et al.*, 1998), each composed of identical monomers, designated O, X, Y and Z (Figure 11a).

Each monomer consists of two major domains (I and II) separated by a deep active site cleft. *Ec*GK contains six subdomains, each unique and defined in the enzyme's structure and function. Subdomains IA and IIA form the ATPase core and subdomain IB forms the glycerol-binding site. The dimer (O-Y) interface is formed via interactions between subdomains IIB and IIC whereas the tetramer (O-X) interface is formed via subdomains IA and IC.

*Ec*GK is independently and non-competitively inhibited by two allosteric effectors: (i) $IIA^{GLC}$, a component from the phosphoenolpyruvate:glycose phosphotransferase system (PTS) and (ii) the glycolytic intermediate fructose-1,6-bisphosphate (FBP) (Ormo *et al.*, 1998). The orthophosphate (or FBP)-binding site is a glycine-rich region between residues 233-236 and is formed by subdomain IC.

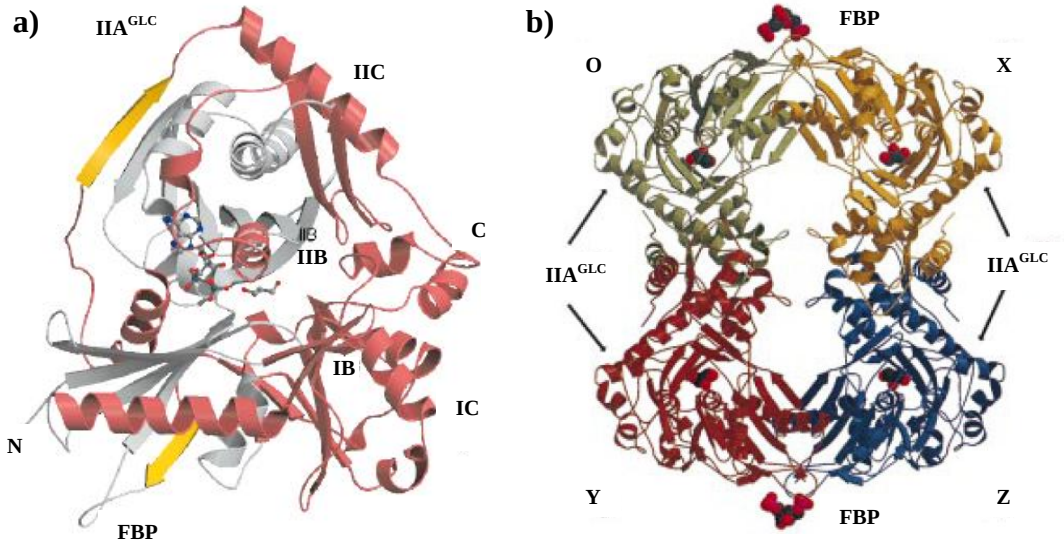



**Figure 11: The *E. coli* glycerol kinase structure.**
Illustration of the **(a)** *Ec*GK monomer. The ATPase core subdomains (IA and IIA) are shown in grey; IB, IIB, IC and IIC subdomains, which are part of the ATPase core domain, are labelled and shown in red. The extra β strands that contribute to the ATPase core β sheets in subdomains IC and IIC are shown in yellow. The orthophosphate (FBP) and $IIA^{GLC}$ binding regions are indicated. Bound ADP and glycerol are shown as stick and ball structures, **(b)** *Ec*GK tetramer in complex with $IIA^{GLC}$ and with FBP. The structure of $IIA^{GLC}$, which binds to subdomain IIC, was omitted for clarity. Each monomer (O, X, Y, Z) is designated with a different colour (Ormo *et al.*, 1998, Feese *et al.*, 1998).

The phosphocarrier protein $IIA^{GLC}$-binding site (residues 472-481) is part of subdomain IIC (Hurley *et al.*, 1993). $IIA^{GLC}$ binding changes the conformation of the *Ec*GK tetramer, however, the catalytic site is unaffected since the inhibitor binds 30Å from and distal to the catalytic domain. The mechanism of $IIA^{GLC}$ inhibition is unclear. The FBP binding site is created when *Ec*GK dimers interact to form a tetramer (Figure 11b). FBP binds to and stabilizes the tetramer, rendering it inactive (Ormo *et al.*, 1998).

## 1.5.3 Structure of *Pf*GK

High resolution structures of *Pf*GK, the first of a eukaryotic GK, have recently been resolved (Schnick *et al.*, 2009). The enzyme structure, in complex with its substrates, glycerol and ADP (*Pf*GK-glycerol and *Pf*GK-ADP), was solved by

molecular replacement using *Ec*GK as a search model. *Pf*GK is topologically identical to the *Ec*GK orthologue and shares common catalytic residues. The 501 amino acid polypeptide consists of two domains separated by a deep cleft and as predicted, each domain is constructed around the ßßßαßαßα ATPase core structure.

*Pf*GK exists as a dimer (Figure 12) with the majority of an otherwise 1750Å accessible area buried by dimer-forming segments in domain II. The orientation and features of the loop structure at the dimer-dimer interface implies that FBP may not be an inhibitor of *Pf*GK. The loop is orientated towards the catalytic core and lacks the glycine-rich $^{229}$GGKGGTR$^{236}$ signature sequence. Similarly, the structure and sequence of the IIA$^{GLC}$ binding site in *Ec*GK is significantly different in *Pf*GK and suggests that IIA$^{GLC}$ may also not be an allosteric inhibitor of *Pf*GK. An interesting observation is that the domain opening is 27º, much larger than the 7º of orthologous proteins. This may be exploited for selective inhibitor design should the enzyme prove to be important for parasite development (Schnick *et al.*, 2009).

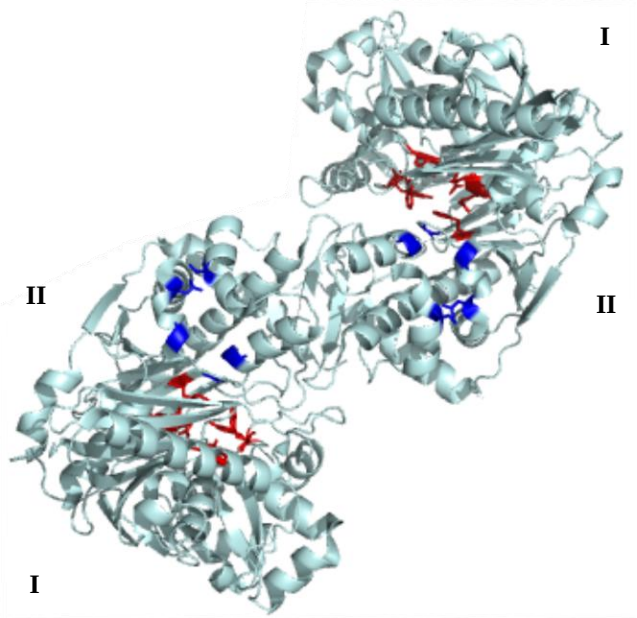



**Figure 12: The *P. falciparum* glycerol kinase structure.**
A ribbon representation of the crystallized *Pf*GK protein. *Pf*GK exits as functional dimers. Each monomer contains 2 domains separated by a catalytic cleft. Glycerol-binding residues are present in domain I and are highlighted red. ADP-binding residues are located in domain II and are highlighted in blue. The *Pf*GK monomer is topologically identical to the *Ec*GK monomer (Schnick *et al.*, 2009).

### 1.5.4 Glycerol acquisition in *Plasmodium* parasites

Aquaporins (AQPs) are a family of small membrane spanning water and solute channels. They are divided into two groups: (i) the orthodox aquaporins, which are water-specific channels and (ii) the aquaglyceroporins, which are highly permeable to glycerol and other physiological solutes such as urea, and are variably permeable to water (Agre *et al.*, 2002). Orthodox AQPs are thought to maintain homeostasis in the cell whereas aquaglyceroporins are involved in nutrient uptake during metabolism.

Erythrocytes are known to be highly permeable to glycerol and in humans, AQP3 is the major glycerol facilitator. Red blood cells cannot metabolize glycerol and the physiological significance of glycerol in the cell is unclear (Lin, 1977), but it may play a role in osmoregulation (Brisson *et al.*, 2001). In *P. falciparum*-infected RBCs, AQP3 is recruited to the PVM, during or soon after invasion, in an oligomeric state (possibly dimerized AQP3) that is absent in uninfected red cells (Bietz *et al.*, 2009). Its role in the PVM still has to be evaluated, however, the internalized AQP3 may facilitate nutrient transport (glycerol) from the RBC cytoplasm to the parasite. *P. falciparum* contains a single bi-functional aquaglyceroporin (*Pf*AQP; PlasmoDB ID PF11_0338) with permeability to both water and glycerol. It is 50.4% similar to the *E. coli* glycerol facilitator (Hansen *et al.*, 2002) (*Ec*AQPZ; UniProtKB ID P60844), 47.9% similar to the human orthologue (*Hs*AQP3; UniProtKB ID Q92482) (Bietz *et al.*, 2009) and is expressed on the parasite membrane throughout blood-stage development (Beitz *et al.*, 2004).

## 1.6 Aims and objectives

Glycerol kinase is a key enzyme that functions at the interface of carbohydrate and lipid metabolism. The aim of this study is to evaluate the role of the enzyme during intra-erythrocytic blood stage development of *P. falciparum* parasites. More specifically, the objectives are to:

- Perform bioinformatic and phylogenetic analysis of the *Pf*GK gene.
- Perform molecular modelling and structural analysis of the *Pf*GK protein.
- Clone, express and purify recombinant *Pf*GK.
- Examine the catalytic and biochemical properties of the recombinant enzyme and its behaviour in the presence of possible inhibitors.
- Generate a glycerol kinase knockout parasite line via double crossover homologous recombination and evaluate its growth during *in vitro* blood stage development.

Results from these studies will enhance our current knowledge of the basic biology of the malaria parasite.

# Chapter 2: Materials and Methods

## 2.1 Bioinformatic analysis of *P. falciparum* glycerol kinase

### 2.1.1 *P. falciparum* glycerol kinase sequence and structure analysis

A putative *Pf*GK gene was annotated during the *Plasmodium* genome project (Gardner *et al.*, 2002). *Pf*GK nucleotide and amino acid sequences (Appendix A1), as well as transcriptional profiles were obtained from PlasmoDB (www.plasmodb.org) (Bahl *et al.*, 2002, Aurrecoechea *et al.*, 2009) versions 4.4-7.2. Primary sequence analysis was performed using algorithms available on the European Bioinformatics Institute (EBI; http://www.ebi.ac.uk/) and ExPASy http://www.expasy.org/) websites. Tertiary structure analysis was performed using the putative human GK (*Hs*GK) protein structure, based on the characterized *Ec*GK crystal structure, and the recently crystallized *Pf*GK 3D structures (Protein Data Bank, PDB, identification numbers 2w40 and 2w41; (Schnick *et al.*, 2009), available at the USCS Genome Bioinformatics (http://genome.ucsc.edu/) and PDB (http://www.rcsb.org/) databases respectively.

### 2.1.2 Evolutionary analysis of *Pf*GK

The evolutionary relationships between GK proteins were analyzed as previously described (Durand *et al.*, 2008), with slight modifications. Briefly, 27 orthologous GK protein and DNA coding sequences, with laboratory or bioinformatic evidence, were retrieved from the UniprotKB database (http://www.uniprot.org/uniprot/; Appendix A2). Protein multiple sequence alignments (MSAs) were performed using MUSCLE (multiple sequence comparison by log expectation) (Edgar, 2004), with default parameters. Insertions and deletions in the alignment were manually removed using BioEdit, version 7.0.9.0 (Hall, 1999) and phylogenetic analyses were performed using the PAUP* version 4.0b10 software (Swofford, 2002). Parsimony-informative amino acids

were used to generate an unweighted parsimony tree. Parsimony analyses were performed by heuristic search with the tree bisection-reconnection (TBR) branch swapping algorithm to generate a single most-parsimonious tree (MPT). The phylogenetic signal (g1), consistency (CI) and retention indices (RI) were used to assess the fit-of-character data to phylogenetic trees. The g1 statistic differentiates random noise from information regarding the evolutionary history of the taxa. Data sets with a phylogenetic signal produce tree length distributions that are skewed, whereas data sets with a high degree of random noise produce distributions that are closer to symmetrical (Hillis and Huelsenbeck, 1992). The CI is the ratio of the minimum possible number of steps for a data set in a tree to the actual number of steps achieved and is simply a measure of homoplasy (traits that were derived independently and not present in their common ancestor) in a data set. For example, if there are 25 characters, the minimum possible number of steps is 25; i.e. each character has to change at least once. No tree, however, will allow for the absolute minimum as some characters will change more than once by way of convergence or reversals. CI values closer to 1.0 therefore suggest no homoplasy in the tree. CI does, however, increase with addition of uninformative data. For example, if the shortest tree for a 25 informative character data set is 50 steps, then the CI = 25 ÷ 50 = 0.5. An addition of 15 uninformative characters will generate the same tree, however, the CI changes to (25 + 15) ÷ 50 = 0.8, suggesting a better fit-of-character data to a tree. The rescaled CI, which takes RI into consideration, was used to validate the tree (Farris, 1989). The RI is the ratio of the number of extra-steps achieved to the maximum number of extra steps possible for a particular character on a given tree, in other words, a measure of the proportion of similarities on a tree. RI is sensitive to both the number of informative characters and the number of taxa and is insensitive to addition of autapomorphies (a distinctly derived trait present only in one member of a clade) in the data. RI values closer to 1.0 reflect the degree to which similarities apparent in the data can be retained as homologies on a tree. Maximum parsimony analyses were performed with 1,000 bootstrap (Felsenstein, 1985) replicates. TREEVIEW version 1.6.6 (Page, 1996) was used to view the trees. *Ec*GK was used as the outgroup (Appendix A2).

### 2.1.3 Identification of potential *Pf*GK drug target sites

A maximum likelihood substitution matrix was used to estimate the ratio of non-synonymous to synonymous substitutions at individual codons in the GK coding sequence (Durand *et al.*, 2008). PAML (Yang, 1997) and SLR (Massingham and Goldman, 2005) analysis on protozoa, Apicomplexa, metazoan and vertebrate groups, from phylogenetic analysis of 28 orthologous GK nucleotide and amino acid sequences, was used to determine the evolutionary rates at individual codons within the GK gene (Durand *et al.*, 2008). Using *Pf*GK and *Hs*GK as representatives of the protozoan and metazoan clades respectively, it was possible to identify evolutionary conserved codons that are different to the human enzyme and may represent potential drug target sites (DTS) in the *Pf*GK orthologue. A pairwise alignment of the two protein sequences was performed with the Needleman and Wunsch algorithm (Needleman and Wunsch, 1970). The coordinates for evolutionary conserved residues different to the human orthologue were mapped on *Pf*GK's tertiary structure and visualized using Pymol (http://www.delsci.com/rel/099/).

## 2.2 Parasite culture

The 3D7 *P. falciparum* parasite strain was cultured as described by Trager and Jensen (1976), with slight modifications. All relevant work was performed in a Labotec® Bio Flow laminar hood (Labotec, Midrand, South Africa) and sterility was maintained by aseptic techniques. Cultures were maintained at 37°C in a Nuaire™ IR Autoflow water-jacketed incubator (Nuaire™, Plymouth, USA). Spent medium was aspirated using a Savant VP100 2 Stage vacuum pump (Savant Instruments, Inc., England). Centrifuge tubes, Pasteur pipettes, pipette tips and phosphate buffered saline (PBS; Appendix A3) were sterilized using the Speedy Autoclave (Merck Chemicals (Pty) Ltd., Johannesburg, South Africa) at 120°C, 1kg/cm$^2$. Culture medium and reagents were sterilized by filtration through 0.22µm Millex®-GP Syringe Driven Filter Units (Millipore Corporation, Bedford, USA). Milli-Q water was collected from the Direct-Q® 3 UV system (Millipore

SAS, Molsheim, France) and used for the preparation of all reagents and culture medium.

### 2.2.1 Preparation of parasite cultures from frozen stocks

Cryotubes (Nunc™, Roskilde, Denmark) containing 1ml frozen 3D7 *P. falciparum* parasite cultures were removed from liquid nitrogen storage and thawed in a preheated Lauda B/MA6 water bath (Lauda, Lauda-Königshofen, Germany) at 37°C. One-tenth volume (100µl) 12% (w/v) NaCl (Saarchem (Pty) Ltd., Wadeville, South Africa) was added drop wise to the tubes, whilst swirling, and incubated for five minutes at room temperature. Suspensions were transferred to sterile 15ml centrifuge tubes (Greiner Bio-One GmbH, Baden-Württemberg, Germany) and 10 volumes (10ml) 1.6% NaCl were added. The tubes were gently inverted and centrifuged in an Eppendorf 5702R centrifuge (Eppendorf AG, Hamburg, Germany) at 500g for five minutes at room temperature. The supernatant was aspirated and the parasite pellets resuspended in 10 volumes (10ml) incomplete RPMI culture medium supplemented with 20% AB plasma or 1% Albumax II (Appendix A3). The tubes were centrifuged at 500g for five minutes at room temperature and freshly washed 50% erythrocyte suspension (Appendix A3) was added to the cell pellet to a final volume of 750µl. Incomplete RPMI culture medium supplemented with 20% AB plasma or 1% Albumax II (Appendix A3) was pre-warmed at 37°C and 4.25ml was added drop wise to the cells. The culture was transferred to a sterile 25cm$^2$ small tissue culture flask (Nunc™, Roskilde, Denmark), gassed for 30 seconds with a 93% $N_2$, 5% $CO_2$ and 2% $O_2$ gas mixture (African Oxygen Ltd., Johannesburg, South Africa), tightly closed and incubated at 37°C (Haeggström and Schlichtherle, 2004).

### 2.2.2 Continuous parasite culture

Cultures with 5% haematocrit were incubated at 37°C in 25cm$^2$, 80cm$^2$ or 175cm$^2$ sterile tissue culture flasks (Nunc™, Roskilde, Denmark). Spent medium was gently removed from the cultures, without disturbing the layer of uninfected and

parasitized erythrocytes, and replaced daily with 5ml ($25cm^2$ small culture flask), 20ml ($80cm^2$ medium culture flask) or 35ml ($175cm^2$ large culture flask) incomplete RPMI culture medium supplemented with 10% AB plasma or 0.5% Albumax II (complete RPMI culture medium; Appendix A3). Cultures were gassed with 93% $N_2$, 5% $CO_2$ and 2% $O_2$ for 30 seconds ($25cm^2$ small culture flask), two minutes ($80cm^2$ medium culture flask) or 3.5 minutes ($175cm^2$ large culture flask). Parasitaemia was monitored daily via microscopic analysis of blood smears stained with a Rapid Haematology Stain (Diagnostic Media Products, NHLS, Sandringham, South Africa; Appendix A3). Erythrocytes were added to maintain cultures at a 5% haematocrit and parasitaemia at ≤5% (Haeggström *et al.*, 2004).

### 2.2.3 Freezing parasite cultures

Predominantly ring stage parasites were frozen when cultures reached >5% parasitaemia. Cultures were transferred to sterile 50ml centrifuge tubes (Greiner Bio-One GmbH, Baden-Württemberg, Germany) and centrifuged using an Eppendorf 5702R centrifuge at 1,000g at room temperature for five minutes. The supernatant was aspirated and the cell pellet resuspended to a 40% haematocrit with RPMI either containing 10% plasma or 5% Albumax II. For example, 375µl medium was added to the 250µl cell pellet from a 5ml culture. An equal volume of sterile freezing solution (Appendix A3) was added drop wise, while gently swirling, and 1ml of the mixture was aliquoted into sterile cryotubes and stored in liquid nitrogen (Haeggström and Schlichtherle, 2004).

### 2.2.4 Synchronization of parasite cultures

Parasites were synchronized using D-sorbitol, which lyses trophozoite-infected red blood cells (Fernandez, 2004, Lambros and Vanderberg, 1979). Synchronization was performed during the ring-stage of parasite development. Cultures were transferred to sterile 50ml tubes and centrifuged in an Eppendorf 5702R centrifuge at 1,000g for five minutes at room temperature. The supernatant

was aspirated and the pellet resuspended in 20 volumes of 5% D-sorbitol (Appendix A3). The suspension was incubated at room temperature for 10 minutes, with gentle shaking every 2-3 minutes, and centrifuged at 1,000g for five minutes at room temperature. The supernatant was aspirated and the cell pellet washed once in complete RPMI culture medium or PBS (Appendix A3). Pelleted cells were resuspended in complete RPMI culture medium and the suspension transferred to a culture flask, gassed and incubated at 37°C.

### 2.2.5 Culturing gametocytes

Exposure to certain stressful conditions induces the formation of gametocytes in the 3D7 strain of *P. falciparum*. An increase in parasitaemia, a reduced surface area and no addition of fresh erythrocytes are some of the factors that can induce the formation of gametocytes *in vitro*. Five millilitre cultures were initiated at a parasitaemia of 2% and a haematocrit of 5% in a 25cm$^2$ tissue culture flask. Parasites were prepared from stock cultures that had been recently thawed, as gametocyte conversion rate decreases with time in prolonged cultures. Flasks were incubated in an upright position to reduce the surface area available to the parasites. The medium was replaced daily, but only 4ml medium was removed each day to avoid aspirating cultured cells. The final volume of the medium was increased to 7.5ml (~3% haematocrit) when the parasitaemia reached 5%. Culture flasks were gassed daily for 30 seconds. Microscopic analysis was performed every alternate day, for 14-18 days (Bhattacharyya and Kumar, 2008, Carter *et al.*, 1993).

## 2.3 Subcloning *Pf*GK DNA into pGEXT-4T-2 and pTriEx-3 vectors

### 2.3.1 *P. falciparum* DNA extraction and analysis

3D7 parasite DNA was extracted as described by Schlichtherle and Wahlgren (2004), with slight modifications. Using an Eppendorf 5702R centrifuge, cells from a 20ml culture with >5% parasitaemia (predominantly mature stage parasites) were pelleted at 1,000g for five minutes at 4°C. Cells were washed once with 50ml PBS and pelleted at 1,000g for five minutes at 4°C. Cells were resuspended to a final volume of 2ml with PBS and the red blood cells lysed with 40µl 0.5% saponin (final concentration of 0.01%; USB Corporation, Cleveland, USA) for five minutes at room temperature. Parasites were pelleted via centrifugation at 1,000g for five minutes at 4°C. Saponin lysis was repeated to lyse residual red blood cells. Parasites were pelleted and washed three times with 50ml PBS. The parasites were resuspended in 1ml lysis buffer (40mM Tris-HCl, pH 8.0, 80mM EDTA, pH 8.0, 2% SDS), supplemented with 0.1mg/ml Proteinase K (Roche Diagnostics Gmbh, Mannheim, Germany), and incubated at 37°C for three hours, with occasional mixing. DNA was purified using standard methods of phenol/chloroform extraction and ethanol precipitation (Appendix A4). The DNA pellet was air-dried and resuspended in 20-50µl nuclease-free water (Fermentas International Inc., Burlington, Canada). DNA quantity and quality were evaluated using spectrophotometric and agarose gel analysis with the Fermentas MassRuler$^{TM}$ DNA Ladder Mix as a reference (Fermentas International Inc., Burlington, Canada; Appendix A5).

### 2.3.2 Polymerase chain reaction (PCR) amplification of the *Pf*GK gene

The 1,506bp full length *Pf*GK gene was amplified via PCR using synthesized oligonucleotides (primer set PfK03/PfK04; Inqaba Biotechnical Industries, Pretoria, South Africa) containing appropriate restriction enzyme sites to facilitate

the in-frame cloning of the gene into the pGEX-4T-2 expression vector (Amersham Biosciences Ltd., Buckinghamshire, UK; Appendix A6). A 1,503bp *Pf*GK* fragment without the stop codon was amplified (primer set PfK08/PfK06) and used for cloning into the pTriEx-3 expression vector (Novagen, Inc., Madison, USA; Appendix A6). The stop codon was removed to enable translation of the C-terminal 8-histidine residues (His-tag) necessary for protein purification. A guanidine nucleotide was added downstream of the *Bam*HI restriction site in the PfK08 forward primer for in-frame translation of the recombinant protein. Each 50µl amplification reaction mixture consisted of 100ng *P. falciparum* genomic DNA, 0.2µM of each primer, 200µM each of the four deoxynucleotide triphosphates (dNTPs), 2.5U Roche Expand High Fidelity *Taq* polymerase (Roche Diagnostics Gmbh, Mannheim, Germany) and its reaction buffer. *Pf*GK was amplified using PCR conditions as summarized in Table 2 and using an Eppendorf Mastercycler® Gradient thermocycler (Eppendorf AG, Hamburg, Germany). A two-step PCR proceedure was used: i) a 35°C annealing temperature was used for the first five amplification cycles to allow the complementary *Pf*GK primer sequences to bind to the *Pf*GK gene followed by ii) a 55°C annealing temperature for amplification of the full length *Pf*GK gene including the 5'-*Bam*HI and 3'-*Xho*I restriction sites. Amplicons were analysed via agarose gel electrophoresis, with the Fermentas MassRuler™ DNA Ladder Mix as a reference (Appendix A5).

**Table 2: PCR cycling conditions for the amplification of *Pf*GK.**

| Segment | Cycles | Temperature | Time |
|---|---|---|---|
| 1 | 1 | 94ºC | 2 min |
| 2 | 1-5 | 94ºC | 1 min |
| | | 35ºC | 1 min |
| | | 72ºC | 2 min |
| 3 | 5-30 | 94ºC | 1 min |
| | | 55ºC | 1 min |
| | | 72ºC | 2 min |
| 4 | 1 | 72ºC | 10 min |
| 5 | 1 | 15ºC | Hold |
| ***Pf*GK PCR product sizes** | | | |
| Cloning into pGEX-4T-2 | | 1,518bp | |
| Cloning into pTriEx-3 | | 1,516bp | |

### 2.3.3 Plasmid vector preparation

Plasmid DNA extraction was performed using the alkaline lysis method which relies on the selective denaturation of high molecular weight bacterial DNA (Birnboim, 1983). Two milliliters Luria Bertani (LB; Appendix A7) medium, supplemented with 100µg/ml ampicillin (Roche Diagnostics Gmbh, Mannheim, Germany), were inoculated with a single pGEX-4T-2 or pTriEx-3 transformed DH5α *E. coli* colony and cells were grown in a Labotec® Orbital Shaker (Labotec, Midrand, South Africa) at 250rpm at 37°C for 16-18 hours. *E. coli* cultures were pelleted at 16,100g for one minute using an Eppendorf 5415R centrifuge (Eppendorf AG, Hamburg, Germany) and the supernatant aspirated. Cells were resuspended in 100µl ice-cold resuspension buffer (25mM Tris-HCl pH 8.0, 10mM EDTA, 50mM D-glucose) and incubated at room temperature for five minutes. Two hundred microlitres of freshly prepared 0.2M NaOH/1% SDS was added and the solution gently mixed by inversion to lyse the cells. One hundred and fifty microlitres ice-cold freshly prepared 3.3M acetic acid (Merck Chemicals (Pty) Ltd., Johannesburg, South Africa) and 1.6M potassium acetate (Saarchem (Pty) Ltd., Wadeville, South Africa) solution (two volumes 5M acetic acid and one volume 5M potassium acetate) pH 4.8 was added to the lysate and the suspension incubated on ice water for 15 minutes. The solution was centrifuged at 16,100g for five minutes at 4°C to pellet cell debris. DNA in the supernatant was extracted using phenol-chloroform and purified using ethanol precipitation (Appendix A4). The DNA was resuspended in 100µl TE buffer (100mM Tris-HCl, 1mM EDTA, pH 8.0) containing 1µg DNase-free RNaseA (Fermentas International Inc., Burlington, Canada) and incubated at 37°C for 30 minutes. DNA was extracted and purified using phenol-chloroform and ethanol precipitation. The DNA pellet was air-dried and resuspended in 20-50µl nuclease-free water.

For rapid plasmid DNA extraction from *E. coli*, a commercially available silica membrane and spin column format kit, the Sigma GenElute™ Plasmid Miniprep Kit (Sigma-Aldrich Corporation, St. Louis, USA), was used. Plasmid DNA was

extracted from transformed *E. coli* cells as per manufacturer's instruction. Briefly, pelleted cells from 1-5ml cultures were resuspended in an RNase solution and lysed using an alkaline-SDS lysis buffer for five minutes at room temperature. A neutralization solution was added to precipitate cell debris, protein, lipids, SDS and chromosomal DNA. The precipitate was pelleted via centrifugation at 16,100g for 10 minutes at 4°C and the DNA in the supernatant transferred to the column assembly. Following centrifugation at 16,100g for one minute at 4°C, and in the presence of a high salt concentration, the DNA was absorbed and trapped within the silica membrane. A wash-spin step with absolute ethanol was used to remove residual salt contaminants and the DNA was eluted from the column with 50µl nuclease-free water.

DNA quality and quantity were determined via spectrophotometric analysis and agarose gel electrophoresis (Appendix A5). DNA was stored in 50µl aliquots at a concentration of 100ng/µl at -20°C.

### 2.3.4 Preparation of *Pf*GK gene insert and vector for ligation

*Pf*GK and *Pf*GK* DNA amplicons were purified using ethanol precipitation (Appendix A4), the QIAquick® PCR Purification Kit (Qiagen GmbH, Hilden, Germany) or the QIAquick® Gel Extraction Kit (Qiagen GmbH, Hilden, Germany) as per manufacturer's instruction. The QIAquick® nucleic acid purification system utilizes spin-column technology to trap DNA in a silica membrane, a wash step with ethanol to remove impurities, and an elution step with nuclease-free water to recover DNA. One microgram *Pf*GK, *Pf*GK*, pGEX-4T-2 and pTriEx-3 were each digested in a 20µl reaction containing one FastDigest® unit (FDU, 1µl) FastDigest® Fermentas *Bam*HI (Fermentas International Inc., Burlington, Canada) and 1FDU FastDigest® *Xho*I (Fermentas International Inc., Burlington, Canada) restriction enzymes in 1x FastDigest® buffer at 37°C for 10 minutes. One unit (1µl) FastAP™ Thermosensitive Alkaline Phosphatase (Fermentas International Inc., Burlington, Canada) was added to the plasmid digest reactions to dephosphorylate the 5'-end and prevent plasmid re-

circularization during the ligation reaction. Digested products were purified with a QIAquick® PCR Purification Kit and the quality and quantity of the purified digested products were determined spectrophotometrically and with agarose gel electrophoresis. The SynGene GeneGenius gel documentation system with the GeneSnap version 7.04 GeneTools version 4.0 analysis software (Syngene, Cambridge, UK) was used to quantitate DNA using the Fermentas MassRuler™ DNA Ladder Mix as a reference (Appendix A5).

### 2.3.5 Cloning *Pf*GK into pGEX-4T-2 and pTriEx-3 vectors

*Pf*GK was ligated into the *Bam*HI/*Xho*I restriction sites, downstream of the Glutathione S-transferase (GST) gene sequence, in the pGEX-4T-2 plasmid to generate a recombinant pGEX-4T-2-*Pf*GK plasmid construct and *Pf*GK* was ligated into the *Bam*HI/*Xho*I restriction sites, upstream of the 8xHis tag sequence, in the pTriEx-3 plasmid to generate a recombinant pTriEx-3-*Pf*GK* plasmid construct (Appendix A6). Ligation reactions were assembled using the Roche Rapid DNA Ligation Kit (Roche Diagnostics Gmbh, Mannheim, Germany) as per manufacturer's instruction. Briefly, digested vectors were combined with a 3-fold molar excess of digested *Pf*GK insert, to a maximum of 200ng total DNA. The DNA was resuspended in 1x DNA dilution buffer and the volume adjusted to 10µl with nuclease-free water. Ten microlitres 2x T4 Ligase buffer and 1µl T4 DNA Ligase were added to the mixture and the ligation reaction was performed at 16°C for 30 minutes. To transform DH5α *E. coli* (Invitrogen Ltd., California, USA), 50µl of competent cells were inoculated with 5µl of the ligation mix and incubated on ice for 30 minutes. Cells were heat-shocked for 30 seconds at 42°C, incubated on ice for two minutes and added to 0.5ml LB medium. Cells were incubated at 37°C for one hour, with shaking, and 50-100µl of the culture was spread on LB agar plates supplemented with 100µg/ml ampicillin (Appendix A7). Plates were incubated at 37°C overnight (16-18 hours) and transformed colonies were verified with PCR, plasmid extraction, restriction enzyme digestion and DNA sequencing analysis.

## *Gene- and vector-specific PCR*

An initial PCR screen for transformed *E. coli* colonies was performed prior to plasmid extraction and restriction analysis. Ten random colonies were each resuspended in 10µl Milli-Q water. Two milliliters LB, supplemented with 100µg/ml ampicillin, was inoculated with 5µl of the bacterial solution and cultures were incubated at 37°C overnight for subsequent plasmid DNA extraction. The remaining 5µl suspension was incubated at 94°C for five minutes to lyse the bacteria, denature bacterial proteins and release bacterial and plasmid DNA. Colonies were screened for the GK inserts using PCR with pGEX-4T-2/pTriEx-3 vector specific primers or GK gene-specific primers (Appendix A6 and Appendix A17). PCR reactions were assembled in a final volume of 20µl containing 10µl 2x PCR Master Mix (Promega Corporation, Madison, USA), 0.2µM each primer and 100ng DNA under PCR conditions as summarized in Table 3. Amplicons were analysed using agarose gel electrophoresis (Appendix A5). Plasmid DNA extractions were performed only on colonies containing either the pGEX-4T-2-*Pf*GK or pTriEx-3-*Pf*GK* recombinant plasmids.

**Table 3: Colony PCR cycling conditions for the detection of *Pf*GK in transformed bacteria.**

| Segment | Cycles | Temperature | Time |
|---|---|---|---|
| 1 | 1 | 94ºC | 2 min |
| 2 | 1-29 | 94ºC | 1 min |
| | | 55ºC | 1 min |
| | | 72ºC | 2 min |
| 3 | 1 | 72ºC | 10 min |
| 4 | 1 | 15ºC | Hold |
| **Template** | | **Primers** | **PCR product sizes** |
| pGEX-4T-2 | | SE-pGEX-F/SE-pGEX-R | 174bp |
| pGEX-4T-2-*Pf*GK | | SE-pGEX-F/SE-pGEX-R | 1,661bp |
| pTriEx-3 | | SE-pTriEx-F/SE-pTriEx-R | 496bp |
| pTriEx-3-*Pf*GK* | | SE-pTriEx-F/SE-pTriEx-R | 1,885bp |
| *Pf*GK/*Pf*GK* | | PfK11/PfK05; PfK16/PfK17* | 150bp |

## *Restriction enzyme analysis*

One microgram purified pGEX-4T-2-*Pf*GK or pTriEx-3-*Pf*GK* recombinant plasmids were digested in a 20µl reaction with *Bam*HI and *Xho*I (section 2.3.4), separating the *Pf*GK insert from the plasmid backbone. Digested products were analyzed using agarose gel electrophoresis on a 0.8% agarose gel (Appendix A5). Digested pGEX-4T-2 or pTriEx-3 vectors without inserts were used as negative controls.

## *DNA sequencing analysis*

Recombinant plasmids were either sent to Inqaba Biotechnical Industries (Pty) Ltd for automated sequencing analysis or sequenced using the ABI PRISM® 3100 Genetic Analyzer (Applera Corporation, Foster City, USA) and the Applied Biosystems Big Dye® Terminator v1.1/3.1 sequencing kit (Life Technologies Corporation, California, USA). Ten microlitres BigDye terminator sequencing reactions were assembled by adding 500ng recombinant plasmid DNA to 2µl Ready Reaction Premix, 1µl BigDye sequencing buffer and 3.2pmol primer (either vector- or *Pf*GK gene-specific). Sequencing reactions were performed under conditions as summarized in Table 4.

**Table 4: Sequencing cycling conditions for the detection of *Pf*GK in transformed bacteria.**

| Segment | Cycles | Temperature | Time |
|---|---|---|---|
| 1 | 1 | 96ºC | 1 min |
| 2 | 1-25 | 96ºC | 10 sec |
| | | 55ºC | 5 sec |
| | | 60ºC | 4 min |
| 3 | 1 | 4ºC | Hold |

Unincorporated dye terminators were removed using ethanol/EDTA precipitation as per manufacturer's instruction, with slight modifications. Briefly, 5µl 125mM EDTA pH 8.0 and 60µl 100% ethanol were added to the 10µl sequencing reaction and incubated at room temperature for 15 minutes. The DNA was pelleted via

centrifugation at 16,100g for 30 minutes at 4°C and the supernatant aspirated. Sixty microliters 70% ethanol was added to the DNA pellet to remove residual salts and unincorporated terminator dyes. The DNA was pelleted via centrifugation at 16,100g for five minutes at 4°C and air-dried for 15 minutes at room temperature, in the dark. DNA was sent to the HIV genotyping laboratory in the Department of Molecular Medicine and Haematology for sequence determination.

## 2.4 Expression of recombinant GST-*Pf*GK (r*Pf*GK) in *E. coli*

A statistical model was used to predict the expression of soluble recombinant proteins in *E. coli* (http://www.biotech.ou.edu/). Rosetta2 (DE3) *E. coli* (Novagen, Inc., Madison, USA) were transformed with the pGEX-4T-2-*Pf*GK construct. Rosetta2 (DE3) cells were designed to enhance expression of eukaryotic proteins that contain codons rarely used in *E. coli* (Appendix A8). The strain expresses seven rare tRNAs for arginine (AGG, AGA and CGG), isoleucine (AUA), leucine (CUA and CCC) and glycine (GGA) on a pRARE plasmid containing a chloramphenicol-resistance cassette (Novagen, 2009a User Protocol TB009).

*Pf*GK was expressed as a fusion protein with an N-terminal GST following induction with the inducing agent, isopropyl-β-D-thiogalactopyranoside (IPTG, Promega Corporation, Madison, USA), or the Overnight Express™ Instant TB Medium (Novagen, Inc., Madison, USA; Appendix A10). Overnight Express™ medium is a modified terrific broth that allows the bacteria to autoinduce in the presence of a secondary energy source, lactose, once the primary glucose energy source has been depleted (Novagen, 2005 User Protocol TB383, Novagen, 2009b User Protocol TB526). Proteins were expressed under conditions as reviewed by Sorensen and Mortensen (2005b) where the expression of soluble recombinant proteins may be improved following induction at high cell densities and at lower temperatures.

Two millilitres LB medium supplemented with 100µg/ml ampicillin and 50µg/ml chloramphenicol (Sigma-Aldrich Corporation, St. Louis, USA; Appendix A7) was inoculated with transformed colonies and allowed to grow to a cell density with an optical absorbance at 600nm ($A_{600nm}$) of greater than 1.0 by incubation in a Labotec® Orbital Shaker at 250rpm overnight (16-18 hours) at 37°C. In all cases, a 1:10 (v/v) dilution was used to measure the density of the cultures. 1:100 (v/v, 250µl) of the antibiotic-selected culture was added to 25ml LB without antibiotics, in a 125ml flask, and allowed to grow to an $A_{600nm}$ cell density between 0.8-1.0 (typically 6-8 hours) in the Labotec® Orbital Shaker at 250rpm at 37°C. Recombinant protein expression was induced with 0.4-1mM IPTG, with shaking on a Red Rotor PR70-230V shaker (Hoefer Scientific Instruments, San Francisco, USA) at 250rpm for 16-18 hours at room temperature. For optimal aeration, cultures were grown in flasks at a ratio of one volume culture to ≥5 volumes culture vessel. Alternatively, Overnight Express™ Instant TB Medium was inoculated with 1:100 (v/v) of an antibiotic-selected culture and incubated on the Red Rotor PR70-230V shaker at 250rpm for 20-24 hours at room temperature. Final cell suspensions had an $A_{600nm}$ value between 4.0-8.0.

Cells were pelleted at 3,000g for 10 minutes at 4°C using an Eppendorf 5702R centrifuge and resuspended in 2ml BugBuster® HT Protein Extraction Reagent (Novagen, Inc., Madison, USA) supplemented with 2µl Protease Inhibitor Cocktail Set III (100mM 4-(2-Aminoethyl) benzenesulphonyl fluoride hydrochloride, 80µM Aprotinin, 5mM Bestatin; 1.5mM E-64; 2mM Leupeptin, 1mM Pepstatin A; Calbiochem®, San Diego, USA). Suspensions were incubated for 20 minutes, with shaking on a Red Rotor PR70-230V shaker. One hundred and fifty microlitres of the suspension was collected and solubilized for analysis of the total protein fraction (Appendix A10). The lysate was centrifuged at 16,100g for 20 minutes at 4°C using an Eppendorf 5415R centrifuge to separate the soluble protein fraction in the supernatant from the insoluble protein in the pellet. One hundred and fifty microlitres of the soluble protein fraction was collected for analysis. The pellet was resuspended in 1.85ml PBS (the original volume prior to separation of the soluble and insoluble protein fractions) and

150µl was collected for analysis of the insoluble protein fraction. Similarly, rGST was expressed from Rosetta2 (DE3) *E. coli* transformed with the pGEX-4T-2. The total, soluble and insoluble protein fractions from each clone were analyzed using SDS-PAGE and immunoblot analysis with a 1:50,000 Anti-GST-HRP Conjugated antibody (Amersham Biosciences, Ltd., Buckinghamshire, UK; Appendix A11). Five microlitres soluble rGST was used as a control in the immunoblot analysis.

### 2.4.1 Large scale r*Pf*GK protein expression

Two millilitres (1:100, v/v) of transformed cells with an $A_{600nm}$ greater than 1.0 was added to 200ml LB or Overnight Express™, in a 1L flask, without antibiotics. Recombinant protein expression was induced with IPTG or the Overnight Express™ system, as described (section 2.4).

Cells expressing r*Pf*GK were lysed using sonication. Cells were resuspended in 10ml 1x GST buffer (4.3mM $Na_2HPO_4$, 1.47mM $KH_2PO_4$, 137mM NaCl and 2.7mM KCl, pH 7.3) supplemented with 5µl Protease Inhibitor Cocktail Set III, in a 50ml tube and sonicated with a Bandelin Sonoplus UW 2070 sonicator (Bandelin Electronics, Berlin, Germany) on ice water at 75% power for 15 seconds. The sonication cycle was repeated five times, with a 45 second cooling period on ice between each cycle. The cell lysate was centrifuged at 16,100g for 20 minutes at 4°C using a Beckman Coulter J2-21 or Avantji® J-E centrifuge with a JA17 rotor (Beckman Coulter, Inc., Fullerton, USA) to separate insoluble protein in the pellet from soluble protein in the supernatant. The supernatant was collected and soluble r*Pf*GK was purified. Protein fractions were collected for analysis as described (section 2.4, Appendix A10 and Appendix A11).

### 2.4.2 Purification of r*Pf*GK

Soluble r*Pf*GK was purified from endogenous *E. coli* proteins using the Promega MagneGST™ Protein Purification System (Promega Corporation, Madison, USA) as per manufacturer's instruction. The MagneGST™ agarose beads are supplied

as a 25% slurry and have a binding capacity of 5-10mg of protein per 1ml settled particles. They are used for affinity purification of small amounts of proteins. The particles contain reduced glutathione covalently attached to agarose beads with an iron oxide core. Twenty microlitres GST agarose beads were added to a sterile 1.5ml centrifuge tube (Eppendorf AG, Hamburg, Germany) and washed three times with five volumes MagneGST$^{TM}$ binding/wash buffer (4.2mM $Na_2HPO_4$, 2mM $K_2HPO_4$, 500mM NaCl, 10mM KCl). One volume is equivalent to the volume of packed beads and the wash procedure involved adding buffer, inverting the suspension five times, collecting the beads with a Magnetic Particle Separator (Roche Diagnostics Gmbh, Mannheim, Germany) and aspirating the supernatant. The beads were resuspended in one volume MagneGST$^{TM}$ binding/wash buffer and the soluble protein fraction, following cell lysis (section 2.4.1), was added. r*Pf*GK was allowed to bind to the beads for 6-18 hours at 4°C, with shaking using the Elmi RM-2M Intelli-Mixer (Elmi, Riga, Latvia). Non-specific binding of endogenous *E. coli* proteins was removed by washing the beads five times with 50 volumes of MagneGST$^{TM}$ binding/wash buffer, each wash for five minutes with gentle agitation. r*Pf*GK was eluted from the beads with 200µl elution buffer (500mM reduced glutathione, 50mM Tris-HCl, pH 8.25) for 15 minutes at room temperature, with gentle agitation every two minutes. Residual r*Pf*GK was eluted from the beads a second time with 100µl elution buffer. The fractions were kept separate for enzyme activity analysis. Residual protein still bound to the beads was analyzed by resuspension in 100µl elution buffer and boiling for one minute in the SDS suspension solution (Appendix A10).

All protein fractions were collected and stored at -20°C. Fractions (5-25µl) were analyzed using SDS-PAGE. Recombinant proteins were detected using immunoblotting with a 1:50,000-1:100,000 diluted anti-GST-HRP conjugated antibody (Amersham Biosciences, Ltd., Buckinghamshire, UK) (Appendix A11).

## 2.5 r*Pf*GK protein refolding

### 2.5.1 Inclusion body purification and quantitation

Incorrectly folded r*Pf*GK trapped in inclusion bodies in bacteria was extracted using the BugBuster® HT Protein Extraction Reagent. Insoluble proteins from induced cultures were resuspended in the same volume of BugBuster® HT used to lyse the original harvested cell pellet (1 volume). One kilo unit lysozyme (Roche Diagnostics Gmbh, Mannheim, Germany) per millilitre BugBuster® HT was added to the solution, mixed and incubated at room temperature for five minutes. The lysozyme hydrolyses the glycosidic linkage between *N*-acetylmuramic acid and *N*-acetylglucosamine in the peptidoglycan of bacterial cell walls and facilitates extraction of incorrectly folded r*Pf*GK trapped within the inclusion bodies. Six volumes 1:10 BugBuster® HT dilution were added and vortexed vigorously for one minute. The suspension was centrifuged at 5,000g at 4°C for 15 minutes using a Beckman Coulter J2-21 or Avantji® J-E centrifuge with a JA17 rotor. The inclusion body pellet was resuspended in 0.5 volumes 1:10 BugBuster® HT, vortexed for one minute and centrifuged at 5,000g at 4°C for five minutes. The supernatant was discarded and the process repeated twice. Inclusion bodies were resuspended in 0.5 volumes 1:10 BugBuster® HT and pelleted at 15,000g at 4°C for 15 minutes. The purified inclusion body pellet was resuspended in 2ml Guanidine-hydrochloride (GnHCl; Merck KGaA, Darmstadt, Germany) solution (6M GnHCl, 50mM Tris-HCl pH 8.0, 100mM NaCl, 10mM EDTA and 10mM DTT) and incubated at 4°C overnight, on a Red Rotor PR70-230V shaker, to dissolve the pellet. The suspension was centrifuged at 17,000g at 4°C for 20 minutes and the supernatant collected (Novagen, 2004 User Protocol TB245). Solubilized inclusion body protein concentration was estimated at 280nm under ultraviolet (UV) light using the NanoDrop® spectrophotometer and Beer-Lambert's equation:

$$A_{280nm} = ecl$$

e = r*Pf*GK protein extinction coefficient ($M^{-1}.cm^{-1}$)

c = concentration (mol/L)

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

The r*Pf*GK protein extinction coefficient was obtained from the ProtParam database (Gasteiger E. *et al.*, 2005) (http://web.expasy.org/protparam/protpar-ref.html).

A 1:10, 1:50, 1:100, 1:200 and 1:500 r*Pf*GK protein dilution series, using PBS as the diluent, was prepared and absorbance measured. The absorbance against protein dilution was plotted and absorbance values within the linear range were used to calculate protein concentration:

$$C = \frac{A280_{nm}}{el} \times \text{dilution factor}$$

Concentration in g/L was calculated by multiplying the protein concentration in $mol.L^{-1}$ with the protein's molecular weight ($g.mol^{-1}$). Solubilized inclusion bodies were stored in approximately 10mg aliquots at -20°C.

## 2.5.2 Refolding r*Pf*GK from inclusion bodies with guanidine hydrochloride

Five millimolar 1,4-dithiothreitol (DTT; Roche Diagnostics Gmbh, Mannheim, Germany) was added to 10mg defrosted, purified and denatured inclusion bodies and incubated at 4°C for 60 minutes to reduce disulphide bonds. Two hundred milliliters refolding buffer (200mM Tris-HCl pH 8.0, 10mM EDTA, 1M L-arginine, 0.1mM phenylmethylsulfonyl fluoride (PMSF), 2mM reduced glutathione, 0.2mM oxidized glutathione) was prepared and cooled to 4°C. The positively charged L-arginine (Merck KGaA, Darmstadt, Germany) in the refolding buffer reduces protein aggregation, the glutathione redox couple (Sigma-Aldrich Corporation, St. Louis, USA) shuffles disulphide bond formation during refolding and the PMSF (Roche Diagnostics Gmbh, Mannheim, Germany)

protease inhibitor protects the protein from proteases. The solubilized inclusion bodies were diluted to 3M GnHCl with one volume refolding buffer on ice over a period of 30 minutes. The solution was transferred to dialysis tubing (Pierce Biotechnology Incorporated, Rockford, USA), with a 10kDa molecular weight protein size limitation, and dialyzed against 100 volumes of refolding buffer at 4°C for 48 hours (refold database http://refold.med.monash.edu.au) (Chow *et al.*, 2006a, Chow *et al.*, 2006b). The protein sample in the dialysis tubing was transferred to a 2.0ml centrifuge tube (Eppendorf AG, Hamburg, Germany) and centrifuged in an Eppendorf 5415R centrifuge at 16,100g for 10 minutes at 4°C to remove insoluble protein aggregates. The supernatant was collected and the refolded r*Pf*GK protein purified via affinity chromatography (section 2.4.2). Immunoblot analysis with an anti-GST-HRP conjugated antibody was used to analyze protein fractions (Appendix A11).

### 2.5.3 r*Pf*GK protein refolding from inclusion bodies with urea

r*Pf*GK was refolded from solubilized inclusion bodies by removing GnHCl present in the denatured protein sample through a serial dilution series using urea (Promega Corporation, Madison, USA) and PBS. One hundred micrograms purified inclusion bodies in a final volume of 100µl were added to a Pierce Slide-A-Lyzer mini dialysis unit (Pierce Biotechnology Incorporated, Rockford, USA) with a 10kDa molecular weight cut-off. The solution was dialyzed against 100ml 7M urea in PBS for 15 minutes at room temperature. The dialysis unit was transferred to 100ml 6M urea in PBS and dialyzed for 15 minutes at room temperature. The serial dilution series was repeated until the solution was dialyzed against PBS only. Proteins were transferred to a tube and centrifuged in an Eppendorf 5415R centrifuge at 16,100g for 10 minutes at 4°C to remove aggregates. The supernatant was collected and the refolded r*Pf*GK purified by affinity chromatography (section 2.4.2). Protein fractions were analyzed using immunoblot analysis with an anti-GST-HRP conjugated antibody (Appendix A11).

## 2.6 *Pf*GK enzyme kinetics

### 2.6.1 r*Pf*GK enzyme assay

r*Pf*GK was quantitated using densitometry of Coomassie blue stained SDS-PAGE gels and 250ng, 500ng and 1,000ng bovine serum albumin (BSA, Pierce Biotechnology Incorporated, Rockford, USA; Appendix A10) as a reference. Enzyme activity was measured via an ADP-coupled spectrophotometric assay using the Biomate 5 spectrophotometer (Thermo Electron Corporation, Cambridge, UK) and the VISIONlite$^{TM}$ version 2.2 kinetics software (Thermo Fisher Scientific, Inc., Ueberlingen, Germany). When r*Pf*GK is added to a reaction mixture containing ATP, NADH, phosphoenolpyruvate (PEP), pyruvate kinase (PK), lactate dehydrogenase (LDH) and glycerol the following sequence of events occurs:

$$\text{Glycerol} + \text{ATP} \xrightarrow{\text{r}Pf\text{GK}} \text{Glycerol-3-phosphate} + \text{ADP}$$

$$\text{ADP} + \text{PEP} \xrightarrow{\text{PK}} \text{ATP} + \text{Pyruvate}$$

$$\text{Pyruvate} + \text{NADH} + \text{H}^{+} \xrightarrow{\text{LDH}} \text{Lactate} + \text{NAD}^{+}$$

By measuring the decrease in NADH absorbance at 340nm, the enzyme activity of r*Pf*GK can be calculated. The spectrophotometer was connected to a waterbath to measure enzyme activity at 37°C. One milliliter Milli-Q water was used to equilibrate the instrument and as a blank. Thirty three microlitres of purified r*Pf*GK (65.8-172.6ng) was added to a 0.967ml pre-warmed reaction mixture, mixed via inversion and transferred to pre-warmed cuvettes in the spectrophotometer. The reaction mixture contained 0.233ml reagent solution (8.5mM ATP, 1.22mM NADH, 2mM PEP, 7U/ml (v/v) PK, 15U/ml (v/v) LDH, 28mM $MgSO_4$, 26mM reduced glutathione, pH 7.4), 0.7ml 0.1M Triethanolamine-HCl buffer pH 7.4 or 0.7ml 0.1M Tris-HCl pH 7.4 and 33µl 0.1M glycerol. Final concentrations of the reagents were 2mM ATP, 0.3mM NADH, 0.5mM PEP, 1.6 units/ml PK, 3.5 units/ml LDH, 6.5mM $MgSO_4$, 6.0mM

reduced glutathione, 70mM Triethanolamine-HCl pH 7.4 or 70mM Tris-HCl pH 7.4 and 3.3mM glycerol. Purified rGST and elution buffer (500mM reduced glutathione, 50mM Tris-HCl, pH 8.25) were used as negative controls. The linear decrease in NADH absorbance at 340nm was measured for 20 minutes, at 30 second intervals, and used to calculate enzyme velocity as follows:

$$\text{Enzyme velocity (nmol/min)} = \frac{\Delta A_{340nm} \text{ x } 1{,}000 \text{ x reaction volume (ml)}}{6.22}$$

$\Delta A_{340nm}$ = change in NADH absorbance at 340nm per minute

6.22 = Extinction coefficient of NADH ($mM^{-1}cm^{-1}$ or $\mu mol^{-1}.ml\ cm^{-1}$)

1,000 = conversion of μmol NADH to nmol NADH

The specific activity of the enzyme was calculated using the equation:

$$\text{Enzyme activity} = \frac{\text{Enzyme velocity}}{\text{mg enzyme}}$$

One microlitre of a 1:100 dilution of recombinant *E. coli* glycerol kinase (Sigma-Aldrich Corporation, St. Louis, USA) in 70mM Triethanolamine-HCl pH 7.4 was used as a positive control. A glycine-carbonate buffer system (0.7ml 0.4M glycine, 0.45mM carbonate, pH 8.9) was used for the *E. coli* glycerol kinase reaction (Worthington and Worthington, 1993). All enzyme kinetic graphs and transformation data were generated with GraphPad Prism® version 5.02.

### 2.6.2 r*Pf*GK enzyme stability

Purified r*Pf*GK, in elution buffer, was stored at 4°C between activity assays. Enzyme activity was monitored in duplicate at 37°C and at 0, 24, 48 and 72 hours post protein purification. Three independent experiments were performed and a total of six samples were used in the analysis.

### 2.6.3 r*Pf*GK temperature optimum

r*Pf*GK enzyme activity was measured in duplicate at 28°C, 37°C and 40°C. Two independent experiments were performed and a total of four samples were used in the analysis.

### 2.6.4 r*Pf*GK pH optimum

Enzyme activity was measured in duplicate at pH 5.0-9.0 at 37°C to determine the r*Pf*GK pH optimum. A 0.1M Tris-HCl buffer system was used to measure enzyme activity at a pH 6.8-8.4 and a 0.1M sodium acetate-acetic acid buffer system was used for pH 5.0-6.0. Buffer pH was adjusted at 37°C and final concentrations of each buffer in the enzyme assay were 70mM. Three independent experiments were performed and a total of six samples were acquired for pH 6.8-8.2 enzyme activity analysis. Four samples were used to examine enzyme activity at pH 5.0, 5.5 and 8.4.

### 2.6.5 $V_{max}$ and $K_m$ for glycerol and ATP

To determine the $V_{max}$ and $K_m$ for both glycerol and ATP substrates, enzyme activity at saturating concentrations of each partner substrate was measured, in duplicate. The Michaelis-Menten constants for glycerol were determined using glycerol concentrations between 0.005-3.3mM at a final and saturated 2mM ATP concentration. $V_{max}$ and $K_m$ for ATP were determined using ATP concentrations between 0.005-2mM at a final and saturated 3.3mM glycerol concentration.

### 2.6.6 r*Pf*GK activity in the presence of Fructose-1,6-bisphosphate

Fructose-1,6-bisphosphate (FBP) was evaluated as non-competitive inhibitor against r*Pf*GK. The Michaelis-Menten constants for glycerol were determined in the presence of 0mM, 5mM and 10mM of FBP. Enzyme activity in the absence or

presence of FBP was measured in duplicate at 37°C using glycerol concentrations between 0.005-3.3mM at a final and saturated 2mM ATP concentration.

## 2.7 Expression and purification of soluble *Pf*GK-His (r*Pf*GK*) in insect cells

### 2.7.1 Insect cell culture

Recommended cell lines for insect cell expression of recombinant proteins using pTriEx are the *Spodoptera frugiperda* (*Sf*9; Novagen, Inc., Madison, USA) and TriEx™ *Sf*9 cells (Novagen, Inc., Madison, USA). *Sf*9 and TriEx *Sf*9 cells were maintained in optimized serum-free BacVector® Insect Cell Medium (Novagen, Inc., Madison, USA) and serum-free TriEx™ Insect Cell Medium (Novagen, Inc., Madison, USA) respectively, with no additives. Where necessary, cultures were supplemented with irradiated fetal bovine serum (FBS; Lonza Group Ltd., Basel, Switzerland) to improve cell growth. Healthy *Sf*9 cells settle at the bottom of the flask and are round with distinct boundaries, bright and uniform in size. Unhealthy cells are granular, dark and float in the medium. *Sf*9 cells were used to generate recombinant baculoviruses containing the *PfGK* gene. Recombinant *Pf*GK protein was expressed in TriEx™ *Sf*9 cell infected cells (Novagen, 2008a, Novagen, 2008b).

#### *Preparation of insect cells from frozen cultures*

Cryotubes containing 1ml frozen *Sf*9 or TriEx™ *Sf*9 cells were removed from liquid nitrogen storage and thawed in a preheated water bath at 28°C. Cells were aseptically transferred to a sterile 50ml centrifuge tube. Five milliliters of pre-warmed BacVector® insect cell medium or TriEx™ insect cell medium were added, drop wise, to *Sf*9 or TriEx *Sf*9 insect cells respectively. Cells were suspended by repeated pipetting and transferred to a 25cm$^2$ tissue culture flask. Cultures were incubated for 30-60 minutes to allow the cells to settle/attach to the

bottom of the flask. The medium was aspirated and replaced with fresh pre-warmed medium. Cultures were incubated at 28°C and were monitored using an Olympus CKX41 inverted microscope (Olympus Optical Corporation Ltd., Tokyo, Japan), every other day until the monolayer reached 85-95% confluency.

## *Continuous culture of insect cells*

### Monolayer cultures

*Sf*9 and TriEx™ *Sf*9 monolayer cultures were maintained in either 25cm$^2$ (5ml culture) or 75cm$^2$ (10ml culture) culture flasks. Cultures were diluted when monolayers were 85-95% confluent. Spent medium was aspirated and replaced with pre-warmed fresh medium. Using a sterile cell scraper (TPP®, Techno Plastic Products, Trasadingen, Switzerland), cells were resuspended in solution and transferred to a sterile centrifuge tube. One hundred microlitres of the culture was used to determine cell concentration and viability using the trypan blue exclusion method (Appendix A14). Cultures were diluted to a seed density of 1-2 x 10$^6$ cells/ml for a 5ml culture or 3 x 10$^6$ cells/ml for a 10ml culture and maintained in a closed 25cm$^2$ or 75cm$^2$ flask as a monolayer at 28°C in a 150L Scientific incubator (Separations Scientific (Pty) Ltd., Honeydew, South Africa).

### Suspension cultures

*Sf*9 and TriEx™ *Sf*9 were grown in suspension for preparation of high titre viral stocks and recombinant protein expression respectively. Exponentially growing cells were maintained in a 150L temperature controlled orbital shaker at 150rpm and 28°C (Separations Scientific (Pty) Ltd., Honeydew, South Africa) and passaged every 2-3 days, depending on cell density. To initiate suspension cultures, cells from monolayer cultures were counted and seeded at a density of 0.3-0.5 x 10$^6$ cells/ml in a sterile disposable 125ml Corning® polycarbonate Erlenmeyer flask (VWR International Ltd, West Sussex, UK). For proper aeration, the liquid culture was not more that 20% of the culture vessel and caps of the flasks were kept slightly loose. Usually, 20ml cultures were maintained in a

125ml flask at a concentration between 0.3-5 x $10^6$ cells/ml. Cells were passaged not more than 20-25 times.

### *Freezing insect cells*

Exponentially growing *Sf*9 and TriEx™ *Sf*9 cells (0.5-6 x $10^6$ cells/ml) were counted and cell viability assessed using the trypan blue exclusion method (Appendix A14). Cells with >90% viability were pelleted via centrifugation at 1,000g for five minutes at 4°C using an Eppendorf 5702 centrifuge. The supernatant was aspirated and cells resuspended in pre-warmed insect cell medium to a concentration of 4 x $10^6$ cells/ml. An equal volume of freezing solution (BacVector® or TriEx™ insect cell medium, 20% DMSO and 5% FBS) was added, drop wise, to the suspension. One milliliter aliquots were added to sterile cryogenic vials. Vials were placed at -20°C for two hours, -70°C for 16-18 hours and in liquid nitrogen for long term storage.

## 2.7.2 Generation of recombinant baculovirus (Bac-*Pf*GK*)

The basis for the production of recombinant baculoviruses, using the BacMagic™ DNA Kit (Novagen, Inc., Madison, USA), is homologous recombination between sequences in the baculoviral backbone DNA and the recombinant plasmid containing the gene of interest. The BacMagic™ DNA is an *Autographa californica* nuclear polyhedrosis virus (AcNPV) genome with a portion of the (ORF) 1629 open reading frame deleted. Homologous recombination between the modified baculovirus DNA and the transfer plasmid (pTriEx in this study) results in a baculovirus genome with the promoter and foreign gene sequence from the transfer pTriEx plasmid. The transfer plasmid also contains segments of the baculovirus genome flanking the promoter and this restores the function of the viral ORF1629, enabling the recombinant baculovirus to replicate and produce viruses which are released into the medium.

Transfections were performed using the BacMagic™ DNA kit as per manufacturer's instruction (Novagen, 2007 User Protocol TB459). For each transfection, three 35mm plates (Nunc™, Roskilde, Denmark) were seeded with 1 x $10^6$ *Sf*9 cells in 2ml BacVector® insect cell medium. Cells were incubated for 30-60 minutes to attach/settle to the bottom of the plate. Components from the BacMagic™ DNA Kit were assembled, in order, in a sterile 1.5ml tube: 1ml BacVector® medium, 5µl Insect GeneJuice®, 5µl (100ng) BacMagic Virus DNA and 5µl (500ng) of recombinant pTriEx-*Pf*GK*. The negative control had BacVector® medium instead of the recombinant plasmid and the blank had none of the transfection reaction components. Reagents were gently mixed and incubated at room temperature for 30 minutes to allow complexes to form. Medium from the pre-attached *Sf*9 monolayer cells was aspirated and the transfection solution added, drop wise, to the cells. Cultures were incubated at 28°C for five hours in a humidified chamber. One milliliter BacVector® medium was added, gently mixed and incubated at 28°C for a maximum of five days. Cultures were monitored daily for cell growth using microscopy. After five days, cultures were transferred to a sterile tube, centrifuged at 1,000g for five minutes and recombinant baculovirus in the supernatants were collected and stored at 4°C, in the dark.

### 2.7.3 Recombinant baculovirus amplification

Three successive rounds of reinfection with Bac-*Pf*GK* were performed in order to increase recombinant baculoviral titers. Six-well culture dishes (Nunc™, Roskilde, Denmark) were seeded with 1 x $10^6$ *Sf*9 cells in 2ml BacVector® medium and 500µl recombinant baculoviruses were added to each well. Plates were incubated at 28°C. After five days, cultures were centrifuged at 1,000g for five minutes in an Eppendorf 5702R centrifuge and the supernatants collected and stored at 4°C, in the dark. Successive rounds of reinfection were performed until supernatants with recombinant baculoviruses were positive for *Pf*GK gene expression (section 2.3.5).

### 2.7.4 Baculoviral DNA extraction

Bac-*Pf*GK* DNA was extracted using the Roche High Pure Viral Nucleic Acid Kit (Roche Diagnostics Gmbh, Mannheim, Germany) as per manufacturer's instruction. Briefly, 200µl of a guanidine-TritonX-100 based lysis/binding buffer supplemented with Proteinase K was added to 200µl culture supernatant and incubated at 72°C for 10 minutes to lyse baculoviruses. The lysate was added to a column assembly pre-packed with a layer of glass fibers and centrifuged. Only viral DNA binds to the glass fibers. The DNA was purified from salts, proteins and other impurities using an ethanol-based washing step and eluted from the glass fibers using 50µl nuclease-free water. All centrifugation steps were at 8,000g for one minute. DNA quality and quantity were analyzed using spectrophotometry and agarose gel electrophoresis (Appendix A5).

### 2.7.5 PCR screening for Bac-*Pf*GK* recombinant viruses

PCR reactions were assembled in a final volume of 20µl containing 10µl 2x PCR Master Mix, 0.2µM each primer (*Pf*GK gene-specific: PfK11/PfK05 or PfK16/PfK17; pTriEx-3 vector specific: SE-pTriEx-F/SE-pTriEx-R) and 500ng Bac-*Pf*GK* DNA. One hundred nanograms pTriEx-3-*Pf*GK* and pTriEx-3 plasmid DNA were used as a positive and negative control, respectively (Appendix A6, Appendix A13 and Appendix A17). PCR was performed under the conditions summarized in Table 5 and Table 6. Amplicons were analyzed using agarose gel electrophoresis (Appendix A5).

**Table 5: *Pf*GK* amplification from recombinant baculoviral DNA using *Pf*GK gene-specific primers.**

| Segment | Cycles | Temperature | Time |
|---|---|---|---|
| 1 | 1 | 94ºC | 2 min |
| 2 | 1-29 | 94ºC | 30 sec |
| | | 55ºC | 15 sec |
| | | 72ºC | 30 sec |
| 3 | 1 | 72ºC | 5 min |
| 4 | 1 | 15ºC | Hold |
| ***Pf*GK PCR product size = 150bp** | | | |

**Table 6: *Pf*GK* amplification from recombinant baculoviral DNA using pTriEx-vector specific primers.**

| Segment | Cycles | Temperature | Time |
|---|---|---|---|
| 1 | 1 | 94ºC | 2 min |
| 2 | 1-29 | 94ºC | 1 min |
| | | 55ºC | 30 sec |
| | | 72ºC | 2 min |
| 3 | 1 | 72ºC | 10 min |
| 4 | 1 | 15ºC | Hold |
| **PCR product sizes** | | | |
| pTriEx-3 | | 496bp | |
| pTriEx-3-*Pf*GK* | | 1,885bp | |

## 2.7.6 Determining Bac-*Pf*GK* viral titre

### *Viral plaque assay*

A plaque assay was performed to determine Bac-*Pf*GK* viral titre (Novagen, 2007 User Protocol TB459). Eight 35mm plates were each seeded with 0.9 x $10^6$ *Sf*9 cells and incubated at room temperature for one hour to allow the cells to form a sub-confluent monolayer. During the incubation period, a $10^{-1}$ to $10^{-7}$ Bac-*Pf*GK* serial dilution series containing 500µl of each dilution in BacVector® insect cell medium, was prepared. The medium was gently aspirated from each plate and 100µl $10^{-5}$, $10^{-6}$ and $10^{-7}$ diluted recombinant baculoviruses was added, in duplicate, to each set of plates. Two plates were each inoculated with 100µl BacVector® insect cell medium as controls. Plates were incubated at room

temperate for one hour, with gentle agitation/rocking every 15 minutes. The inoculum was gently removed and 2ml of the BacPlaque™ Agarose-medium-serum mixture (Appendix A14) was gently overlaid over the cells and incubated at room temperature for 20 minutes or until the agarose had solidified. One milliliter BacVector® insect cell medium was added to each plate and cultures were incubated in a humidified incubator at 28°C for 3-4 days.

### *Plaque identification and viral titre determination*

After 3-4 days incubation, plates were stained with Neutral Red (Sigma-Aldrich Corporation, St. Louis, USA). Neutral Red stains live cells red whereas plaques remain clear in the cell monolayer. The liquid overlay in each plate was gently aspirated, without disturbing the cell monolayer. One milliliter freshly diluted Neutral Red staining solution (Appendix A14) was added to the centre of each plate and cultures were incubated at 28°C for two hours. The stain was gently aspirated and cultures incubated in the dark at room temperature for 18 hours. Plaques were counted and the viral titre, in plaque forming units per ml (pfu/ml), was calculated as follows:

$$\text{Viral titre (pfu/ml)} = \frac{\text{number of plaques x dilution factor}}{\text{volume plates (ml)}}$$

The multiplicity of infection (MOI) is defined as the number of viral particles used to infect a single cell. An MOI of 1-10 pfu/cell is typically used for optimal baculovirus-based insect cell protein expression. The viral inoculum required for protein expression analysis was calculated as:

$$\text{Inoculum required (ml)} = \frac{\text{MOI x number of cells}}{\text{viral titre (pfu/ml)}}$$

Bac-*Pf*GK* viral titres were determined from an average of three plated samples.

### 2.7.7 *Pf*GK* protein expression in TriEx™ insect cells

For protein expression, TriEx™ *Sf*9 cells were resuspended in 20ml of TriEx™ insect cell medium in a 125ml sterile Erlenmeyer flask at a concentration of 5 x $10^5$ cells/ml. Cells were incubated, with shaking at 150rpm, at 28°C until they reached a density of 2 x $10^6$ cells/ml. Bac-*Pf*GK* recombinant viruses were added to the culture at an MOI of 2, 5 and 10. Cells were cultured for 24-96 hours post-infection and protein expression analysis performed at each 24-hour interval. Cell density and viability were monitored using the trypan blue exclusion method (Appendix A14). Uninfected TriEx™ *Sf*9 insect cells were used as a control. Under a phase-contrast inverted microscope, infected cells appear uniformly rounded and enlarged, with distinct large nuclei, compared to uninfected cells.

### 2.7.8 r*Pf*GK*-His protein extraction

Insect cells were pelleted at 1,000g for five minutes and supernatants containing Bac-*Pf*GK* recombinant viruses were collected and stored at 4°C, in the dark. Cells were resuspended in 2ml histidine wash/bind buffer (100mM HEPES, 10mM imidazole, 500mM NaCl, pH7.5), supplemented with 2µl Protease Inhibitor Cocktail Set III. Samples were freeze-thawed twice using acetone and dry ice. Samples were added to the acetone-dry ice solution and incubated until the cell suspension froze (~5 minutes). Samples were then incubated at 37°C until they were completely thawed. The process was repeated once before sonication (section 2.4.1).

### 2.7.9 r*Pf*GK*-His protein purification

The sonicated cell pellet (section 2.4.1) was centrifuged at 16,100g for five minutes at 4°C to separate the soluble and insoluble cell fractions. The supernatant was collected and the pellet resuspended in 2ml histidine wash/bind buffer. Soluble r*Pf*GK* was purified using the Promega MagneHis™ Protein

Purification System (Promega Corporation, Madison, USA). MagneHis™ nickel particles are supplied as a 50% slurry and 1ml packed particles has a binding capacity of 1mg protein. Thirty microlitres MagneHis™ beads were added to the soluble protein fraction and r*Pf*GK* was allowed to bind to the beads overnight at 4°C with shaking using the Elmi RM-2M Intelli-Mixer. Beads were collected using a magnetic particle separator and non-specific binding of endogenous *E. coli* proteins was removed by washing the beads three times with 150µl histidine wash/bind buffer, each time incubating for five minutes with gentle agitation. r*Pf*GK* was eluted from the beads with 100µl elution buffer (100mM HEPES, 500mM imidazole, pH 7.5) during a 15 minute incubation at room temperature, with gentle agitation every two minutes. Residual r*Pf*GK* was eluted from the beads with 50µl elution buffer.

One hundred microlitre protein samples were collected for analysis at each cell lysis and purification step. All protein fractions were stored at -20°C. Fractions (5-20µl) were analyzed using SDS-PAGE electrophoresis (Appendix A10). The recombinant protein was detected using immunoblotting with a 1:2,000 diluted Penta•His™ HRP conjugated antibody (Qiagen GmbH, Hilden, Germany) (Appendix A11).

## 2.8 *Pf*GK gene knockout parasites

### 2.8.1 Generation of *Pf*GK gene knockout parasites (3D7Δ*Pf*GK)

*Pf*GK was disrupted using double-crossover homologous recombination DNA technology (Crabb and Cowman, 1996). Primers (PFK12/PfK13, Inqaba Biotechnical Industries, Pretoria, South Africa) were designed to amplify a 519bp 5'-*Pf*GK fragment (gene position 25-544bp), which was cloned into the *Spe*I/*Afl*II restriction sites in the pCC-1 plasmid, upstream of the human dihydrofolate reductase (hDHFR) positive selection cassette (Appendix A15). Similarly, a 515bp 3'-*Pf*GK fragment (gene position 966-1,481bp) was amplified with primers

PfK14/PfK15 and cloned into the *Eco*RI/*Nco*I restriction sites, downstream of hDHFR, to generate a pCC-1-*Pf*GK knockout plasmid.

Genetic manipulation of 3D7 *P. falciparum* parasites was performed following transfection of the pCC-1-*Pf*GK plasmid as described (Fidock and Wellems, 1997). One hundred micrograms purified pCC-1-*Pf*GK plasmid DNA in 30µl TE buffer was added to 370µl cytomix buffer (Appendix A15) and electroporated into predominantly ring-stage parasites (>5% parasitaemia) at 0.31kV and 950µF capacitance using a Bio-Rad GenePulse Xcell™ electrophorator (Bio-Rad Laboratories, Hercules, USA). WR99210 (Jacobus Pharmaceuticals, Princeton, USA; Appendix A15) was added 48 hours post-transfection to a final concentration of 5nM to select for transformed parasites (Fidock and Wellems, 1997). Fresh medium, supplemented with WR99210, was added to the cultures for the next three days and every alternate day thereafter. Following establishment of WR99210-resistant parasites, negative selection using 0.38µM 5-fluorocytosine (Sigma-Aldrich Corporation, St. Louis, USA; Appendix A15) against the cytosine deaminase/uracil phosphoribosyl transferase (CD/UPRT) cassette was performed to remove parasites with episomal pCC-*Pf*GK plasmid DNA (Maier *et al.*, 2006). Only parasites that had integrated the hDHFR cassette into the targeted *Pf*GK gene survived during negative selection. Parasites transfected with a pCC-1-EBA knockout plasmid was used as a control. pCC-1-EBA targets and disrupts the non-essential *P. falciparum* invasion protein, erythrocyte binding antigen-175 (EBA-175). Parasites were monitored every alternate day via Giemsa-stained smears (Appendix A3).

### 2.8.2 Verification of *Pf*GK knockout parasites

Genotyping was performed using both PCR-based and Southern hybridization analysis.

## *PCR*

Genomic DNA was isolated from both 3D7 and 3D7Δ*Pf*GK parasites as described (section 2.3.1) and various primer combinations were used to screen for knockout parasites via PCR. Integration at the 5'-end of the *Pf*GK gene was detected using primer set PfK03/Cam-R. Integration at the 3'end of the gene was detected using primers set HrpII-F/PfK04. Episomal plasmids and/or 3D7Δ*Pf*GK knockout parasites were detected using primer sets PfK12/Cam-R and HrpII-F/PfK15 for detection of the 519bp 5'- and 515bp 3'-*Pf*GK homologous regions in the pCC-1-*Pf*GK knockout plasmid respectively (Figure 36, Appendix A15 and Appendix A17). PCR was performed using the GoTaq® Green PCR Master Mix (Promega Corporation, Madison, USA). Each 20µl PCR reaction mixture consisted of 100ng purified DNA, 0.5µM of each primer and 10µl 2x GoTaq® Green PCR Master Mix. Fragments were amplified under conditions summarized in Table 7. pCC-1 and pCC-1-*Pf*GK plasmid DNA were used as controls. Amplicons were analysed using 1% agarose gel electrophoresis (Appendix A5).

**Table 7: PCR cycling conditions for the detection of 3D7Δ*Pf*GK parasites.**

| Segment | Cycles | Temperature | Time |
|---|---|---|---|
| 1 | 1 | 94ºC | 1 min |
| 2 | 1-29 | 94ºC | 1 min |
| | | 50ºC | 30 sec |
| | | 72ºC | 1 min |
| 3 | 1 | 72ºC | 5 min |
| 4 | 1 | 15ºC | Hold |

| PCR product sizes | | |
|---|---|---|
| Native | Full-length *Pf*GK | 1506bp |
| Episomal and/or integrated pCC-1-*Pf*GK | 5'-*Pf*GK | 519bp |
| | 3'-*Pf*GK | 515bp |
| Integrated pCC-1-*Pf*GK | 5'-*Pf*GK | 665bp |
| | 3'-*Pf*GK | 580bp |

## *Southern hybridization Analysis*

Southern hybridization analysis was performed using the Roche Digoxigenin (DIG) system as per manufacturer's instruction. A DIG-labelled probe was

prepared using the PCR DIG Probe synthesis Kit (Roche Diagnostics Gmbh, Mannheim, Germany) and primers PfK14/PfK15, which bind to the 515bp 3'-region of the *Pf*GK gene (Figure 36, Appendix A15 and Appendix A17). A PCR reaction was performed using the Roche Expand High Fidelity$^{PLUS}$ PCR System (Roche Diagnostics Gmbh, Mannheim, Germany). Each 50µl amplification reaction consisted of 100ng pCC-1-*Pf*GK plasmid DNA, 0.2µM each primer, 2.5 units Expand High Fidelity$^{PLUS}$ enzyme blend and 10µl 5x Expand High Fidelity$^{PLUS}$ reaction buffer (final concentrations: 1.5mM $MgCl_2$, 14µM DIG-dUTP, 26µM dUTP and 40µM each dATP, dCTP and dGTP). The probe was amplified under the conditions as summarized in Table 8 and analyzed using spectrophotometry and agarose gel electrophoresis (Appendix A5).

**Table 8: PCR cycling conditions for amplification of the DIG-labelled *Pf*GK probe.**

| Segment | Cycles | Temperature | Time |
|---|---|---|---|
| 1 | 1 | 94ºC | 1 min |
| 2 | 1-34 | 94ºC | 45 sec |
| | | 60ºC | 1 min |
| | | 68ºC | 1 min |
| 3 | 1 | 68ºC | 10 min |
| 4 | 1 | 15ºC | Hold |
| **3'-*Pf*GK PCR product size = 515bp** | | | |
| **DNA fragment sizes detected by the probe in the Southern hybridization analysis** | | | |
| pCC-1-*Pf*GK plasmid DNA | | 2,577bp and 3,534bp | |
| 3D7 wild type parasites | | 861bp and >10kb | |
| 3D7Δ*Pf*GK parasites and/or parasites with episomal pCC-1-*Pf*GK plasmid DNA | | 2,577bp, 3,534bp and >10kb | |
| 3D7Δ*Pf*GK parasites | | 2,577bp and >10kb | |

Similarly, a DIG-labelled probe was prepared using primers *Pf*EBA-175-sc_F/*Pf*EBA-175-sc_R (Appendix A15) and used to detect 3D7ΔEBA parasites. The primers bind to an internal 564bp fragment within the cloned 579bp 5'EBA-175 homologous recombination region that was used to generate the pCC-1-EBA knockout plasmid (Appendix A15). The probe was amplified under the conditions summarized in Table 9.

**Table 9: PCR cycling conditions for amplification of the DIG-labelled EBA-175 probe.**

| Segment | Cycles | Temperature | Time |
|---|---|---|---|
| 1 | 1 | 94ºC | 1 min |
| 2 | 1-29 | 94ºC | 1 min |
| | | 55ºC | 30 sec |
| | | 72ºC | 1 min |
| 3 | 1 | 72ºC | 5 min |
| 4 | 1 | 15ºC | Hold |
| **5'-EBA-175 PCR product size = 564bp** | | | |
| **DNA fragment sizes detected by the probe in the Southern hybridization analysis** | | | |
| pCC-1-EBA plasmid DNA | | 1,581bp | |
| 3D7 wild type parasites | | 8,585bp | |
| 3D7ΔEBA parasites and/or parasites with episomal pCC-1-EBA plasmid DNA | | 1,581bp, 7,190bp and 8,585bp | |
| 3D7ΔEBA parasites | | 7,190bp | |

**Southern blot**

Five micrograms parasite DNA was digested with five FastDigest® units of the appropriate FastDigest® enzyme (Fermentas International Inc., Burlington, Canada) in 1x FastDigest® buffer for two hours at 37°C. *Pst*I- and *Eco*RI/*Bam*HI-digested DNA was used to verify 3D7Δ*Pf*GK and 3D7ΔEBA parasites respectively. The digested DNA was resolved on a 30cm, 0.8% agarose gel in TEA buffer (Appendix A5) at 25V overnight. DIG-labelled DNA Molecular Weight Marker VII (Roche Diagnostics Gmbh, Mannheim, Germany) was used as a reference. The resolved DNA fragments were depurinated in 200ml 250mM HCl for 20 minutes and denatured in 200ml denaturing solution (0.5M NaOH, 1.5M NaCl) for 30 minutes, with shaking on a Hoefer Red Rotor PR70-230V platform. The gel was incubated in 200ml neutralization solution (0.5M Tris-HCl, pH 7.5, 1.5M NaCl) for 30 minutes, with shaking, and equilibrated in 200ml 20x SSC buffer (Appendix A16) for 10 minutes. DNA was transferred onto a 100cm$^2$ Hybond™-N Nylon membrane (Amersham Biosciences, Ltd., Buckinghamshire, UK) via capillary action for 16-18 hours in 20x SSC transfer buffer. The DNA was cross-linked to the membrane by exposure to UV light at 260nm (using the SynGene GeneGenius gel documentation system) for five minutes to immobilize the DNA. The membrane was placed in a 100ml Hybaid hybridization tube

(ThermoFisher Scientific, Inc., Waltham, USA) and 10ml 42°C pre-warmed DIG Easy Hyb hybridization buffer (Roche Diagnostics Gmbh, Mannheim, Germany) was added. The membrane was incubated in a Hybaid Shake 'n' Stack hybridization oven (ThermoFisher Scientific, Inc., Cambridge, UK) at 42°C for 30 minutes. The hybridization solution was aspirated and 10ml 42°C pre-warmed DIG Easy Hyb hybridization buffer supplemented with 20µl PCR-DIG-labelled probe (2µl DIG-labelled PCR probe per millilitre hybridization solution) was added to the membrane. The membrane was incubated with the probe at 42ºC for 16 hours. The probe/hybridization solution was collected and stored at -20°C (the probe can be re-used 3-5 times). The membrane was incubated with 200ml low stringency buffer (2x SSC, 0.1% SDS) at room temperature for 10 minutes, with shaking. The membrane was incubated with 200ml 65°C pre-warmed high stringency buffer (0.5x SSC, 0.1% SDS) at 65°C for 30 minutes, with shaking.

**DIG Chemiluminescent detection**

The high stringency buffer was discarded and fragments were detected and visualized using the DIG Luminescent Detection Kit (Roche Diagnostics Gmbh, Mannheim, Germany). The membrane was equilibrated in 100ml washing buffer (Appendix A16) at room temperature for two minutes, with shaking on a Hoefer Red Rotor PR70-230V platform. The buffer was discarded and 100ml freshly prepared blocking solution (Appendix A16) was added to the membrane and incubated for 30 minutes, with shaking. The buffer was discarded and 20ml freshly prepared 75mU/ml anti-DIG alkaline phosphatase conjugated antibody solution (1:10,000 antibody in blocking solution) was added to the membrane and incubated for 30 minutes, with shaking. The antibody solution was discarded and the membrane was washed twice, each time with 100ml washing buffer for 15 minutes with shaking. The membrane was equilibrated with 20ml detection buffer (0.1M Tris-HCl, 0.1M NaCl, pH 9.5) for three minutes. The membrane was placed on a plastic sheet, overlaid with 1ml CSPD (the alkaline phosphatase chemiluminescence substrate) and covered with a second plastic sheet to spread the CSPD substrate. The cassette was incubated for five minutes. Excess CSPD

was removed by gently pressing on the assembly and absorbing the liquid with paper towel. The membrane was incubated at 37°C for 15 minutes to enhance the luminescent reaction. Images were captured using the SynGene GeneGenius Chemidoc system, following a 1-20 minute exposure.

### 2.8.3 *Pf*GK mRNA expression in 3D7Δ*Pf*GK knockout parasites

#### *RNA extraction and purification*

Parasites were isolated from red cells using saponin lysis and resuspended in 1ml TRI Reagent™ (Sigma-Aldrich Corporation, St. Louis, USA) per 5ml parasite culture. RNA was extracted as per manufacturer's instruction and reverse transcribed into cDNA using SuperScript™ III Reverse Transcriptase (Invitrogen Ltd., California, USA).

Using an Eppendorf 5702R centrifuge, cells from a 5ml culture (>5% parasitaemia, predominantly trophozoites) were pelleted at 1,000g for five minutes at 4°C. Cells were washed once with 50ml PBS and pelleted at 1,000g for five minutes at 4°C. Cells were resuspended to a final volume of 2ml with PBS and the red cells lysed with 40µl 0.5% saponin (final concentration of 0.01%) for five minutes at room temperature. Parasites were pelleted via centrifugation at 1,000g for five minutes at 4°C and washed once with 15ml PBS. The parasites were resuspended in 1ml TRI Reagent™ and incubated for five minutes at room temperature. TRI Reagent™ is a mixture of phenol and guanidine isothiocyanate that disrupts cells and dissolves cellular components whilst maintaining the integrity of the RNA. Two hundred microlitres chloroform was added to the suspension, vortexed for 15 seconds and incubated at room temperature for 15 minutes. The solution was centrifuged in an Eppendorf 5415R centrifuge at 12,000g for 15 minutes at 4°C to separate the RNA (clear aqueous phase), DNA (white interphase) and protein (red organic layer) fractions. The aqueous phase was transferred to a sterile 1.5ml tube and 0.5ml isopropanol was added to precipitate the RNA. The solution was incubated for 10 minutes at room

temperature and the RNA was pelleted via centrifugation at 12,000g for 10 minutes at 4°C. RNA was washed once with 1ml ice-cold 75% isopropanol and pelleted via centrifugation at 12,000g for five minutes at 4°C. The RNA was air-dried for 10 minutes, resuspended in 20-30µl nuclease-free water and dissolved at 60°C for 15 minutes before being stored at -70°C.

To ensure that all *Pf*GK mRNA transcripts amplified in the RT-PCR reactions originated from RNA, the RNA template had to be devoid of any contaminating DNA. *Pf*GK exists as a single exon and residual DNA from the RNA extraction procedure would result in amplification of the *Pf*GK fragment from both the mRNA and the gene. Therefore, residual DNA, if any, was removed via digestion with RNase-free DNaseI (Fermentas International Inc., Burlington, Canada). Briefly, 1µg RNA was assembled in a 20µl reaction containing 1µl (1 unit) RNase-free-DNaseI, 2µl 10x reaction buffer (final concentrations: 10mM Tris-HCl pH 7.5, 2.5mM $MgCl_2$ and 0.1mM $CaCl_2$) and incubated for 10 minutes at 37°C. DNaseI was inactivated by adding 1µl 50mM EDTA followed by incubation at 65°C for 10 minutes. RNA integrity and yield were evaluated using spectrometry and agarose gel electrophoresis (Appendix A5). The DNA-free RNA was stored at -70°C.

## *PfGK mRNA expression*

To assess *Pf*GK mRNA expression, real-time PCR was used to amplify an internal 150bp *Pf*GK region that should have been displaced by the hDHFR cassette following the double crossover recombination event (Figure 36, Appendix A15 and Appendix A17). One microgram total RNA, 2pmol PfK05 *Pf*GK-gene specific reverse primer, 50µM each dATP, dTTP, dCTP and dGTP, 5mM DTT, 4µl 5x First-Strand buffer, 40 units RNase Inhibitor (Fermentas International Inc., Burlington, Canada) and 200units SuperScript™ III reverse transcriptase (Invitrogen Ltd., California, USA) were assembled in a 20µl reaction. cDNA was synthesized at 55°C for 60 minutes and the enzyme heat-inactivated at 70°C for 15 minutes. Four microlitres cDNA was added to a reaction mixture

containing 2µM PfK11, 2µM PfK05 and 10µl 2x SYBR® Green JumpStart Taq Readymix™ (Sigma-Aldrich Corporation, St. Louis, USA). PCR was performed on a LightCycler® platform (Roche Diagnostics Gmbh, Mannheim, Germany) under the conditions as summarized in Table 10. Single acquisition of fluorescence was detected using channel F1 and data were acquired after each extension step. The rate of temperature change was 20°C/second.

Following amplification, melting temperature analysis was performed as summarized in Table 11. Fluorescence was detected using channel F1 and data were acquired continuously at a change in temperature rate of 0.2°C/second. Melting temperature analysis was performed using the LightCycler® Software version 5.32. Primers designed to amplify a 206bp region of the *P. falciparum* rifin gene and a 68bp fragment from *P. falciparum* 18s rRNA gene were used as controls. Transcripts were resolved and analysed on a 1% agarose gel (Appendix A5).

**Table 10: Real time RT-PCR cycling conditions for the detection of *Pf*GK.**

| Segment | Cycles | Temperature | Time |
|---|---|---|---|
| 1 | 1 | 95ºC | 30 sec |
| 2 | 1-29 | 95ºC | 10 sec |
| | | 50ºC | 10 sec |
| | | 72ºC | 30 sec* |
| 3 | 1 | 72ºC | 5 min |
| 4 | 1 | 15ºC | Hold |

| RT-PCR product sizes | |
|---|---|
| *Pf*GK | 150bp |
| Rifin | 206bp |
| 18s rRNA | 68bp |

*single fluorescence acquisition

**Table 11: *Pf*GK melting temperature conditions.**

| Segment | Cycles | Temperature | Time |
|---|---|---|---|
| 1 | 1 | 95ºC | 5 sec |
| 2 | 1 | 60ºC | 30 sec* |
| | | 95ºC | |
| 3 | 1 | 15ºC | Hold |

*continuous fluorescence acquisition

## 2.9 *In vitro* growth analysis of 3D7Δ*Pf*GK parasites

Ring-stage parasites were tightly synchronized as previously described (Lambros and Vanderberg, 1979), twice per cycle at 6-8 hours apart for three consecutive 48-hour cycles. Five millilitres 3D7, 3D7Δ*Pf*GK and 3D7ΔEBA parasites were each cultured in 25cm$^2$ flasks, in duplicate. Cultures were measured at a starting parasitaemia between 0.5-2.0% rings and samples were acquired over one 48-hour cycle for analysis. Three independent experiments were performed and a total of 14 samples were used in the analysis. Parasitaemia was monitored using flow cytometry and thiazole orange staining as previously described (Makler *et al.*, 1987), with modifications. Thiazole orange stains RNA in reticulocytes and DNA in peripheral blood mononuclear cells. A 10mM thiazole orange (Sigma-Aldrich Corporation, St. Louis, USA) stock solution was prepared in methanol and stored at room temperature, away from light. Thiazole orange was diluted to 1µM in Sorensen's phosphate buffer pH 7.2 (Appendix A3) and 1ml was added to 5µl parasite culture samples. Uninfected red cells were used as a control. The suspensions were mixed and incubated at room temperature for 20 minutes, in the dark. Parasitaemia was measured using a Beckman Coulter FC500 Flow Cytometer (Beckman Coulter, Inc., Fullerton, USA), with excitation by argon laser at 488nm and detection in the FL1 channel. Data were analysed using the accompanying CXP software. Post analysis was performed using the Kaluza Flow Cytometry Analysis Software (Beckman Coulter, Inc., Fullerton, USA). A population of erythrocytes was delineated by an inclusion gate on a forward-scatter vs side-scatter two-parameter dot blot. The selected population was analysed on a single-parameter histogram for FL1 fluorescence (458V, 1.0 gain, 4.99mW laser power). Inclusion gates for uninfected erythrocytes, ring and trophozoite/schizont stages were established using an asynchronous culture (Figure 43), with morphological confirmation via microscopy. Background counts from the erythrocyte control culture were subtracted from parasitized samples. All scales were logarithmic and a minimum of 50,000 events were acquired.

Giemsa-stained smears were examined and the number of schizonts formed per 3D7-, 3D7ΔEBA- and 3D7Δ*Pf*GK-infected red blood cell was counted. An unpaired t-test with a two-tailed p-value was used to compare the means between the data sets.

# Chapter 3: Results

## 3.1 Expression of r*Pf*GK in *E. coli*

### 3.1.1 Subcloning *Pf*GK into pGEX-4T-2

The full length *Pf*GK gene (Appendix A1) was amplified as a 1,518bp fragment (primers PfK03/PfK04; Figure 13) and cloned into the *Bam*HI/*Xho*I restriction sites of the pGEX-4T-2 plasmid vector (Appendix A6 and Appendix A17).

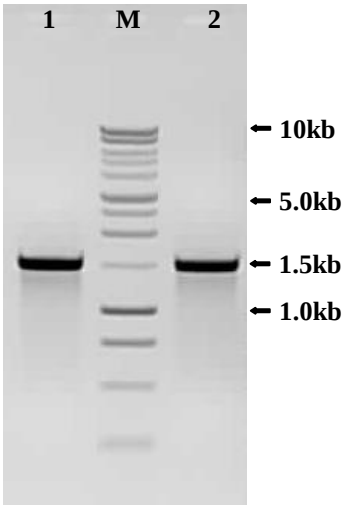



**Figure 13: Amplification of *Pf*GK.**
High fidelity PCR amplification of the full length glycerol kinase gene. Five microlitres of each *Pf*GK amplicon (lanes 1 and 2) were resolved on a 1% agarose gel. The PCR products migrated at approximately 1,500bp. Lanes are numbered and sizes of selected fragments in the 2.5µl Promega 1kb DNA marker (M) are indicated.

Transformed DH5α *E. coli* colonies were screened using primers specific to regions flanking the pGEX-4T-2 multiple cloning site (primers SE-pGEX-F/SE-pGEX-R; Appendix A6). Successful ligation of *Pf*GK into pGEX-4T-2 resulted in amplification of a 1,661bp fragment (Figure 14a, lanes 4-6) in pGEX-4T-2-*Pf*GK recombinant constructs.

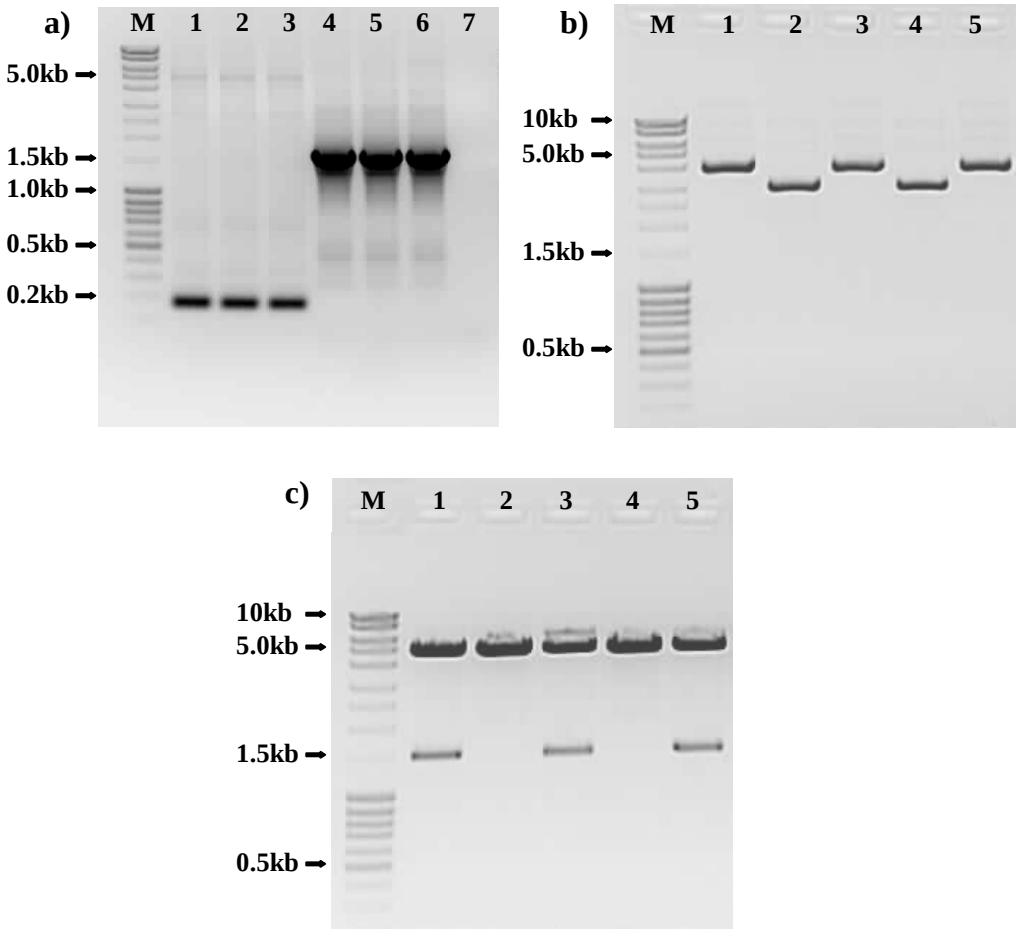



**Figure 14: Verification of pGEX-4T-2- *Pf*GK transformed *E. coli*.**
**(a)** Colony PCR using pGEX-4T-2 vector-specific primers (SE-pGEX-F/SE-pGEX-R). One hundred nanograms pGEX-4T-2 plasmid DNA was used as a positive control (lane 1) and 100ng *P. falciparum* DNA was used as a negative control (lane 7). Five microlitres of each sample were resolved on a 1% agarose gel. Lanes 2-3 are transformed DH5α colonies containing the pGEX-4T-2 plasmid. The 174bp amplicons represent re-circularized pGEX-4T-2 plasmid DNA during the ligation reactions. Lanes 4-6 represent clones with recombinant pGEX-4T-2-*Pf*GK plasmid DNA (1,661bp amplicons); **(b)** plasmid DNA was extracted from transformed colonies and 500ng of each plasmid was resolved on a 0.8% agarose gel. Lanes 2 and 4 represent pGEX-4T-2 plasmid DNA whereas lanes 1, 3 and 5 represent plasmid DNA extracted from pGEX-4T-2-*Pf*GK transformed colonies; **(c)** restriction analysis of extracted recombinant plasmids. Five hundred nanograms pGEX-4T-2 and pGEX-4T-2-*Pf*GK plasmid DNA were digested with *Bam*HI and *Xho*I restriction enzymes and resolved on a 0.8% agarose gel. The *Pf*GK insert was separated from pGEX-4T-2-*Pf*GK DNA and resolved as 4,945bp (pGEX-4T-2) and 1,512bp (*Pf*GK) fragments (lanes 1, 3 and 5) respectively. pGEX-4T-2 without the insert migrated as 4,945 and 25bp fragments. Lanes are numbered and sizes of selected fragments in the 2.5µl Fermentas MassRuler$^{TM}$ DNA Ladder Mix (M) are indicated.

Unsuccessful DNA ligation reactions resulted in amplification of a 174bp fragment, representing re-circularized pGEX-4T-2 plasmid DNA (Figure 14a, lanes 2-3). pGEX-4T-2-*Pf*GK transformed colonies were also screened using PfK03 and PfK04 *Pf*GK gene specific primers (data not shown) for further verification.

Plasmid DNA was extracted from transformed colonies (Figure 14b) and digested with *Bam*HI and *Xho*I restriction enzymes to separate the pGEX-4T-2 plasmid backbone from the inserted *Pf*GK gene (Figure 14c). DNA sequencing analysis verified that the gene was in-frame (Appendix A9). Despite a few errors in the sequence (T255C, A279T and T430C) the amino acids remained unchanged.

### 3.1.2 Expression of r*Pf*GK

A summary of the recombinant proteins used in this study, as well as selected molecular features, is presented in Table 12. Bioinformatic analysis predicted that the native *Pf*GK recombinant protein would be predominantly expressed as an insoluble aggregate in *E. coli*. Fusion to the N-terminal GST tag slightly improves the solubility profile. Codon optimization of the *Pf*GK gene (*Pf*GK$^{Sf}$) for expression in insect cells also predicts that the C-terminal His-tagged *Pf*GK$^{Sf}$ protein would be expressed as an insoluble protein in *E. coli*.

**Table 12: Primary sequence analysis of *Pf*GK.**

| Protein | Amino acids | Molecular Weight (kDa)[1] | pI[1] | Extinction coefficient ($M^{-1}cm^{-1}$)[1] | Vector tag | Stability[1] | Percentage solubility[2] |
|---|---|---|---|---|---|---|---|
| Native *Pf*GK | 501 | 56.65 | 5.92 | 80,330 | N/A[3] | Stable | 35.1 |
| r*Pf*GK | 727 | 82.96 | 6.00 | 123,190 | N-GST | Stable | 40 |
| rGST | 220 | 25.70 | 5.90 | 42,860 | N/A | Stable | 58 |
| *Pf*GK$^{Sf}$ | 521 | 59.11 | 6.13 | 80,330 | C-8xHis | Stable | 39.2 |

[1]Parameters predicted using the ProtParam Tool (http://web.expasy.org/protparam)
[2]Predicted soluble recombinant protein expression in *E. coli* (http://www.biotech.ou.edu/)
[3]N/A = not applicable

r*Pf*GK was expressed as a GST-fusion protein following induction of a 10ml culture with 1mM IPTG. Aliquots of total, soluble (a total volume of ~1.8ml) and insoluble (pellet resuspended in PBS to ~1.8ml) protein fractions were loaded on an SDS-PAGE gel (Figure 15).

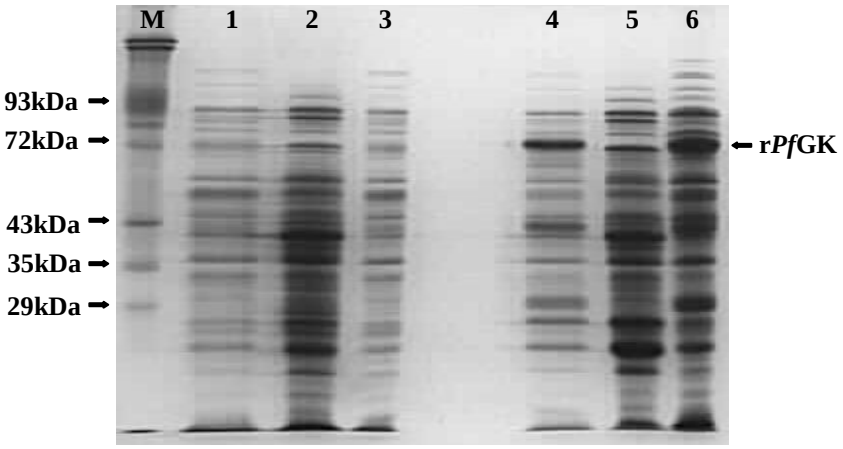



**Figure 15: Expression of r*Pf*GK in *E. coli*.**
Rosetta2 (DE3) *E. coli* were grown at 37°C for six hours and r*Pf*GK protein expression was induced with IPTG for 16-18 hours. Cells were lysed using BugBuster® HT protein extraction reagent. Ten microlitres total, 15µl soluble (from a total of 1.8ml) and 10µl insoluble protein samples were loaded onto an SDS-PAGE gel and r*Pf*GK protein expression was analyzed using Coomassie blue staining. Lanes 1-3 and lanes 4-6 are uninduced and induced r*Pf*GK cultures respectively. Lanes 1 and 4 are total protein fractions, Lanes 2 and 5 are soluble protein fractions and Lanes 3 and 6 are the insoluble protein fractions. M is 5µg of the red cell membrane protein ladder. r*Pf*GK migrated at approximately 72kDa. The recombinant protein was predominantly present in the insoluble protein fraction.

As predicted, r*Pf*GK was expressed mainly as an insoluble protein (Figure 15, lane 6). This is not an uncommon phenomenon as expression of heterologous proteins, especially *Plasmodium* proteins, in the cytoplasm of *E. coli* is often accompanied by the formation of misfolded proteins trapped as insoluble aggregates (inclusion bodies). Soluble r*Pf*GK could not be detected when stained with Coomassie dye (Figure 15, lane 5), probably due to the low sample volume loaded on the Laemmli gel. r*Pf*GK migrated at 72.8kDa (Appendix A12), which is smaller than the predicted 83kDa GST-fusion protein. The coding sequence, however, was correct and in frame. The general assumption is that SDS-coated polypeptides exist as random coils, with constant charge to mass and mass to

length ratios. SDS-coated polypeptides should therefore separate according to size. However, a protein with incompletely disrupted secondary structures will not unfold to its full length and, assuming constant charge, will migrate faster than the expected size in the gel. The reproducible smaller size of the r*Pf*GK protein may be due to the consistent preparation of samples for analysis (Hjelmeland *et al.*, 1981). The size difference may also imply a truncated r*Pf*GK protein. ADP-binding residues essential for glycerol kinase activity are situated at the C-terminal end of the protein sequence. An approximately 10kDa (~91 amino acids) deletion at the C-terminus will include some residues involved in ADP-binding. The truncated protein therefore should have no enzyme activity and an enzyme assay should reflect this. The calculated Mw of the BSA (66.8kDa) and GST polypeptides (26.5kDa) from the Coomassie-stained SDS-PAGE gels were similar to their actual Mw of 66.5kDa and 26kDa, respectively (Appendix A12).

Four pGEX-4T-2-*Pf*GK transformed Rosetta2 (DE3) colonies were screened for soluble r*Pf*GK protein expression by SDS-PAGE and immunoblotting (Figure 16 and Appendix A11).

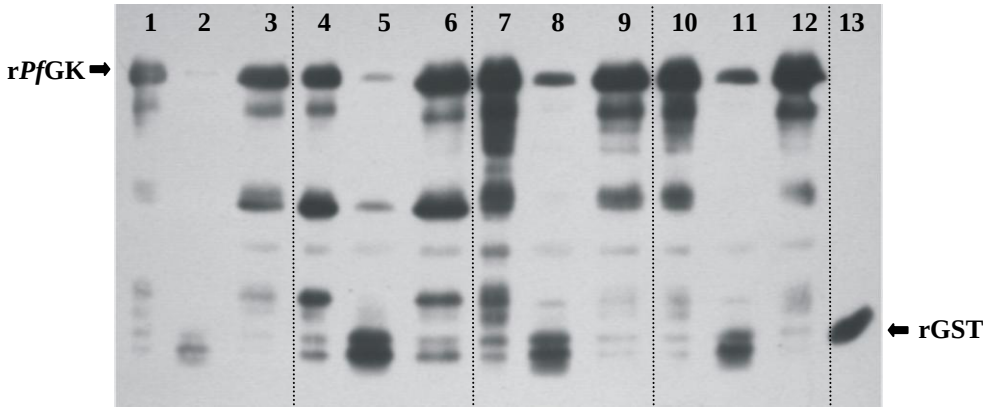



**Figure 16: Immunoblot analysis of r*Pf*GK.**
Ten microlitres total, 15µl soluble (from a total of 1.8ml) and 10µl insoluble protein fractions from a 10ml pGEX-4T-2-*Pf*GK induced culture were loaded onto an SDS-PAGE gel and analyzed via immunoblotting using an anti-GST-HRP antibody. Five microlitres of the soluble fraction of a pGEX-4T-2 induced clone was used as a rGST positive control (lane 13). Four colonies were analysed: 1 (lanes 1-3), 2 (lanes 4-6), 3 (lanes 7-9) and 4 (lanes 10-12). Soluble r*Pf*GK (lanes 2, 5, 8 and 11) was detected in all colonies, although the band in lane 2 was very faint. Consistent with the Coomassie-stained SDS-PAGE gels, most of the expressed r*Pf*GK protein was insoluble (lanes 3, 6, 9 and 12).

r*Pf*GK expression was induced in all four colonies. Immunoblot analysis confirmed that r*Pf*GK is predominately expressed as an insoluble protein. Soluble r*Pf*GK was however detected in each colony, with lower amounts in colonies 1 (very faint band) and 2 compared to colonies 3 and 4. Differences in the amount of soluble r*Pf*GK proteins between each colony may be due to variations in cell lysis or cell densities prior to extraction. Several other protein bands were also present (Figure 16). These were probably truncations of r*Pf*GK due to premature termination of translation or products from proteolysis during sample preparation, although protease inhibitors had been used. Two additional bands at approximately the same size of rGST (~26kDa) were also present in each soluble r*Pf*GK fraction. Expression of genes containing rare codons, like *P. falciparum* proteins, may result in ribosomal stalling leading to translational errors such as amino acid substitutions, frameshift mutations or premature translational termination. *Pf*GK was cloned downstream of the *Schistosoma japonicum* GST tag in the pGEX-4T-2 vector (GE-Healthcare, 2007) and the difference in the codon usage of the GST and *Pf*GK genes when expressed in *E. coli* may result in translational termination, just before translation of the *Pf*GK protein. This may result in the expression of rGST instead of the full length rGST-tagged protein (Sorensen and Mortensen, 2005a). Alternatively, the ~26kDa band could represent endogenous *E. coli* GST.

### 3.1.3 Refolding r*Pf*GK from inclusion bodies

r*Pf*GK was predominantly expressed as insoluble proteins trapped as inclusion bodies in *E. coli*. Proteins were purified from the insoluble protein fraction, denatured using either urea or guanidine-HCl and refolded via dialysis. Most of the denatured r*Pf*GK precipitated during the refolding procedures, implying incorrectly folded protein. Depending on the refolding method, ~80-800ng refolded soluble protein per 1mg of denatured inclusion bodies was purified (Figure 17 and Figure 18). Enzyme activity analysis, however, showed that the refolded protein was not active (data not shown).

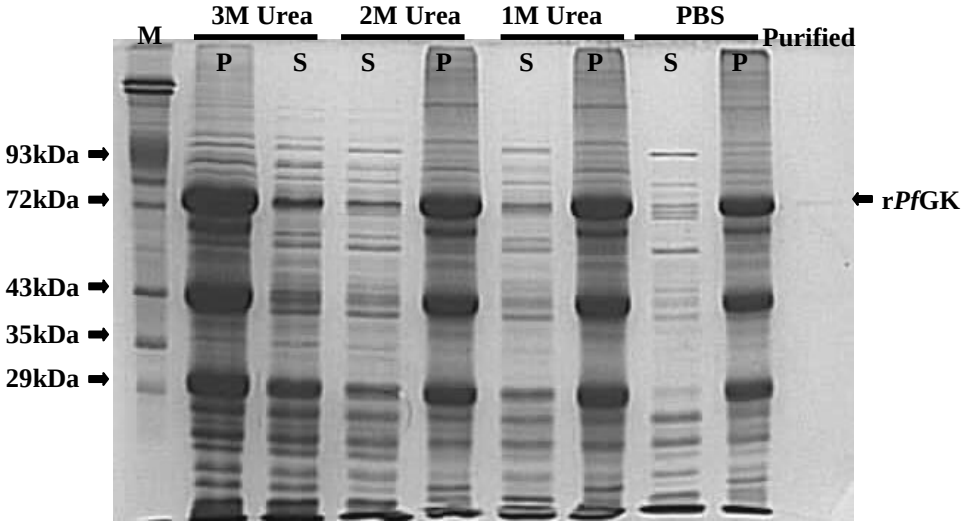



**Figure 17: Refolding r*Pf*GK using urea.**
One hundred micrograms denatured r*Pf*GK in 100µl in Guanidine-hydrochloride solution (section 2.5.1) was dialyzed against a serial dilution series of urea in PBS. Samples were collected after each dialysis step and centrifuged to separate precipitated (P) from soluble (S) protein. Refolded r*Pf*GK was purified using MagneGST™ agarose beads. Five microlitres P, 10µl S and 25µl (from a total of 200µl) purified protein samples were resolved and analyzed using SDS-PAGE. The majority of denatured r*Pf*GK precipitated during the refolding process. Approximately 80ng soluble r*Pf*GK was purified from 100µg of denatured protein. M is 5µg of the red cell membrane marker.

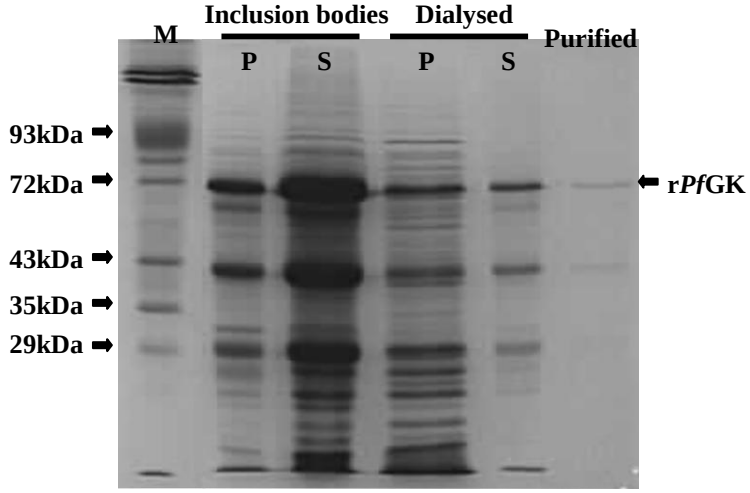



**Figure 18: Refolding r*Pf*GK using GnHCl.**
Ten milligrams denatured r*Pf*GK in 6M GnHCl was dialyzed against a 200ml Tris-based buffer for 48 hours at 4°C. Samples were collected before and after dialysis and centrifuged to separate precipitated (P) from soluble (S) protein. Refolded r*Pf*GK was purified using MagneGST™ agarose beads. Five microlitres P, 10µl S and 25µl (from a total of 200µl) purified protein samples were resolved and analyzed using SDS-PAGE. The majority of denatured r*Pf*GK precipitated during the refolding process. Approximately 800ng soluble r*Pf*GK was purified from 10mg of denatured protein. M is 5µg of the red cell membrane marker.

### 3.1.4 Expression of soluble r*Pf*GK

Recombinant proteins were induced with either IPTG or Novagen's Overnight Express™ autoinduction medium. In both cases, the expressed recombinant protein was predominantly insoluble (comparison data not shown). r*Pf*GK was purified from a 25ml culture (Figure 19), using 20µl MagneGST™ beads (4µl packed beads). However, the low yield (~800ng r*Pf*GK per 25ml culture; Appendix A10) usually accompanied by poor purity, was insufficient for enzyme activity analysis.

To improve r*Pf*GK yield and purity, conditions such as IPTG concentration used for induction, time of induction, cell growth volumes and density, induction temperature, cell lysis and protein purification protocols were modified.

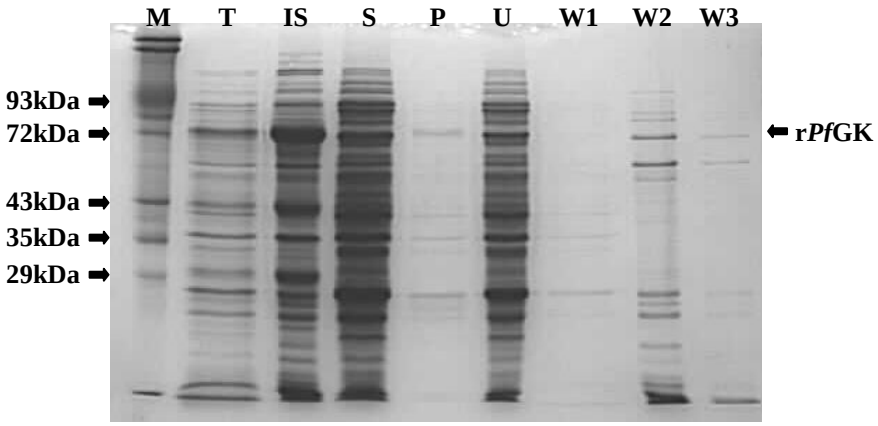



**Figure 19: Small scale expression and purification of r*Pf*GK in *E. coli*.**
r*Pf*GK expression was induced in a 25ml culture using Novagen's Overnight Express™ medium. Cells were lysed using BugBuster® HT and r*Pf*GK purified from the soluble protein fraction using MagneGST™ agarose beads. SDS-PAGE analysis shows induction of r*Pf*GK in the total (T; 10µl) protein fraction. Most of the protein was in the insoluble fraction (IS; 10µl) and visually undetectable in the soluble fraction (S; 15µl from a total of 1.8ml). r*Pf*GK was purified (P; 25µl from a total of 200µl), although the purity was poor. Samples from the unbound fraction (U; 10µl) and from each wash step (W1-3; 10µl) during the purification procedure are shown. M is 5µg of the red cell membrane marker.

To obtain sufficient protein for analysis, the culture volume was scaled-up to 800ml. Four 1L Erlenmeyer flasks each containing 200ml LB without antibiotics

were each inoculated with 2ml of an antibiotic-selected culture and induced for 16-18 hours (Figure 20).

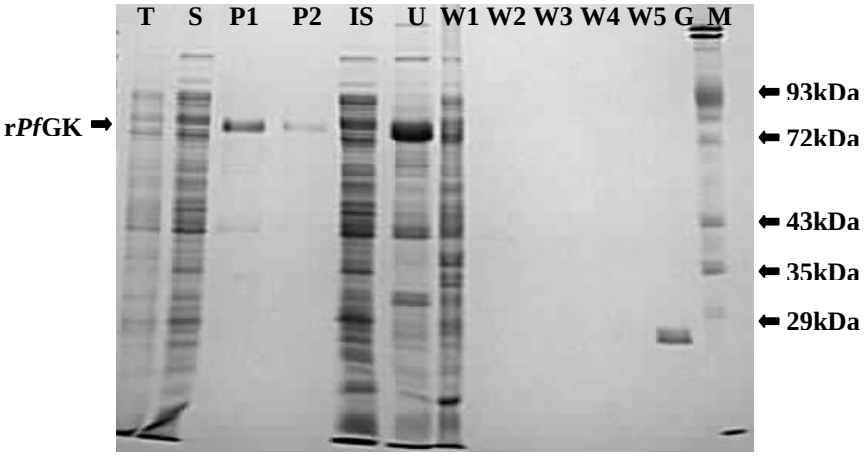



**Figure 20: Large-scale expression and purification of soluble r*Pf*GK in *E. coli*.**
r*Pf*GK was induced in an 800ml culture using Novagen's Overnight Express™ medium. Cells were lysed using sonication and r*Pf*GK purified using MagneGST™ agarose beads. The total (T; 10 µl), soluble (S; 15µl from a total of 10ml) and insoluble (IS; 10µl) protein fractions were analyzed using SDS PAGE. r*Pf*GK was eluted twice from the GST beads using reduced glutathione (P1; 25µl from a total of 200µl and P2; 25µl from a total of 100µl). Samples from the unbound fraction (U; 10µl) after binding to the GST-beads and from each wash step (W1-5; 10µl) during the purification procedure are shown. rGST (G; 5µl) was used as a control for the MagneGST™ beads. M is 5µg of the red cell membrane marker.

No difference was observed when cultures were induced with 0.4 or 1mM IPTG or Overnight Express™ medium (data not shown). Soluble r*Pf*GK protein yield and purity were improved and were more consistent by (data not shown):

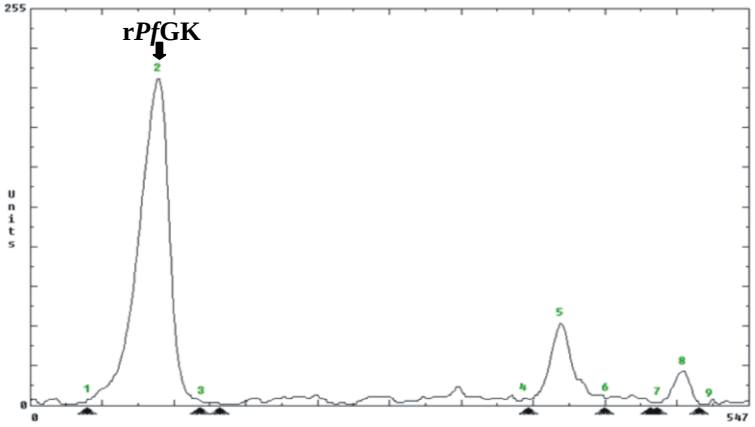
1. inducing protein expression at 18-22ºC instead of at 37ºC,
2. increasing the induction time (when using IPTG) from three hours to 16-18 hours,
3. using sonication instead of the BugBuster® HT chemical reagent to lyse cells,
4. increasing the binding time of the soluble protein fraction, following cell lysis, to the MagneGST™ beads from 30 minutes to 6-18 hours at 4ºC,
5. increasing the stringency of the wash conditions with the GST wash/bind buffer during the protein purification process. The wash volume was increased

from five volumes to 50 volumes per packed volume of beads and the number of wash steps was increased from three to five washes,

6. increasing the concentration of reduced glutathione in the MagneGST$^{TM}$ elution buffer from 50mM to 500mM.

As a result, approximately 3µg of soluble r*Pf*GK at approximately 80-90% purity (Figure 20, Figure 21 and Appendix A10) was purified from an 800ml culture. The yield was relatively low but sufficient for downstream analysis. Immunoblot analysis (Figure 22 and Appendix A11) confirmed the presence of r*Pf*GK.

**a)**




**b)**

| N | Height | Width | Area | %Area | %Protein |
|---|---|---|---|---|---|
| 2 | 210 | 86 | 5790 | 65.8 | 78.9 |
| 5 | 53 | 58 | 1199 | 13.6 | 16.3 |
| 8 | 22 | 32 | 351 | 4.0 | 4.8 |

**Figure 21: Densitometric analysis of r*Pf*GK.**
**(a)** Coomassie blue stained proteins resolved on SDS-PAGE gels were scanned on a Hoefer GS300 transmittance/reflectance scanning densitometer. The r*Pf*GK protein band is indicated; (**b)** the raw volumes or percentage area of each protein band was determined using the Hoefer GS365W software. A tabulated summary of the results is presented. In this example, analysis showed a 78.9% r*Pf*GK purity.

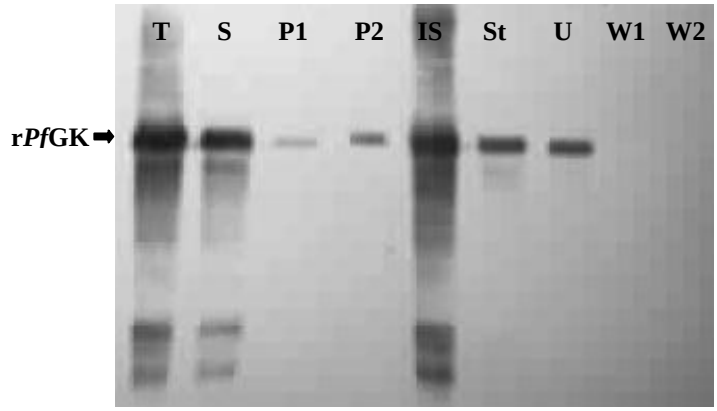



**Figure 22: Immunoblot analysis of expressed and purified r*Pf*GK in *E. coli*.**
Proteins were transferred onto a nitrocellulose membrane and an anti-GST-HRP conjugated antibody was used to detect the rGST-*Pf*GK fusion protein. The total (T; 10µl), soluble (S; 15µl from a total of 10ml) and insoluble (IS; 10µl) protein fractions as well as purified r*Pf*GK after each elution step (P1; 10µl from a total of 200µl and P2; 20µl from a total of 100µl) are indicated. Residual proteins bound to the purification beads were "stripped" (St; 15µl) from the beads for analysis. Samples from the unbound fraction (U; 10µl) and the wash steps (W1 and W2; 20µl each sample) during the purification process are shown. In this example, the poor binding efficiency of the MagneGST$^{TM}$ agarose beads resulted in a reduced purification efficiency since excess r*Pf*GK was present in the unbound fraction.

# 3.2 Expression of r*PfGK* in insect cells

## 3.2.1 Subcloning *Pf*GK* and *Pf*GK$^{Sf}$ into pTriEx-3

The full length *Pf*GK* without the stop codon was amplified as a 1,516bp fragment (primers PfK08/PfK06) and cloned into the pTriEx-3 plasmid vector (Appendix A6). Colony PCR using pTriEx-3 vector- and *Pf*GK* gene-specific primers, plasmid DNA extraction and *Bam*HI/*Xho*I restriction enzyme analysis verified the generation of pTriEx-3-*Pf*GK* recombinant plasmid DNA (data not shown). However, DNA sequencing analysis showed a nucleotide deletion in a 19bp adenine-rich region between positions 1,441-1,459 of the gene (Figure 59 and Appendix A9). The adenine deletion was seen in the same region after three independent cloning attempts and using alternative high fidelity *Taq* polymerases (Roche Expand High Fidelity$^{PLUS}$ PCR System or the Fermentas High Fidelity PCR Enzyme Mix; Fermentas International Inc., Burlington, Canada). Even

though the glycerol- and ATP- binding sites were unaffected, the rest of the sequence was not in-frame. The translated *Pf*GK* protein may be functional; however it would not have the C-terminal 8x histidine tag necessary for purification.

A decision was taken to codon optimize and synthesize the *Pf*GK$^{Sf}$ gene for expression in insect cells. GeneArt codon optimized, synthesized and cloned the gene into the pMK plasmid vector (Appendix A13). The supplied pMK-*Pf*GK$^{sf}$ plasmid DNA (3,810bp; Appendix A13) was resuspended in 50µl 10mM Tris-HCl pH 8.0 to a concentration of 100ng/µl and stored at -20°C. Fifty microlitres competent DH5α *E. coli* was transformed with 50ng (0.5µl) of the pMK-*Pf*GK$^{sf}$ plasmid (section 2.3.5). Five millilitres LB supplemented with 20µg kanamycin was inoculated with a pMK-*Pf*GK$^{sf}$-transformed colony and incubated in a Labotec® Orbital Shaker at 250rpm at 37°C for 16-18 hours. pMK-*Pf*GK$^{sf}$ plasmid DNA was extracted using the Sigma GenElute™ Plasmid Miniprep Kit (section 2.3.3). One microgram pMK-*Pf*GK$^{sf}$ plasmid DNA was digested in a 20µl reaction containing 1FDU (1µl) FastDigest® Fermentas *Bam*HI and 1FDU (1µl) FastDigest® *Xho*I restriction enzymes in 1x FastDigest® buffer at 37°C for 10 minutes. The pMK (2,294bp) and *Pf*GK$^{sf}$ (1,516bp) fragments were resolved on a 0.8% agarose gel and the *Pf*GK$^{sf}$ band was excised from the gel using a scalpel under a UV transilluminator (UVP, Upland, USA). *Pf*GK$^{sf}$ was purified using the QIAquick® Gel Extraction Kit as per manufacturer's instruction (section 2.3.4) and ligated into the *Bam*HI/*Xho*I restriction sites of the pTriEx-3 vector (sections 2.3.3 and 2.3.4) to generate the pTriEx-3-*Pf*GK$^{Sf}$ recombinant plasmid. PCR analysis (Figure 23a) using pTriEx-3 vector primers (SE-pTriEx-F/SE-pTriEx-R; Appendix A6) and *Pf*GK$^{sf}$ gene-specific primers (PfK16/PfK17; Appendix A13), plasmid DNA extraction and *Bam*HI/*Xho*I restriction enzyme analysis (Figure 23b) verified the generation of pTriEx-3-*Pf*GK$^{Sf}$ recombinant plasmid DNA.

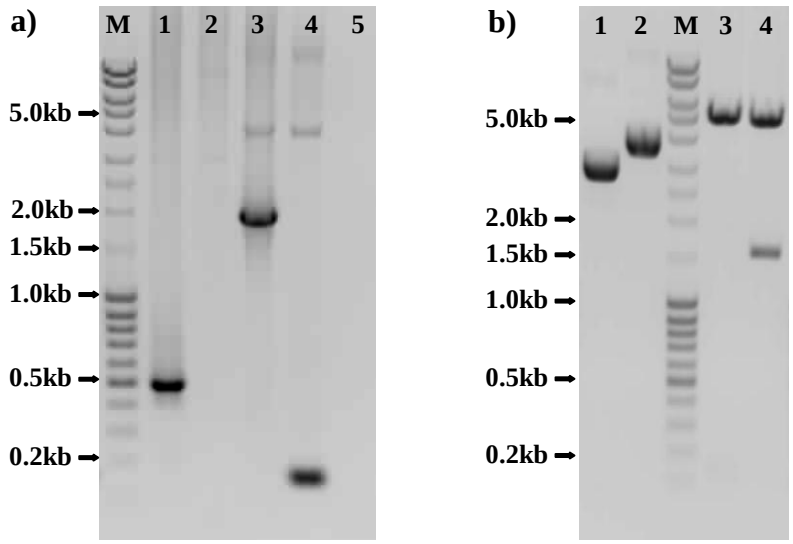



**Figure 23: Generation of pTriEx-3-*Pf*GK$^{Sf}$ recombinant plasmid DNA.**
The *Pf*GK gene was cloned into the pTriEx-3 multiple expression vector. Cloning was verified by **(a)** PCR of transformed DH5α colonies and **(b)** analysis of recombinant plasmids. **(a)** PCR using pTriEx-3 vector-specific primers (SE-pTriEx-3-F/SEpTriEx-3-R) and *Pf*GK$^{Sf}$ gene-specific primers (PfK16/17). Samples were resolved on a 1% agarose gel. Lanes 1 and 2 represent colonies containing the re-circularized pTriEx-3 plasmid. Vector-specific primers amplified a 496bp amplicon (lane 1) and gene-specific primers did not amplify the pTriEx-3 plasmid (lane 2). Lanes 3 and 4 are colonies containing recombinant pTriEx-3-*Pf*GK$^{Sf}$ plasmid DNA. PCR using vector-specific primers resulted in amplification of a 1,885bp fragment (lane 3) and a 150bp fragment (lane 4) when using gene-specific primers. A control reaction with no template DNA is shown (lane 5); **(b)** plasmid DNA was extracted from pTriEx-3 (lane 1) and pTriEx-3-*Pf*GK$^{sf}$ (lane 2) transformed colonies and resolved on a 0.8% agarose gel. Plasmid DNA was digested with *Bam*HI and *Xho*I restriction enzymes. The *Pf*GK insert was excised from pTriEx-3-*Pf*GK$^{Sf}$ DNA and resolved as 4,961bp (pTriEx-3) and 1,510bp (*Pf*GK$^{Sf}$) fragments respectively (lane 4). pTriEx-3 without the insert migrated as 4,961bp and 121bp (not visible) fragments (lane 3). Lanes are numbered and sizes of selected fragments in the 2.5µl Fermentas MassRuler$^{TM}$ DNA Ladder Mix (M) are indicated.

## 3.2.2 Generation of Bac-*Pf*GK$^{Sf}$ recombinant baculoviruses

Recombinant baculoviruses were generated using the BacMagic$^{TM}$ DNA Kit and after 3 rounds of amplification, viral DNA was extracted. Cell morphology and viability were used as indicators for the presence of replicative baculoviruses. Infected cells appeared grainy, uniformly rounded and enlarged with distinct enlarged nuclei compared to healthy cells. The growth rate of infected cells was significantly reduced compared to the 24-hour doubling rate of a healthy culture. A 2 x $10^6$ *Sf*9 cells/ml culture typically had a cell density of ~10 x $10^6$ *Sf*9 cells/ml

and viability of >90% after 72 hours. Comparatively, baculovirus-infected cells had a density of ~3 x $10^6$ *Sf*9 cells/ml and ~70% viability after 72 hours. PCR analysis using *Pf*GK$^{Sf}$ gene-specific primers (PfK16/PfK17; Figure 24, lane 2) verified the generation of recombinant Bac-*Pf*GK$^{Sf}$ baculoviruses.

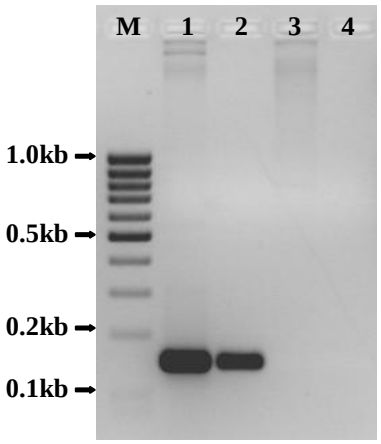



**Figure 24: Screening for Bac-*Pf*GK$^{Sf}$ recombinant baculoviruses.**
DNA was extracted from Bac-*Pf*GK$^{Sf}$ baculoviruses and PCR was performed using 100ng template DNA and *Pf*GK$^{Sf}$-gene specific primers (PfK16/PfK17). Five microlitres of each amplicon was resolved on a 1% agarose gel. Amplification of a 150bp internal *Pf*GK$^{Sf}$ gene fragment verified the generation and presence of Bac-*Pf*GK$^{Sf}$ baculoviruses (lane 2). Amplification from pTriEx-3-*Pf*GK$^{Sf}$ (lane 1) and pTriEx-3 plasmid DNA (lane 3) were used as positive and negative controls respectively. A blank reaction with no template DNA is shown (lane 4). Lanes are numbered and sizes of selected fragments in the 2.5µl Fermentas MassRuler$^{TM}$, Low Range DNA Ladder Mix (M) are indicated.

A plaque assay was performed and an average baculoviral titre of 1.23 x $10^9$ pfu/ml was determined (Figure 25). For optimal protein expression, insect cells were infected with 1-10 viral particles per cell (MOI between 1-10).

## 3.2.3 Expression of r*Pf*GK$^{Sf}$

A preliminary protein expression screen was performed by infecting 20ml TriEx$^{TM}$ *Sf*9 cells, at a concentration of 2 x $10^6$ cells/ml and a viability of 95%, with Bac-*Pf*GK$^{Sf}$ baculoviruses at an MOI of 10.

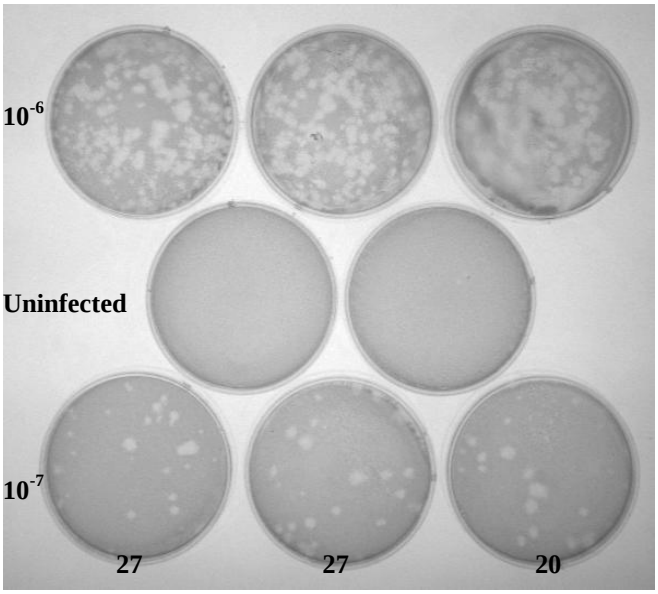



**Figure 25: Determining Bac-*Pf*GK$^{Sf}$ baculoviral titre.**
A viral plaque assay was used to determine Bac-*Pf*GK$^{Sf}$ titre. Each 35mm plate was inoculated with 0.9 x $10^6$ *Sf*9 cells, infected with 200µl of a $10^{-6}$ or $10^{-7}$ Bac-*Pf*GK$^{Sf}$ viral dilution (in triplicate) and incubated at 28°C for four days. Viable cells were stained with Neutral Red and unstained, clear plaques were counted on the plates with a $10^{-7}$ viral dilution. The number of plaques in each plate is indicated. There was an average of 25 plaques per plate and 1.23 x $10^9$ pfu/ml viral titre. There were too many plaques to count when cells were infected with the $10^{-6}$ viral dilution. No plaques were visible in uninfected *Sf*9 control cells.

After a 72 hour incubation, trypan blue staining (Appendix A14) showed that there were 2.6 x $10^6$ Bac-*Pf*GK$^{Sf}$-infected TriEx™ *Sf*9 cells/ml at a 69% viability compared to the 10.1 x $10^6$ cells/ml and a 94% viability in the uninfected TriEx™ *Sf*9 control cells. Proteins were extracted from the cells and r*Pf*GK$^{Sf}$ protein purification was performed using MagneHis™ nickel particles (section 2.7.9). SDS-PAGE and immunoblot analysis, using a Penta•His™ HRP Conjugate antibody, showed no difference between the uninfected and recombinant baculovirus-infected cells (Figure 26). r*Pf*GK$^{Sf}$ protein expression was not induced and the expected ~59kDa recombinant protein could not be purified. As an additional verification for the lack of r*Pf*GK$^{Sf}$ protein expression, enzyme analysis of the purified fraction showed no *Pf*GK activity (data not shown).

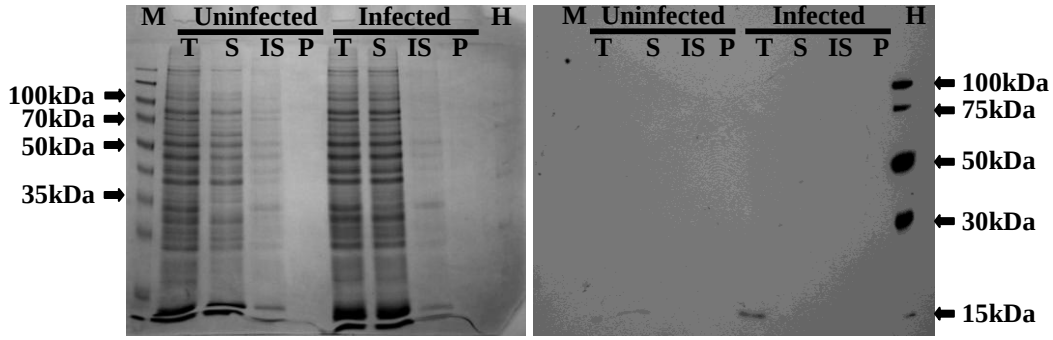



**Figure 26: r*Pf*GK$^{Sf}$ expression in insect cells infected with an MOI = 10.**
TriEx$^{TM}$ *Sf*9 cells (2 x 10$^6$ cells/ml; 20ml) were infected with Bac-*Pf*GK$^{Sf}$ baculoviruses at an MOI = 10. Insect cells were lysed and r*Pf*GK$^{Sf}$ was purified from the soluble protein extract. Total (T; 15µl), soluble (S; 15µl from a total of 1.8ml), insoluble (IS; 15µl) and purified (P; 15µl from a total of 200µl) protein fractions were analyzed using SDS-PAGE **(left)**. No difference in protein expression was observed between uninfected- and Bac-*Pf*GK$^{Sf}$-infected TriEx$^{TM}$ *Sf*9 cells. Immunoblot analysis **(right)**, using a Penta•His™ HRP Conjugate antibody, verified that there was no expression of His-tagged proteins. M is 5µl of the Spectra$^{TM}$ Multicolor Broad Range Protein Ladder (Fermentas International Inc., Burlington, Canada) and H is 2.5µl of the 6xHis Protein Ladder (Qiagen GmbH, Hilden, Germany).

In an attempt to optimize r*Pf*GK$^{Sf}$ protein expression, TriEx$^{TM}$ *Sf*9 cells were infected with Bac-*Pf*GK$^{Sf}$ at a MOI = 2 and MOI = 5. Cell density and viability were monitored at 24 hour intervals (Table 13).

**Table 13: Monitoring cell density and viability of Bac-*Pf*GK$^{Sf}$-infected cells.**

| [1]Bac-*Pf*GK$^{Sf}$ | 0hrs | | 24hrs | | 48hrs[2] | | 72hrs[2] | |
|---|---|---|---|---|---|---|---|---|
| | Cells/ml | Viability | Cells/ml | Viability | Cells/ml | Viability | Cells/ml | Viability |
| MOI 0 | 2.0x10$^6$ | 95% | 3.8x10$^6$ | 97% | [3]6.7x10$^6$ | 95% | 3.3x10$^6$ | 93% |
| MOI 2 | | | 2.3x10$^6$ | 95% | 2.8x10$^6$ | 86% | 2.0x10$^6$ | 73% |
| MOI 5 | | | 2.2x10$^6$ | 98% | 1.9x10$^6$ | 82% | 1.7x10$^6$ | 69% |

[1]Data summary from 20ml cultures
[2]Time points at which cultures were monitored for recombinant protein expression analysis
[3]Diluted uninfected cells 1:4 with fresh medium to limit cell death

After 72 hours, trypan blue staining showed that the growth of Bac-*Pf*GK$^{Sf}$-infected TriEx$^{TM}$ *Sf*9 was reduced and the cell densities of infected cultures did not change significantly from day one. In contrast, uninfected cells reached the

recommended maximum suspension concentration of 5-6 x $10^6$ cells/ml after 48 hours and were diluted 1:4 with TriEx™ Insect Cell Medium to limit cell death due to accumulation of toxic by-products and lack of oxygen. After 72 hours, cell viability was reduced to ~70% in the Bac-*Pf*GK$^{Sf}$-infected cells compared to the 93% viability in uninfected TriEx™ *Sf*9 cells (Table 13).

Under phase contrast microscopy, infected cells displayed characteristics reminiscent of baculoviral infection. r*Pf*GK$^{sf}$ protein expression was evaluated at 48 and 72 hours post-infection (Figure 27). A prominent band between 70-100kDa was purified, which did not correspond to the expected size of ~59kDa for r*Pf*GK$^{sf}$ (Figure 27a; gels not shown for cultures infected with viruses at an MOI = 2). Immunoblot analysis with an anti-His antibody revealed no difference between uninfected and infected cells and several proteins >50kDa were detected in the purified fractions (Figure 27b). These could be post-translationally modified r*Pf*GK$^{sf}$ or endogenous histidine-containing insect or baculoviral proteins.

Further analysis of purified protein fractions verified that there was no difference in protein expression between uninfected and recombinant baculovirus infected cells (Figure 28). The purified proteins were therefore, likely to be insect or baculoviral proteins with stretches of histidine residues. Similarly, no r*Pf*GK$^{Sf}$ protein was expressed or purified when larger 200ml cultures were infected with Bac-*Pf*GK$^{sf}$ baculoviruses at a MOI = 2 and MOI = 10.

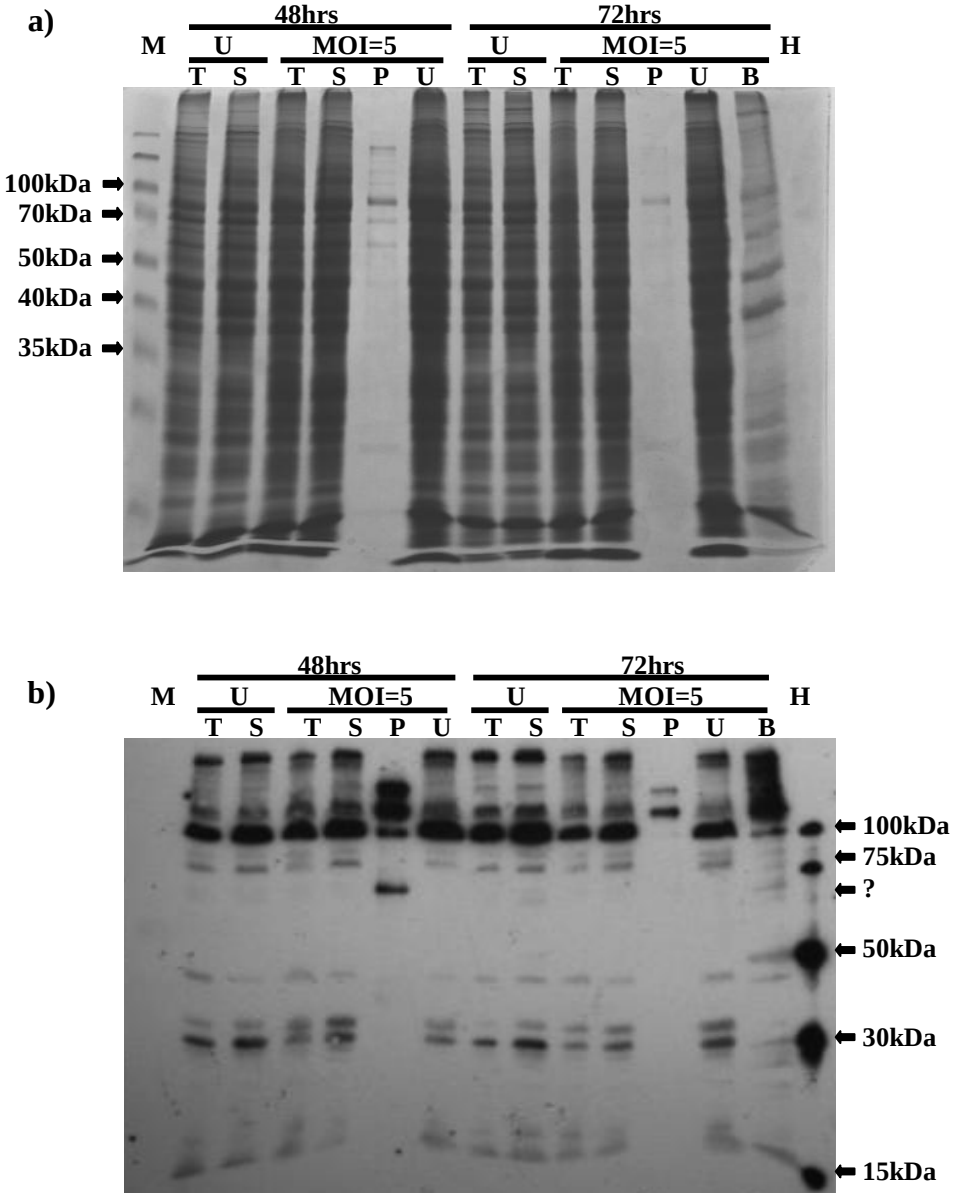



**Figure 27: Expression of r*Pf*GK$^{Sf}$ in insect cells infected with an MOI = 5.**
TriEx$^{TM}$ *Sf*9 cells (2 x $10^6$ cells/ml; 20ml) were infected with Bac-*Pf*GK$^{Sf}$ baculoviruses at an MOI = 5. Insect cell cultures were lysed after 48 and 72 hours post infection and r*Pf*GK$^{Sf}$ was purified from the soluble protein extract. Total (T; 10µl), soluble (S; 50µl from a total of 1.8ml), purified (P; 20µl from a total of 200µl), unbound (U; 20µl) and MagneHis$^{TM}$ beads (B; 20µl) protein fractions were analyzed using **(a)** SDS-PAGE. Total and soluble protein fractions from an uninfected culture (U) were used as controls; **(b)** immunoblot analysis, using a Penta•His™ HRP Conjugate antibody, showed a band between 50-70kDa in size in the Bac-*Pf*GK$^{Sf}$ infected 48-hour purified sample. M is 5µl of the Spectra$^{TM}$ Multicolor Broad Range Protein Ladder and H is 2.5µl of the 6xHis Protein Ladder.

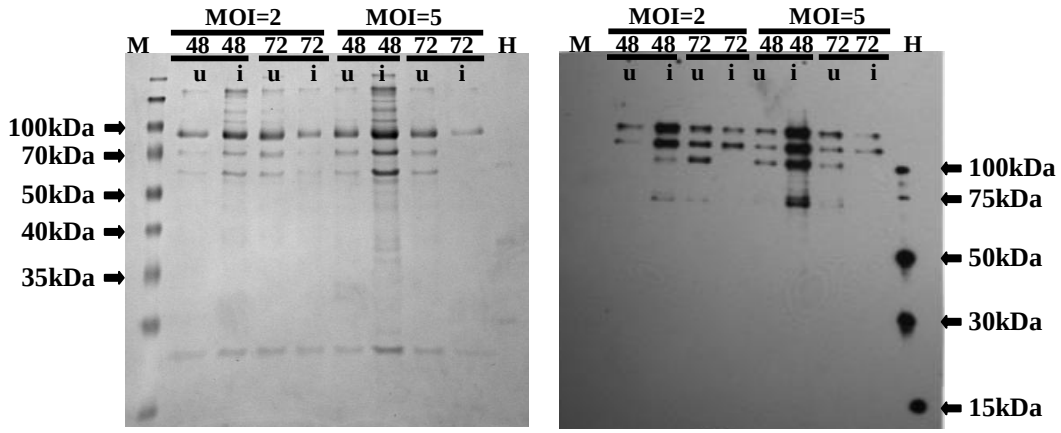



**Figure 28: Purified fractions from uninfected and Bac-*Pf*GK$^{sf}$-infected cells.**
Twenty microlitres of each purified fraction was resolved using SDS-PAGE and Coomassie blue staining **(left)**. Immunoblot **(right)** analysis showed no clear difference in proteins between the uninfected (u) and Bac-*Pf*GK$^{Sf}$-infected cells (i). The MOI and time intervals are indicated. M is 2.5µl of the Spectra$^{TM}$ Multicolor Broad Range Protein Ladder. H is 2.5µl of the 6xHis Protein Ladder.

# 3.3 r*Pf*GK enzyme activity

## 3.3.1 r*Pf*GK enzyme assay

An automated spectrophotometric enzymatic assay was used to measure glycerol kinase activity. In the presence of an active kinase, the ADP product from the primary reaction acts as a substrate in the secondary PK and LDH enzymatically coupled reaction that ultimately results in the oxidation of NADH to $NAD^+$. This decrease in NADH absorbance can be measured at 340nm (Figure 29) and the enzyme activity calculated (section 2.6.1). Enzyme assays were performed at 37°C and three independent purifications showed that r*Pf*GK was active.

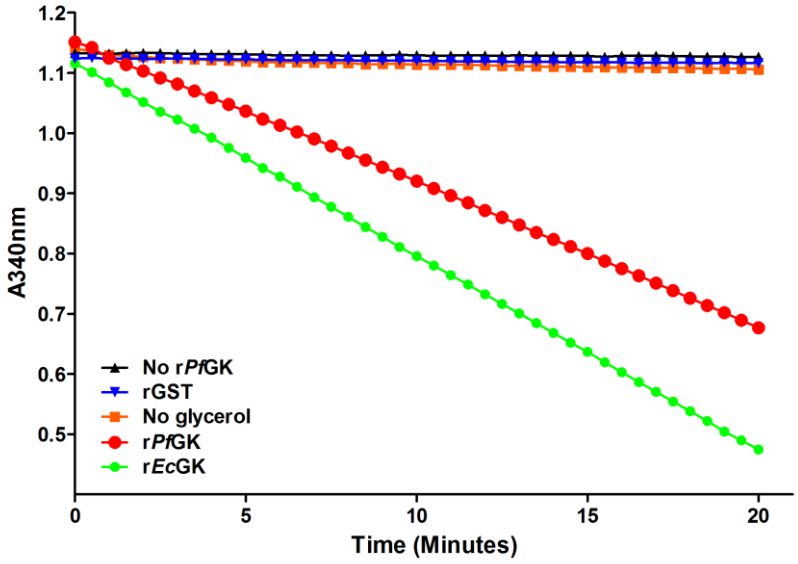



**Figure 29: Measuring glycerol kinase activity.**
Glycerol kinase ( ) activity was determined by measuring the decrease in NADH absorbance in a coupled reaction with LDH and PK as auxiliary enzymes. Data were acquired at 30 second intervals over 20 minutes using the Biomate 5 spectrophotometer and VISIONlite™ version 2.2 software. The linear range of the line was used to calculate enzyme activity. No enzyme activity was recorded in the 'no r*Pf*GK' ( ), rGST ( ) and 'no glycerol' ( ) control reactions. r*Ec*GK ( ) was used as a positive control.

## 3.3.2 r*Pf*GK enzyme stability

Storage of r*Pf*GK at 4°C in elution buffer (500mM reduced glutathione, 50mM Tris-HCl, pH 8.25) resulted in a daily decrease in activity (Figure 30). The greatest loss occurred during the first two days of storage, at a rate of ~30% loss in activity per day.

## 3.3.3 The effect of temperature on r*Pf*GK activity

r*Pf*GK activity was measured at all physiologically relevant temperatures: 28°C (temperature in the mosquito - 28-33°C (Alano and Billker, 2005)); 37°C (human host) and 40ºC (fever temperature in the human host). r*Pf*GK was active at all three temperatures, with the highest activity at temperatures of *Plasmodium*-induced fever (Figure 31).

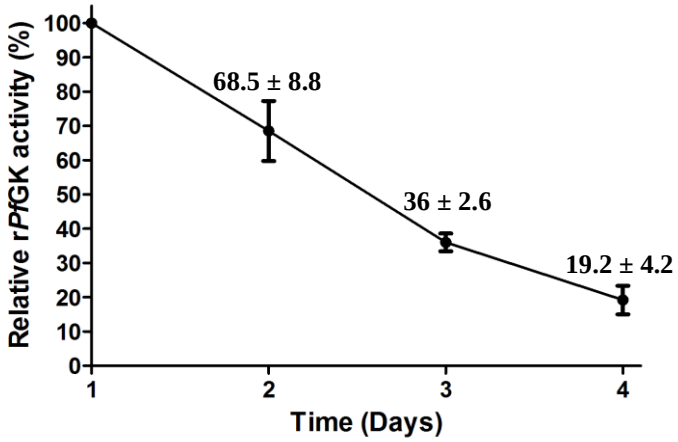



**Figure 30: r*Pf*GK stability.**
Specific enzyme activity was measured over a period of four days and normalized to the highest enzyme activity recorded at day one. Activity was measured in duplicate from three separate experiments (n = 6). Standard deviation of the mean is indicated.

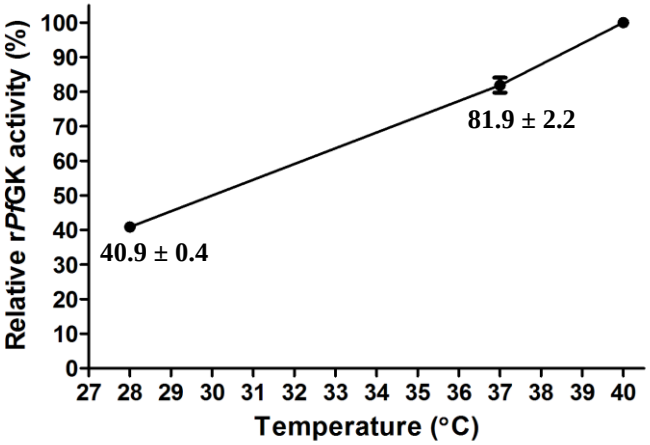



**Figure 31: r*Pf*GK activity at different temperatures.**
Specific enzyme activity was measured 28°C, 37°C and 40°C and normalized to the highest enzyme activity, which was recorded at 40°C. Activity was measured in duplicate from three separate experiments (n = 6). Standard deviation of the mean is indicated.

## 3.3.4 The effect of pH on r*Pf*GK activity

r*Pf*GK activity was measured at the pH in the *P. falciparum* food vacuole (pH 5.0-5.5), the human erythrocyte host (pH 7.0-7.8) and the alkaline environment of the mosquito midgut (pH 8.0-9.0). Combined data indicated that r*Pf*GK activity was

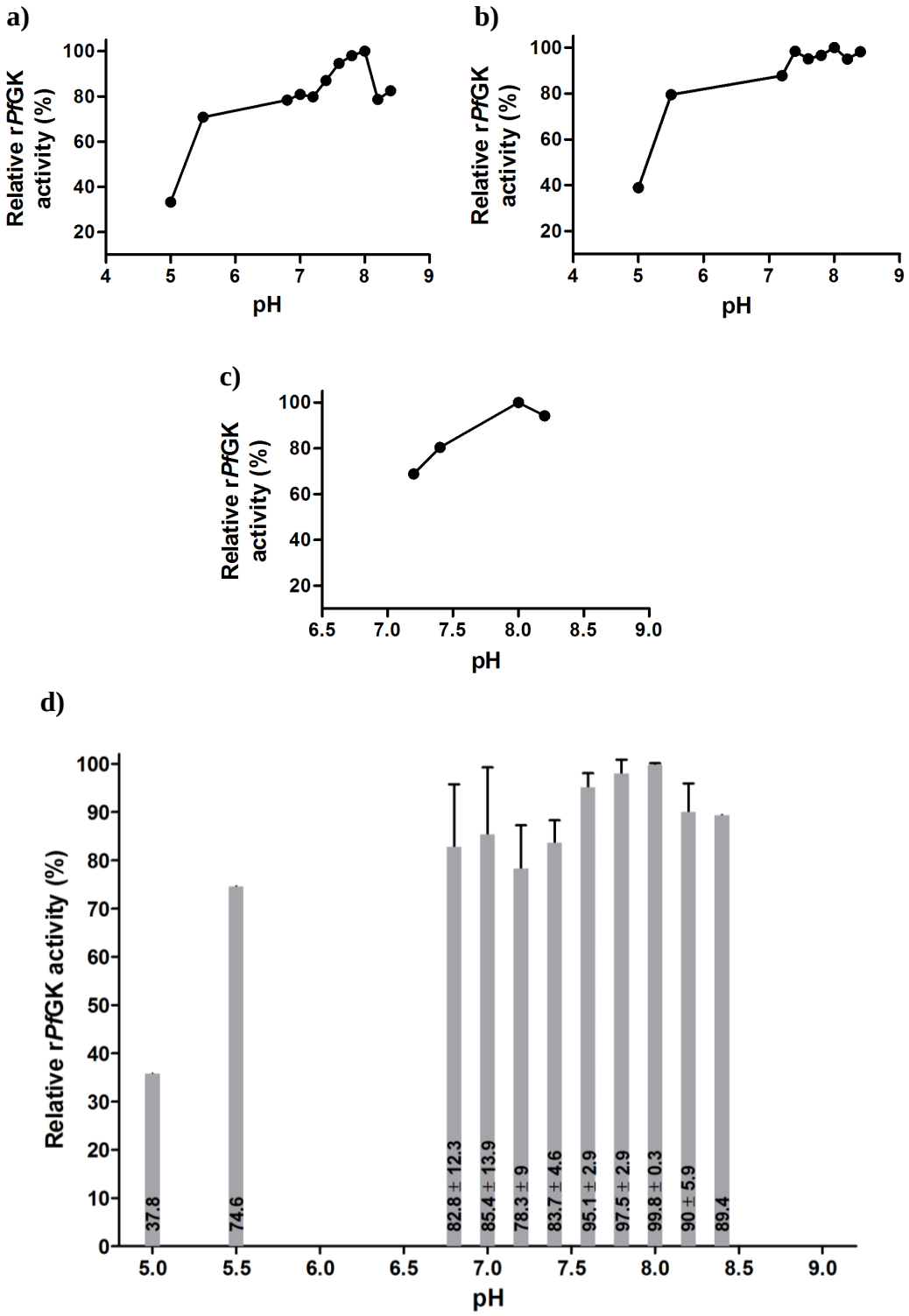



**Figure 32: Relative r*Pf*GK activity at different pH.**
Combined data **(d)** from three independent and normalized experiments **(a-c)**. A 0.1M sodium-acetate and a 0.1M Tris-HCl buffer system was used to measure specific enzyme activity at pH 5.0-5.5 and pH 6.8-8.4 respectively. Relative specific r*Pf*GK enzyme activity at pH 6.8-8.2 was measured in duplicate from three separate experiments (n = 6), whereas activity was measured in duplicate twice at pH 5.0, 5.5 and 8.4 (n = 4). Standard deviation of the mean is indicated.

relatively uniform between pH 6.8-8.4, with highest activity recorded at pH 8.0 (Figure 32).

### 3.3.5 Kinetic parameters of r*Pf*GK

Michaelis-Menten (Figure 33) and Lineweaver-Burk double reciprocal plots (Figure 34) were used to determine the $V_{max}$, $K_m$ and r*Pf*GK specific activity for glycerol and ATP. Non-linear regression analysis from the Michaelis-Menten equation was used to determine the average specific activities and the results are summarized in Table 14.

**Table 14: Summary of the kinetic parameters of r*Pf*GK.**

| r*Pf*GK activity | $V_{max}$ (nmol NADH/min)* | $K_m$ (µM)* | Specific activity (µmol NADH/min/mg)* |
|---|---|---|---|
| Glycerol | 2.15 | 15.7 | 32.7 |
| ATP | 1.66 | 15.9 | 19.1 |

*Representation of the mean value of two r*Pf*GK enzymatic reactions

### 3.3.6 r*Pf*GK activity in the presence of FBP

In the *E. coli* orthologue, a fructose-1,6-bisphosphate site is created when two *Ec*GK dimers associate to form a tetramer. FBP binds to and stabilizes the tetrameric form, inactivating the enzyme in a non-competitive mechanism of enzyme inhibition ($K_{i(FBP)}$ = 0.25mM) (Ormo *et al.*, 1998, Feese *et al.*, 1998). To evaluate whether FBP acts as a non-competitive inhibitor of *Pf*GK, its activity was measured in the presence of 5mM and 10mM FBP. Analysis of Michaelis-Menten kinetic parameters (Figure 35a) is summarized in Table 15. Linear regression analysis from the Lineweaver-Burk plot (Figure 35b) suggests an uncompetitive mechanism of r*Pf*GK inhibition. However, r*Pf*GK activity was only ~10% lower in the presence of 20-40 times the $K_i$ of FBP in *Ec*GK. This implies that FBP is not a physiological inhibitor of r*Pf*GK.

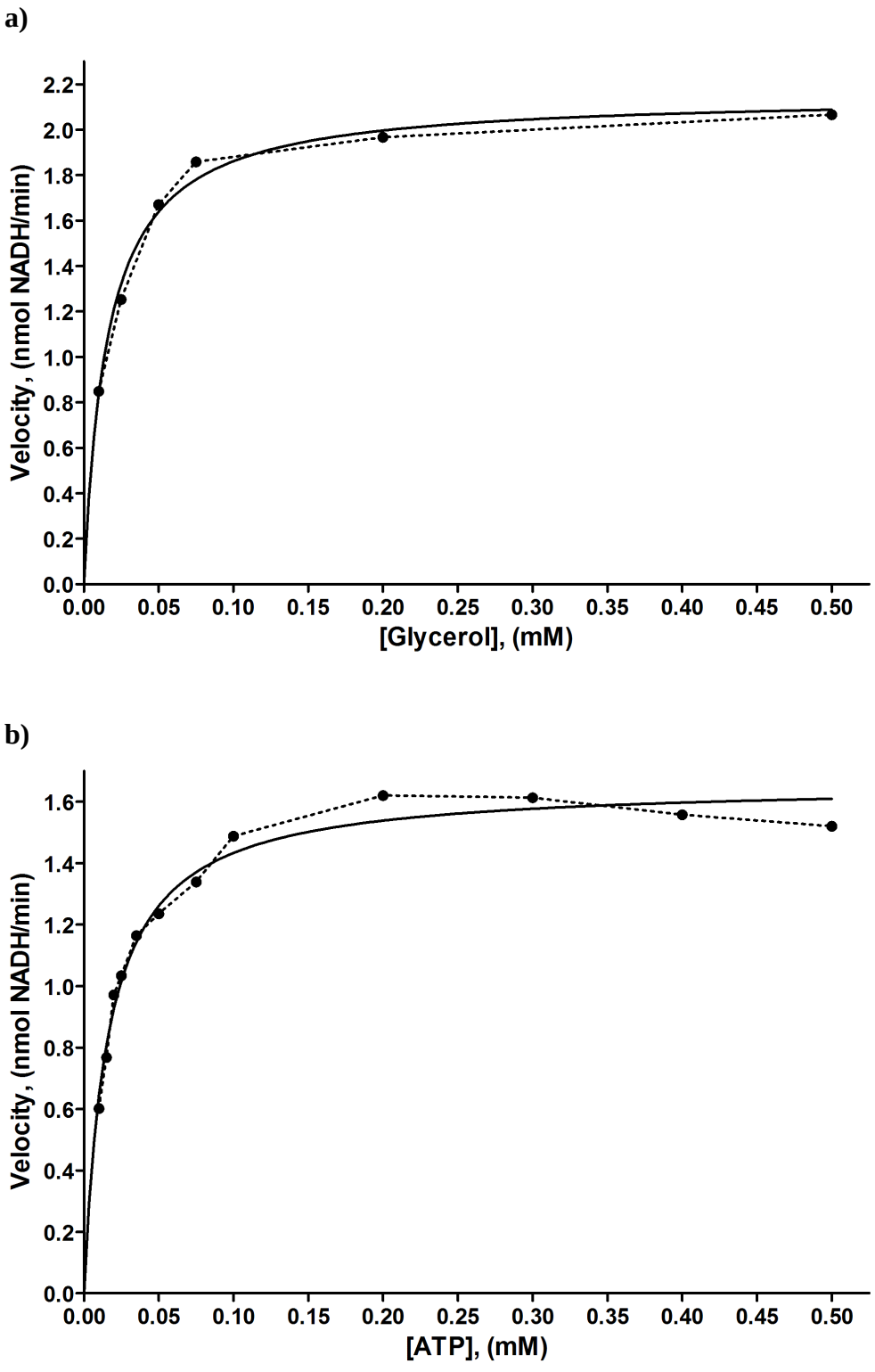



**Figure 33: Michaelis-Menten plot for r*Pf*GK.**
Plot of velocity against glycerol and ATP substrates according to the Michaelis-Menten equation for r*Pf*GK; **(a)** kinetic parameters of glycerol were measured in the presence of a 2mM saturated ATP concentration; **(b)** kinetic parameters of ATP were measured in the presence of a 3.3mM saturated glycerol concentration. Dotted lines connect data points and solid lines are the "best-fit" lines. The mean values of r*Pf*GK duplicate enzyme reactions are illustrated.

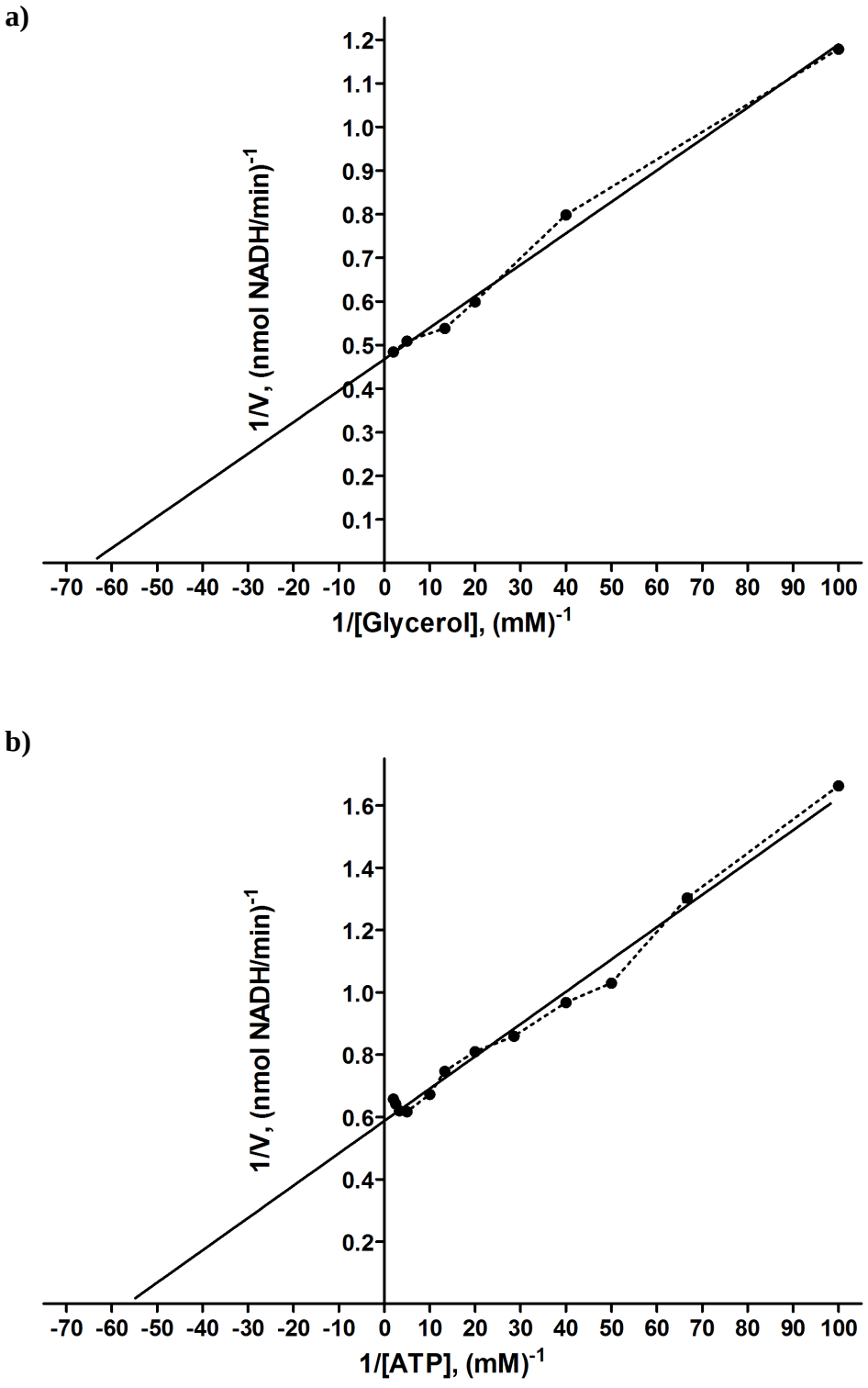



**Figure 34: Lineweaver-Burk plot for r*Pf*GK.**
Double-reciprocal plot of 1/Velocity against 1/[Glycerol] and 1/[ATP] in mM for r*Pf*GK. Kinetic parameters of **(a)** glycerol and **(b)** ATP were measured in the presence of saturating concentrations of each partner substrate. Dotted lines connect data points and solid lines are the "best-fit" lines. The mean values of r*Pf*GK duplicate enzyme reactions are illustrated.

**Table 15: Relative r*Pf*GK activity in the presence of inhibitors.**

| Inhibitor | Concentration (mM) | $V_{max}$ (nmol NADH /min)* | $K_m$ (µM)* | Specific activity (µmol NADH/min/mg)* | Relative enzyme activity (%)*[1] |
|---|---|---|---|---|---|
| None | 0 | 2.15 | 15.7 | 32.7 | 100 |
| FBP | 5 | 1.97 | 13.1 | 29.8 | 91.1 |
| | 10 | 1.93 | 13.6 | 29.3 | 89.6 |

*Mean value of r*Pf*GK enzymatic reactions
[1] Expressed as a percentage of the enzyme's specific activity relative to that of r*Pf*GK without any inhibitors

# 3.4 *Pf*GK gene disruption

## 3.4.1 Generation of *P. falciparum* glycerol kinase knockout parasites (3D7Δ*Pf*GK)

The gene knockout was performed using a double crossover homologous recombination strategy (Figure 36). A 519bp 5' *Pf*GK gene region, representing part of the glycerol-binding domain, was cloned upstream of the hDHFR cassette and a 515bp 3' *Pf*GK gene region, representing part of the ADP-binding domain, was cloned downstream of hDHFR. The internal 420bp *Pf*GK region (nucleotides 545-965, amino acids 182-322) was replaced with the hDHFR cassette and as a result, residues involved in binding glycerol (D246, Q247 and F271) and ADP (T268 and G313) were removed following recombination (Figure 36, Figure 37 and Appendix A15). The disrupted *Pf*GK gene will not be able to produce a functional *Pf*GK protein. To verify the construction of the pCC-1-*Pf*GK knockout plasmid, (i) *Spe*I/*Afl*II and *Eco*RI/*Nco*I restriction enzymes were used to excise the cloned 5'-*Pf*GK (Figure 38a, lane 2) and 3'-*Pf*GK (Figure 38a, lane 5) fragments from the pCC-1-*Pf*GK knockout plasmid backbone and (ii) PCR analysis was performed with 5'-*Pf*GK (Figure 38b, lane 2) and 3'-*Pf*GK (Figure 38b, lane 6) specific primers.

*P. falciparum* parasites were transfected with the pCC-1-*Pf*GK knockout plasmid (Appendix A16). pCC-1-*Pf*GK transformed parasites, under WR99210 positive drug selection, first appeared 51 days post-transfection.

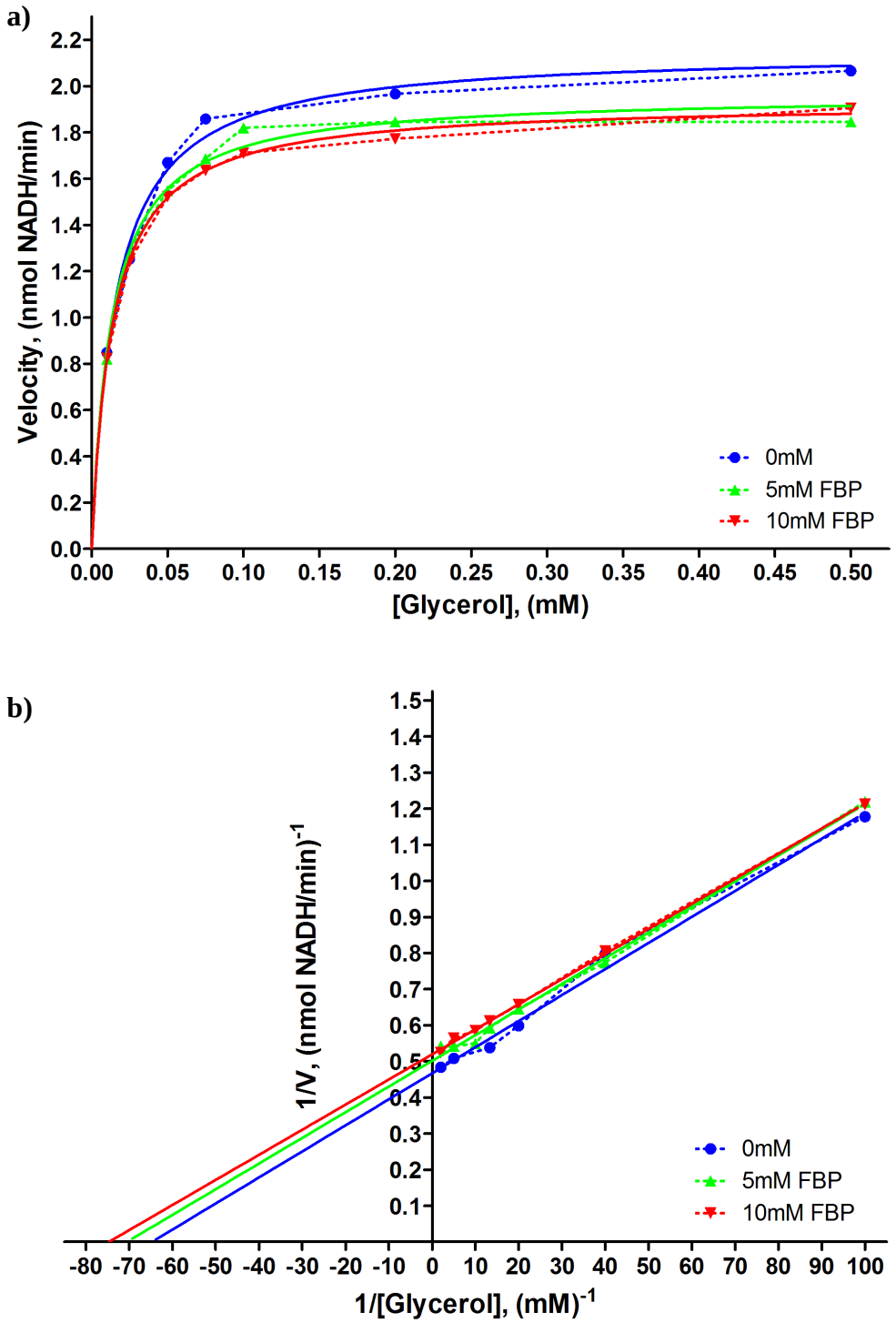



**Figure 35: The effect of fructose-1,6-bisphosphate on r*Pf*GK activity.**
Michaelis-Menten **(a)** and Lineweaver-Burk **(b)** plots for glycerol in the presence and absence of fructose-1,6-bisphosphate (FBP). Kinetic parameters of glycerol were measured in the presence of a saturated 2mM ATP concentration. Dotted lines connect data points and solid lines are the "best-fit" lines. The mean values of r*Pf*GK duplicate enzyme reactions are illustrated.

Parasites with the integrated hDHFR selection cassette, and disrupted *Pf*GK gene, were visualized 11 days thereafter, following both positive WR99210 and negative 5-fluorocytosine selection (dual selection). *P. falciparum* parasites transfected with the pCC-1-EBA knockout plasmid appeared after 32 days under WR99210 positive drug selection and those with the disrupted EBA-175 gene, 6 days thereafter, following dual selection.

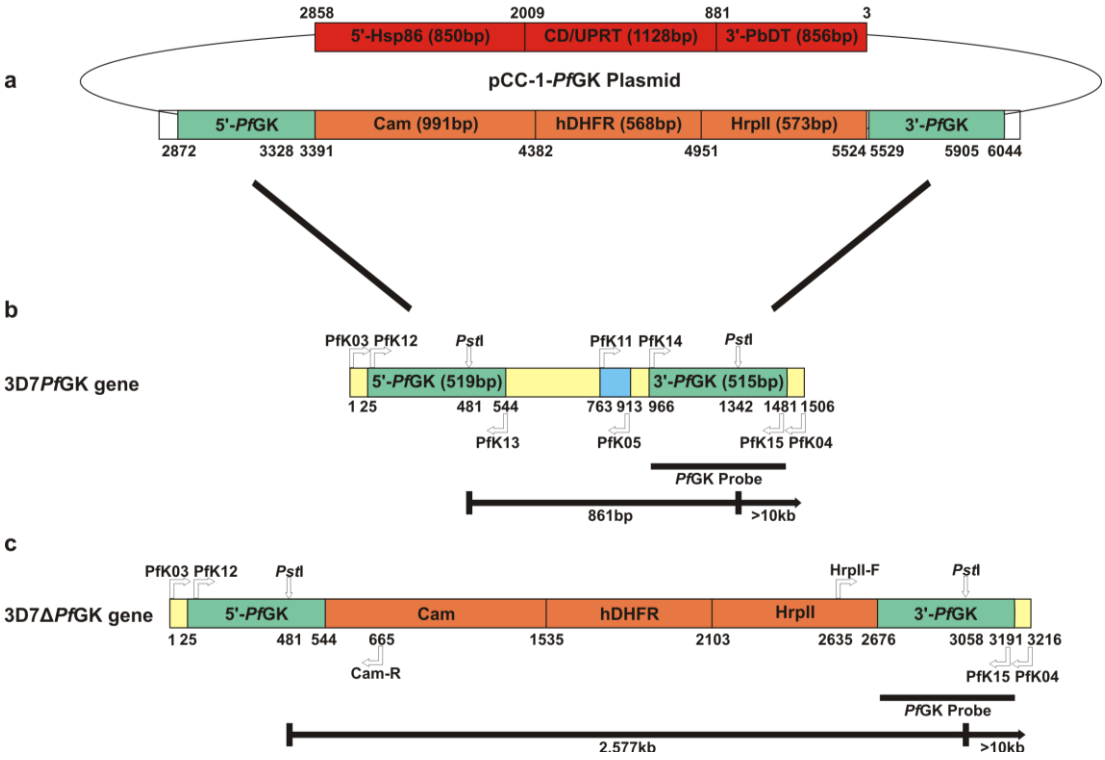



**Figure 36: Disruption of the *Pf*GK gene.**
**(a)** The pCC-1-*Pf*GK knockout plasmid was created following sequential cloning of 5'- and 3'-*Pf*GK fragments (green) into the pCC-1 vector. The positive hDHFR cassette, including the Cam promoter and the HrpII terminator regions, is indicated (orange) and the negative CD/UPRT selection cassette, with its Hsp86 promoter and PbDT terminator regions are highlighted in red; **(b)** the full length *Pf*GK gene showing the positions and sizes of the fragments that were cloned into pCC-1. The homologous double crossover recombination between the 5'- and 3'-*Pf*GK regions of the gene and the vector are indicated. Primers (PfK11/PfK05) designed against an internal *Pf*GK fragment (blue) and used to examine *Pf*GK mRNA expression levels are indicated. This region is displaced by the hDHFR cassette following homologous recombination; **(c)** the disrupted *Pf*GK gene following insertion of the hDHFR cassette between the 5'- and 3'- flanks of the gene. Primer binding and *Pst*I restriction enzyme sites are indicated by arrows. Numbers represent sequence lengths and fragment/gene positions. The 3'-*Pf*GK probe position and DNA fragment sizes, following Southern blot analysis of *Pst*I-digested DNA, are indicated by bars.

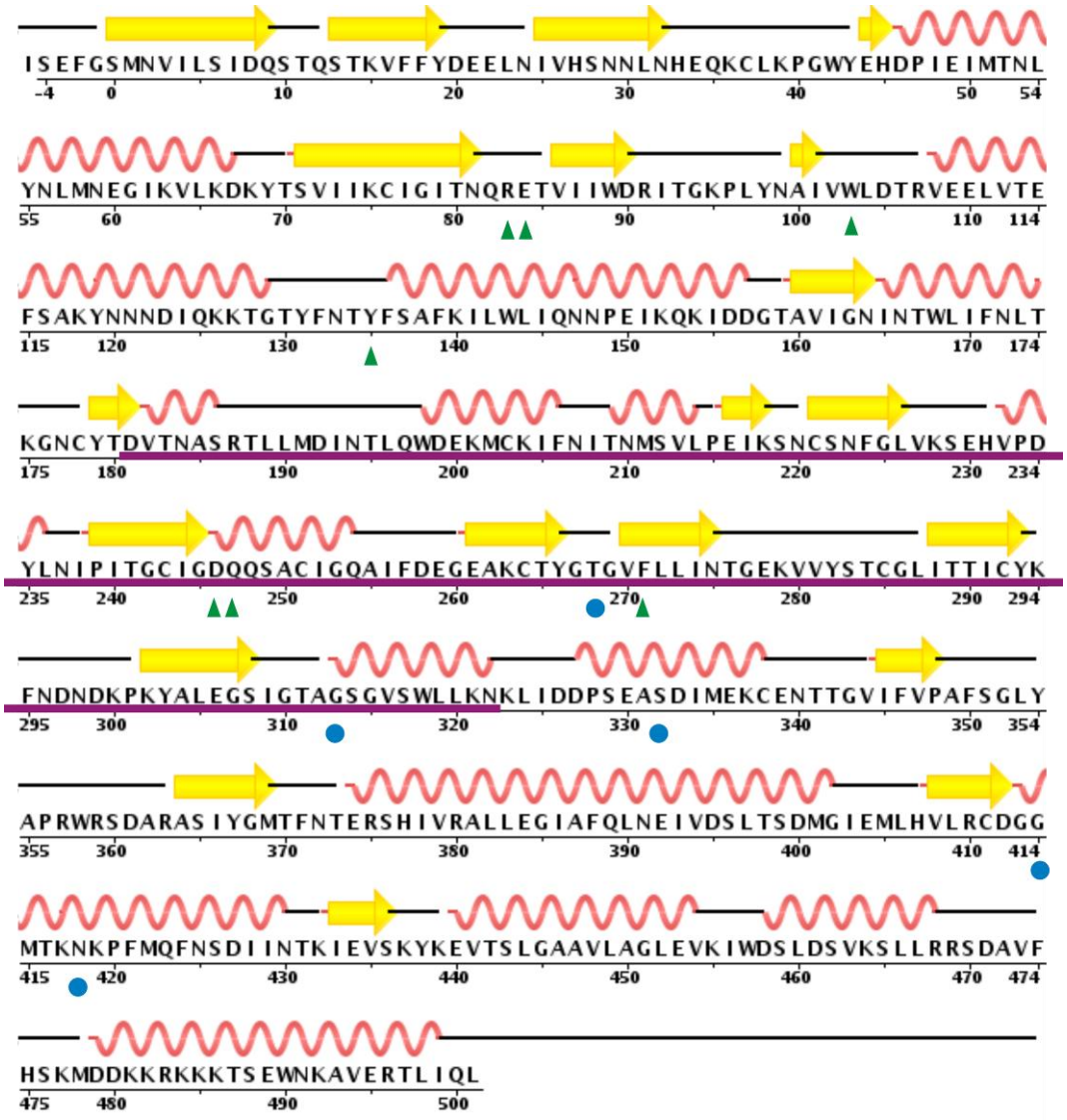



**Figure 37: The *Pf*GK protein primary and secondary structural features.**
Residues involved in glycerol- (▲) and in ADP- (●) binding are represented. α-helix (red ribbons) and β-sheet (yellow arrows) secondary structural elements are indicated. Numbers represent amino acids (www.pdb.org; PDB I.D 2W40 and 2W41) (Berman *et al.*, 2000, Kabsch and Sander, 1983). The deleted *Pf*GK region after double-crossover homologous recombination is underlined in purple.

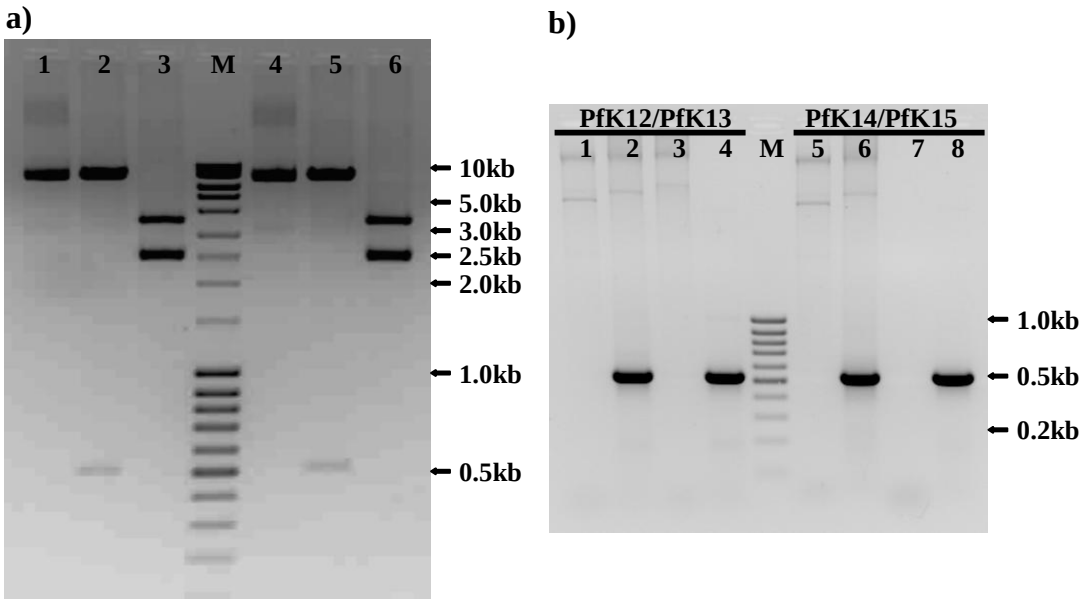



**Figure 38: Cloning *Pf*GK homologous recombination fragments into pCC-1.**
A 519bp 5'- and a 515bp 3'-*Pf*GK fragment were amplified and cloned into the *Spe*I/*Afl*II and *Eco*RI/*Nco*I restriction sites respectively to generate a pCC-1-*Pf*GK knockout plasmid; **(a)** 500ng plasmid DNA was digested and resolved on a 1% agarose gel. *Spe*I/*Afl*II restriction analysis showed the 5'-*Pf*GK fragment and the 8,048bp plasmid backbone (lane 2). Similarly, *Eco*RI/*Nco*I restriction analysis showed the 3'-*Pf*GK fragment and the 8,052bp plasmid backbone (lane 5). Digestion with *Pst*I resulted in 2,445, 2,577 and 3,534bp fragments (lanes 3 and 6). The 2,445 and 2,577bp fragments did not resolve completely on the gel. Undigested pCC-1-*Pf*GK DNA is shown in lanes 1 and 4; **(b)** the 5'-*Pf*GK (lane 2) and 3'-*Pf*GK (lane 6) fragments were amplified from the *Pf*GK knockout plasmid. pGEX-4T-2-*Pf*GK (lanes 4 and 8) and pCC-1 (lanes 1 and 5) plasmid DNA were used as a positive and negative control respectively. PCR with no template DNA (lanes 3 and 7) was used as a blank reaction. Lanes are numbered and sizes of selected fragments in the 2.5µl Fermentas MassRuler$^{TM}$ DNA ladder Mix (M) are indicated.

## 3.4.2 Verification of glycerol kinase knockout parasites

### *PCR screening*

A PCR screen was performed on knockout parasites to evaluate whether they had integrated the pCC-1-*Pf*GK knockout plasmid into their genome or whether the plasmid was also present as an episomal construct (Figure 39, Appendix A15 and Appendix A17). Amplification of a 665bp fragment using primer PfK03, 25bp upstream of the 5'-*Pf*GK homologous region, and primer Cam-R, specific to the pCC-1-*Pf*GK plasmid backbone (121bp downstream of the 5'-*Pf*GK homologous

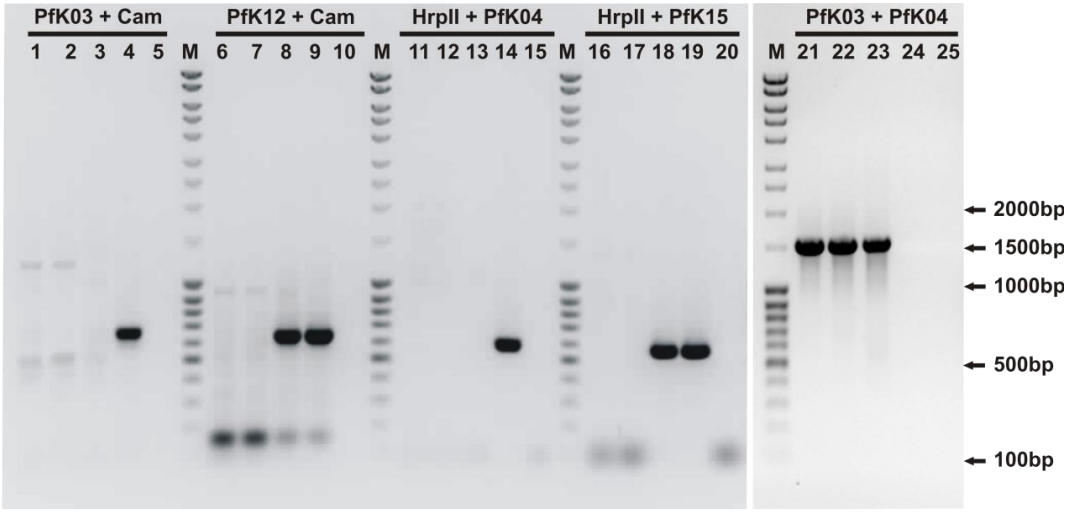



**Figure 39: PCR screening for knockout parasites.**
PCR analysis of DNA extracted from wild type, 3D7ΔEBA and 3D7Δ*Pf*GK parasites. Various primer combinations (indicated above lanes) were used to detect integration at the 5'- and 3'-*Pf*GK regions. Lanes are numbered and sizes of selected fragments in the 2.5µl Fermentas MassRuler$^{TM}$ DNA ladder Mix DNA ladder (M) are indicated. Samples were loaded between each DNA ladder as follows: 3D7, 3D7ΔEBA, 3D7Δ*Pf*GK (+WR99210 only), 3D7Δ*Pf*GK (+WR99210 and 5-fluorocytosine dual selection) and pCC-1 plasmid as a negative control. Lanes 3 and 13: a parasite population with the disrupted *Pf*GK gene was not detected in 3D7Δ*Pf*GK parasites under positive selection. Lanes 4 and 14: gene disruption at the 5'end of *Pf*GK (665bp, PfK03/Cam-R) and at the 3'end of *Pf*GK (580bp, HrpII/PfK04) was only detected after dual selection. Lanes 8 and 9 (5'-end) and lanes 18 and 19 (3'-end): parasites with episomal pCC-1-*Pf*GK plasmid DNA and/or 3D7Δ*Pf*GK knockout parasites were detected during positive selection (519bp, PfK12/Cam-R and 515bp, HrpII/PfK15). Lane 24: the full length *Pf*GK gene (1,506bp) was not amplified in 3D7Δ*Pf*GK knockout parasites (PfK03/PfK04). Lane 23: the undisrupted *Pf*GK gene was detected in 3D7Δ*Pf*GK parasites under positive selection only.

region) showed integration at the 5'-end after dual drug selection (Figure 39, lane 4). Similarly, amplification of a 580bp fragment with primers HrpII/PfK04 at the 3'-*Pf*GK region showed integration at the 3'-end (Figure 39, lane 14). The integrated *Pf*GK fragments were not detected in 3D7ΔEBA (Figure 39, lanes 2, 7, 12 and 17) and wild type (Figure 39, lanes 1, 6, 11 and 16) parasite DNA. Episomal constructs and/or 3D7Δ*Pf*GK parasites were detected using primers specific to the pCC-1-*Pf*GK knockout plasmid (Figure 39, lanes 8, 9, 18 and 19). The 1,506bp full length *Pf*GK gene was present in wild type, 3D7ΔEBA and in 3D7Δ*Pf*GK parasites under positive drug selection only (Figure 39, lanes 21-23). The latter is due to a parasite population with an undisrupted *Pf*GK gene but the pCC-1-*Pf*GK knockout plasmid present as an episomal construct, which provides

resistance against WR99210. Parasites with a disrupted *Pf*GK gene were not detected with primers flanking the homologous recombination regions under positive drug selection (Figure 39, lanes 3 and 13). The full length *Pf*GK gene was absent in 3D7Δ*Pf*GK parasites following dual selection (Figure 39, lane 24).

## *Southern hybridization analysis*

Southern blot analysis was used to verify the presence of knockout parasites. A DIG-labelled probe (Figure 40), prepared using the 3'-*Pf*GK homologous region in the pCC-1-*Pf*GK knockout plasmid as a template, was hybridized to wild type and 3D7Δ*Pf*GK DNA digested with *Pst*I (Figure 41).

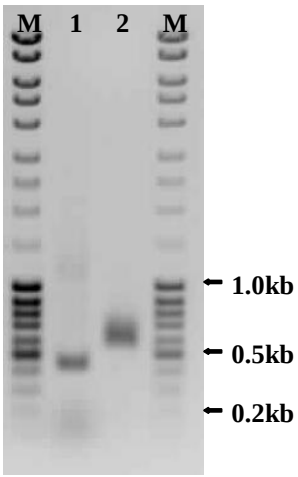



**Figure 40: Amplification of the 3'-*Pf*GK DIG-labelled probe.**
The pCC-1-*Pf*GK plasmid was used as the DNA template in the amplification reaction. Primers (PfK14/PfK15) designed to amplify the 3'-*Pf*GK homologous DNA fragment were used to generate a DIG-labelled probe. Two microlitres of the DIG-labelled 3'-*Pf*GK probe (lane 2) was resolved on a 1% agarose gel. The molecular weight of the DIG-labelled DNA is increased and thus it migrates slower than the unlabelled 515bp fragment (lane 1). Lanes are numbered and sizes of selected fragments in the 2.5µl Fermentas MassRuler$^{TM}$ DNA ladder Mix DNA ladder (M) are indicated.

The 3'-*Pf*GK homologous region has an internal *Pst*I site such that the DNA probe binds to 2,577bp and 3,534bp fragments in *Pst*I-digested pCC-1-*Pf*GK plasmid DNA (Figure 36 and Figure 41b, lane 1) and knockout parasites with episomal plasmid DNA (data not shown). The probe recognizes a 861bp fragment and a larger fragment (>10kb) in *Pst*I-digested 3D7 DNA (Figure 36, Figure 41a,

lane 1 and Figure 41b, lane 2). Following *Pf*GK gene disruption, a 2,577bp fragment and a larger fragment (>10kb) were detected in *Pst*I-digested 3D7Δ*Pf*GK DNA (Figure 36, Figure 41a, lane 3 and Figure 41b, lane 3).

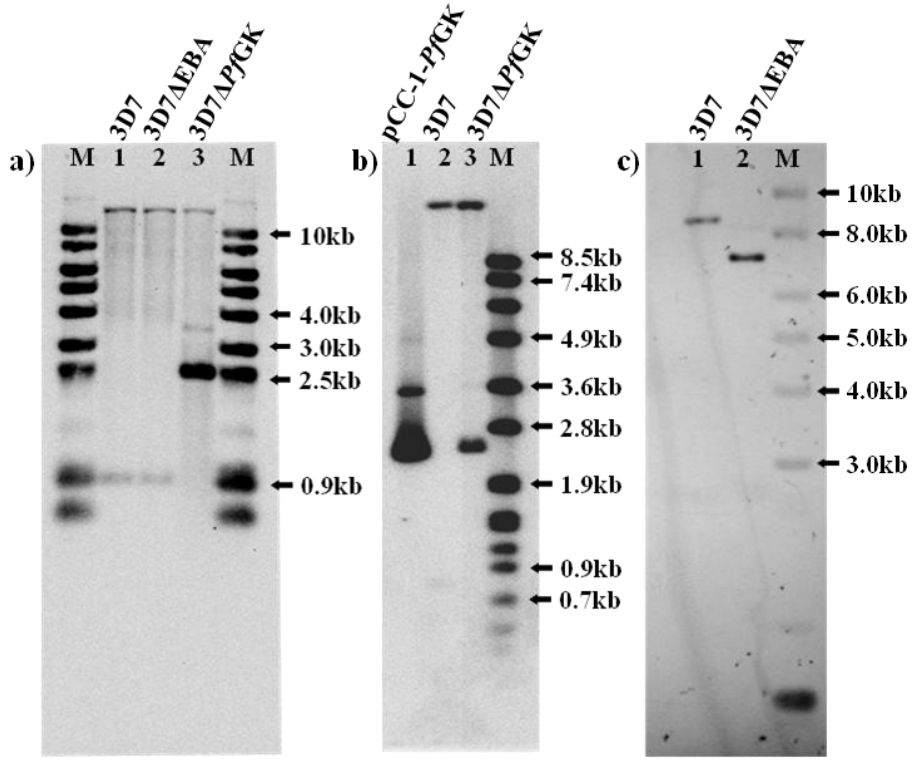



**Figure 41: Southern blot verification of 3D7Δ*Pf*GK parasites.**
A 515bp 3'-*Pf*GK DIG-labelled probe was used in the Southern hybridization analysis of 5µg *Pst*I-digested DNA. Lanes are numbered and labelled and sizes of selected fragments in the 5µl of the Roche DIG-labelled DNA Molecular Weight Marker VII (M) are indicated. **(a)** In parasites with an undisrupted *Pf*GK gene, the probe binds to 861bp and >10kb fragments, as seen in the 3D7 wild type (lane 1) and 3D7ΔEBA knockout parasites (lane 2). The probe binds to 2,577bp and >10kb fragments in 3D7Δ*Pf*GK knockout parasites (lane 3); **(b)** a second Southern hybridization experiment confirmed the generation of 3D7Δ*Pf*GK knockout parasites (lane 3). Lane 2 is the 3D7 control. The probe binds to 2,577 and 3,534bp fragments in the pCC-1-*Pf*GK plasmid DNA positive control (lane 1); **(c)** a 3D7ΔEBA knockout parasite line was used as a control in the targeted gene disruption experiments. Southern hybridization analysis verified that the EBA gene had been disrupted in 3D7ΔEBA parasites (lane 2; 7,190bp) compared to wild type parasites (lane 1; 8,585bp).

## *mRNA analysis*

*Pf*GK mRNA expression levels, using primers (PfK11/PfK05) against an internal 150bp *Pf*GK gene region that would have been displaced by the hDHFR cassette

(Appendix A16), showed that *Pf*GK mRNA was absent in 3D7Δ*Pf*GK knockout parasites (Figure 42, lane 9). 18s rRNA and rifin genes were used as controls and produced mRNA in 3D7Δ*Pf*GK parasites (Figure 42, lanes 3 and 6). 3D7ΔEBA knockout parasites have an intact *Pf*GK gene and produced *Pf*GK mRNA similar to wild type parasites (Figure 42, lanes 7 and 8). In summary, PCR, Southern blot hybridization and *Pf*GK mRNA transcript analysis confirmed that the *Pf*GK gene had been disrupted.

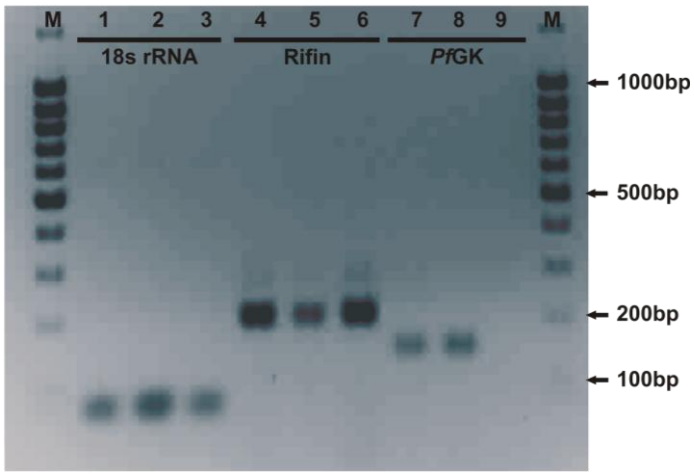



**Figure 42: mRNA expression in 3D7Δ*Pf*GK knockout parasites.**
RNA was extracted and RT-PCR analysis from 1µg RNA showed that *Pf*GK mRNA was absent in 3D7Δ*Pf*GK knockout parasites (lane 9). Ten microlitres of each amplicon was resolved on a 2% agarose gel. The loading order is: 3D7, 3D7ΔEBA and 3D7Δ*Pf*GK. mRNA was present for the 18s rRNA (lane 3, 68bp fragment) and rifin (lane 6, 206bp fragment) template controls in 3D7Δ*Pf*GK parasites. Lanes are numbered and sizes of selected fragments in the 2.5µl Fermentas MassRuler™ DNA ladder Mix DNA ladder (M) are indicated.

## 3.5 Growth analysis of 3D7Δ*Pf*GK knockout parasites

Growth rates of wild type, 3D7ΔEBA and 3D7Δ*Pf*GK knockout parasites were analysed using thiazole orange staining and flow cytometry. An asynchronous culture was used to set the gating parameters for uninfected erythrocytes, ring and trophozoite/schizont stage parasites (Figure 43). These were used for subsequent growth analysis from highly synchronized cultures. Starting parasitaemia was measured at 0.5-2.0% rings and growth rates were monitored over 48 hours.

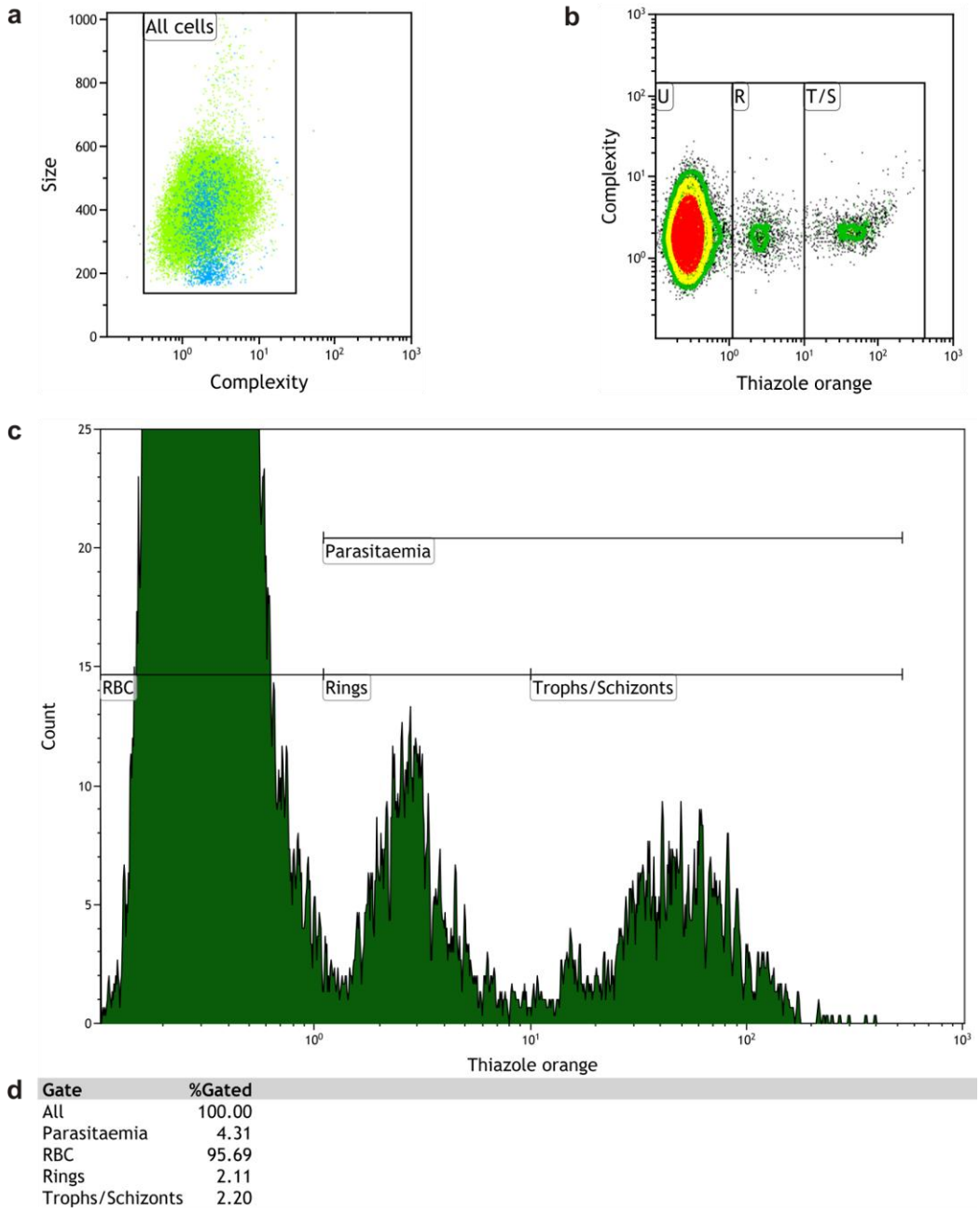



**Figure 43: Flow cytometry gate parameters for ring and trophozoite and/or schizont populations.**

An asynchronous culture with similar number of rings and trophozoite/schizont stage parasites was used to set the inclusion gate parameters of each population; **(a)** total erythrocyte population was delineated by an inclusion gate on a forward- vs side-scatter two-parameter dot blot; **(b,c)** the selected population was analysed on a single-parameter histogram for thiazole orange fluorescence to distinguish parasite stages from uninfected red cells; **(d)** analysis of the gated cell populations and percentage parasitaemia. Results correlated with microscopy analysis of Giemsa-stained smears.

The ring form (0-20 hours post invasion; hpi) develops into a trophozoite (20-36 hpi) and schizont (36-48 hpi), completing the asexual replication cycle (Dvorin *et al.*, 2010). Since the starting time period of data acquisition was ring-stage parasites between 0-20 hpi, the invasion of erythrocytes by daughter parasites for the next developmental cycle (i.e. rings → trophozoites → schizonts → rings; Figure 44) would have been monitored. Analysis of the first experimental data set showed that wild type and 3D7ΔEBA parasites had a 4.74- and 4.36-fold increase in parasitaemia, respectively (Figure 45a and Table 16). In contrast, 3D7Δ*Pf*GK parasites only showed a 2.62-fold increase in parasitaemia, approximately 45% less than the wild type growth rate after one cycle. Results were consistent over three experiments (Figure 45a-c and Table 16) where the parasitaemia of 3D7Δ*Pf*GK parasites were an average of 56.5 ± 1.8% of the wild type parasites following one re-invasion cycle (Figure 45d).

**Table 16: Summary of the change in parasitaemia after one growth cycle.**

| Data set | Parasite line | % Parasitaemia | | Parasitaemia fold change[1] | Relative parasite growth (%)[2] |
|---|---|---|---|---|---|
| | | t = 0 hr | t = 48 hrs | | |
| 1 | 3D7 | 0.69 | 3.27 | 4.74 | 100 |
| | 3D7ΔEBA | 0.69 | 3.01 | 4.36 | 91.57 |
| | 3D7Δ*Pf*GK | 0.58 | 1.52 | 2.62 | 54.91 |
| 2 | 3D7 | 1.00 | 7.36 | 7.36 | 100 |
| | 3D7ΔEBA | 1.00 | 7.89 | 7.89 | 107.20 |
| | 3D7Δ*Pf*GK | 1.00 | 4.3 | 4.30 | 58.46 |
| 3 | 3D7 | 1.00 | 8.91 | 8.91 | 100 |
| | 3D7ΔEBA | 1.00 | 7.98 | 7.98 | 89.58 |
| | 3D7Δ*Pf*GK | 1.97 | 9.86 | 5.01 | 56.20 |

[1]Fold change calculated by dividing the parasitaemia at time = 48hrs by the parasitaemia at time = 0hr
[2]Relative parasite fold change was normalized to 3D7 parasites by dividing the fold change in parasitaemia for each of the 3D7ΔEBA and 3D7Δ*Pf*GK knockout parasites by the fold change in 3D7 parasites. Results are expressed as a percentage.

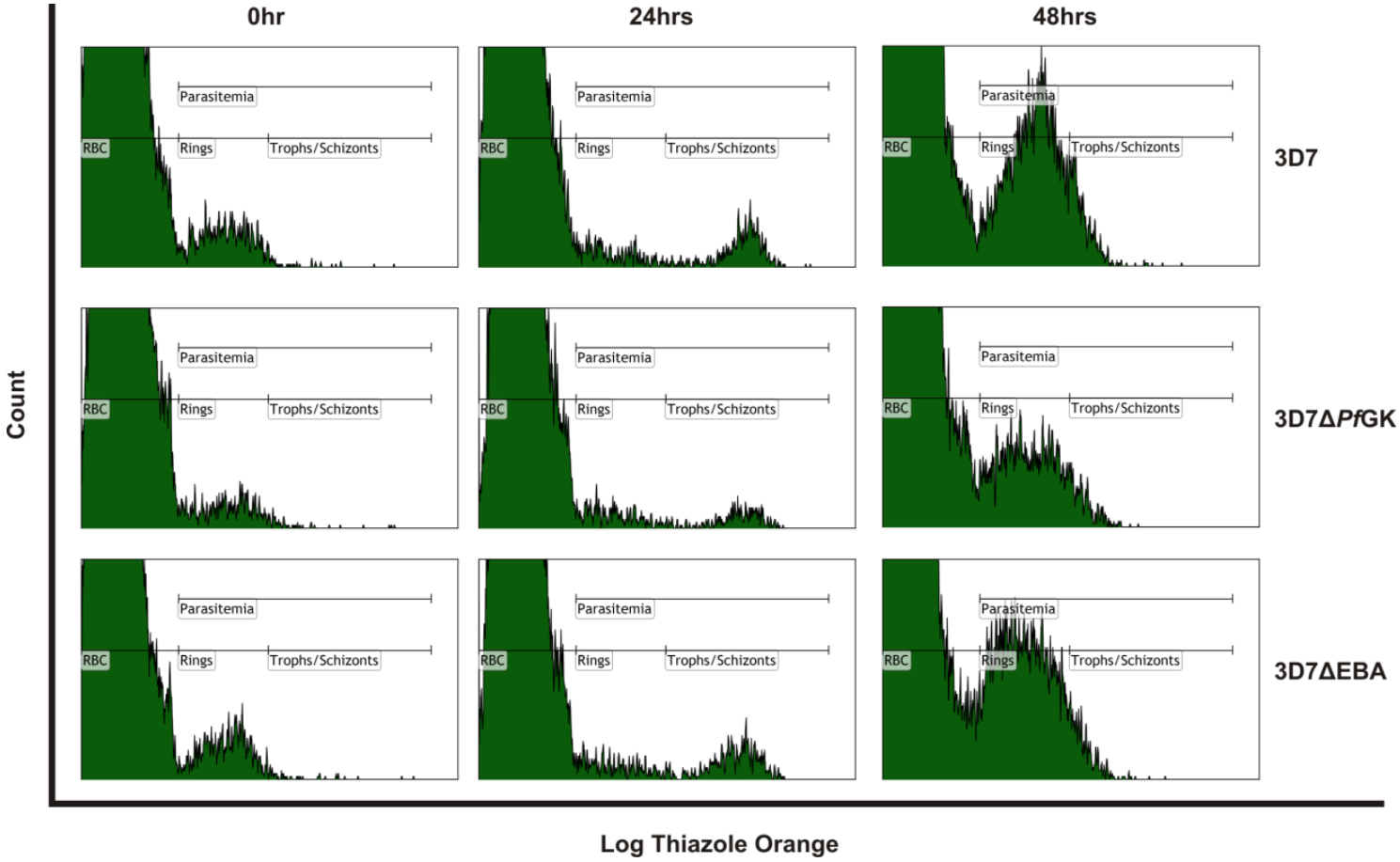



**Figure 44: Monitoring parasite growth using flow cytometry.**
The percentage parasitaemia at each indicated time point was measured using a single-parameter histogram for thiazole orange fluorescence. The ring and trophozoite/schizont stage parasite populations are indicated.

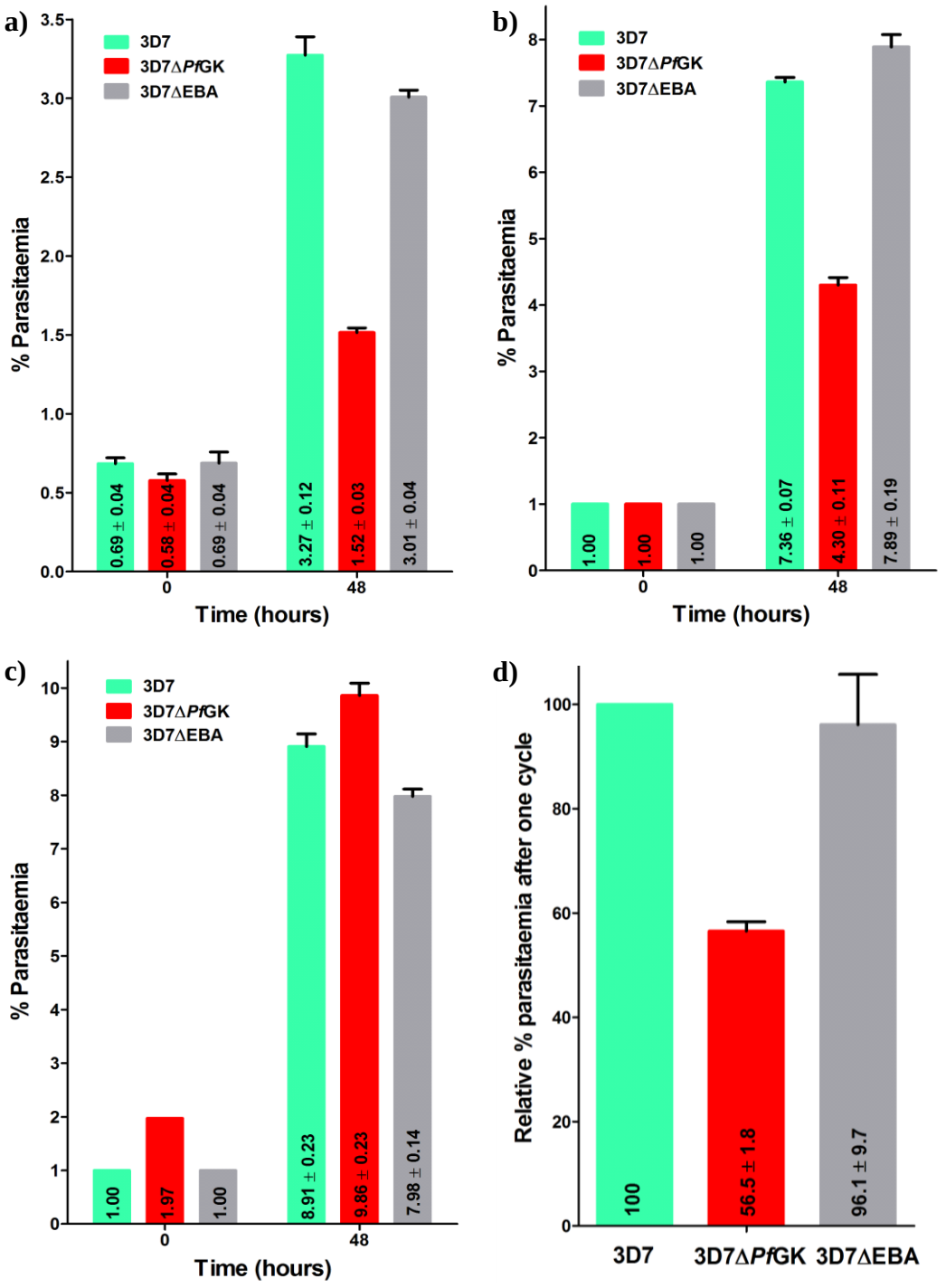



**Figure 45: Relative 3D7Δ*Pf*GK growth rate analysis.**
Monitoring the growth rates of wild type and 3D7Δ*Pf*GK knockout parasites; **(a)** summary of the results from the first (n = 4), **(b)** second (n = 4) and **(c)** third (n = 6) sets of flow cytometry experiments. Combined data analysis **(d)** showed that the parasitaemia fold change over ring- → ring-stage cycle is significantly lower in 3D7Δ*Pf*GK parasites compared to the wild type and 3D7ΔEBA control groups. Parasitaemia mean ± SD are indicated.

Bright field microscopy of Giemsa-stained blood smears showed no difference in the morphology of wild type and 3D7Δ*Pf*GK asexual parasites and gametocytes (Figure 46). However, the average number of schizonts formed per 3D7Δ*Pf*GK-infected red blood cell was 14.2 ± 2.5 (n = 108), which was approximately 20% ($p < 0.0001$) less than 17.8 ± 2.6 in the control parasites (n = 100) (Table 17).

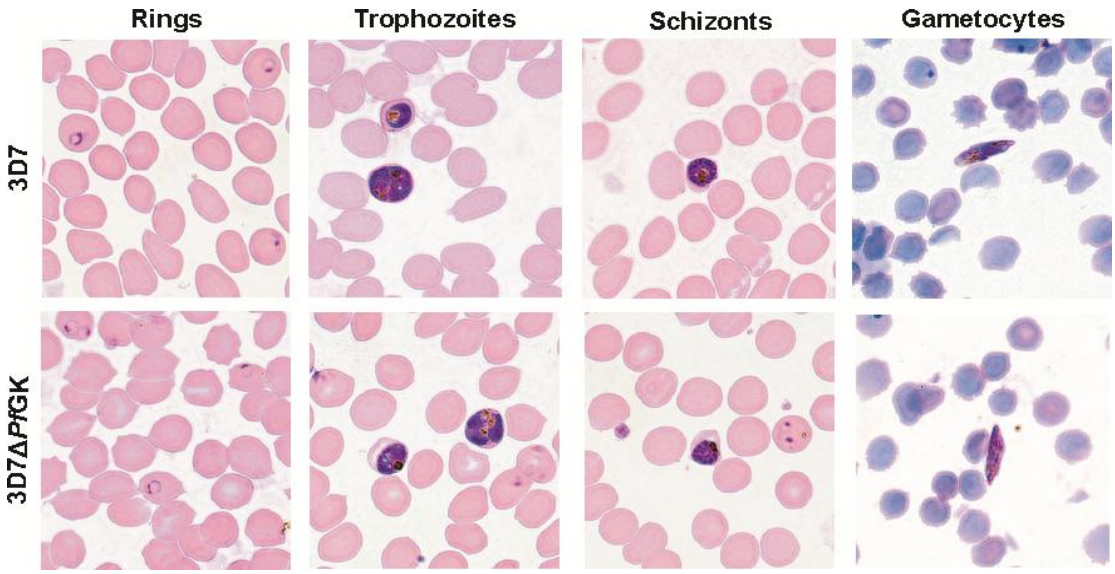



**Figure 46: Morphology of wild type and 3D7Δ*Pf*GK parasites.**
Asexual parasites and gametocytes are indistinguishable between wild type and glycerol kinase knockout *P. falciparum* parasites. Giemsa-stained smears were observed under oil immersion at 100x magnification using an Olympus IX70 inverted microscope with a SC30 colour camera (Olypmus Optical Corporation Ltd., Tokyo, Japan) and images were captured using the accompanying analySIS LS Research software, version 3.4.

**Table 17: Number of schizonts formed per parasite-infected erythrocyte.**

| Parasite line | Infected erythrocytes | Number of schizonts | |
|---|---|---|---|
| | | Range | Mean ± SD[1] |
| 3D7 | 49 | 14-26 | 17.9 ± 2.8[2] |
| 3D7ΔEBA | 51 | 14-24 | 17.7 ± 2.3[2] |
| Controls[3] | 100 | 14-26 | 17.8 ± 2.6[*] |
| 3D7Δ*Pf*GK | 108 | 10-20 | 14.2 ± 2.5[*] |

[1]Statistical analysis was performed using an unpaired t-test with a two-tailed p value
[2]No significant difference between the means ($p = 0.74$)
[3]Combined data from the 3D7 and 3D7ΔEBA parasites
[*]Significant difference between the means ($p < 0.0001$)

## 3.6 Identification of potential *Pf*GK drug target sites (DTS)

### 3.6.1 Evolutionary patterning

To limit resistance, targeting protein regions containing several evolutionary conserved amino acids that are different to the host orthologue is preferable. Evolutionary patterning (EP), developed by Dr Pierre Durand (2008), uses the ratio (ω) of non-synonymous (dN) nucleotide substitutions which change an amino acid, to synonymous (dS) nucleotide substitutions which do not change amino acids, to measure the substitution rates at a site. If a residue is critical to the structure or function of a protein and important to the fitness of the organism, natural selection will remove any changes that occur at that site at a rate that reflects its relative importance. If $\omega > 1.0$, there are more non-synonymous substitutions than synonymous substitutions. The site is therefore under positive selection and amino acid changes may be retained to increase the fitness of the organism. These sites are poor DTS. If $\omega < 1.0$, the site is under purifying selection; mutations are retained infrequently because they decrease the fitness of the organism. These may be regarded as potential DTS. For $\omega < 0.1$, the number of non-synonymous substitutions is at least 10 times less than synonymous substitutions. These residues are under extreme purifying selection, where mutations are unlikely to be retained as they are detrimental to the organism and these represent good DTS. EP was applied to *Pf*GK to identify potential DTS.

### 3.6.2 Identification of the *Pf*GK human orthologue

Five human glycerol kinase orthologues (*Hs*GK) have been identified (Appendix A2). To perform comparative analysis with *Pf*GK, the true *Hs*GK orthologue had to be identified. By examining the evolutionary relationship between 27 specifically selected GK proteins within the metazoa (Martinez Agosto and McCabe, 2006) and protozoa groups, which included six *Plasmodium* GKs in the

Apicomplexan subgroup, (section 2.1.2 and Appendix A2), the true human *Pf*GK orthologue could be identified.

Nucleotide MSAs were performed and DAMBE version 5.0.72 (Xia and Xie, 2001, Xia *et al.*, 2003) was used to determine saturation, defined as multiple nucleotide substitutions at a particular site within a codon (data not shown). In divergent nucleic acid sequences, multiple transition (pyrimidine ↔ pyrimidine or purine ↔ purine) and transversion (pyrimidine ↔ purine) nucleotide substitutions at a particular site result in saturation of that site. A consequence thereof is reflected in the low resolution of the phylogram, such as increased variance in branch lengths or an obscure picture of evolutionary relationships (Xia *et al.*, 2003). Analysis showed that the GK sequence was saturated (data not shown). Nucleic acid sequences were therefore not used in evolutionary analysis; protein MSAs were used instead. The aligned GK amino acid sequences were 494 amino acids in length, of which 423 were parsimony-informative, 67 were parsimony-uninformative and four amino acids were constant. A single most-parsimonious tree with a tree length of 423 amino acids, g1 = -0.631, CI = 0.697 and RI = 0.673 was generated with a >75% bootstrap support (log files not shown). *Ec*GK was used as the outgroup.

**Table 18: Identity and similarity of *Ec*GK, *Pf*GK and *Hs*GK orthologues.**

| **Protein*** | **Amino acids** | ***Ec*GK** | | ***Pf*GK** | | ***Hs*GK** | |
|---|---|---|---|---|---|---|---|
| | | **Identity** | **Similarity** | **Identity** | **Similarity** | **Identity** | **Similarity** |
| *Ec*GK (P0A6F3) | 502 | | | 39.5% | 57.5% | 30.1% | 47.7% |
| *Pf*GK (Q8IDI4) | 501 | 39.5% | 57.5% | | | 33.2% | 51.4% |
| *Hs*GK (Q6ZS86) | 529 | 30.1% | 47.7% | 33.2% | 51.4% | | |

*UniprotKB IDs are indicated in brackets

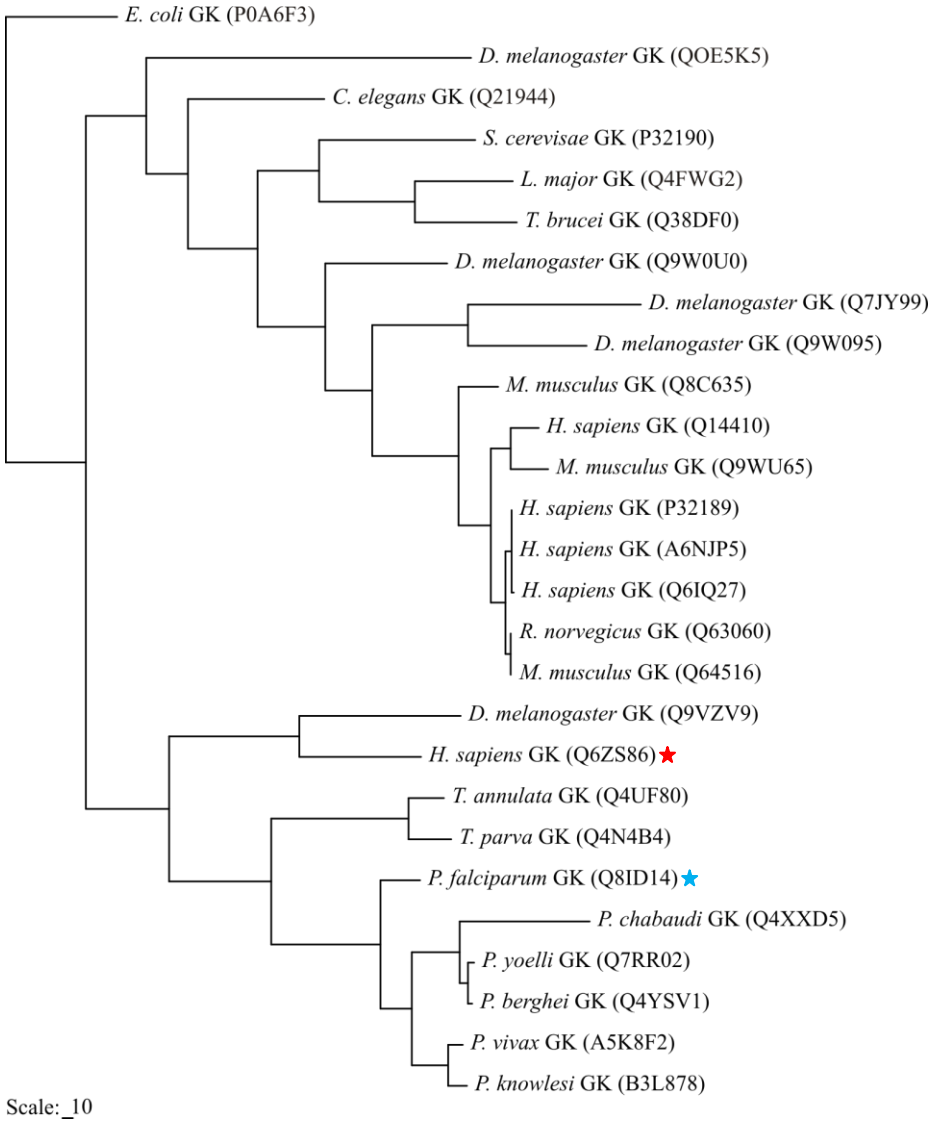



**Figure 47: Evolutionary relationship between GK proteins.**
Twenty-seven GK protein sequences were used to generate an unrooted phylogram (Appendix A2). Analysis was performed using the maximum parsimony method, with bootstrap analysis. *Ec*GK was used as the outgroup. Branch lengths represent estimated evolutionary distances, according to the scale bar (number of substitutions per site). GK Sequences were extracted from the UniprotKB database and UniProt IDs are indicated in brackets. The *Pf*GK (★) human orthologue (★) is indicated (Q6ZS86).

A 39.5% identity and 57.5% similarity to the *Ec*GK orthologue (Table 18) suggests that *Pf*GK is ancient and as a consequence of multiple nucleotide and amino acid changes over time, the evolutionary relationship of GKs could not be resolved completely. From the phylogram (Figure 47), *Hs*GK (Q6ZS86) is the only human enzyme that clusters with *Pf*GK and is therefore, the human *Pf*GK orthologue. These results were verified using phylogenic analysis by maximum

likelihood (data not shown) (Yang, 1997). Interestingly, *Pf*GK clusters separately from other *Plasmodium* GKs. This is consistent with the divergence of *P. falciparum* from most *Plasmodium* species (Hayakawa *et al.*, 2008) and further highlights differences between *P. falciparum* and the other species.

### 3.6.3 Structural features of *Pf*GK

The *Pf*GK orthologue is topologically identical in both the modelled (ID: PF13_0269-1Bu6) (Bahl *et al.*, 2002) and crystallized protein (Schnick *et al.*, 2009), with conserved catalytic residues (Figure 48). The *Pf*GK structural model, with functionally important residues based on homology to *Ec*GK, is shown in Figure 48b. A multiple sequence alignment between the crystallized *E. coli*, *P. falciparum* and the modelled human glycerol kinase proteins, as well as protein functional and structural residues, is summarized in Figure 49 and Table 19.

### 3.6.4 *Pf*GK drug target sites

To identify evolutionary conserved *Pf*GK amino acids, and potential drug target sites, a multiple sequence alignment between *Pf*GK and the human orthologue was performed (Figure 50). For illustrative purposes, an arbitrary length of five amino acids was chosen as a *Pf*GK DTS. Several stretches of residues under extreme purifying selection that may represent potential *Pf*GK DTS were identified: amino acids 82-86 (QRETV), 99-102 (NAIV), 244-247 (GDQQ), 266-270 (GTGVF) and 345-349 (FVPAF; indicated by black bars in Figure 50). However, most of these amino acids were identical in *Hs*GK and would thus not be good DTS. Other regions (indicated by green bars in Figure 50) were therefore considered, based on the following selection criteria: (i) they contained one or more residues that were under extreme purifying selection ($\omega \leq 0.1$), (ii) the adjacent residues were under strong purifying selection ($\omega < 0.2$), and (iii) *Pf*GK and *Hs*GK shared no more than two identical residues (Durand *et al.*, 2008).

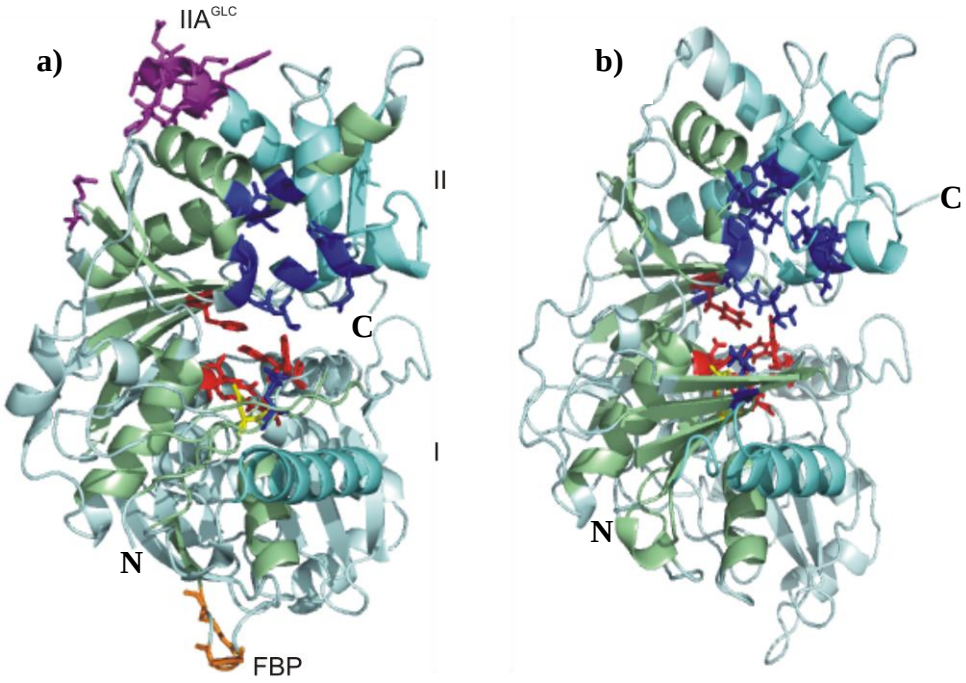



**Figure 48: *Ec*GK and *Pf*GK structures.**
Ribbon structure representations of the crystallized *Ec*GK and *Pf*GK monomers. Functional residues are illustrated as sticks; **(a)** an *Ec*GK monomer with residues involved in interactions with ADP (dark blue), $Mg^{2+}$ (yellow), glycerol (red), FBP (orange) and $IIA^{GLC}$ (purple) (Hurley *et al.*, 1993, Feese *et al.*, 1998). Domains I and II are labelled and the N- and C-terminus of each protein are indicated. α Helices and ß sheets within Domains I and II form the ATPase domain, shaded in light green and areas involved in subunit interactions are shaded in aquamarine; **(b)** the *Pf*GK model showing interdomain interactions and functional residues based on *Ec*GK. The modelled *Pf*GK structure is topologically identical to the crystallized *Pf*GK structure (Schnick *et al.*, 2009). Images were constructed using Pymol Molecular Graphics System, version 1.3, Schrödinger, LLC.

Six regions (labelled I to VI) were selected as potential DTS (Figure 50). To evaluate the structural and functional significance of the six potential DTS, the regions were mapped to the *Pf*GK structure (Figure 51).

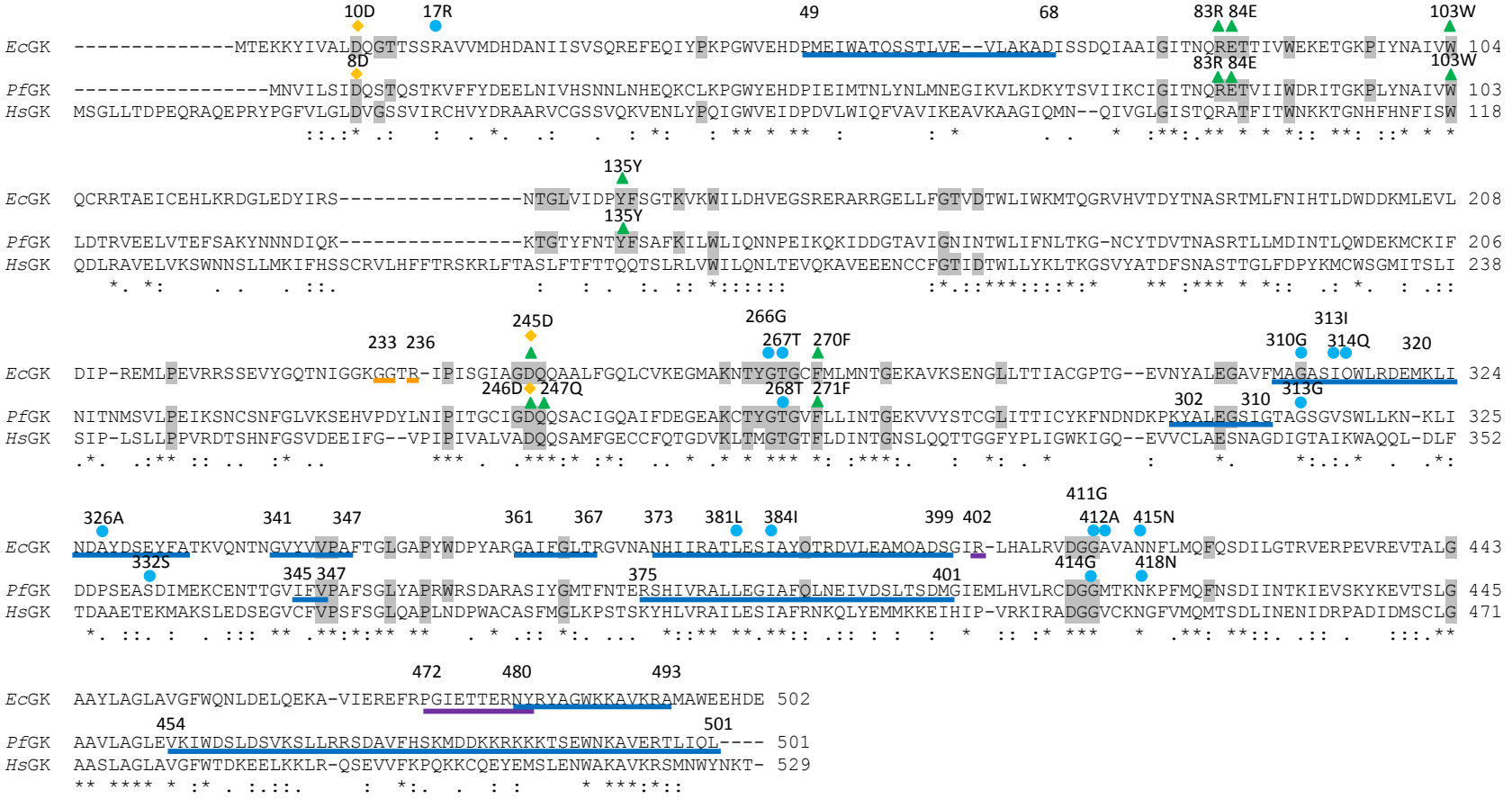



**Figure 49: Amino acid sequence alignment and functional and structural features of *Ec*GK, *Pf*GK and *Hs*GK.**
A ClustalW2 (Larkin *et al*., 2007) multiple sequence alignment of *Ec*GK (UniprotKB ID: P0A6F3), *Pf*GK (UniprotKB ID: Q8IDI4) and *Hs*GK (UniprotKB ID: Q6ZS86) orthologues. Identical (*), conserved (:) and semi-conserved (.) residues are indicated. Conserved residues from an MSA analysis of 19 GK orthologues (Kralova *et al*., 2000) are shaded grey. Functional $Mg^{2+}$- (◆), glycerol- (▲) and ADP- (●) residues from the crystallized *Ec*GK (Hurley *et al*., 1993) and *Pf*GK (Schnick *et al*., 2009) structures are numbered and indicated above each sequence. Residues involved in intersubunit interactions within the *Ec*GK tetramer and the *Pf*GK homodimer are underlined (▬). Fructose-1,6-bisphosphate (▬) and the phosphocarrier protein $IIA^{GLC}$- (▬) *Ec*GK regulatory binding domains are illustrated.

**Table 19: Summary of the functional amino acids in *Ec*GK and *Pf*GK (Hurley *et al.*, 1993, Schnick *et al.*, 2009).**

| Glycerol kinase | Functional domain and location | Amino acids |
|---|---|---|
| ***Ec*GK** | **Domain I** | 1-254 |
| | Glycerol-binding | Arg83, Glu84, Trp103, Tyr135, Asp245 and Phe270* |
| | $Mg^{2+}$-binding | Asp10 and Asp245 |
| | Intersubunit interactions | Monomer O: 49-68 and Monomer X |
| | **Domain II** | 255-502 |
| | ADP-binding | Arg17, Gly266, Thr267, Gly310, Ile313, Gln314, Ala326, Leu381, Ile384, Gly411, Ala412 and Asn415 |
| | $IIA^{GLC}$-binding | Arg402, Pro472, Gly473, Leu474, Glu475, Thr476, Thr477, Glu478, Arg479, Asn480, Tyr481 |
| | FBP-binding | Gly233, Gly234 and Arg236 |
| | Intersubunit interactions | Monomer O: 361-367 and Monomer Y: 341-347 |
| | | Monomer O: 308-320 and Monomer Y: 373-399, 480-493 |
| | | Monomer O: 321-333 and Monomer Z |
| ***Pf*GK** | **Domain I** | 1-262 and 436-471 |
| | Glycerol-binding | Arg83, Glu84, Trp103, Tyr135, Asp246, Gln247 and Phe271* |
| | $Mg^{2+}$-binding | Asp8 and Asp246 |
| | **Domain II** | 263-435 and 472-501 |
| | ADP-binding | Thr268, Gly313, Ser332, Gly414 and Asn418 |
| | Intersubunit interactions | 302-310, 345-347, 375-401, 454-501 |

*The only glycerol-binding amino acid situated in domain II

All the target regions, except for Region II, were found in functionally important domains as summarized in Table 20. Region I (amino acids 4-8) represents a β sheet at the N-terminus of *Pf*GK and is located within the catalytic cleft. It shares a common $Mg^{2+}$ binding site (D8) with *Hs*GK. Region II (amino acids 155-159) has two negatively charged residues and represents a loop structure between an α helix and β sheet at the surface of the molecule. It does not represent any known functional or structural GK domain. Region III (amino acids 163-167) forms an α helix-turn-β sheet structure and is located within the hydrophobic core of the molecule, below the glycerol-binding domain. It represents part of the ATPase core in Domain I. Region IV (amino acids 326-330) is part of a loop between two α helices and is located at the entrance to the catalytic cleft. Interestingly S332, which participates in ADP-binding, is under positive selection. The D327, P328 and E330 amino acids, however, are in close proximity and are under purifying selection and may confer conformational stability to the ADP-binding site. Region V (amino acids 340-344) represents a loop structure on the surface of the molecule and contains residues involved in intersubunit interactions in *Ec*GK dimerization. Region VI (amino acids 481-485) is situated in an α helix and is the site for *Ec*GK $IIA^{GLC}$-allosteric regulation. In *Pf*GK, Region VI forms part of a low complexity region which loops out on the surface of the molecule. It is part of the C-terminal region (amino acids 454-501) involved in intermolecular interactions with amino acids 302-310 and 375-401 in domain II. All five residues within Region VI are positively charged and share no similarity with the corresponding region of *Hs*GK, which has three negatively charged residues.

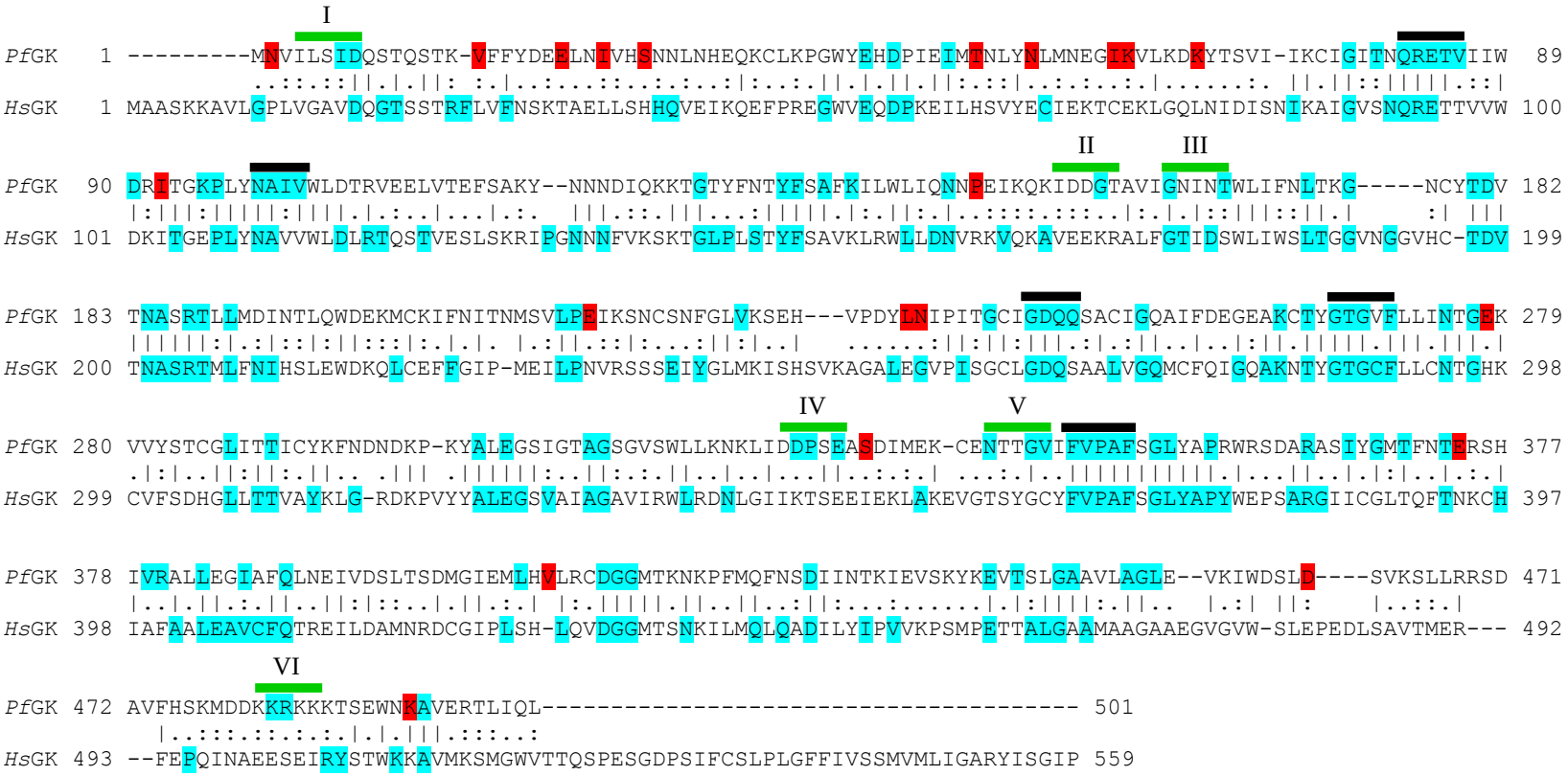



**Figure 50: Identification of potential drug target sites in *Pf*GK.**

A Needleman and Wunsch pairwise sequence alignment between *Pf*GK and *Hs*GK. Identical (|), conserved (:) and semi-conserved (.) residues are indicated. PAML and SLR were used to identify selective constraints in both *Pf*GK and *Hs*GK (Durand *et al.*, 2008). Residues highlighted in red are under positive selection ($\omega > 1.0$; data not shown for *Hs*GK). Residues highlighted in blue are under extreme purifying selection ($\omega \leq 0.1$). Six potential DTS, each composed of five amino acid residues, are indicated (▬). All DTS sites had at least one residue under extreme purifying selection while the rest were under strong purifying selection ($\omega < 0.2$). Five candidate *Pf*GK DTS with 4-5 residues under extreme purifying selection (▬) were excluded due to similarities with the *Hs*GK orthologue.

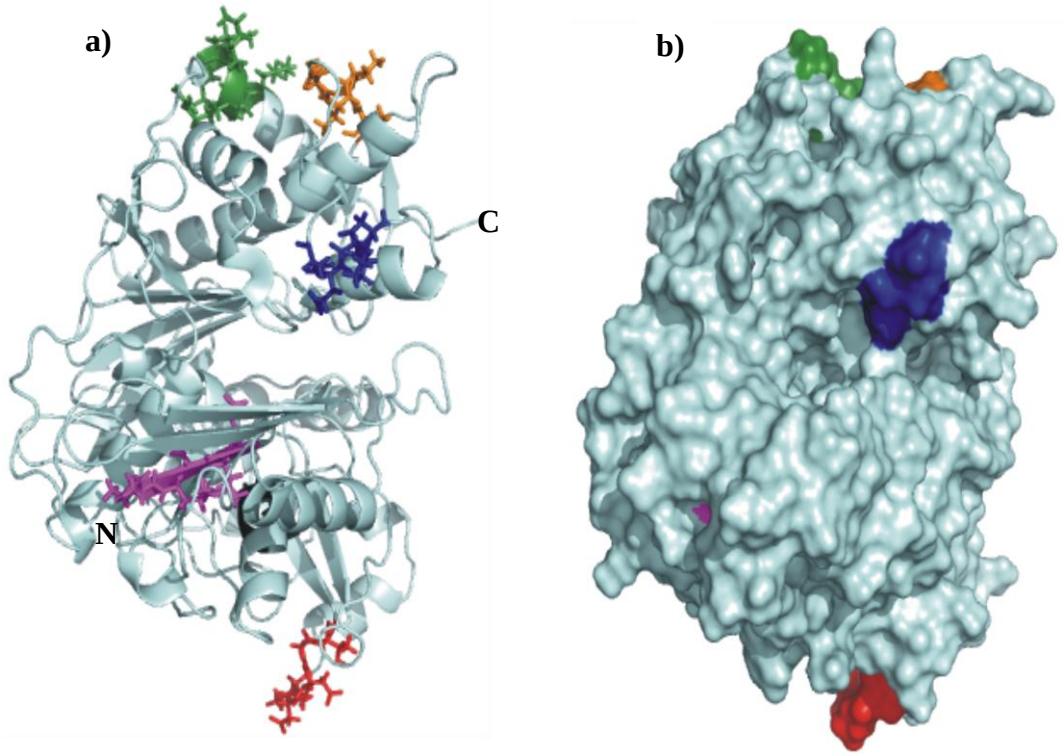



**Figure 51: Structural modelling of potential drug target sites in *Pf*GK.**
**(a)** *Pf*GK ribbon structure and **(b)** surface model showing potential drug target sites. Residues within each region are summarized in Table 20 and are highlighted as follows: Region I (magenta), Region II (red), Region III (black), Region IV (blue), Region V (orange) and Region VI (green). Region III is situated below Region I and is difficult to visualize in the ribbon configuration since it is partly obscured by one of the helices. It is within the core of the molecule and cannot be seen on the surface *Pf*GK model. The N- and C-terminus of the protein are indicated. Images were constructed using Pymol Molecular Graphics System, version 1.3, Schrödinger, LLC.

**Table 20: Summary of the location and function of potential *Pf*GK drug target sites.**

| Region | Colour[1] | Sequence | Functional domain[2] | Location |
|---|---|---|---|---|
| I | Magenta | ILSID | $Mg^{2+}$-binding | Catalytic Cleft |
| II | Red | IDDGT | None | Surface |
| III | Black[3] | GNINT | ATPase core | Core |
| IV | Blue | DDPSE | ADP-binding | Catalytic Cleft |
| V | Orange | NTTGV | Subunit interaction | Surface |
| VI | Green | KKRKK | Subunit interaction | Surface |

[1]Colours of each region correspond to those in Figure 51
[2]GK functional and regulatory domains are illustrated in Figure 48
[3]This region is in the core of the molecule and cannot be seen on the surface model. It is part of the ATPase core domain (Figure 10)

# Chapter 4: Discussion

## 4.1 r*Pf*GK is predominantly expressed as an insoluble protein in *E. coli*

An important requirement for the functional analysis of proteins is the ability to express and purify soluble proteins that closely resemble their native conformation. One of the most commonly used systems for the expression of recombinant proteins is the enterobacterium, *E. coli*. Other less frequently used expression systems such as yeast, mammalian cells, *Xenopus* oocytes and cell-free protein expression methods may also be used (Birkholtz *et al.,* 2008). However, the ease of genetic manipulation, speed, reproducibility and low cost has made *E. coli* the expression host of choice. Therefore, in this study, the bacterial protein expression system was used.

Although *Pf*GK has a high homology to *Ec*GK (39.5% identity; Table 18), bioinformatic analysis predicted that the protein would be expressed as a predominantly insoluble aggregate in *E. coli* (~65% insoluble; Table 12). Based on the known ability of certain affinity tags to improve the expression, solubility and native folding of a fusion protein, in addition to their use in protein purification (Sorensen and Mortensen, 2005a), *Pf*GK was expressed as a fusion protein. The technology relies on fusion of a foreign protein to a soluble partner protein, which drives and stabilizes the expression of that foreign protein. The rationale is attributed to the probable mRNA stability of the transcribed gene (Makrides, 1996). For example, an mRNA fragment encoding the C-terminal region of the *E. coli* $F_0$ ATPase subunit was stabilized following fusion to the green fluorescent protein (GFP). The GFP transcript provided protective mRNA structural elements which led to high yields of protein production (Arechaga *et al.,* 2003). Another example of solubility enhancing tags is the *E. coli* 40kDa maltose-binding protein (MBP), which was used in the expression and crystallization of *Pf*GK (Schnick *et al.,* 2009). Other examples are the 54.8kDa N-utilizing substance A (NusA), 11.6kDa thioredoxin and the *S. japonicum* 26kDa GST

protein. Although no single fusion tag can increase the expression and solubility of all target proteins, some fusion tags have been more successful than others. In a systematic comparison of the effectiveness of various fusion tags in increasing the solubility of six diverse proteins that are prone to aggregation, MBP was found to be superior to thioredoxin and GST as a solubilizing partner (Kapust and Waugh, 1999). The precise mechanism of the solubility enhancement of MBP is unclear; although it is thought that MBP acts as a chaperone which stabilizes an otherwise insoluble passenger protein (Kapust and Waugh, 1999).

In this study, the pGEX-4T-2 expression vector, which contains a GST tag upstream of the multiple cloning site, was used. GST is known to protect against intracellular proteolysis and stabilize soluble recombinant proteins. It has been used as an N- or C-terminal tag in a variety of expression systems and proved significant in studies on protein-DNA interactions, protein-protein interactions and as antigens for vaccine studies (Young *et al.*, 2012). Bioinformatic analysis predicted an improvement of ~5% in the production of soluble *Pf*GK when fused to GST. Protein expression analysis, however, showed that the fusion protein was expressed mainly as an insoluble aggregate in *E. coli* (section 3.1.2).

Several factors may have been responsible for the poor expression of soluble r*Pf*GK. Generally, the expression of *P. falciparum* proteins in *E. coli*, and in other expression systems, is problematic. Indeed, successful expression of proteins in *E. coli* is influenced by (i) the structural features of the inserted gene, (ii) transcription efficiency and stability of the transcribed mRNA, (iii) correct and efficient translation, (iv) correct folding of the protein and (v) toxicity of the protein (Schumann and Ferreira, 2004).

With an average A+T content of 76%, the genes coding for *P. falciparum* proteins are unique and there are a number of factors that contribute to the poor expression and solubility in heterologous protein expression systems (Aravind *et al.*, 2003): *P. falciparum* proteins (i) lack homology to *E. coli* proteins, (ii) are generally larger than their homologues in other species, (iii) can contain long repetitive

amino acid stretches, for example, a *P. falciparum* cyclin dependent kinase contains a segment of 83 consecutive asparagines (Aravind *et al.*, 2003) and (iv) often possess low-complexity regions (LCRs) (DePristo *et al.*, 2006). The *P. falciparum* proteome is one of the most LCR-rich among eukaryotes, being present in 87% of genes compared to the average of 65-70% (DePristo *et al.*, 2006). None of these features, however, are present in *Pf*GK and therefore cannot account for the poor solubility. *Pf*GK is an evolutionary well conserved enzyme (section 3.6.2 and Appendix A2), with a strong homology to *Ec*GK and to other Apicomplexan GK orthologues. The gene and amino acid sequence (1,506bp, 501 amino acids) are almost identical in length to the *E. coli* orthologue (1,509bp, 502 amino acids). *Pf*GK has no striking long repetitive amino acid repeats, although it contains a small lysine-rich LCR at the C-terminus of the protein. Although there is no universal functional explanation for LCRs, they are thought to (i) enhance mRNA stability (Xue and Forsdyke, 2003), (ii) be involved in protein-protein interactions (Karlin *et al.*, 2002), (iii) act as recombination hotspots (Zilversmit *et al.*, 2010) and (iv) generate antigenic diversity where LCRs act as decoys for the host immune system (Hughes, 2004). The *Pf*GK LCR is an A+T-rich segment (26 out of 30 bases) that codes for amino acids $^{477}$KMDDKKRKKK$^{486}$. It is distal to the catalytic domains and located on α-helix 16, the last helical structure which is involved in interdomain protein interactions (section 3.6.3).

The codon usage and frequency of *P. falciparum* is different and limiting in *E. coli*. This can result in translational stalling, whereby the bacterium produces low yields or truncated forms of the protein (Sorensen and Mortensen, 2005a). Several methods can be used to compensate for the codon bias in *E. coli*. The *E. coli* intracellular tRNA pool, for example, can be expanded via the use of plasmids that encode tRNAs for rare codons. In this study, r*Pf*GK was expressed in Rosetta2 (DE3) *E. coli* cells. These cells contain a pRARE plasmid which expresses tRNAs for seven rare codons in *E. coli*, including the AUA codon for isoleucine. AUA is present at a frequency of five per 1,000 codons in *E. coli*, which is low compared to 50 per 1,000 codons in *P. falciparum* (Appendix A8).

Despite the use of Rosetta2 (DE3) *E. coli* cells, soluble r*Pf*GK protein production was poor.

Other strategies that may improve soluble protein expression include codon optimization, where a *P. falciparum* sequence with codons representative of the tRNA pool in *E. coli* is synthesized. Expression of the *P. falciparum* FALVAC-1 multistage vaccine candidate, for example, was increased 3-fold in a codon optimized construct (Zhou *et al.*, 2004). Codon harmonization via site-directed mutagenesis can also be used to modify a *P. falciparum* sequence at low frequency codons that are responsible for translational stalling. Codon harmonization was applied to express the full-length *Pf*s48/45 *P. falciparum* transmission blocking vaccine candidate in a conformation that elicited an immune response in mice and primates (Chowdhury *et al.*, 2009). Despite these successes, two independent large-scale studies on *Plasmodium* protein expression showed no correlation between codon usage and soluble protein expression in *E. coli* (Mehlin *et al.*, 2006, Vedadi *et al.*, 2007).

In an attempt to define some of the features that influence the expression and solubility of recombinant parasite proteins, and indirectly illustrate the difficulties of expression of *P. falciparum* proteins, Mehlin *et al* (2006) expressed 1,000 open reading frames that were representative of the *P. falciparum* genome. From the 1,000 genes, 337 recombinant proteins were expressed of which only 63 were soluble, producing 0.9-407mg soluble protein from 1L of induced culture. The study showed that proteins with a higher molecular weight, greater protein disorder, more basic pI and low homology to *E. coli* proteins were negatively correlated with protein expression and solubility.

Of the 98 target proteins that were <20kDa in size, 41 were expressed and 14 were soluble. In contrast, out of the 100 target proteins that were >60kDa in size, only 20 were expressed and three were soluble (Mehlin *et al.*, 2006). The molecular weight of the native *Pf*GK protein is ~57kDa, whilst the GST-fusion recombinant protein is ~83kDa. Statistically, this suggests that r*Pf*GK was (i) unlikely to be

expressed and (ii) if expressed, it may be poorly soluble. The latter was observed in our experiments.

Disordered regions are usually highly flexible protein regions with no secondary structures and were not conducive to the production of soluble proteins (Mehlin *et al.*, 2006). However, analysis using DISOPRED2 (Ward *et al.*, 2004) predicted no disordered regions in *Pf*GK (data not shown).

The correlation between the protein's pI and solubility in the Mehlin *et al* study (2006) was striking: of the 288 target proteins with a pI > 10, only one protein was soluble. Proteins with a high pI were correlated with insolubility whereas those with a low pI (< 6.0) showed a decrease in protein expression, but increased solubility. The pI of the native *Pf*GK protein is 5.92, whilst the GST-fusion recombinant protein is 6.0. Although the pI favoured the production of soluble r*Pf*GK, the yield of soluble protein was very low.

Mehlin *et al* (2006) observed that soluble protein expression was independent of codon usage and the A+T bias of the *P. falciparum* genome. This was contrary to the general assumption. The genes coding for 12 proteins which did not express in the bacterium were codon optimized for expression in *E. coli* and synthesized. Three of these constructs produced the full length protein and four produced truncated versions of the protein, however, none were soluble. This was verified in another study by Vedadi *et al* (2007). Hence, *Pf*GK was not codon optimized for expression in *E. coli*. In the Vedadi *et al* (2007) study, 468 *P. falciparum* reference genes which code for haemoglobin metabolism and biosynthesis, hypothetical, exported and apicoplast-targeted proteins were selected for analysis. With *P. falciparum* as the model organism, orthologues from six other Apicomplexan species: *P. yoelii*, *P. vivax*, *P. berghei*, *P. knowlesi*, *C. parvum* and *T. gondii* were identified and a total of 1,008 genes were analyzed. Only 304 (30.2%) proteins from the seven genomes were expressed in soluble form, of which 97 were crystallized and 36 distinct crystal structures resolved. A significant finding was that by regarding the seven genomes as a 'super

Apicomplexan genome' and treating orthologous genes as alternate expression constructs, at least twice the percentage of purified proteins, crystals and structures were obtained when compared to those from *P. falciparum* alone.

Another limitation of the *E. coli* protein expression system, and not unique to *P. falciparum* proteins, is that *E. coli* is sensitive to the production of foreign proteins and over-expressed proteins often aggregate into insoluble inclusion bodies. One advantage of inclusion bodies is that they contain very little host protein, ribosomal components and nucleic acid fragments and because they are highly concentrated with misfolded over-expressed protein aggregates, these inclusion bodies can be purified and the protein refolded under certain conditions (Singh and Panda, 2005). Indeed, there has been some success when working with refolded *P. falciparum* proteins. For example, refolding procedures were applied to successfully yield 12mg of purified and functional *P. falciparum* falcipain-2, a cysteine protease, from a 5L induced culture. The ~30kDa refolded protein was biochemically identical to the native falcipain-2 protein and the large amounts were sufficient for crystallization of the protein (Sijwali *et al.*, 2001). In this study, attempts to refold the insoluble r*Pf*GK protein were made, however, most of the protein precipitated during the renaturing process. Although 80-800ng of purified protein was refolded from a milligram of insoluble protein, the refolded r*Pf*GK protein was inactive. Refolding then, does not guarantee that the refolded protein will retain its native conformation (Vallejo and Rinas, 2004).

Inclusion body formation is influenced by non-native intermolecular hydrophobic interactions between protein intermediates that have not yet buried their hydrophobic amino acids. Due to the strong temperature dependence of hydrophobic interactions, aggregate formations are favoured at higher temperatures (Kiefhaber *et al.*, 1991). By lowering the temperature, the rate of protein translation is reduced and this favours the slower and correct folding of the protein. To promote soluble protein expression, r*Pf*GK was expressed in *E. coli* at room temperature, with expression induced at post-log phase. These changes slow down the translational machinery of the host. Two protein induction methods were used: IPTG and Overnight Express™ Instant TB Medium. The pGEX vector

is under the control of the *tac* promoter. The vector contains an internal *LacI*$^{q}$ gene which codes for a lac repressor protein. The lac repressor binds to the operator region of the *tac* promoter, preventing recombinant protein expression. Addition of the non-hydrolyzable lactose analog, IPTG, which binds to the lac repressor, allows for regulated expression of the foreign gene. The Overnight Express™ Instant TB Medium, in contrast, is an autoinduction system that combines Terrific Broth with three additional components to formulate a rich culture medium: (i) an induction solution containing a blend of carbon sources optimized for uninduced bacterial growth to a high cell density, followed by induction with lactose, (ii) a buffer solution to maintain pH and provide additional nitrogen necessary for increased protein synthesis and (iii) a magnesium solution necessary for maximum cell density (Novagen, 2005). In this study, approximately 3µg of purified and soluble r*Pf*GK from an 800ml induced culture was obtained. No significant difference in soluble protein yield was observed between the IPTG and Overnight Express™ Instant TB Medium induction methods. The yield was relatively low and in an attempt to improve the amount of soluble r*Pf*GK, an alternative baculovirus-insect cell protein expression system was used.

## 4.2 r*Pf*GK is not expressed in insect cells

Baculoviruses are a diverse group of large DNA viruses that are pathogenic to arthropods. The virions are enveloped and rod-shaped, with a circular double-stranded DNA genome of 80-180kbp, depending on the virus species. Baculoviruses replicate in the nucleus of insect cells and during their biphasic life cycle, two morphologically distinct virion phenotypes are produced: (i) the lytic extracellular or budded viruses which spread infection within the insect and (ii) the occlusion derived viruses which develop and mature in the host nucleus and are responsible for primary infection as well as horizontal viral transmission. Occlusion derived viruses are embedded within a crystalline proteinaceous matrix, termed occlusion bodies which are predominantly composed of a 33kDa polyhedrin protein. Following virus maturation, a fibrillar structure composed of a

10kDa p10 protein is formed. p10 is involved cell lysis and in the release of occlusion bodies from the nucleus. Polyhedrin and p10 are highly expressed during the very late stages of viral replication and assembly and are under the control of the strong *polh* and *p10* viral promoters. Both proteins are required for completion of the virus infection cycle, but are non-essential for budded viruses formation. Therefore, *polh* and *p10* promoters can be used to drive the expression of a foreign protein and this forms the basis of the baculovirus expression vector system (BEVS) (van Oers, 2011).

*Autographa californica* nucleopolyhedrovirus (AcNPV) is the most commonly used baculovirus for the production of foreign proteins in insect cells and its use in the synthesis of soluble and functional recombinant proteins is well established (Kost *et al.*, 2005). Being a eukaryotic-based expression system, the folding, disulfide bond formation, oligomerization, glycosylation, acylation and proteolytic cleavage of newly synthesized proteins may be similar to the natural process in eukaryotes. The BEVS has been successfully applied to several *P. falciparum* proteins. Examples include the expression of *Pf*EBA-175 in a soluble form that mimics its native functional conformation (Liang *et al.*, 2000) and a significantly improved expression, yield and purity of PFD1235w domains of the blood-stage vaccine candidate, PfEMP1, when compared to the expression in *E. coli* (Victor *et al.*, 2010). The BEVS has also been successfully used as a recombinant protein display system where expression of the *Pf*s25 transmission blocking vaccine candidate antigen on the viral envelope elicited an immune response in mice and rabbits (Mlambo *et al.*, 2010). In the Mehlin *et al* (2006) study, 17 proteins that were insoluble in *E. coli* were transferred to a baculovirus-*Sf*21 insect cell expression system. Seven of the 17 genes (41.2%) produced soluble protein, with yields ranging from 0.9-11.7mg of soluble protein per 500ml culture. The proteins were considerably less pure and the yield was likely to be overestimated, however, it was emphasized that the BEVS is a good alternative for the expression of *P. falciparum* proteins.

In this study, the *Pf*GK gene was integrated downstream of the *p10* baculoviral promoter and infective recombinant *Pf*GK-baculoviruses were generated. However, several attempts to express the r*Pf*GK protein in *Sf*9 insect cells were unsuccessful.

The specific reasons for the lack of r*Pf*GK protein expression are unclear. It is possible that synthesized r*Pf*GK protein was incorrectly folded and subsequently degraded by the host's cellular proteases. The lack of expression may also be ascribed to the poor status of the cell following infection; the p10 promoter is expected to drive exogenous protein expression at a very late stage, by which time, the cell machinery and environment may have deteriorated (Li *et al.*, 2012, Vlak *et al.*, 1988). Continuous production of recombinant baculoviruses often results in the formation and accumulation of defective interfering viral particles (Kool *et al.*, 1991). These particles are shorter than the parental baculovirus and cannot propagate autonomously due to a deletion of a large portion of their genomic DNA, which can be up to 43% of the viral genome and include the inserted foreign gene. A consequence of defective interfering viral particles is that the budded viral titre as well as the level of recombinant protein expression declines upon serial viral passage. For example, the expression of GFP in recombinant baculoviruses was reduced to zero after 30 viral passages (van Oers, 2011). r*Pf*GK protein expression experiments, however, were performed shortly after infective viral titres were determined (< 10 passages).

Although the BEVS can perform appropriate post-translational modifications, the high expression level of recombinant proteins in the late stages of the viral life cycle can lead to incorrect processing. For example, N- and O-glycosylation occur at very low levels in *P. falciparum* (Gowda and Davidson, 1999) and the native *Pf*EBA-175 protein is essentially unglycosylated. However, the recombinant baculovirus protein was ~20% glycosylated, even though the immunogenicity of *Pf*EBA-175 was unaffected (Liang *et al.*, 2000). The differences in the complexity of the N-glycosylation patterns between lower eukaryotes and insect cells can also affect the folding of recombinant glycoproteins, and even lead to their degradation

in the ER (Helenius and Aebi, 2004). *Pf*GK has no glycosylation or other post-translational modifications sites, however, whether the recombinant protein is modified and/or degraded in the BEVS is unclear. Generally, the expression of each recombinant protein in the BEVS varies, with cytoplasmic proteins expressed at significantly higher levels compared to secreted or transmembrane proteins (van Oers, 2011). *Pf*GK, however, is a cytoplasmic protein, with no export signal or transmembrane domains. Similar to bacterial expression systems, rare codon usage may result in inefficient translation of the foreign gene; although there is no strong codon bias in insect cells and the correlation between codon optimization and recombinant protein expression in the BEVS is yet to be determined (van Oers, 2011). As a precaution, the *Pf*GK gene was codon-optimized for expression in *Sf*9 insect cells, but this had no effect.

In summary, the absence of r*Pf*GK expression in the BEVS further highlights that there is no universal procedure or system for the expression of soluble *P. falciparum* proteins; each protein has to be assessed independently as its expression varies in different hosts and under different conditions.

## 4.3 r*Pf*GK is active

The recombinant GST-*Pf*GK fusion protein was active. rGST was included as a control reaction during the enzyme activity analysis and had no effect on the reaction. Therefore, it was not necessary to remove the GST-moiety from the *Pf*GK fusion protein prior to activity analysis. The native *Pf*GK protein exists as a dimer and involves the interaction of amino acids 302-310, 344-347 and 454-501 in domain II at the C-terminus of two *Pf*GK monomers (Schnick *et al.*, 2009). The recombinant GST-*Pf*GK protein used in this study has GST conjugated to the N-terminus of *Pf*GK. We cannot exclude the possibility that GST may disrupt *Pf*GK dimerization, however, this is unlikely since the enzyme was active with $K_m$ values similar to that of the characterized *Pf*GK dimer (Schnick *et al.*, 2009). The purified r*Pf*GK enzyme was unstable. r*Pf*GK activity decreased significantly each day, even though the enzyme was stored at 4°C between analyses. A possible

reason for the poor stability is that the enzyme was stored in the elution buffer (500mM reduced glutathione, 50mM Tris-HCl, pH 8.25) and the high glutathione concentration may not have been ideal. Various GK enzyme storage buffers have been reported: (i) the *Ec*GK orthologue is highly stable in saturated ammonium sulphate (Thorner and Paulus, 1971), (ii) recombinant *T. brucei* GK was stored in 100mM Triethanolamine-HCl, pH 7.6, 200mM NaCl and 100mM imidazole during the characterization of the enzyme (Kralova *et al.*, 2000) and (iii) depending on the manufacturer, recombinant *Hs*GK may be supplied in 20mM potassium phosphate buffer, pH 7.5 (Creative Biomart, Shirley, USA) or 50mM Tris, pH 8.0 and 10mM reduced glutathione (Abnova Corporation, Taipei, Taiwan). Therefore, with no uniform glycerol kinase storage buffer, and to prevent a potential loss of an already limited amount of enzyme when removing the reduced glutathione via dialysis, a decision was taken to perform all enzymatic analysis immediately after r*Pf*GK was purified. The glutathione concentration would have been diluted to 5mM in the enzymatic reaction, which is closer to the 0.22-2.6mM physiological concentrations of reduced glutathione in *P. falciparum* trophozoites (Tripathi *et al.*, 2007). Another possible reason for the poor stability was that the purified r*Pf*GK had a concentration of 2-5ng/µl. This is significantly lower that the > 1µg/µl protein concentration that is correlated with protein stability in solution. Protein concentrations at < 1µg/µl are more prone to inactivation and loss as a result of low-binding to the storage tube (Pierce, 2009).

r*Pf*GK exhibited a broad temperature and pH activity range. Interestingly, the enzyme activity was higher at febrile temperatures. r*Pf*GK was active between pH 6.8-8.4, with a significant loss of activity at pH 5.0, indicating that the enzyme plays no role in the parasite food vacuole. The activity was highest at pH 8.0 which corresponds to the environment in the mosquito midgut. The mRNA expression profile of *Pf*GK (Figure 52) is highest in the transmissible-forms of the parasite, the gametocytes (Le Roch *et al.*, 2003, Aurrecoechea *et al.*, 2009). This suggests that the enzyme may play a more significant role in mosquito-stage development. mRNA expression does not necessarily indicate the presence of the protein (Moreira *et al.*, 2004, Le Roch *et al.*, 2004), however, 14 *Pf*GK peptides

were identified in early and late gametocytes (Le Roch *et al.*, 2003, Aurrecoechea *et al.*, 2009). The amount of protein the the sexual mosquito stages is not known. Interestingly, the activity of r*Pf*GK was ~60% lower at temperatures similar to the mosquito environment (28°C). This implies that more *Pf*GK molecules may be required to perform a comparable glycerophospholipid turnover to the human blood stages. The role of *Pf*GK in gametogenesis and the remaining mosquito stages is uncertain.

r*Pf*GK obeyed Michaelis-Menten kinetics. The $V_{max}$ was 2.15nmol NADH/min for glycerol and 1.66nmol NADH/min for ATP. The $K_m$ values of 15.7µM for glycerol and 15.9µM for ATP were comparable to the *Pf*GK kinetic analysis from Schnick *et al* (2009) ($K_{m(glycerol)}$ = 18 ± 2µM; $K_{m(ATP)}$ = 21 ± 1µM) and to that of *Ec*GK ($Km_{(glycerol)}$ = 16µM; $K_{m(ATP)}$ = 16µM) (Feese *et al.*, 1998). Unlike the *Ec*GK orthologue, r*Pf*GK was not inhibited by the allosteric regulator, fructose-1,6-bisphosphate (FBP). FBP traps the *Ec*GK tetramer in an inactive state at a $K_i$ = 0.25mM (Ormo *et al.*, 1998). At the highest concentrations of FBP tested (10mM), the $V_{max}$, $K_m$ and specific activity of r*Pf*GK decreased approximately 10%. This suggests that FBP is not a physiological inhibitor of *Pf*GK. Sequence analysis and the crystallized *Pf*GK structure (Schnick *et al.*, 2009) showed that *Pf*GK lacks the features associated with the FBP binding site: the loop structure to which FBP binds (i) is orientated towards the core of the dimer making it inaccessible and (ii) lacks a glycine-rich segment, as well as the essential arginine residue, which is involved in ionic interactions with phosphates from FBP.

## 4.4 Parasite growth is reduced in glycerol kinase knockout *P. falciparum* parasites

The role and importance of glycerol metabolism during parasite development are controversial (Dechamps *et al.*, 2010b, Daily *et al.*, 2007, Olszewski and Llinas, 2010). To investigate this, the glycerol kinase gene in *P. falciparum* was disrupted and the growth rates of 3D7Δ*Pf*GK knockout parasites were evaluated over one 48-hour developmental cycle using thiazole orange DNA staining coupled to flow

cytometry. 3D7Δ*Pf*GK knockout parasites proliferated significantly slower, at 56 ± 1.8%, compared to the wild type and 3D7ΔEBA control parasite lines. In support of this, 3D7Δ*Pf*GK parasites appeared after 51 days on positive selection which was approximately 59% longer than the 32 days it took for 3D7ΔEBA parasites to appear. Parasites lacking *Pf*GK also formed ~20% fewer schizonts compared to controls, presumably due to an insufficient supply of PLs which impaired membrane biogenesis.

Glycerol transport in *Plasmodium* and red cells has been studied extensively. Erythrocytes are known to be highly permeable to solutes, including glycerol. The primary glycerol facilitator in human red cells is aquaglyceroporin 3 (*hs*AQP3), which is internalized and integrated into the parasitophorous vacuole membrane following *Plasmodium* infection (Bietz *et al.*, 2009). Together with the highly permeable bi-functional water and glycerol *P. falciparum* aquaglyceroporin (*Pf*AQP) (Hansen *et al.*, 2002), parasites can easily acquire glycerol from the host. A defect in glycerol transport is known to reduce *Plasmodium* virulence. Two mouse model studies have shown that *P. berghei* proliferation was reduced when glycerol transport to the parasite was compromised. *Pf*AQP is 62% identical to *Pb*AQP and AQP-null *P. berghei* (Δ*Pb*AQP) parasites proliferated significantly slower relative to wild-type (WT) parasites. Although Δ*Pb*AQP parasites were viable and indistinguishable from WT parasites, mice infected with Δ*Pb*AQP parasites survived longer; WT parasites reached a parasitaemia of ~40% nine days post-infection whereas Δ*Pb*AQP parasites reached a comparable parasitaemia at day 18 (Promeneur *et al.*, 2007). In the second study, it was established that the primary glycerol facilitator in mouse erythrocytes is AQP9 (Liu *et al.*, 2007). Apart from a defect in transport of glycerol across their membranes, erythrocytes of AQP9-null mice appeared to be morphologically and physiologically normal. However, mice with AQP9-null RBCs survived longer during the early stages of *P. berghei* parasite infection compared to WT mice. Although both groups of mice eventually succumbed to infection; at 12 days post-infection, all of the AQP9-null mice were alive whereas only 50% of the WT mice survived (Liu *et al.*, 2007). These studies highlight the contribution of glycerol during intra-

erythrocytic *Plasmodium* growth and development. Interestingly, *Pb*GK is refractory to targeted gene disruption in *P. berghei* parasites (personal communication referenced in Schnick *et al* (2009)) indicating that the gene may be essential in parasite development in the rodent model and further highlighting metabolic differences between *P. falciparum* and the rodent parasites (Dechamps *et al.*, 2010a).

Findings from the current study correlate with *in vivo* human studies performed by Daily *et al* (2007) who profiled parasite gene expression in 43 Senegalese patients infected with *P. falciparum*. Despite similar host clinical features, three *in vivo* states, based on parasite transcription patterns superimposed onto the yeast transcriptome, were identified: (i) an active growth based on glycolytic metabolism and similar to known *in vitro* ring stage parasite profiles (Llinas *et al.*, 2006), (ii) a starvation response accompanied by metabolism of alternative carbon sources and (iii) an environmental stress response in individuals with increased inflammatory markers. *Pf*GK and *Pf*AQP were amongst those transcripts that were significantly upregulated during the starvation response. In addition, glycerol-3-phosphate dehydrogenase was markedly elevated, implying that glycerol could replenish glycolysis at the dihydroxyacetone level, possibly in hypoglycaemic patients (Hansen *et al.*, 2002, Daily *et al.*, 2007). Re-analysis of the Daily data (2007) by Lemieux *et al* (2009) showed that there was no segregation of genes into separate groups and they suggested that the discrete clusters were due to varying proportions of sexually committed but phenotypically indistinguishable parasites. However, Wirth *et al* (2009) defended their initial observations and ruled out gametocyte contamination in their samples. The absence of differences (Lemieux *et al.*, 2009) was probably due to the cultivation of parasites *ex vivo* in enriched medium prior to analysis (LeRoux *et al.*, 2009).

During *Plasmodium* infection, glucose metabolism is increased up to 100-fold to meet the demanding energy requirements for growth and development. The culture medium in this study had a glucose concentration of approximately 10mM, which is almost twice the physiological glucose concentration in a healthy

individual and sufficient to meet the parasite's energy and PL demands, if G3P is derived solely from glycolysis. Our findings of drastically reduced parasite proliferation when glycerol metabolism was disrupted imply that glycerol is an important alternative source of G3P. In support of this, $^{14}C$ labelling studies showed that glycerol was readily and efficiently incorporated into the PL backbone of *P. knowlesi* membranes (Rock, 1971). Furthermore, the physiological concentration of glycerol in healthy individuals ranges between 50-100μM (Lin, 1977); but in *Plasmodium*-infected patients, basal glycerol concentrations (133 ± 65μM) and turnover are doubled (Pukrittayakamee *et al.*, 1994). Recently, Lian *et al* (2009) identified glycerol as a major end-product of glycolysis in *P. falciparum*-infected red cells, and proposed a G3P shuttle mechanism to account for this. However, this could have been an end-product of a non-specific phosphatase (Olszewski and Llinas, 2010) or *Pf*GK may have phosphatase activity, like *T. brucei* GK (Kralova *et al.*, 2000). The reverse *Pf*GK reaction could be measured spectrophotometrically using hexokinase and glucose-6-phosphate dehydrogenase as auxiliary enzymes and monitoring the change in NADPH absorbance at 340nm (Kralova *et al.*, 2000). Regardless of the mechanism, glycerol can be generated by the malaria parasite and the glycerol pool in the human host is sufficient as an additional carbon and PL source for the parasite.

The level of *Pf*GK mRNA expression during asexual development is controversial. In this study, RT-PCR analysis showed that *Pf*GK mRNA was present in wild type and 3D7ΔEBA parasites, consistent with previously published 3D7 transcriptional analysis (Le Roch *et al.*, 2003) (Figure 52) and supports data from the rodent model (personal communication as referenced in Schick *et al* (2009)).

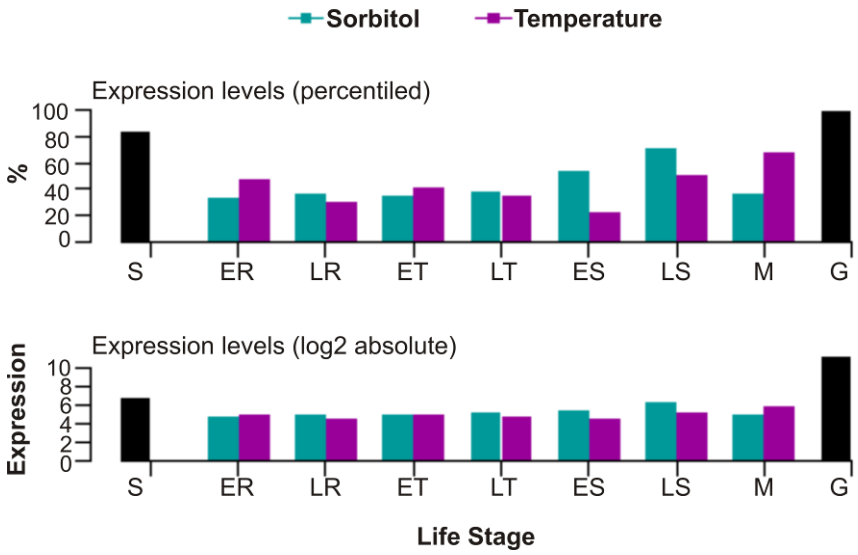



**Figure 52: *Pf*GK mRNA expression**
*Pf*GK transcripts are present throughout asexual development. Expression is significantly elevated in gametocyte (G) and sporozoite (S) stages. Cultures were synchronized using sorbitol- (blue) and temperature- (purple) based methods. ER = early rings, LR = late rings, ET = early trophs, LT = late trophs, ES = early schizonts, LS = late schizonts and M = merozoites (Le Roch *et al.*, 2003, Aurrecoechea *et al.*, 2009) (http://PlasmoDB.org).

This contradicts data from the study by Schnick and co-workers (2009) who found no expression of *Pf*GK mRNA in asexual stages of *P. falciparum*. They also showed no effect on the growth of blood stage parasites when the *Pf*GK gene was disrupted, which is contrary to the findings from the current study. The authors did not indicate how parasitaemia had been measured and did not show the data, but presumably, they performed manual quantification of Giemsa-stained smears.

In our study we utilized a highly sensitive and objective flow cytometry method to monitor parasitaemia and this could explain the discrepancy. We agree with Schnick *et al* (2009) that *Pf*GK is non-essential since disruption of the gene is not lethal, however, parasite growth is compromised. The reduced growth of 3D7Δ*Pf*GK parasites implies that *P. falciparum* utilizes glycerol as an alternative carbon source during asexual development and that it is essential for optimal proliferation. The energy requirements of the parasite are at a peak during blood stage development and these are met through glucose metabolism by glycolysis to produce ATP, whereas it is postulated that glycerol is used for PL synthesis and membrane biogenesis. Further studies such as measuring the incorporation of

glycerol into phospholipid species will provide insights into the significance of glycerol metabolism during *P. falciparum* parasite growth and development.

*Pf*GK mRNA and protein are expressed in gametocytes (Le Roch *et al.*, 2003, Young *et al.*, 2005) but in the 3D7Δ*Pf*GK parasites, the lack of the *Pf*GK mRNA and enzyme had no apparent effect on the production of gametocytes when cultures were stressed to undergo gametocytogenesis, although an in depth analysis such as the viability of macrogametocytes or gametogenesis was not performed. Schnick *et al* (2009) made the same observation. Expression of *Pf*GK mRNA is also elevated in sporozoites (Le Roch *et al.*, 2003, Young *et al.*, 2005) and mass spectrometry analysis has identified *Pf*GK peptides in the mosquito salivary gland sporozoites (Aurrecoechea *et al.*, 2009). This implies that *Pf*GK is required during the exo-erythrocytic hepatic stage of development in the human host. However, to fully understand the role of *Pf*GK during the parasite life cycle, future studies on the mosquito and liver stages, such as monitoring the effect on sporozoite formation and sexual reproduction and development in mosquitoes infected with 3D7Δ*Pf*GK parasites, will be explored.

## 4.5 *Pf*GK as a model to identify potential DTS

The complex life cycle of *P. falciparum* offers several points of intervention when targeting the malaria parasite. Classical curative drugs, such as chloroquine or the antifolates, target the asexual blood-stage parasites. Prophylactic drugs that target the asymptomatic liver stages prevent blood-stage infection. 8-Aminoquinoline primaquine for example, kills liver-stage parasites, including dormant hypnozoites that reside in *P. vivax*- and *P. ovale*-infected hepatocytes (Hill *et al.*, 2006). Compounds that interfere with parasite sexual development, either through the inhibition of gametocytogenesis in the human host or by blocking sexual development and sporogony in the mosquito vector, can reduce malaria transmission. Artemisinin for example, kills young gametocytes in addition to its schizonticidal properties (Kumar and Zheng, 1990). Strategies aimed at eradicating malaria will have to make use of drugs, and vaccines, that target all

the different stages of parasite development (Greenwood *et al.*, 2008). However, the spread of parasite resistance to almost all antimalarials, including the highly effective artemisinins (Dondorp *et al.*, 2009, Phyo *et al.*, 2012), is worrying and threatens current and future malaria control measures. Therefore, there is an urgent need not only to identify new drug targets and produce new drugs, but also to develop novel strategies to limit the onset of drug resistance.

Kinases are generally regarded as attractive drug target candidates since substrate phosphorylation regulates most aspects of a cell's life. The choline analogue, bis-thiazolium for example, inhibits both *P. falciparum* choline- and ethanolamine kinases via a competitive inhibition mechanism and impairs the essential *de novo* phosphatidylcholine synthesis pathway. Bis-thiazolium is currently being explored as a potential antimalarial for treating severe malaria (Alberge *et al.*, 2010). The success of inhibitors against protein kinases (PKs) in human cancers has made PKs attractive drug targets. In chronic myelogenous leukaemia (CML), the *bcr-abl* translocation results in the constitutive activation of the chimeric Abelson (ABL) tyrosine kinase in white blood cells. This leads to excessive proliferation and a high white blood cell count, characteristic of CML. Gleevec was the first important drug designed to specifically target a PK. It is an ATP competitive inhibitor that binds to the ABL tyrosine kinase, inhibiting its activity (Cohen, 2002).

PKs are essential for growth and development throughout the *P. falciparum* life cycle. A collection of structural data, inhibitory profiles and target validation has laid the foundation and support for targeting the *P. falciparum* PKs, also referred to as the malarial kinome (Doerig *et al.*, 2008, Doerig *et al.*, 2010). *Pf*PKG, for example, is a cGMP-dependent PK that is essential for gametogenesis (McRobert *et al.*, 2008). Its description led to a model of the signalling pathway controlling exflaggelation, which is at present, the best-characterized signalling pathway in the life cycle of the malaria parasite (Doerig *et al.*, 2010). The *P. falciparum* kinome also includes PKs that do not cluster with any of the established eukaryotic PK groups or families. These are referred to as "orphan" kinases.

*Pf*PK7, for example, is involved in cell proliferation and development pathways and its disruption not only decreases the number of schizonts formed per asexual developmental cycle, similar to the effect of a disrupted *Pf*GK, but also blocks oocyst development in the mosquito (Dorin-Semblat *et al.*, 2008).

As a blood stage drug target candidate, *Pf*GK does not fulfill the necessary criteria since parasites with a disrupted *Pf*GK gene, which produce no *Pf*GK mRNA and therefore no protein, can still proliferate during the blood stages, albeit at a reduced rate. *Pf*GK may, however, still prove essential for mosquito- or liver-stage development and may therefore play a role in transmission-blocking strategies.

In this study, *Pf*GK was used as a model for the identification of potential DTS that would limit the development of resistance. An evolutionary bioinformatics approach was developed to identify amino acid residues under purifying selection, where mutations would markedly decrease the fitness of the parasite and therefore unlikely to be retained. The rapid progress in whole genome sequencing and structural biology means that, for many protein families, structural and evolutionary information is readily available. During evolution, functional sites are known to retain fewer mutations than other parts of the protein. Evolutionary-based methods used to identify protein features, functional sites and potential DTS have relied on (i) amino acid sequence homology for the identification of functional or structural domains (Altschul *et al.*, 1997), (ii) phylogenetic tracing of residues implicated in structural and functional roles (Lichtarge *et al.*, 1996) and (iii) gene orthology (Brenner, 1999). However, as useful as these methods are, they have limitations. Multiple sequence alignments or homology-based methods may identify functional amino acids that are present in large contact areas that are only apparent in the protein's tertiary structure and may not be evident in the primary sequence. Functional residues may also be organism specific and therefore not detected in MSAs, particularly if the sequence homology is low like in many of the *Plasmodium* proteins. If the sequence identity is less than 40% for example, functional differences are more common and homologues may have

unrelated functions or different functional sites. Finally, homology-based methods do not provide evolutionary information from nucleic acid sequences (Lichtarge and Sowa, 2002). Phylogenetic tracing methods such as evolutionary tracing (ET) (Lichtarge *et al.*, 1996) generate a trace sequence from MSAs of various functional classes of a protein family. ET is based on the hypothesis that architecture-defining amino acids are constant and those clusters of amino acids can be identified and mapped onto known tertiary structures to identify the most suitable DTS in terms of conservation, functional and structural importance. ET predictions have been verified experimentally (Sowa *et al.*, 2001, Slep *et al.*, 2001) and have been shown to identify statistically significant and functionally important regions in more than 80% of the proteins tested (Lichtarge and Sowa, 2002). However, ET relies on the availability of functional similarities of proteins within a family and conserved amino acids do not necessarily imply residues under purifying selection.

Glycerol kinase is an ancient and highly conserved protein and therefore provides a good model for evolutionary analysis. In this study, a novel method named Evolutionary Pattering (EP) was developed by Dr Pierre M. Durand (Durand *et al.*, 2008) and applied to *Pf*GK to identify amino acids within the protein that are under evolutionary constraints and are unlikely to undergo viable mutations. EP utilizes a maximum likelihood substitution matrix to estimate the ratio of non-synonymous to synonymous substitutions at individual codons in a coding sequence. This approach allows for the identification of favourable DTS and sites that should be avoided. If there are more non-synonymous substitutions than synonymous substitutions ($\omega > 1.0$), the site is under positive selection and amino acid changes may be retained to maintain or increase the fitness of the organism. These residues are poor DTS. An example is resistance of the *P. falciparum* dihydrofolate reductase thymidylate synthase (DHFR-TS) bi-functional enzyme to pyrimethamine. The primary S108N mutation is essential for pyrimethamine resistance and EP analysis showed that the serine is under weak purifying selection (Durand *et al.*, 2008). Drug pressure induced a serine → asparagine mutation at the drug binding site, which was retained, resulting in resistance to the

drug. If $\omega < 1.0$ or $\omega < 0.1$, the site is under purifying or extreme purifying selection, respectively; mutations are retained infrequently because they decrease the fitness of the organism. These may be regarded as potential DTS.

The true *Pf*GK human orthologue (*Hs*GK, Q6ZS86) gene is located on chromosome three. The gene codes for a 59.2kDa protein (529 amino acids) and two additional splice variants that produce (i) a 33.9kDa (amino acids 1-301) and (ii) a 27.8 kDa (amino acids 1-250) protein. Transcriptional evidence showed that *Hs*GK (Q6ZS86) is not present in erythrocytes, but is expressed in 64 tissues and in nine developmental stages, with elevated expression in the bone marrow, brain, heart, kidneys, liver, lung, pancreas, prostate, spinal cord, spleen, skeletal muscle and the thymus ((UniProt, 2012) (www.uniprot.org) and GeneCards (Safran *et al.*, 2003) (www.genecards.org)). Differences between *Hs*GK and *Pf*GK sequences were favoured when selecting potential DTS to prevent targeting *Hs*GK in other tissues.

Analysis of the evolutionary patterns of *Pf*GK residues compared to the human orthologue (Q6ZS86) led to the selection of six regions that included sites that are not prone to viable mutations (Figure 53 and Table 21).

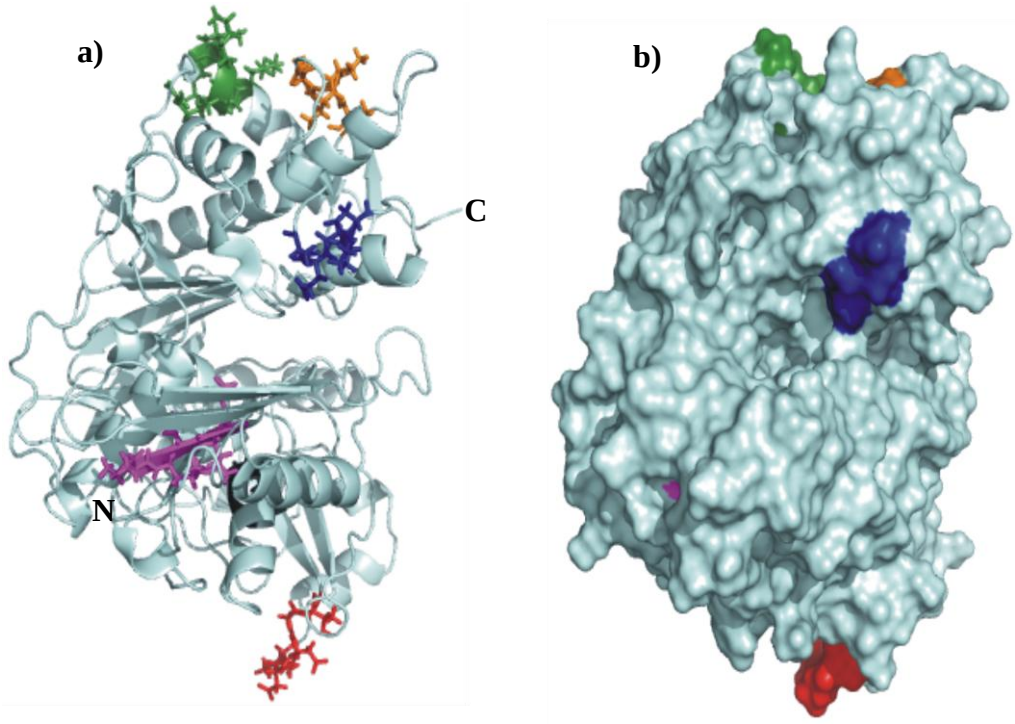



**Figure 53: Location of potential drug target sites in *Pf*GK.**
**(a)** *Pf*GK ribbon structure and **(b)** surface model showing potential drug target sites. Residues within each region are highlighted as follows: Region I (magenta), Region II (red), Region III (black), Region IV (blue), Region V (orange) and Region VI (green). Region III is situated below Region I and is difficult to visualize in the ribbon configuration since it is partly obscured by one of the helices. It is within the core of the molecule and cannot be seen on the surface *Pf*GK model. The N- and C-terminus of the protein are indicated. Images were constructed using Pymol Molecular Graphics System, version 1.3, Schrödinger, LLC.

Primary sequence analysis does not illustrate the functional and structural significance of potential DTS in a protein. Therefore, to strengthen the method, EP was combined with structural modelling. Generally, the drug target binding site should be a concave cavity at the active site of a protein. The exact length of the region would depend on whether the site is druggable, i.e. whether small molecules can bind with high specificity and affinity to a pocket on the surface of a target protein (Hajduk *et al.*, 2005a).

**Table 21: Summary of the features of the potential *Pf*GK drug target sites.**

| Region | [1]Colour | Sequence | [2]Functional domain | Location | [3]Solvent accessible surface area (Å$^2$) | Potential Drug Target |
|---|---|---|---|---|---|---|
| I | Magenta | ILSID | $Mg^{2+}$-binding | Catalytic Cleft | 1,208.9 | No |
| II | Red | IDDGT | None | Surface | 1,042.4 | No |
| III | [4]Black | GNINT | ATPase core | Core | 990.4 | No |
| IV | Blue | DDPSE | ADP-binding | Catalytic Cleft | 993 | Yes |
| V | Orange | NTTGV | Subunit interaction | Surface | 941.5 | Yes |
| VI | Green | KKRKK | Subunit interaction | Surface | 1,285 | Yes |

[1]Colours of each region correspond to those in Figure 53
[2]GK functional and regulatory domains are illustrated in Figure 48
[3]Solvent-protein contact area is defined as the surface traced out by the centre of a water sphere, having a radius of 1.4Å$^2$, rolled over the protein atoms. Parameters were estimated using Pymol Molecular Graphics System, version 1.3, Schrödinger, LLC
[4]This region is in the core of the molecule and cannot be seen on the surface model. It is part of the ATPase core domain (Figure 10)

The small molecule ligand affinity is characteristically dependent on van der Waals interactions and hydrophobic forces; exceptions are covalently bound ligands and interactions with metal ions and a few small anions (Kuntz *et al.*, 1999). Predicting the quality of the potential drug binding site is complex, although an estimate of the number of the surface accessible atoms is useful. The target region may range from a few atoms (Kuntz *et al.*, 1999) to larger protein-protein interaction sites (Cheng *et al.*, 2007) such as disruption of the p53/MDM2 interaction via cis-imidazoline analogs (Vassilev *et al.*, 2004).

The accessible molecular surface area of the six potential DTS in *Pf*GK ranged between 941-1,285Å$^2$ (Table 21). These were greater than that of most passively absorbed oral drugs, which fall within a range of ≤500Da and a polar surface area of ≤140Å$^2$ (Cheng *et al.*, 2007) and therefore, provides a large drug target area. Of the six regions identified, region IV (blue in Figure 53) appears the most favourable. It has a number of features that make it an attractive drug target: (i) one residue (S332) has a predicted function (ADP-binding) as determined from MSAs and the crystal structure of *Ec*GK (Hurley *et al.*, 1993), (ii) the target site is located at the entrance to the catalytic pocket, (iii) D327, P328 and E330 are adjacent to the S332 ADP-binding residue and are under extreme purifying selection in the parasite, indicating that resistance mutations in this region are

unlikely to develop, and (iv) three out of the five residues in *Pf*GK are different to the human enzyme, which favours selective targeting of the parasite.

Analysis of the other potential DTS revealed that regions I, II and III may not be good candidates. Region I (magenta in Figure 53) binds $Mg^{2+}$ within the catalytic pocket. Its function and location make it an attractive target site; however, small $Mg^{2+}$ ions may enter the catalytic pocket whereas a larger potential lead compound may not. Region II (red in Figure 53) is on the surface of the molecule making it accessible to a drug; however, it is relatively isolated from the catalytic domain and is not involved in any functional interactions. Targeting this site may therefore not inhibit the enzyme, although the possibility of long range conformational changes induced by the binding of a drug cannot be ruled out. Region III (black in Figure 53) is within the hydrophobic core of the molecule and may be not be accessible to drugs. Region V (orange in Figure 53) and VI (green in Figure 53) are exposed on the surface of the molecule and in *Ec*GK they participate in inter-subunit and allosteric interactions, respectively. *Pf*GK is not regulated by FBP and $IIA^{GLC}$ and its potential regulation by allosteric effectors, is not known. However, regions V and VI are in close proximity to each other and could collectively form one DTS.

The structural modelling of *Pf*GK DTS illustrates how it complements and refines the EP approach by eliminating potential sites that are not optimally located and hence not accessible to compounds that inhibit the function of the enzyme. This type of combined analysis represents an initial and iterative step in the drug discovery process and may be used as a useful tool by medicinal chemists in evaluating the druggability of sites and in designing lead compounds.

# Chapter 5: Conclusion

One of the fundamental goals in the study of *Plasmodium* is to discover metabolic pathways and key steps that are absent in its human host or different enough such that they may be targeted for drug development. The parasite's requirement for membrane biogenesis and PLs during growth and development implies that the enzymes, proteins and transporters involved in lipid metabolism may be attractive drug target candidates.

Glycerol kinase is an evolutionary well conserved enzyme that functions at the interface of lipid and carbohydrate metabolism. In an effort to understand the role of the annotated *P. falciparum* glycerol kinase orthologue (*Pf*GK), the enzyme was expressed as a GST-fusion protein in *E. coli* and its activity verified. A parasite line containing a disrupted *Pf*GK gene was generated and knockout parasites were able to proliferate and undergo characteristic asexual intra-erythrocytic development *in vitro*. *Pf*GK is therefore non-essential in blood-stage parasite development and is not a suitable drug target candidate. The growth rate of *Pf*GK knockout parasites, however, was significantly reduced, which suggests a role for the enzyme in optimal parasite proliferation. The proposed role of *Pf*GK during asexual development is summarized in Figure 54.

*Pf*GK mRNA is markedly elevated in gametocytes and sporozoites. Parasites with a disrupted *Pf*GK gene were able to differentiate into gametocytes but the enzyme may play a vital role during the mosquito stages of development, where the intracellular environment is different and the parasite may have to rely on the acquisition of alternative carbon sources for growth and survival. The same may be true for the parasite during liver-stage infection. Future studies such as monitoring the development of glycerol kinase knockout parasites in the *Anopheles* mosquito will provide further insight into the role of *Pf*GK in *P. falciparum* growth and development.

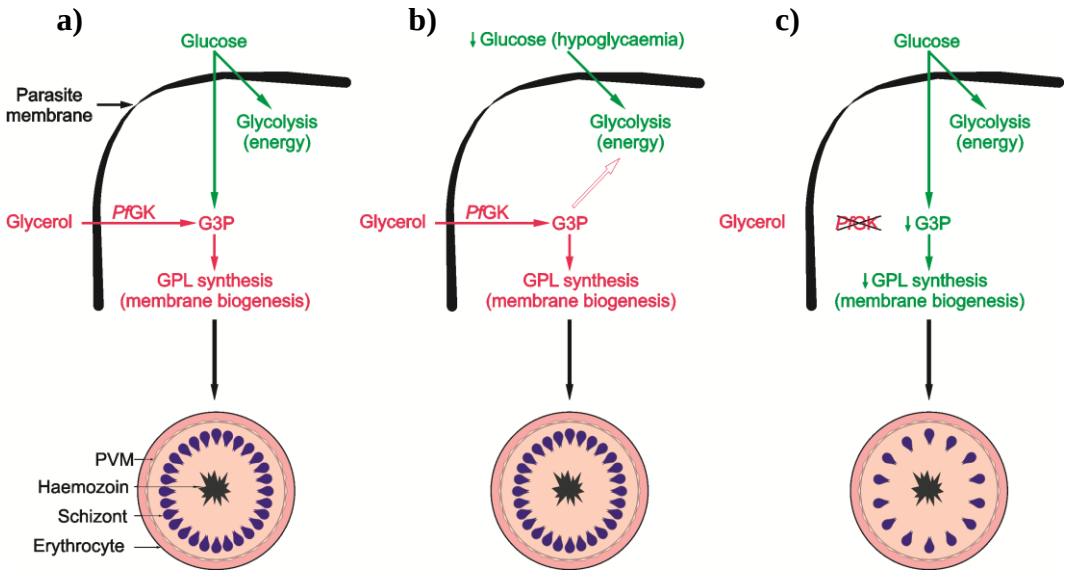



**Figure 54: Schematic of the proposed role of *Pf*GK during blood-stage development.** Glucose and glycerol are the major carbon sources for malaria parasites (Gardner *et al.*, 2002) and their use during intra-erythrocytic asexual parasite development is indicated in green and red, respectively. (**a**) G3P is produced from both glucose and glycerol. Glucose is sufficient for the parasite's energy and glycerophospholipid requirements, however, glycerol is the pathway of choice for membrane biogenesis, with glucose committed to ATP production; (**b**) under hypoglycaemic conditions, glucose is limiting. G3P is therefore mainly produced from glycerol, which may enter the glycolytic pathway. Under these conditions (**a,b**), the number of schizonts formed per 48-hour developmental cycle would be unaffected. (**c**) In parasites without a functional glycerol kinase, the G3P pool is limiting and glucose is utilized for both energy and lipid metabolism, resulting in fewer schizonts formed after each cycle and therefore, reduced growth. **Abbreviations:** G3P, glycerol-3-phosphate; GPL, glycerophospholipid; *Pf*GK, *P. falciparum* glycerol kinase.

# APPENDICES

## A1: *Pf*GK NUCLEOTIDE AND AMINO ACID SEQUENCES

PF13_0269, sequence length: 1,506bp

```
ATGAATGTCATATTAAGTATAGACCAAAGTACACAATCAACCAAAGTGTTTTTCTATGAT
GAAGAATTAAATATTGTTCATTCGAATAATTTAAATCATGAACAAAAATGTTTGAAACCT
GGTTGGTATGAACATGATCCAATAGAGATTATGACCAACTTATATAATTTAATGAATGAA
GGAATAAAAGTGTTAAAAGATAAATATACATCTGTTATAATAAAATGTATAGGTATTACT
AACCAAAGAGAAACTGTAATTATATGGGATAGAATAACAGGAAAACCTTTATACAATGCT
ATAGTTTGGCTAGATACACGTGTAGAAGAACTTGTTACGGAATTTTCTGCGAAATATAAT
AATAATGATATTCAGAAAAAAACAGGAACTTACTTTAATACATATTTTAGTGCTTTTAAA
ATATTGTGGTTAATTCAAAACAATCCAGAAATCAAACAAAAAATTGATGATGGTACTGCA
GTTATAGGAAATATTAATACATGGCTTATTTTTAATTTAACTAAAGGGAATTGTTATACA
GATGTTACCAATGCTTCAAGAACTTTATTAATGGATATTAATACATTACAATGGGATGAA
AAGATGTGTAAAATATTTAATATTACTAATATGTCTGTTTTACCTGAAATAAAAAGTAAT
TGTTCTAATTTTGGTTTAGTTAAGTCAGAACATGTTCCGGACTATTTAAATATTCCTATT
ACTGGATGTATTGGAGATCAACAAAGTGCATGCATTGGTCAAGCTATCTTCGACGAAGGT
GAAGCGAAATGTACATATGGTACAGGTGTATTCCTTCTAATTAATACAGGAGAAAAAGTT
GTATATTCCACATGCGGTTTAATTACTACTATATGTTATAAATTTAATGATAATGATAAA
CCTAAATATGCTCTTGAAGGTTCGATAGGTACTGCCGGATCTGGAGTTTCATGGCTTCTT
AAAAACAAACTAATTGATGATCCAAGTGAAGCTAGCGATATTATGGAAAAATGTGAAAAT
ACAACTGGAGTCATTTTTGTACCAGCTTTCAGTGGATTATATGCACCCAGATGGAGATCT
GATGCACGTGCATCCATATATGGAATGACTTTTAATACGGAAAGAAGCCATATTGTAAGA
GCACTATTAGAAGGTATAGCTTTTCAGTTAAATGAAATTGTGGACTCCTTAACATCGGAT
ATGGGTATTGAAATGTTACATGTCTTACGATGTGATGGAGGCATGACAAAAAATAAACCT
TTTATGCAATTTAACTCAGATATTATTAATACAAAAATTGAAGTTTCAAAATATAAAGAA
GTAACTTCTCTTGGTGCTGCAGTTCTAGCTGGATTAGAAGTTAAAATATGGGACAGTCTA
GATTCTGTTAAAAGTTTATTAAGGAGAAGTGACGCTGTTTTTCATTCTAAAATGGATGAT
AAAAAAAGAAAAAAAAAAACTTCCGAATGGAATAAAGCTGTCGAAAGGACATTAATACAG
TTATAA
```

PF13_0269, sequence length: 501 amino acids

```
MNVILSIDQSTQSTKVFFYDEELNIVHSNNLNHEQKCLKPGWYEHDPIEIMTNLYNLMNE
GIKVLKDKYTSVIIKCIGITNQRETVIIWDRITGKPLYNAIVWLDTRVEELVTEFSAKYN
NNDIQKKTGTYFNTYFSAFKILWLIQNNPEIKQKIDDGTAVIGNINTWLIFNLTKGNCYT
DVTNASRTLLMDINTLQWDEKMCKIFNITNMSVLPEIKSNCSNFGLVKSEHVPDYLNIPI
TGCIGDQQSACIGQAIFDEGEAKCTYGTGVFLLINTGEKVVYSTCGLITTICYKFNDNDK
PKYALEGSIGTAGSGVSWLLKNKLIDDPSEASDIMEKCENTTGVIFVPAFSGLYAPRWRS
DARASIYGMTFNTERSHIVRALLEGIAFQLNEIVDSLTSDMGIEMLHVLRCDGGMTKNKP
FMQFNSDIINTKIEVSKYKEVTSLGAAVLAGLEVKIWDSLDSVKSLLRRSDAVFHSKMDD
KKRKKKTSEWNKAVERTLIQL
```

## A2: PHYLOGENETIC ANALYSIS

### Table 22: Summary of *Pf*GK orthologues.

| Organism | UniProtKB ID | Amino acid sequence length |
|---|---|---|
| **Bacteria** | | |
| *Escherichia coli* | P0A6F3 | 502 |
| **Metazoa** | | |
| *Caenorhabditis elegans* | Q21944 | 502 |
| *Saccharomyces cerevisiae* | P32190 | 709 |
| *Homo sapiens1* | Q14410 | 553 |
| *Homo sapiens2* | P32189 | 559 |
| *Homo sapiens3* | Q6ZS86 | 529 |
| *Homo sapiens4* | Q6IQ27 | 530 |
| *Homo sapiens5* | A6NJP5 | 530 |
| *Drosophila melanogaster1* | Q9VZV9 | 556 |
| *Drosophila melanogaster2* | Q9W095 | 576 |
| *Drosophila melanogaster3* | Q9W0U0 | 538 |
| *Drosophila melanogaster4* | QOE8K5 | 439 |
| *Drosophila melanogaster5* | Q7JY99 | 596 |
| *Mus musculus1* | Q9WU65 | 554 |
| *Mus musculus2* | Q8C635 | 549 |
| *Mus musculus3* | Q64516 | 559 |
| *Rattus norvegicus* | Q63060 | 524 |
| **Protozoa** | | |
| *Theileria parva* | Q4N4B4 | 499 |
| *Theileria annulata* | Q4UF80 | 503 |
| *Plasmodium falciparum* | Q8IDI4 | 501 |
| *Plasmodium chabaudi* | Q4XXD5 | 379 |
| *Plasmodium yoelli* | Q7RR02 | 501 |
| *Plasmodium berghei* | Q4YSV1 | 501 |
| *Plasmodium vivax* | A5K8F2 | 501 |
| *Plasmodium knowlesi* | B3L878 | 501 |
| *Trypanosoma brucei* | Q38DF0 | 512 |
| *Leishmania major* | Q4FWG2 | 512 |

All sequences were retrieved from the UniprotKB database (http://www.uniprot.org/uniprot/)

## A3: PARASITE CULTURE

### Preparation of erythrocytes

Whole blood was collected in Acid Citrate Dextrose vacutainer tubes (ACD; Becton Dickenson, Plymouth, UK) from healthy volunteers and centrifuged using an Eppendorf 5702R centrifuge at 1,000g for 15 minutes at 4°C. The plasma and white cells (buffy coat) were aspirated under sterile conditions. Erythrocytes were washed three times with sterile PBS and diluted with an equal volume of incomplete medium. The 50% red blood cell suspension was stored at 4°C and used within one week.

### Incomplete culture medium

One 15.89g vial RPMI containing 25mM HEPES and 2mM L-Glutamine (Invitrogen Ltd., California, USA) was added to 50mg hypoxanthine (Sigma-Aldrich Corporation, St. Louis, USA) and 50mg gentamycin (Sigma-Aldrich Corporation, St. Louis, USA) in 900ml Milli-Q water and stirred at room temperature for one hour. The pH of the medium was adjusted to 6.72 with 1M sodium hydroxide (Saarchem (Pty) Ltd., Wadeville, South Africa). The volume of the solution was adjusted to one litre with Milli-Q water, sterilized by filtration through a VacuCap® 90PF 0.8/0.2µm Filter Unit (Pall Life Sciences, Michigan, USA) into a sterile 1L Schott bottle and stored at 4°C (Cowman et al., 2004).

### AB Plasma

A unit (225ml) of fresh frozen AB Plasma was obtained from the South African National Blood Service (SANBS; Johannesburg, South Africa). The plasma was thawed at room temperature and heat inactivated at 56°C for two hours. Plasma was centrifuged in an Eppendorf 5702R centrifuge at 1,400g for five minutes and the supernatant was aseptically collected and stored at -70°C in 10ml aliquots in sterile 15ml tubes. Plasma from donors with an AB blood group were also collected, heat inactivated and stored at -70°C.

**Albumax II**

A 5% albumax solution was prepared by dissolving 2.5g Albumax II in 50ml incomplete malaria culturing medium and stirring for 30 minutes. The solution was sterilized by filtration through a Millex®-GP 0.22µm Syringe Driven Filter Unit into a sterile Schott bottle and stored at 4°C.

**Complete RPMI culture medium**

Complete culture medium containing 10% AB plasma was prepared by adding 10ml heat inactivated AB Plasma and 4.2ml 5% sodium bicarbonate to 85.8ml incomplete culture medium. Alternatively, when using Albumax II, complete culture medium containing 0.5% Albumax II was prepared by adding 10ml 5% Albumax and 5.8ml 3.6% sodium bicarbonate to 84.2ml incomplete culture medium. Medium was stored at 4°C and used within three days.

For newly thawed cultures, complete culture medium containing 20% AB plasma or 1% Albumax II was used during the first week of continuous culture.

**Phosphate buffered saline (PBS)**

10mM $Na_2HPO_4$, 1.5mM $KH_2PO_4$, 137mM NaCl, 2.7mM KCl (Saarchem (Pty) Ltd., Wadeville, South Africa) was prepared in one litre Milli-Q water. The pH of the solution should be 7.2-7.4. The solution was sterilized via autoclaving and stored at 4°C.

**Parasite freezing solution**

0.65% NaCl, 3% D-sorbitol (Sigma-Aldrich Corporation, St. Louis, USA) and 28% glycerol (Saarchem (Pty) Ltd., Wadeville, South Africa) were prepared in 100ml Milli-Q water. The solution was sterilized by filtration through a Millex®-GP 0.22µm Syringe Driven Filter Unit into a sterile Schott bottle and stored at 4°C.

**5% D-sorbitol**

5g D-sorbitol was dissolved in 100ml Milli-Q water. The solution was sterilized by filtration through a Millex®-GP 0.22µm Syringe Driven Filter Unit into a sterile Schott bottle and stored at 4°C.

**Microscopic analysis of stained blood smears**

An aliquot (~5µl) of *P. falciparum* culture was removed from a tissue culture flask using a sterile glass Pasteur pipette and a smear prepared on a glass slide. The Rapid Haematology Stain was used to stain the blood smears. The staining kit is supplied with three solutions and each solution was placed into a 50ml tube. Once the smear had dried, the slide was dipped consecutively into the methanol fixative, the eosin stain and methylene blue stain each five times, one second per dip. Excess stain was removed with gentle washing under running water. The stained smear was air-dried and a minimum of five fields were examined with a Zeiss Axiostar (Carl Zeiss Jena GmbH, Jena, Germany) microscope under oil immersion at 1,000X magnification. The parasitaemia was calculated as an average percentage of infected erythrocytes out of the total number of erythrocytes.

**Sorensen's phosphate buffer**

Sorensen's phosphate buffer pH 7.2 was prepared by adding 29.6ml 67mM $KH_2PO_4$ (Saarchem (Pty) Ltd., Wadeville, South Africa) to 70.4ml 67mM $Na_2HPO_4$ (Saarchem (Pty) Ltd., Wadeville, South Africa) to give a 20mM $KH_2PO_4$ and 47mM $Na_2HPO_4$ working solution.

## A4: DNA PRECIPITATION

An equal volume of a 1:1 solution of Tris-saturated phenol pH 7.6 (Sigma-Aldrich Corporation, St. Louis, USA) and chloroform (Saarchem (Pty) Ltd., Wadeville, South Africa) was added to a DNA solution, mixed via inversion and centrifuged

at 16,100g for five minutes at 4°C. The aqueous phase was collected, an equal volume of chloroform was added, mixed via inversion and centrifuged at 16,100g for five minutes at 4°C. The aqueous phase was collected and 1:10 (v/v) 3M sodium acetate pH 5.2 and 2.5 volumes ice-cold 100% ethanol (Saarchem (Pty) Ltd., Wadeville, South Africa) were added. The DNA was precipitated at -70°C for 30 minutes or overnight. DNA was pelleted via centrifugation at 16,100g for 30 minutes at 4°C. One millilitre ice-cold 70% ethanol was added to the DNA pellet and centrifuged at 16,100g for five minutes at 4°C to remove residual salt and small molecules. The pellet was air-dried and dissolved in 20-50µl nuclease-free water. DNA quality and quantity were spectrophotometrically determined before being stored at -20°C (Appendix A5).

## A5: ASSESSING NUCLEIC ACID QUANTITY AND QUALITY

### Spectrophotometry analysis

A DNA solution with a concentration of 50µg/ml in water has an absorbance of 1.0 at a wavelength of 260nm. Similarly, a RNA solution with a concentration of 40µg/ml has an absorbance of 1.0 at 260nm. The absorbance of 1µl DNA or RNA was measured at 260nm ($A_{260}$), 230nm ($A_{230}$) and 280nm ($A_{280}$) using the NanoDrop® ND-1000 UV/VIS spectrophotometer and the integrated ND-1000 Version 3.5.2 software (NanoDrop Technologies, Inc., Wilmington, USA). The spectrophotometer was calibrated with water or TE buffer, depending on the DNA diluent. For accuracy, the nucleic acid concentrations were measured at an $A_{260}$ absorbance range between 0.1-1.0 absorbance units. Pure DNA and RNA have an $A_{260}$:$A_{280}$ ratio of ≥1.8 and ≥2.0 respectively. An $A_{260}$:$A_{280}$ <1.8 indicates protein contamination, the presence of aromatic compounds (such as phenol) or other contaminants that strongly absorb at 280nm. A DNA $A_{260}$:$A_{280}$ ratio >2.0 indicates RNA contamination. The $A_{260}$:$A_{230}$ ratio is a secondary measure of nucleic acid purity, with values generally greater than that of the $A_{260}$:$A_{280}$ ratio. The $A_{260}$:$A_{230}$ ratio ranges between 1.8-2.2 and significantly lower values may indicate co-purified contaminants. Where necessary, phenol/chloroform extraction

and ethanol precipitation (Appendix A4) were performed to improve the quality and/or concentration of the nucleic acid sample.

**Agarose gel electrophoresis**

A 50x Tris EDTA Acetate (TEA) buffer was prepared containing 2M Tris (Roche Diagnostics Gmbh, Mannheim, Germany), 50mM EDTA pH 8.0 (Saarchem (Pty) Ltd., Wadeville, South Africa) and 57.1ml acetic acid (Saarchem (Pty) Ltd., Wadeville, South Africa) in 1L Milli-Q water.

A 0.8-2% agarose gel was prepared by dissolving agarose (Lonza Group Ltd., Basel, Switzerland) in 1x TEA buffer (40mM Tris-acetate, 1mM EDTA, pH8.0) by boiling. The solution was cooled and ethidium bromide (10mg/ml; Sigma-Aldrich Corporation, St. Louis, USA) was added to a final concentration of 0.5µg/ml. 6x MassRuler$^{TM}$ DNA Loading dye (10mM Tris-HCl pH 7.6, 0.03% bromophenol blue, 60% glycerol, 60mM EDTA; Fermentas International Inc., Burlington, Canada) was added to 250-1,000ng of each sample and nucleic acids were resolved in 1x TEA buffer at 100V for 45-60 minutes. The MassRuler$^{TM}$ DNA Ladder Mix (2.5-5µl per lane) was used as a reference. Images were captured using the SynGene GeneGenius gel documentation system with the GeneSnap version 7.04 analysis software (Syngene, Cambridge, UK).

Genomic *P. falciparum* DNA resolves as a single band with a molecular weight >10kb. A smear implies that the DNA is degraded whereas the presence of a low molecular weight band (<200bp) indicates degraded RNA contamination.

The 28S (4,104 bases), 18S (1,384 bases) and 5S (120 bases) ribosomal RNA (rRNA) subunits are used as an indication of the integrity of the RNA (Daily *et al.*, 2004). The intensity of the 28S rRNA band should be twice that of the 18S rRNA. A smear and the absence of discrete rRNA imply that the sample is degraded. The presence of a high molecular weight fragment (>10kb) indicates DNA contamination.

**Quantitating DNA using densitometry**

The Fermentas MassRuler$^{TM}$ DNA Ladder Mix and the DNA sample were resolved on a 1% agarose gel. A non-saturated image was captured using the SynGene GeneGenius gel documentation system with the GeneSnap version 7.04 analysis software. The image was exported to the GeneTools version 4.0 analysis software. Each band in the MassRuler$^{TM}$ DNA Ladder Mix has a known quantity of DNA and the raw volume of each band was determined using the GeneTools densitometry software. The raw volume of the DNA sample was compared to that of the MassRuler$^{TM}$ DNA Ladder Mix and the relative amount and concentration of the DNA sample was calculated.

## A6: VECTOR MAPS AND CLONING CASSETTES

**pGEX-4T-2 and pGEX-4T-2-*Pf*GK**

The pGEX-4T-2 (Figure 55a and Table 23; Genebank accession numbers M21676 and M97937) vector system contains a *tac* promoter for chemically inducible, high level expression of recombinant proteins from cloned constructs. A *lac* operator is situated downstream of the promoter followed by a region encoding Glutathione S-transferase (GST). The multiple cloning site (MCS, Figure 55b, nucleotides 930-967) contains six restriction sites used for unidirectional cloning of inserts. A thrombin cleavage site located downstream of the GST region allows cleavage of GST from GST-fused proteins with thrombin. The β lactamase gene (lac), ampicillin resistance gene (AMPr) and origin of replication (ori) are indicated in Figure 55a. *Bam*HI and *Xho*I restriction sites were used to clone *Pf*GK into the vector (pGEX-4T-2-*Pf*GK; Figure 55c) and their positions in the vector maps are shown. PfK03, PfK04 and PfK05 *Pf*GK gene specific primers and SE-pGEX-F and SE-pGEX-R vector specific primers were used to sequence the cloned *Pf*GK gene.

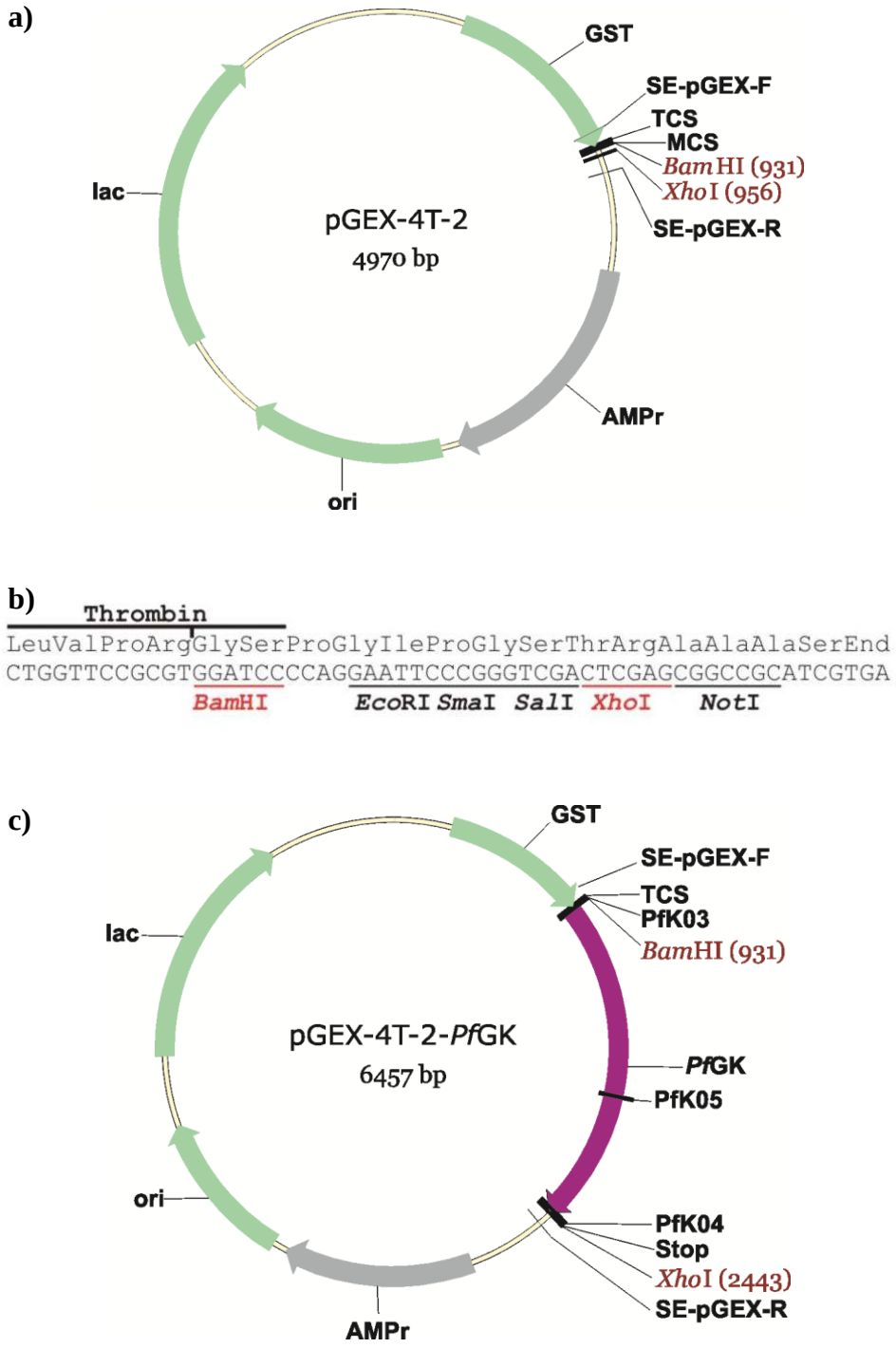



**Figure 55: pGEX-4T-2 expression vectors.**
Structure and features of the **(a)** pGEX-4T-2; **(b)** pGEX-4T-2 multiple cloning cassette and **(c)** pGEX-4T-2-*Pf*GK vectors. The full length *Pf*GK (1,506bp, purple) was cloned into the *Bam*HI/*Xho*I restriction sites (maroon), downstream of the GST cassette. The lac operator and origin of replication (ori) are indicated in green. The ampicillin resistance cassette (AMPr) is indicated in grey. The thrombin cleavage- (TCS) and multiple cloning sites (MCS) are indicated. Primers used in sequencing and PCR analysis are labelled. Plasmid names and sizes are indicated in each figure. Vector maps were adapted from GE Healthcare (2007) and figures were generated using Vector NTi advance, version 10.

**Table 23: Primers used in constructing and sequencing pGEX-4T-2-*Pf*GK.**

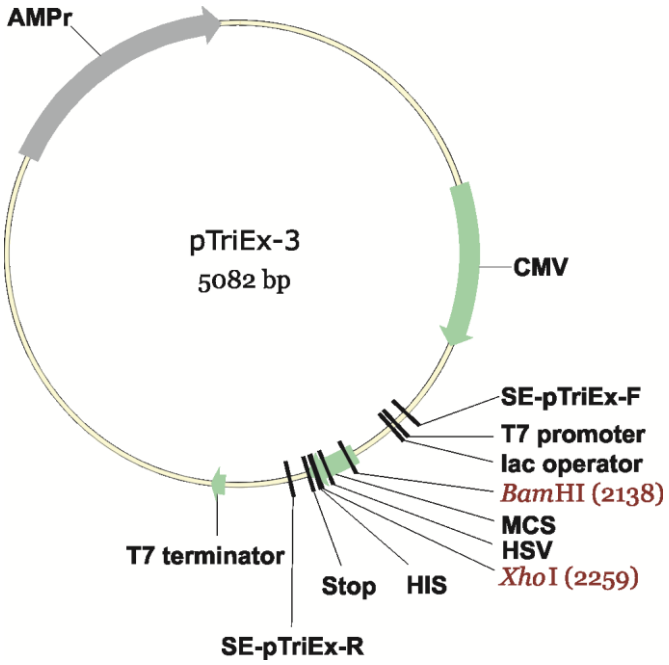
| Primer Name | Orientation | Sequence 5'-3' | Restriction sites |
|---|---|---|---|
| PfK03 | pGEX-*Pf*GK-F | gAT ggA TCC ATg AAT gTC ATA TTA AgT | *Bam*HI |
| PfK04 | pGEX-*Pf*GK-R | gAT CTC gAg TTA TAA CTg TAT TAA TgT | *Xho*I |
| PfK05 | pGEX-*Pf*GK-R-Internal | AgC ATA TTT Agg TTT ATC ATT | None |
| SE-pGEX-F | Forward | ggg CTg gCA AgC CAC gTT Tgg Tg | None |
| SE-pGEX-R | Reverse | CCg ggA gCT gCA TgT gTC AgA gg | None |

Restriction sites are underlined
Nucleotides highlighted in blue are additional nucleotides required for efficient restriction enzyme digestion
Nucleotides highlighted in red represent a stop codon

**pTriEx-3 and pTriEx-3-*Pf*GK**

pTriEx-3 (Figure 56a and Table 24) has three different promoters allowing protein expression in mammalian, insect and bacterial cells. A cytomegalovirus (CMV) immediate early enhancer and promoter allows protein expression in vertebrate cells and a T7 *lac* promoter regulates expression in *E. coli*. For expression in insect cells, pTriEx-3 contains flanking baculovirus sequences for generation of recombinant baculoviruses via homologous recombination. Expression in baculovirus-infected insect cells is regulated by the late p10 promoter. The MCS (Figure 56b, nucleotides 2,112-2,290) contains 24 restriction sites used for unidirectional cloning of inserts. The vector has two C-terminus purification tags, a Herpes Simplex Virus (HSV) and an 8 residue Histidine tag. *Pf*GK was cloned into the *Bam*HI and *Xho*I restriction sites to generate a pTriEx-3-*Pf*GK recombinant plasmid (Figure 57).

**a)**




**b)**

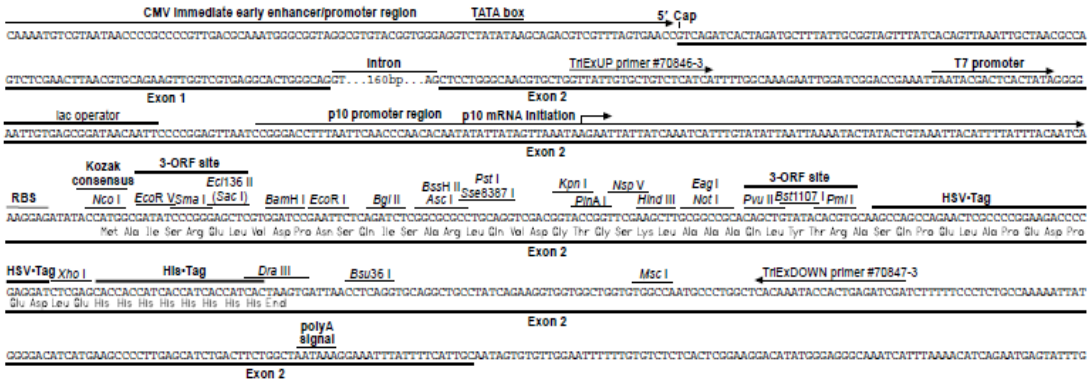



**Figure 56: pTriEx-3 multiple expression vector.**
Structure and features of the **(a)** pTriEx-3 expression vector and **(b)** pTriEx-3 multiple cloning cassette. The *Bam*HI/*Xho*I restriction sites (maroon), upstream of the 8x histidine tag, were used in this study. The lac operator and the mammalian CMV promoter are indicated in green. The ampicillin resistance cassette (AMPr) is indicated in grey. T7 promoter and terminator regions are indicated. The multiple cloning site (MCS) and primers used in sequencing are labelled. Plasmid names and sizes are indicated. Vector maps were adapted from GE Healthcare (2007) and figures were generated using Vector NTi advance, version 10.

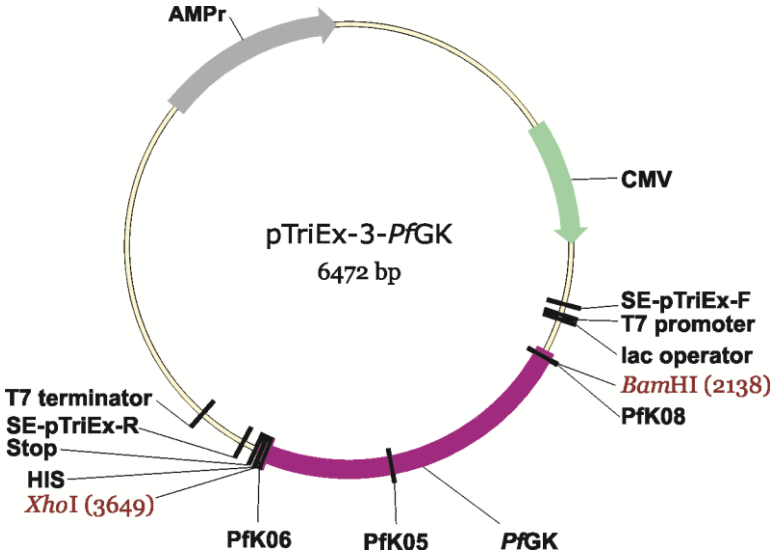



**Figure 57: pTriEx-3-*Pf*GK construct.**
The full length *Pf*GK without the stop codon (1,503bp, purple) was cloned into the *Bam*HI/*Xho*I restriction sites (maroon), upstream of the 8x histidine tag. The lac operator and the mammalian CMV promoter are indicated in green. The ampicillin resistance cassette (AMPr) is indicated in grey. T7 promoter and terminator regions are indicated. Primers used in sequencing and PCR analysis are labelled. Plasmid names and sizes are indicated. Vector maps were adapted from GE Healthcare (2007) and figures were generated using Vector NTi advance, version 10.

**Table 24: Primers used in constructing and sequencing pTriEx-3-*Pf*GK*.**

| Primer Name | Orientation | Sequence 5'-3' | Restriction sites |
|---|---|---|---|
| PfK08 | pTriEx-*Pf*GK-F | gAT ggA TCC **G** ATg AAT gTC ATA TTA AgT | *Bam*HI |
| PfK05 | pGEX-*Pf*GK-R-Internal | AgC ATA TTT Agg TTT ATC ATT | None |
| PfK06 | pTriEx-*Pf*GK-R | gAT CTC gAg TAA CTg TAT TAA TgT CCT | *Xho*I |
| SE-pTriEx3-F | Forward | gTT ATT gTg CTg TCT CAT CA | None |
| SE-pTriEx3-R | Reverse | CAC AAA TAC CAC TgA gAT CgA | None |

Restriction sites are underlined
Nucleotides highlighted in blue are additional nucleotides required for efficient restriction enzyme digestion
The added **G** in PfK08 was necessary for in-frame translation of *Pf*GK in the pTriEx-3 expression system

## A7: SUBCLONING SOLUTIONS

### Luria Bertani (LB)

One percent (10g) tryptone (Becton, Dickinson and Co., Sparks, USA), 0.5% (5g) yeast extract (Oxoid Ltd., Basingstoke, UK), 1% (10g) sodium chloride and 10ml 1M Tris-HCl, pH 7.5 were dissolved in 1L Milli-Q water. The solution was sterilized via autoclaving and stored at room temperature or 4°C.

### Antibiotic preparation

Ampicillin and kanamycin (Roche Diagnostics Gmbh, Mannheim, Germany) were each dissolved in Milli-Q water to a final concentration of 100mg/ml and 20mg/ml respectively. Each solution was sterilized by filtration through a Millex®-GP 0.22µm Syringe Driven Filter Unit into sterile 1.5ml tubes and stored in 500µl aliquots at -20°C. Chloramphenicol (Roche Diagnostics Gmbh, Mannheim, Germany) was dissolved in 100% ethanol to a final concentration of 50mg/ml and stored in 500µl aliquots at -20°C.

### LB plates

Two hundred milliliters LB were added to 1.5% (3g) agar (Oxoid Ltd., Basingstoke, UK) and autoclaved. The solution was cooled to 35-40°C before antibiotics were added. Ampicillin, chloramphenicol or kanamycin (depending on the selectable marker) were added to a final concentration of 100µg/ml, 50µg/ml and 20µg/ml respectively, mixed via swirling and poured into petri dishes. The agar was allowed to solidify by incubation at room temperature for one hour. Plates were stored at 4°C and used within two weeks.

## A8: *E.coli, P. falciparum* AND *S. fruqiperda* CODON USAGE TABLES

***Escherichia coli 536* [gbbct]: 4629 CDS's (1445921 codons)**

fields: [triplet] [frequency: **per thousand**] ([number])

| | | | |
|---|---|---|---|
| UUU 22.6(32725) | UCU 8.8(12736) | UAU 16.6(24004) | UGU 5.3( 7700) |
| UUC 16.3(23512) | UCC 8.8(12703) | UAC 12.1(17425) | UGC 6.4( 9192) |
| UUA 14.2(20514) | UCA 7.7(11067) | UAA 2.0( 2896) | UGA 1.0( 1389) |
| UUG 13.9(20074) | UCG 8.8(12676) | UAG 0.2( 344) | UGG 15.2(21918) |
| CUU 11.5(16623) | CCU 7.4(10695) | CAU 13.3(19209) | CGU 20.8(30107) |
| CUC 11.0(15952) | CCC 5.6( 8159) | CAC 9.6(13895) | CGC 21.1(30550) |
| CUA 4.0( 5840) | CCA 8.4(12206) | CAA 15.5(22441) | CGA 3.7( 5395) |
| CUG 52.1(75338) | CCG 22.5(32573) | CAG 29.0(41888) | CGG 5.6( 8146) |
| AUU 30.3(43849) | ACU 9.2(13234) | AAU 18.3(26426) | AGU 9.3(13376) |
| AUC 24.6(35528) | ACC 22.8(32999) | AAC 21.1(30545) | AGC 15.8(22784) |
| AUA 5.0( 7181) | ACA 7.8(11212) | AAA 33.4(48324) | AGA 2.4( 3417) |
| AUG 27.5(39775) | ACG 14.5(20918) | AAG 10.3(14831) | AGG 1.5( 2144) |
| GUU 18.4(26611) | GCU 15.6(22541) | GAU 32.3(46650) | GGU 24.7(35700) |
| GUC 15.1(21798) | GCC 25.2(36443) | GAC 18.8(27186) | GGC 28.6(41296) |
| GUA 11.1(16030) | GCA 20.7(29960) | GAA 39.1(56532) | GGA 8.6(12413) |
| GUG 25.7(37188) | GCG 32.7(47214) | GAG 17.8(25669) | GGG 11.2(16255) |

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

***Plasmodium falciparum 3D7* [gbinv]: 4023 CDS's (2993721 codons)**

fields: [triplet] [frequency: **per thousand**] ([number])

| | | | |
|---|---|---|---|
| UUU 36.3(108597) | UCU 14.7(44072) | UAU 50.5(151275) | UGU 15.2(45631) |
| UUC 7.1( 21216) | UCC 5.1(15184) | UAC 6.1( 18277) | UGC 2.3( 6813) |
| UUA 47.3(141467) | UCA 16.6(49689) | UAA 0.9( 2762) | UGA 0.3( 856) |
| UUG 10.4( 31055) | UCG 3.0( 8892) | UAG 0.1( 405) | UGG 4.9(14622) |
| CUU 8.7( 25901) | CCU 7.9(23687) | CAU 20.7( 62098) | CGU 3.0( 8929) |
| CUC 1.8( 5366) | CCC 2.0( 6084) | CAC 3.4( 10323) | CGC 0.4( 1260) |
| CUA 6.0( 18000) | CCA 9.1(27384) | CAA 24.2( 72417) | CGA 2.4( 7156) |
| CUG 1.4( 4324) | CCG 0.9( 2802) | CAG 3.7( 11054) | CGG 0.3( 818) |
| AUU 35.9(107358) | ACU 10.6(31753) | AAU 123.5(369612) | AGU 20.3(60737) |
| AUC 6.3( 18786) | ACC 4.8(14295) | AAC 20.0( 59914) | AGC 3.9(11811) |
| AUA 50.0(149610) | ACA 21.6(64810) | AAA 95.7(286523) | AGA 15.9(47461) |
| AUG 21.9( 65616) | ACG 3.7(11222) | AAG 21.4( 64080) | AGG 4.3(12856) |
| GUU 15.5( 46418) | GCU 8.2(24686) | GAU 55.6(166598) | GGU 11.9(35530) |
| GUC 2.4( 7210) | GCC 2.1( 6234) | GAC 8.6( 25791) | GGC 1.3( 3899) |
| GUA 15.9( 47599) | GCA 8.4(25235) | GAA 61.8(184975) | GGA 12.6(37635) |
| GUG 4.8( 14416) | GCG 1.0( 3143) | GAG 10.4( 31177) | GGG 2.8( 8315) |

Coding GC 23.80% 1st letter GC 31.95% 2nd letter GC 22.16% 3rd letter GC 17.28%

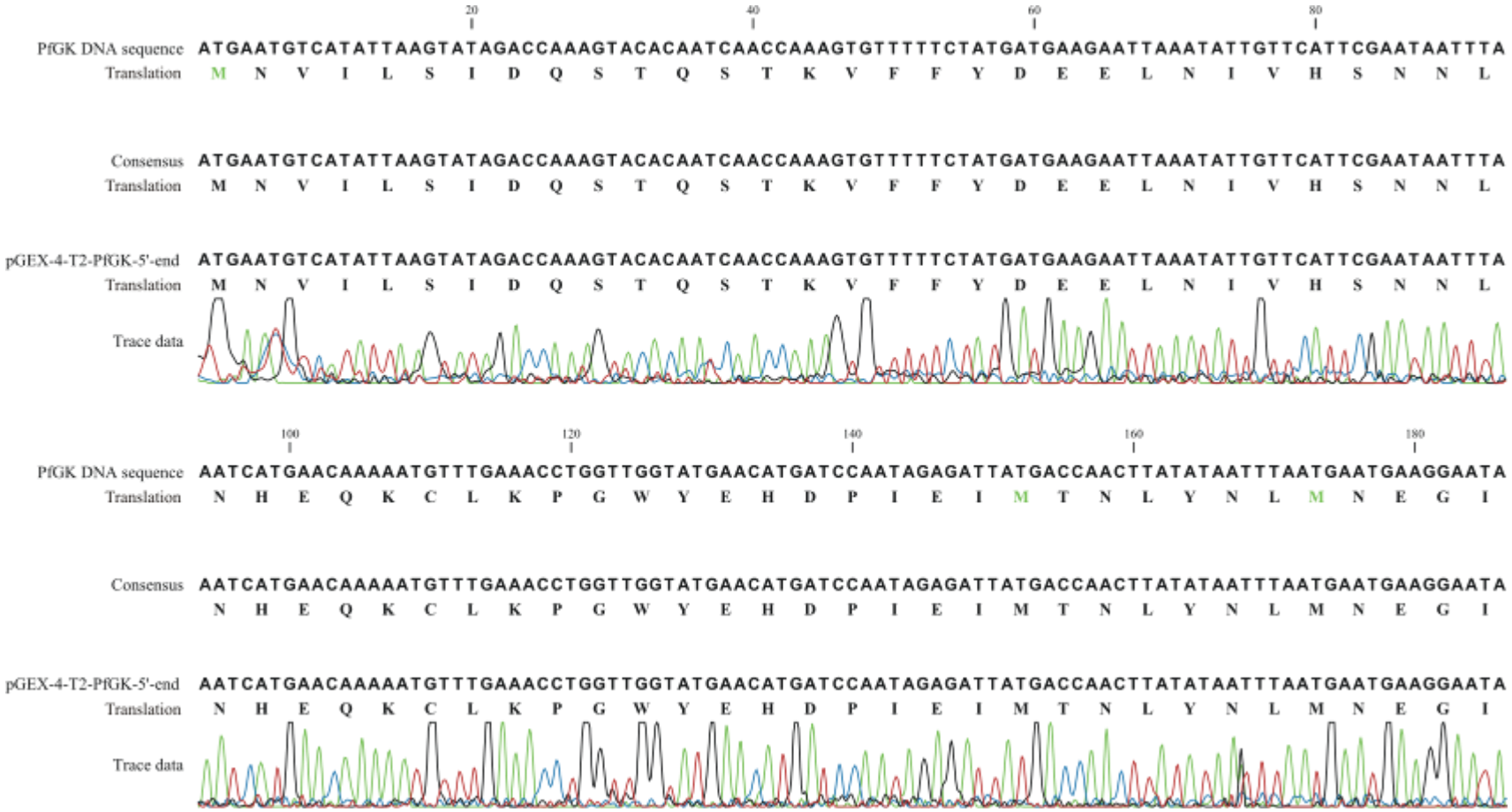
***Spodoptera frugiperda MNPV* [gbvrl]: 151 CDS's (43395 codons)**

fields: [triplet] [frequency: **per thousand**] ([number])

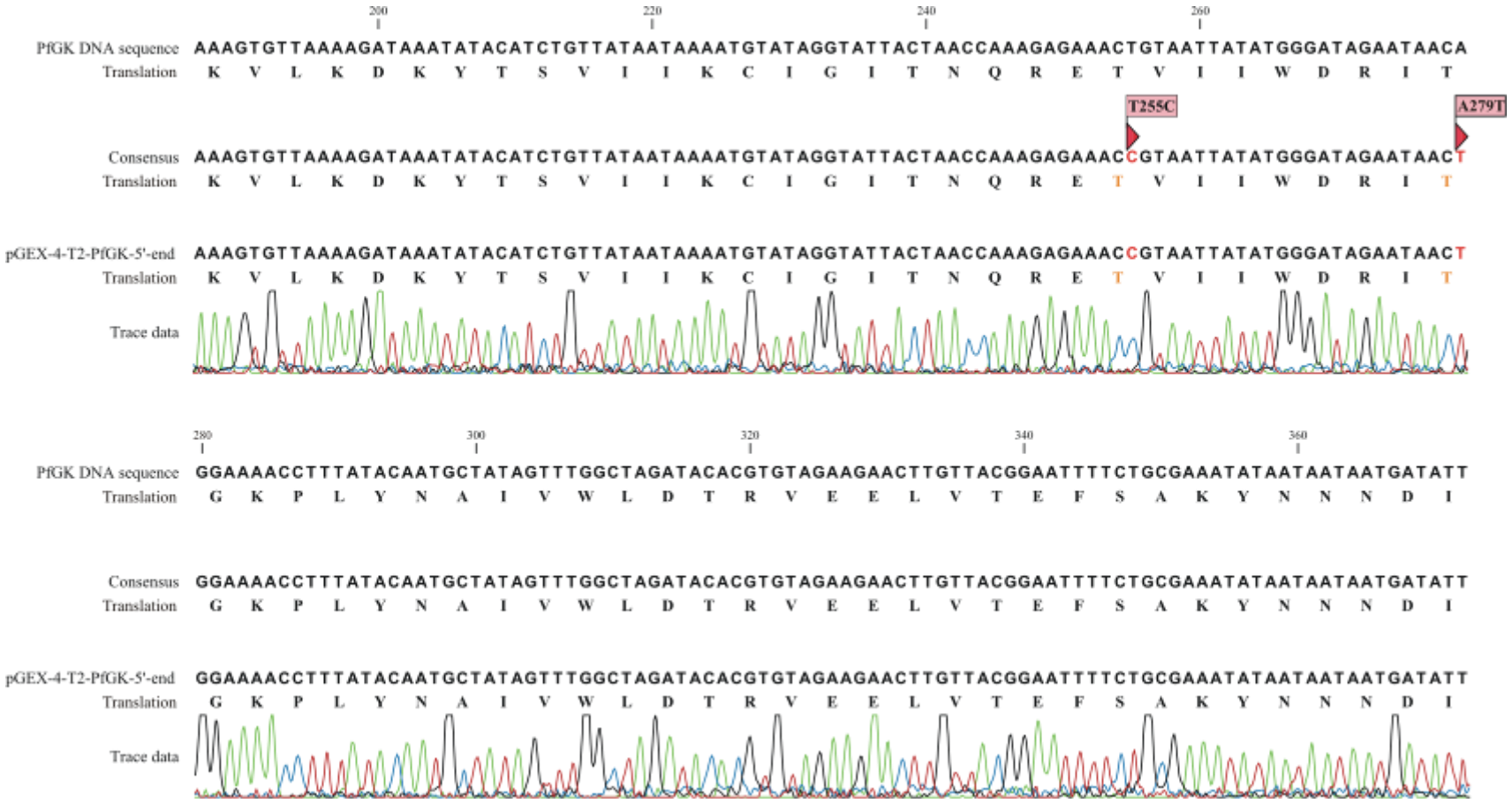
```
UUU 31.1(1351)  UCU 10.6( 461)  UAU 20.6( 896)  UGU 13.5( 585)
UUC 18.7( 813)  UCC  7.0( 303)  UAC 31.1(1348)  UGC  8.7( 377)
UUA 15.6( 678)  UCA  7.2( 311)  UAA  2.3( 101)  UGA  0.6(  28)
UUG 30.8(1337)  UCG 15.7( 682)  UAG  0.5(  22)  UGG  7.3( 316)

CUU  8.1( 352)  CCU  9.0( 389)  CAU 10.9( 474)  CGU 10.8( 468)
CUC 12.1( 524)  CCC  9.5( 412)  CAC 11.9( 515)  CGC  9.7( 420)
CUA 10.7( 463)  CCA  5.2( 227)  CAA 24.9(1079)  CGA  8.5( 369)
CUG 13.0( 564)  CCG 11.0( 476)  CAG  9.8( 425)  CGG  2.0(  88)

AUU 26.8(1161)  ACU 13.6( 591)  AAU 34.4(1494)  AGU 10.2( 444)
AUC 25.4(1103)  ACC 14.8( 642)  AAC 48.3(2096)  AGC 13.3( 575)
AUA 19.8( 859)  ACA 10.9( 472)  AAA 49.2(2135)  AGA 11.0( 477)
AUG 26.7(1157)  ACG 17.7( 768)  AAG 21.5( 934)  AGG  3.2( 137)

GUU 16.1( 699)  GCU 10.5( 456)  GAU 30.5(1322)  GGU  8.4( 366)
GUC 18.2( 790)  GCC 14.9( 645)  GAC 35.0(1518)  GGC 13.6( 591)
GUA 12.5( 544)  GCA  8.6( 373)  GAA 40.7(1766)  GGA  8.0( 347)
GUG 20.9( 907)  GCG 11.5( 498)  GAG 14.2( 618)  GGG  1.3(  56)

Coding GC 41.29% 1st letter GC 43.19% 2nd letter GC 30.76% 3rd letter GC
49.91%
```

(Nakamura *et al.*, 2000) http://www.kazusa.or.jp/codon/).

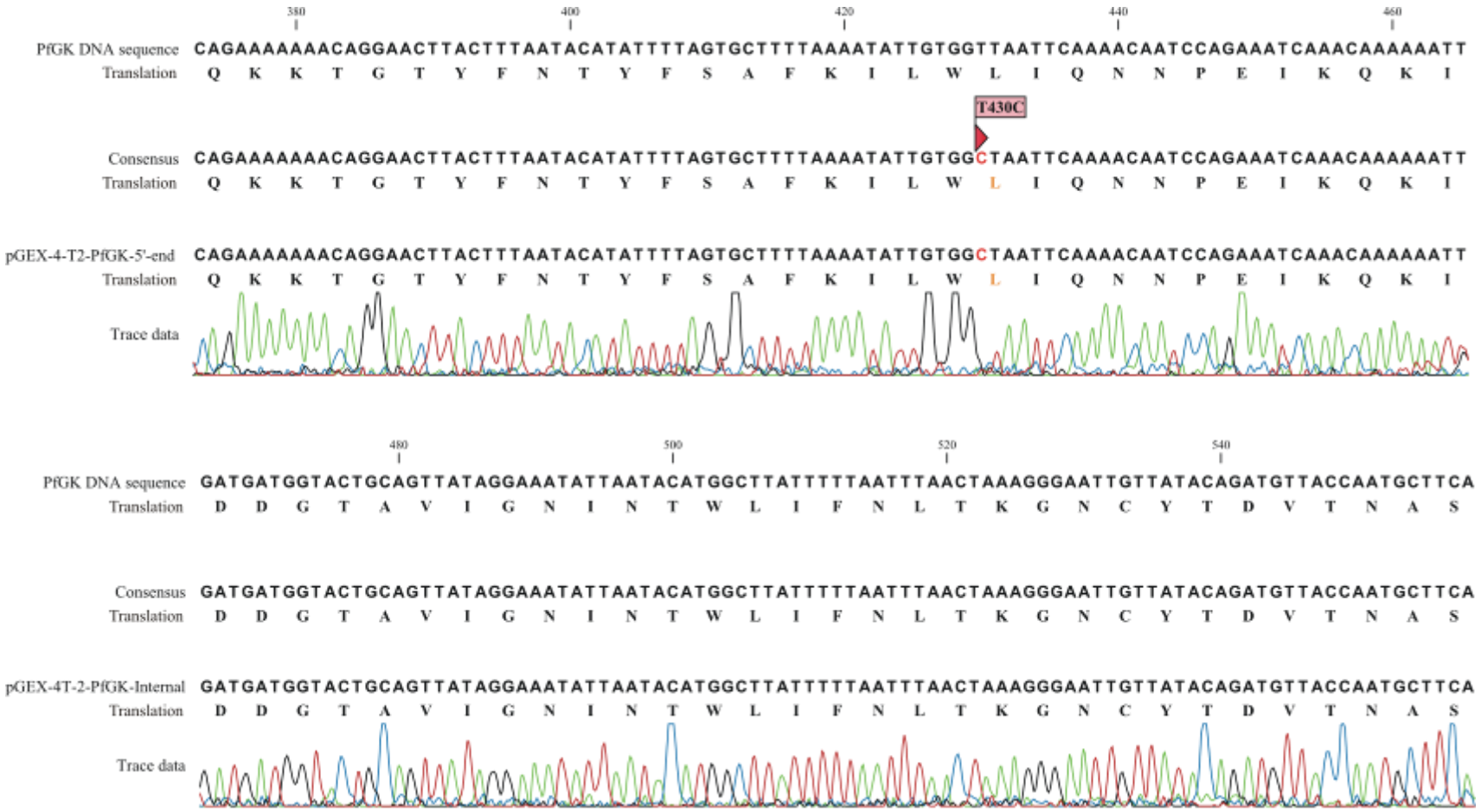
## A9: *Pf*GK DNA SEQUENCING



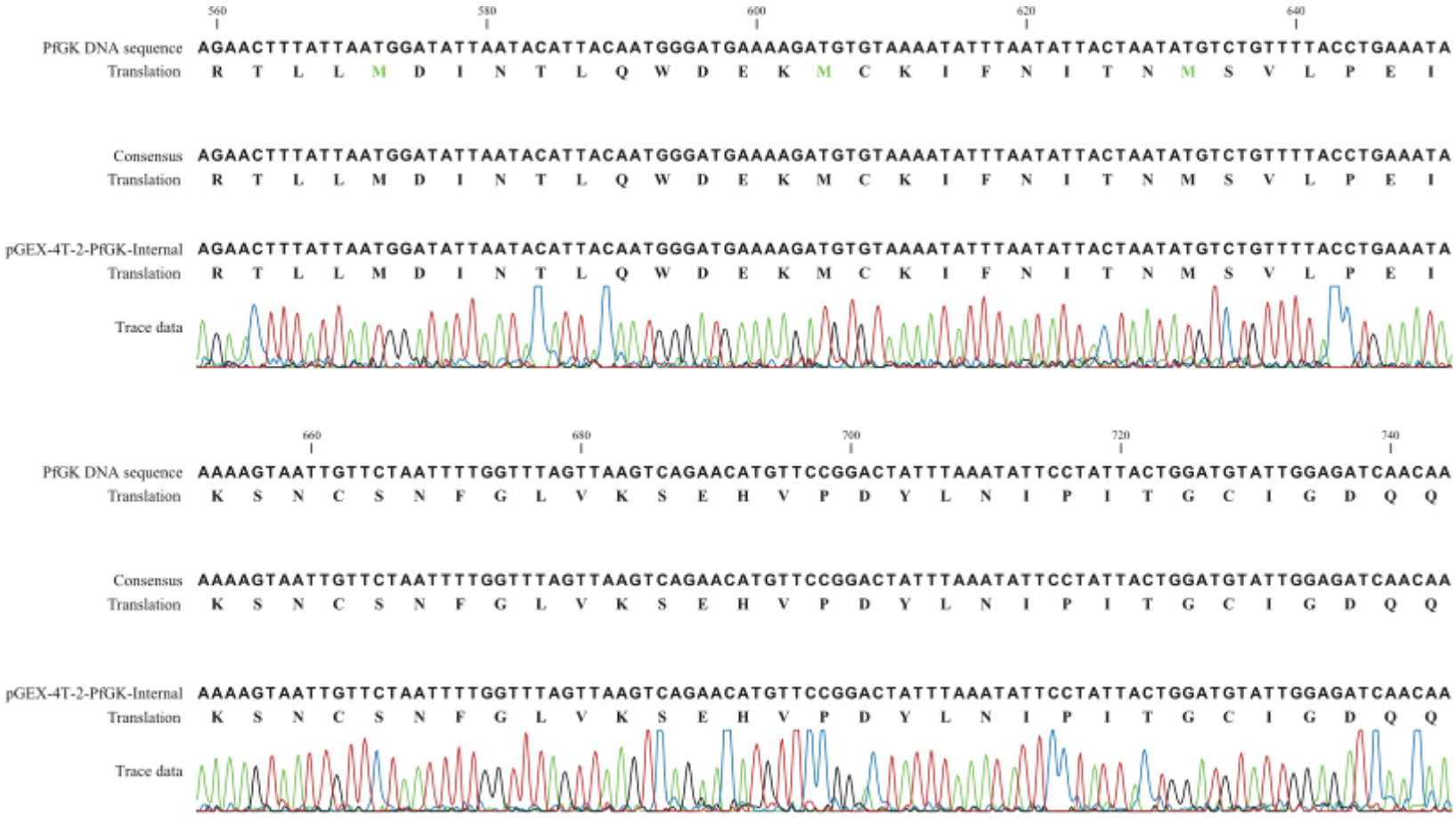
Continued from Page 169



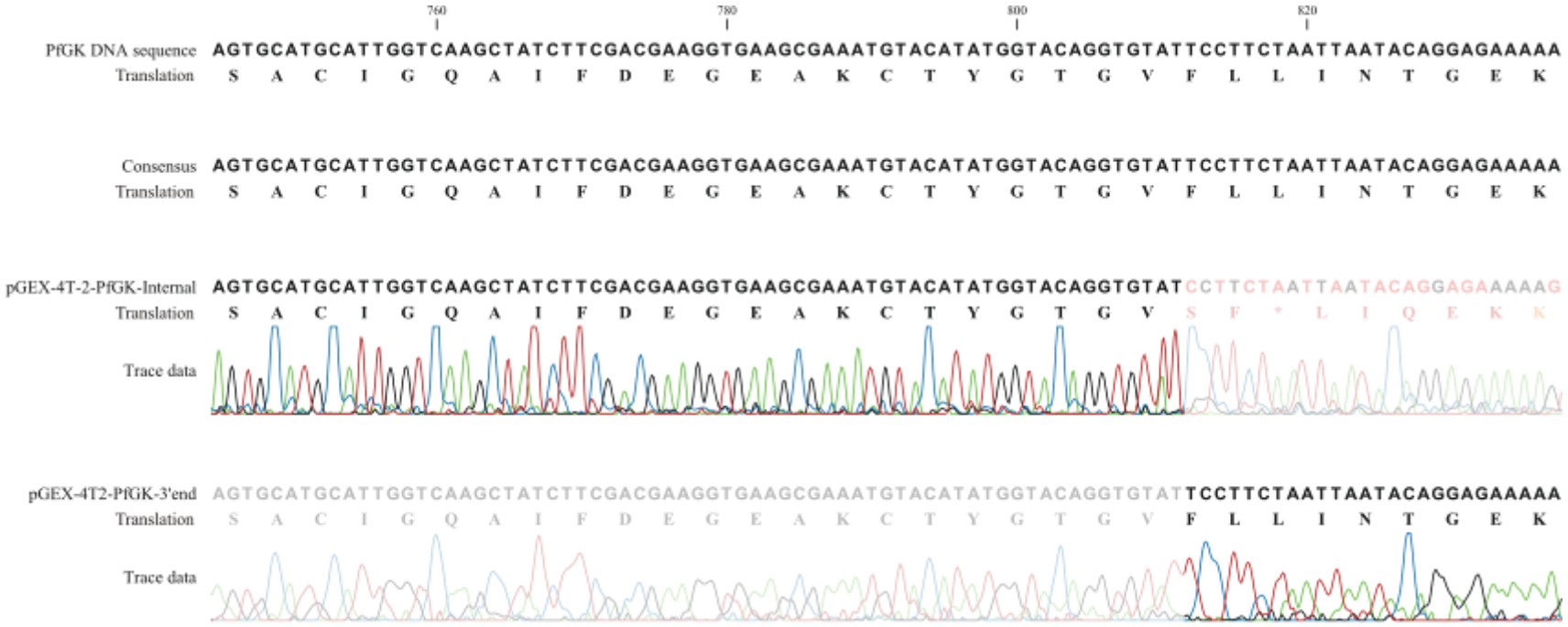
Continued from Page 170



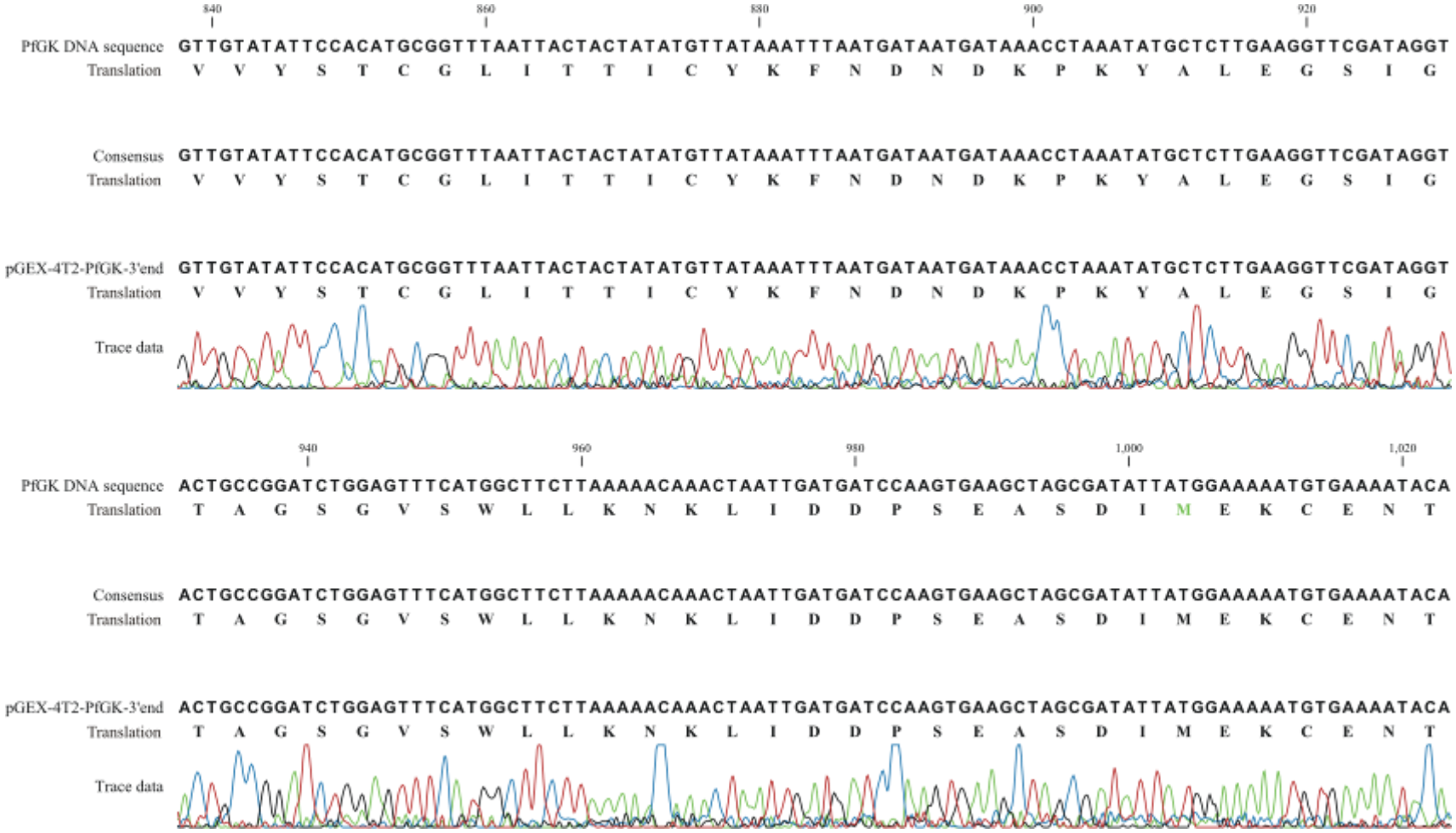
Continued from Page 171



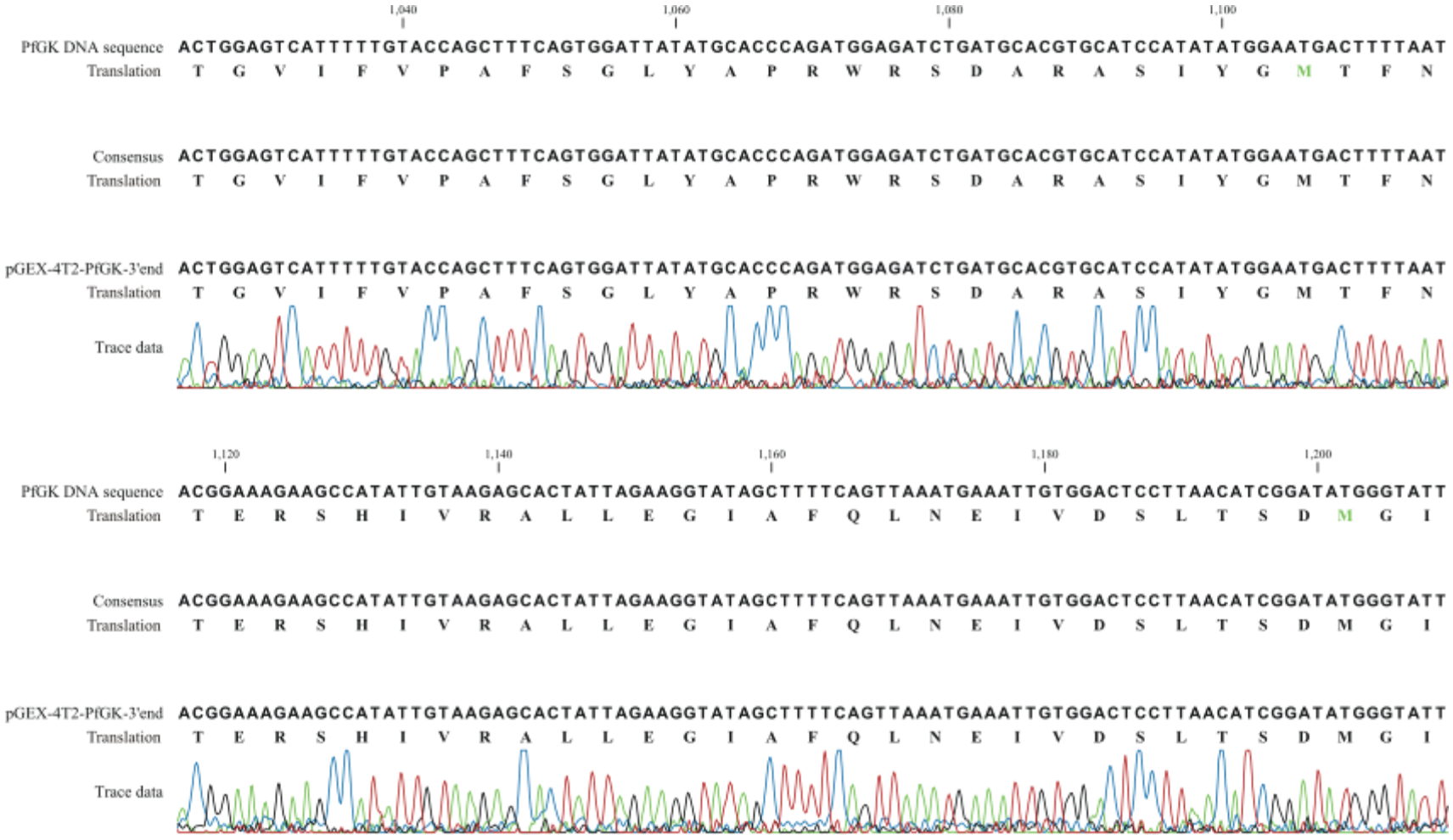
Continued from Page 172



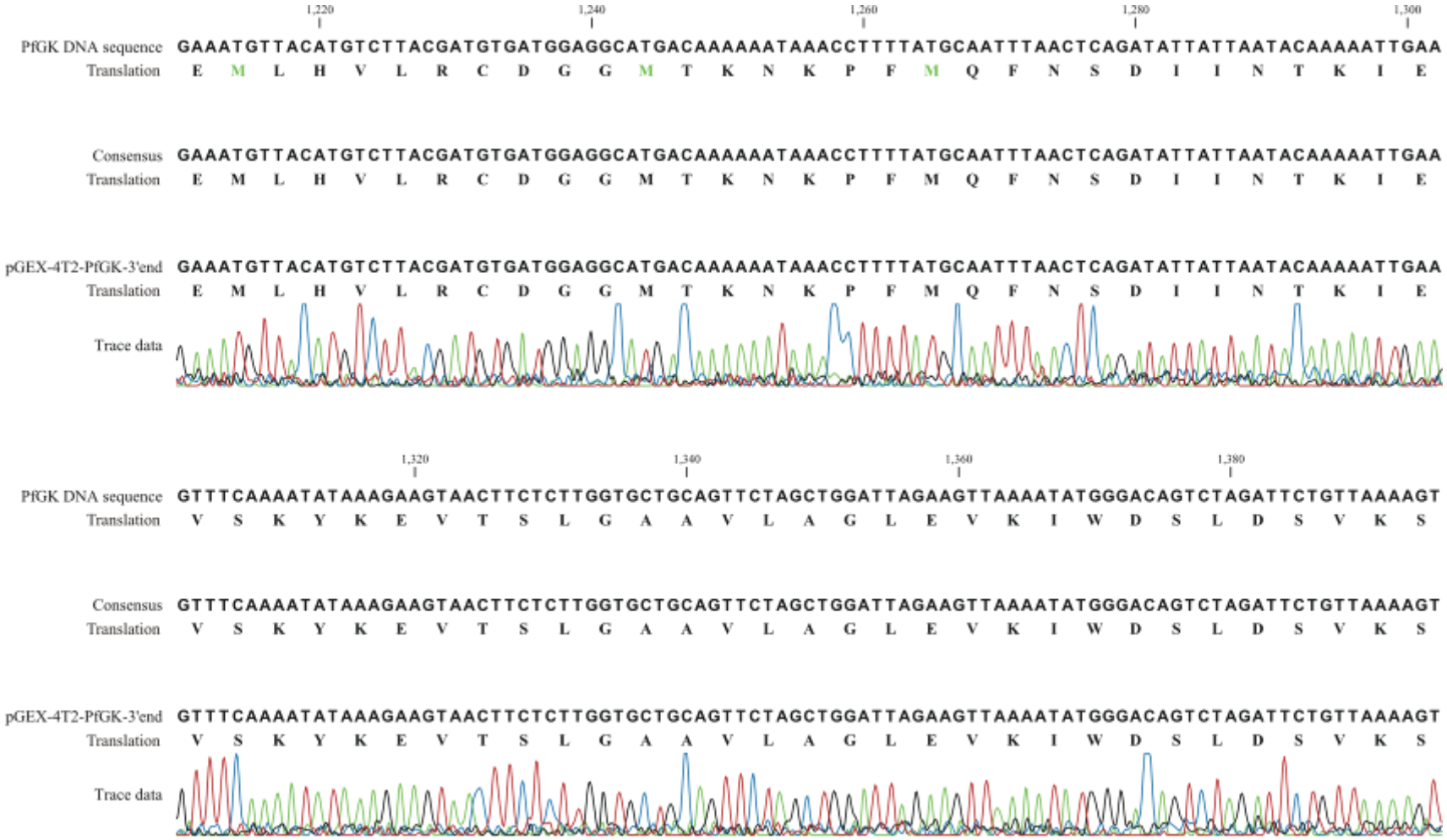
Continued from Page 173



Continued from Page 174



Continued from Page 175



Continued from Page 176

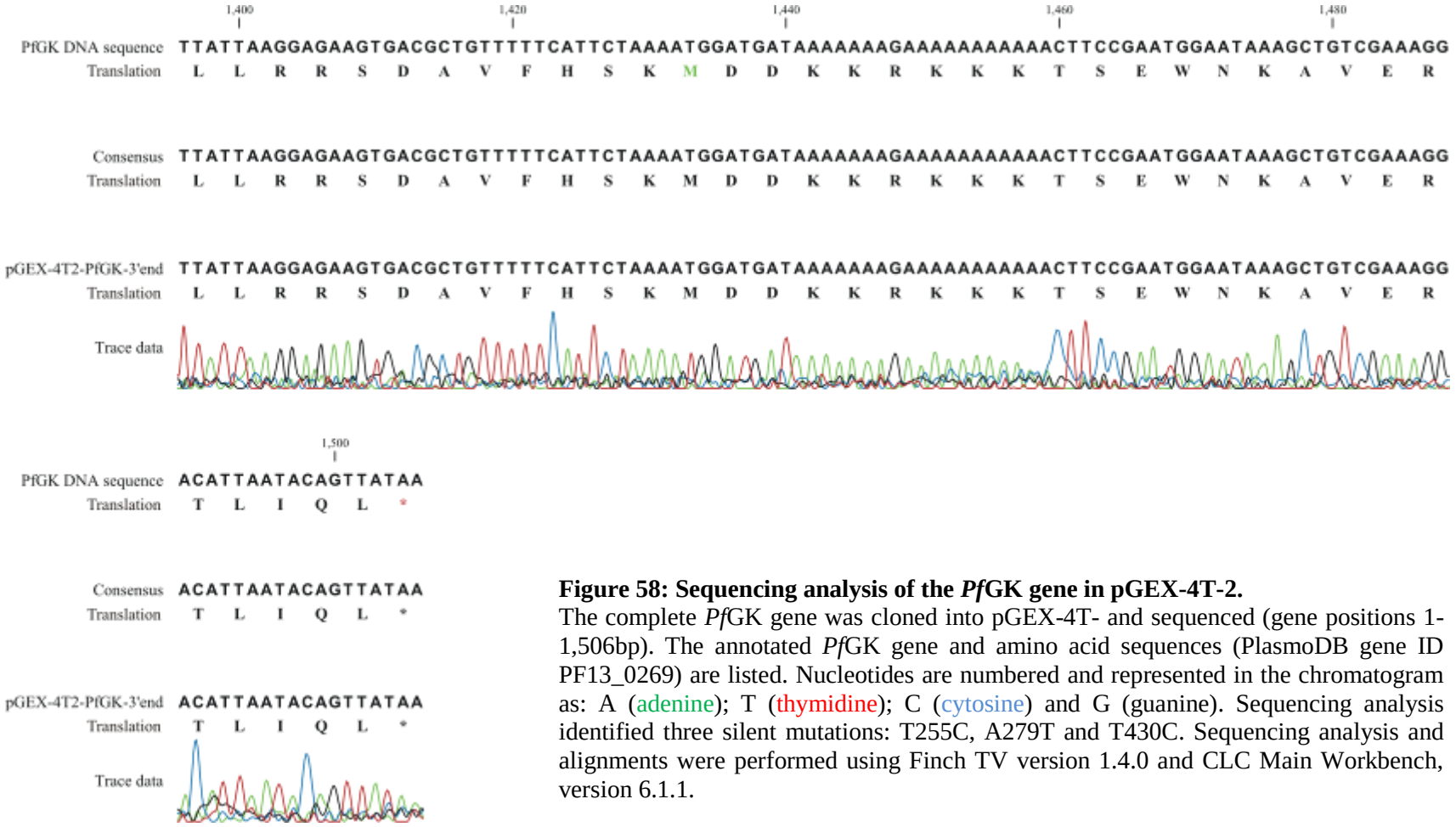



**Figure 58: Sequencing analysis of the *Pf*GK gene in pGEX-4T-2.**
The complete *Pf*GK gene was cloned into pGEX-4T- and sequenced (gene positions 1-1,506bp). The annotated *Pf*GK gene and amino acid sequences (PlasmoDB gene ID PF13_0269) are listed. Nucleotides are numbered and represented in the chromatogram as: A (adenine); T (thymidine); C (cytosine) and G (guanine). Sequencing analysis identified three silent mutations: T255C, A279T and T430C. Sequencing analysis and alignments were performed using Finch TV version 1.4.0 and CLC Main Workbench, version 6.1.1.

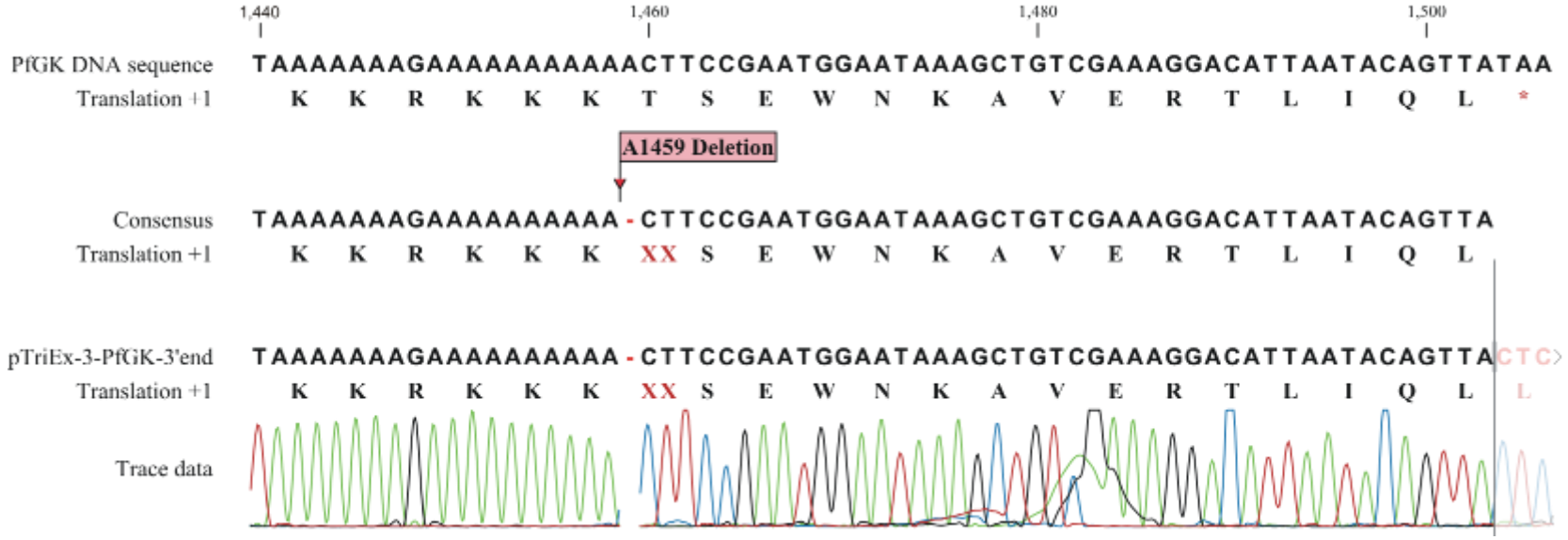



**Figure 59: Sequencing analysis of the 3'end of the *Pf*GK gene in the pTriEx-3-*Pf*GK vector.**
A representation of the sequencing results from the 3'-*Pf*GK gene region (gene positions 1,440-1,503bp). Annotated *Pf*GK gene and amino acid sequences (PlasmoDB gene ID PF13_0269) are listed. Nucleotides are numbered and represented in the chromatogram as: A (adenine); T (thymine); C (cytosine) and G (guanine). The A1459 deletion is illustrated in the trace data and consensus sequence. Following three independent cloning and sequencing attempts, deletions were found in the 19bp adenine-rich region (positions 1441-1459). Sequencing analysis and alignments were performed using Finch TV version 1.4.0 and CLC Main Workbench, version 6.1.1.

## A10: PROTEIN EXPRESSION ANALYSIS

### Overnight Express™ Instant TB Medium

Sixty grams Overnight Express™ Instant TB Medium and 10ml glycerol were dissolved in one litre Milli-Q water. The solution was sterilized via autoclaving and stored at 4°C.

### Preparation of protein samples

One hundred and fifty microlitres of protein extracts were solubilized in 1.5ml Eppendorf tubes containing 40µl 5x suspension solution (50mM Tris-HCl, pH 8.0, 5mM EDTA, 5% (w/v) SDS, 25% (w/v) sucrose), 5µl loading dye (2.5% (w/v) sucrose, 0.05% (w/v) bromophenol blue) and 2% (v/v, 4µl) β-mercaptoethanol. Solutions were mixed and boiled for one minute before being stored at -20˚C.

### Laemmli SDS-PAGE

The Laemmli gel (Laemmli, 1970) consists of a 4% stacking gel layered above a 12% resolving gel. The reagents required for an 8cm Laemmli gel cast in a Hoefer Mighty Small Dual Gel Caster (Hoefer Scientific Instruments, San Francisco, USA) tray are indicated below in Table 25. Numbers in brackets indicate final concentrations.

**Table 25: Laemmli SDS-PAGE gel reagents.**

| Reagent (Stock) | 12% Resolving gel | 4% Stacking gel |
|---|---|---|
| 30% (w/v) Acrylamide | 4ml (12%) | 433µl (4%) |
| 1% (w/v) Bis-acylamide | 1.1ml (0.11%) | 333µl (0.11%) |
| 4X resolving buffer (1.5M Tris, pH 8.8) | 2.5ml (1X) | ---- |
| 4X stacking buffer (0.5M Tris, pH 6.8) | ---- | 833µl (1X) |
| 10% (w/v) SDS | 53µl (0.05%) | 6.7µl (0.02%) |
| MilliQ water | 2.3ml | 1.63ml |
| 10% (w/v) ammonium persulfate (fresh) | 67µl (0.07%) | 67µl (0.2%) |
| TEMED | 5µl | 2.5µl |
| Final volume | 10.0ml | 3.3ml |

A 6cm resolving gel was overlaid with 1ml isopropanol and the gel was allowed to polymerize at room temperature for 30 minutes. The isopropanol was aspirated and a 2cm stacking gel was layered over the resolving gel. A 12- or 15-well comb was inserted into the stacking gel and the gel was allowed to polymerize at room temperature for 30 minutes. Gels were electrophoresed in a Hoefer Mighty Small II Mini Vertical Electrophoresis Unit (Hoefer Scientific Instruments, San Francisco, USA) connected to a cooler and circulator (Labcon (Pty) Ltd., Johannesburg, South Africa) for two hours using pre-cooled Laemmli running buffer (25mM Tris, 192mM glycine, 0.1% (w/v) SDS) for 120 minutes at 20mA at 8°C.

Polyacrylamide gels were stained with Coomassie Blue stain (25% (v/v) isopropanol, 10% (v/v) acetic acid and 0.05% (w/v) Coomassie Brilliant Blue R-250 (BDH Laboratory Supplies, Poole, UK) for 16-18 hours. Gels were destained with 10% methanol and 10% acetic acid. Images were captured using the SynGene GeneGenius gel documentation system with the GeneSnap version 7.04 analysis software.

**Protein quantitation**

Protein quantitation using the Bradford assay, with BSA as a reference or absorbance at 280nm were not applicable in this study. The glutathione in the protein purification elution buffer interferes with the Bradford reagent (Coomassie Plus™ Protein Assay Reagent; Pierce Biotechnology Incorporated, Rockford, USA) and the endogenous *E. coli* GST protein in the purified r*Pf*GK fraction (sections 2.4 and 3.1.4) leads to an overestimation in protein concentration when quantitated via absorbance. To quantitate and assess r*Pf*GK purity, densitometry was used. A 100ng/µl BSA protein stock was prepared from a 2mg/ml stock solution (Pierce Biotechnology Incorporated, Rockford, USA) and 250ng, 500ng and 1,000ng of BSA were resolved on a SDS-PAGE gel. Non-saturated gel images were captured using the SynGene GeneSnap version 7.04 software and raw volumes of the Coomassie-stained protein bands were measured using the

SynGene GeneTools version 4.0 analysis software. Volumes of the purified r*Pf*GK and rGST protein fractions were compared to and quantitated relative to those of the 250ng, 500ng and 1,000ng BSA standards and were calculated as:

$$\text{Recombinant protein (ng)} = \frac{\text{sample volume}}{\text{BSA reference volume}} \times \text{BSA reference amount (ng)} \times \text{dilution factor}$$

dilution factor = 0.75 (Appendix A10 sample preparation)

The concentration (ng/µl) was calculated by dividing the amount of protein in the sample by the volume loaded. An average of the three relative protein amounts was used to calculate protein concentration (Figure 60 and Table 26).

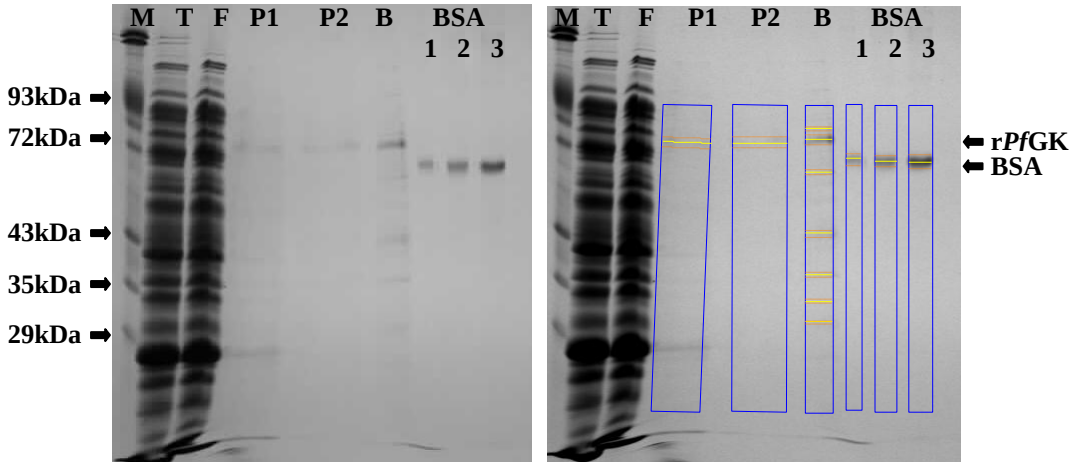



**Figure 60: r*Pf*GK quantitation using densitometry.**
r*Pf*GK was purified from an induced *E. coli* culture; **(left)** 20µl r*Pf*GK (P1; from a 400µl total eluate), 30µl r*Pf*GK (P2; from a 200µl total eluate) from the second elution, as well as 250ng (1), 500ng (2) and 1,000ng (3) BSA were resolved using SDS-PAGE. Samples from the total protein extract (T; 20µl), flow-through (F; 20µl) and beads (B, 20µl) are shown. M is 5µg red cell membrane protein marker; **(right)** a non-saturated image of Coomassie blue stained proteins was captured via SynGene GeneGenius gel documentation system. Raw volumes of each protein band was calculated using the SynGene GeneTools analysis software and recombinant proteins were quantitated relative to known amounts of a BSA reference protein.

**Table 26: r*Pf*GK protein quantitation using densitometry.**

| Protein | Reference | Raw volume | Protein amount (ng) | Average r*Pf*GK in sample (ng) |
|---|---|---|---|---|
| BSA[1] | | 5,583 | 250 | |
| | | 11,337 | 500 | |
| | | 23,013 | 1,000 | |
| r*Pf*GK (Elution 1) | 250ng BSA | 3,936 | 132 | 130 |
| | 500ng BSA | | 130 | |
| | 1,000ng BSA | | 128 | |
| r*Pf*GK (Elution 2) | 250ng BSA | 3,731 | 125 | 123.3 |
| | 500ng BSA | | 123 | |
| | 1,000ng BSA | | 122 | |

[1]BSA was used as a reference and amount was calculated from a 2mg/ml stock solution

Alternatively, gels were scanned on a Hoefer GS300 transmittance/reflectance scanning densitometer (Hoefer Scientific Instruments, San Francisco, USA) and the raw volumes of each protein band measured using the accompanying GS365W software (section 3.1.4).

## A11: WESTERN AND IMMUNOBLOTTING

Proteins from polyacrylamide gels were transferred onto Hybond™-C Extra Nitrocellulose membrane (Amersham Biosciences, Ltd., Buckinghamshire, UK) using a Hoefer TE22 Mighty Small Transfer tank (Hoefer Scientific Instruments, San Francisco, USA) at 35V (130-160mA) for 16-18 hours at 4°C in pre-cooled transblot buffer (25mM Tris, 192mM glycine, 20% (v/v) methanol, 0.1% (w/v) SDS). Nitrocellulose membrane was blocked with 3% BSA Fraction V (Roche Diagnostics Gmbh, Mannheim, Germany) in Tris buffered saline (TBS; 0.05M Tris-HCl, 0.9% NaCl, pH 7.5) and incubated at room temperature for one hour with either a 1:50,000 or a 1:100,000 anti-GST-HRP conjugated antibody for GST tagged proteins (Amersham Biosciences, Ltd., Buckinghamshire, UK) or 1:2,000 Penta•His™ HRP Conjugate antibody (Qiagen GmbH, Hilden, Germany) for histidine tagged proteins. Membranes were washed three times with TBS, 0.25% (v/v) Tween20 (TBST) and once with TBS, each for 10 minutes with shaking. Proteins were visualized with the SuperSignal® West Pico Chemiluminescent Substrate (Pierce Biotechnology Incorporated, Rockford, USA) in the dark room,

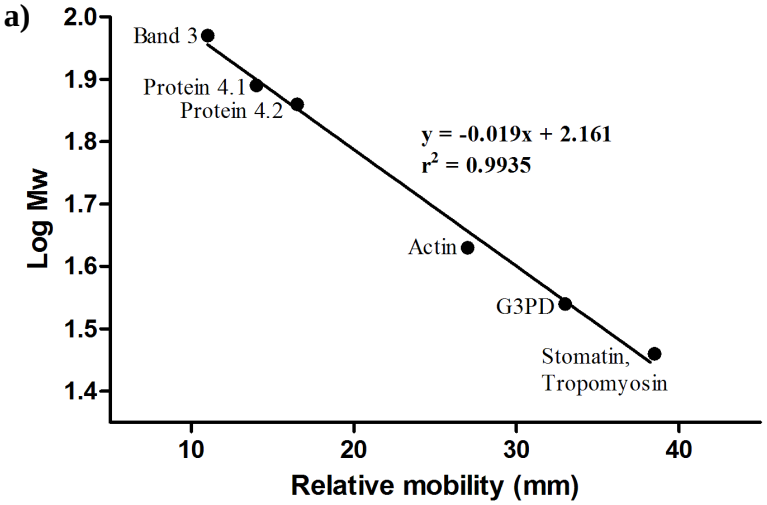
under a red safety light as per manufacturer's instruction. Briefly, the membrane was placed on a plastic sheet, overlaid with 1ml freshly prepared working solution (0.5ml stable peroxide solution, 0.5ml luminol/enhancer solution) and covered with a second plastic sheet to spread the substrate. The assembly was incubated for five minutes and the substrate was removed by gently pressing on the assembly and absorbing the liquid with paper towel. Images were either captured on AGFA CPG-Plus Medical X-ray film (Agfa, Mortsel, Belgium) and using the AXIM developer and fixer (Axim, Midrand, South Africa), following a 5-600 second exposure. Alternatively, images were captured on the SynGene GeneGenius Chemidoc system (Syngene, Cambridge, UK) over a 1-20 minute exposure.

PAGE gels after protein transfer and the nitrocellulose membrane after chemiluminescence were stained with Coomassie blue and amido black staining solution (0.1% (w/v) amido black, 10% (v/v) acetic acid, 25% (v/v) isopropanol) for five minutes, respectively, destained with 10% acetic acid and 10% methanol and used to evaluate transfer efficiency. The destained membrane and gel were captured as an image using the SynGene GeneGenius gel documentation system with the GeneSnap version 7.04 analysis software.

## A12: MOLECULAR WEIGHT DETERMINATION

The molecular weight (Mw) of denatured r*Pf*GK was calculated using the relative mobility ($R_m$) of the protein in relation to that of the red cell membrane protein marker. Protein samples and the protein marker were electrophoresed on a Laemmli SDS-PAGE gel (Appendix A10). BSA was used as an additional size control. The distance migrated by each protein was measured from the start of the 12% resolving gel. The Mw of the red cell membrane proteins were used to plot a $R_m$ against log Mw standard curve (Figure 61). The Mw of r*Pf*GK was calculated using the linear range of the standard curve.

**a)**




**b)**

| Protein | $R_m$ (mm) | LogMw | Mw (kDa) |
|---|---|---|---|
| Band 3 | 11 | 1.97 | 93 |
| Protein 4.1 | 14 | 1.89 | 78 |
| Protein 4.2 | 16.5 | 1.86 | 72 |
| Actin | 27 | 1.63 | 43 |
| G3PD | 33 | 1.54 | 35 |
| Stomatin, tropomyosin | 38.5 | 1.46 | 29 |
| BSA | 18 | 1.825 | 66.8 |
| rGST | 39.5 | 1.423 | 26.5 |
| r*Pf*GK | 16 | 1.862 | 72.8 |

**Figure 61: Determining the molecular weight of r*Pf*GK using SDS-PAGE.**
Five micrograms of the red cell membrane marker and 500ng each of purified r*Pf*GK, rGST and BSA were resolved using SDS-PAGE; **(a)** the relative mobility of red cell membrane proteins was used as references to generate a standard curve and the $R_m$ of unknown proteins was determined from the equation of the line; **(b)** a tabulated summary of the results extrapolated from the Log Mw vs $R_m$ graph.

## A13: *Pf*GK$^{sf}$ CODON OPTIMIZATION FOR EXPRESSION IN INSECT CELLS

**pMK-*Pf*GK$^{Sf}$**

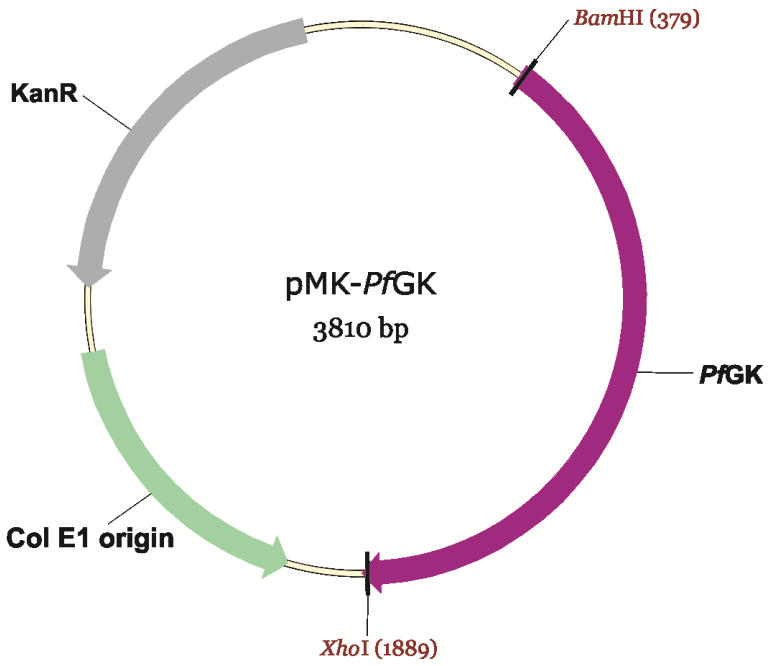



**Figure 62: pMK-*Pf*GK$^{Sf}$ plasmid vector map.**
Structure and features of the pMK-*Pf*GK$^{Sf}$ plasmid. The *Pf*GK gene (purple) was codon-optimized for expression in *Spodoptera frugiperda* insect cells, synthesized and cloned into the pMK plasmid. The *Bam*HI and *Xho*I restriction sites (maroon) necessary for cloning into the pTriEx-3 plasmid are indicated. The kanamycin (KanR) and origin of replication (Col E1 origin) sites are indicated in grey and green, respectively. Vector maps were generated using Vector NTi advance, version 10.

**Codon optimized *Pf*GK gene sequence**

GGCGCGCCGGATCCGATGAACGTGATCCTGTCCATCGACCAGTCCACCCAGTCTACCAAGGTGTTCTTCTACGACGAGGAACTGAACATCGTGCACTCCAACAACCTGAAC
CACGAGCAGAAGTGCCTGAAGCCCGGCTGGTACGAGCACGACCCCATCGAGATCATGACCAACCTGTACAACCTGATGAACGAGGGCATCAAGGTGCTGAAGGACAAGTAC
ACCTCCGTGATCATCAAGTGCATCGGTATCACCAACCAGCGCGAGACTGTGATCATCTGGGACCGTATCACCGGCAAGCCCCTGTACAACGCTATCGTGTGGCTGGACACC
CGTGTGGAGGAACTGGTCACCGAGTTCTCCGCCAAGTACAACAACAACGACATCCAGAAGAAGACCGGCACCTACTTCAACACCTACTTCTCCGCCTTCAAGATCCTGTGG
CTGATCCAGAACAACCCCGAGATCAAGCAGAAGATCGACGACGGCACCGCTGTGATCGGCAACATCAACACCTGGCTGATCTTCAACCTGACCAAGGGCAACTGCTACACC
GACGTGACCAACGCTTCCCGTACCCTGCTGATGGACATCAACACTCTGCAGTGGGACGAGAAGATGTGCAAGATCTTCAACATCACCAACATGTCCGTGCTGCCTGAGATC
AAGTCCAACTGCTCCAACTTCGGTCTGGTCAAGTCCGAGCACGTGCCCGACTACCTGAACATCCCCATCACCGGTTGCATCGGTGACCAGCAGTCCGCTTGCATCGGCCAG
GCTATCTTCGACGAGGGCGAGGCTAAGTGCACCTACGGAACCGGCGTGTTCCTGCTGATCAACACCGGCGAGAAGGTGGTGTACTCCACTTGCGGTCTGATCACCACCATC
TGCTACAAGTTCAACGACAACGACAAGCCCAAGTACGCTCTGGAAGGTTCCATCGGAACTGCTGGTTCCGGTGTCTCCTGGCTGCTGAAGAACAAGCTGATCGACGACCCC
TCCGAGGCTTCCGACATCATGGAAAAGTGCGAGAACACCACCGGCGTGATCTTCGTGCCCGCTTTCTCCGGCCTGTACGCTCCTCGTTGGCGTTCCGACGCTCGTGCTTCC
ATCTACGGCATGACCTTCAACACCGAGCGTTCCCACATCGTGCGTGCTCTGCTGGAAGGTATCGCTTTCCAGCTGAACGAGATCGTGGACTCCCTGACCTCCGACATGGGT
ATCGAGATGCTGCACGTGCTGCGTTGCGACGGTGGCATGACCAAGAACAAGCCCTTCATGCAGTTCAACAGCGACATCATCAACACCAAGATCGAGGTGTCCAAGTACAAG
GAAGTGACCTCCCTGGGCGCTGCTGTGCTGGCTGGCCTGGAAGTGAAGATCTGGGACTCCCTGGACTCCGTGAAGTCCCTGCTGCGTCGCTCTGACGCTGTGTTCCACTCC
AAGATGGACGACAAGAAGCGCAAGAAGAAGACTTCCGAGTGGAACAAGGCTGTGGAGCGTACCCTGATCCAGCTGCTCGAGTTAATTAA

**Figure 63: Codon optimized *Pf*GK$^{Sf}$ gene for expression in *Sf*9 insect cells.**
*Pf*GK was codon optimized and synthesized for expression in *Sf*9 insect cells (GeneArt AG, Regensburg, Germany). Nucleotide substitutions compared to the wild type *Pf*GK gene are underlined. *Bam*HI (blue) and *Xho*I (pink) restriction sites are highlighted. The gene stop codon, upstream of the *Xho*I restriction site, was removed. An additional guanine nucleotide (red) was added upstream of the methionine start codon (green) necessary for in-frame protein synthesis when using the pTriEx-3 expression vector. PfK16 (yellow) and PfK17 (dark yellow) oligonucleotides were used for PCR-based verification of recombinant Bac-*Pf*GK$^{Sf}$ baculoviruses.

## A14: BACULOVIRUS-BASED INSECT CELL PROTEIN EXPRESSION

### Trypan blue exclusion method

Trypan blue stains dead cells only. Dead cells appear blue under the microscope while live cells are clear. For an accurate assessment, it is important to count the cells as soon as the dye is added as live cells lose their capacity to exclude the dye with time. One hundred microlitres of cells was added to 100µl trypan blue solution (0.4% trypan blue in 0.85% saline; Sigma-Aldrich Corporation, St. Louis, USA) and mixed via repeat pipetting (3-5 times). The solution was transferred to both ends of the haemocytometer, with the coverslip on. All cells, both stained (dead) and unstained (living), were counted in the centre grid (25 squares in a volume of $0.1mm^3$) and the cell concentration and viability were determined as:

$$\text{Cell concentration (cells/ml)} = \frac{\text{number of cells x dilution factor}}{10^{-4}}$$

Dilution factor = 2 (sample diluted 1:1 with trypan blue)

$10^{-4}$ = conversion of $0.1mm^3$ volume counted to $cm^3$

$$\text{Cell viability} = \frac{\text{number of viable (stained) cells}}{\text{total number of cells}}$$

### BacPlaque™ agarose

A 10ml 3% BacPlaque™ Agarose (Novagen, Inc., Madison, USA) stock was prepared by dissolving 0.3g in 10ml Milli-Q water via autoclaving. The solution was cooled to 50°C and 20ml 37°C pre-warmed BacVector® insect cell medium supplemented with 5% FBS was added to the BacPlaque™ Agarose solution . The Agarose-medium-serum solution was mixed and incubated at 37°C until use.

### Neutral Red

A 0.33% (w/v) Neutral Red stock solution was prepared by dissolving 0.033g in 10ml Milli-Q water. The solution was sterilized by filtration through a Millex®-

GP 0.22µm Syringe Driven Filter Unit into a sterile 15ml tube and stored at 4°C, away from light. The stock solution was diluted 1:13 with sterile PBS just before plaque staining and used within eight hours.

## A15: GENERATION OF *Pf*GK KNOCKOUT PARASITES

### pCC-1 and pCC-1-*Pf*GK and pCC-1-EBA

The pCC-1 (Figure 64a) vector system is used in targeted gene disruption studies. The vector consists of an ampicillin resistance cassette (AmpR) and a positive selectable marker, human dihydrofolate reductase (hDHFR, resistance to WR99210), under the control of a calmodulin promoter (Cam) and histidine rich protein II terminator (HrpII). The vector contains two MCS: (i) *Spe*I/*Afl*II restriction sites upstream of the Cam promoter and (ii) *Eco*RI/*Nco*I/*Avr*II restriction sites downstream of the HrpII terminator regions. 5'- and 3'-*Pf*GK regions were cloned into the *Spe*I/*Afl*II and *Eco*RI/*Nco*I restriction sites respectively to generate a pCC-1-*Pf*GK knockout plasmid (Figure 64b). Homologous double crossover recombination of the 5'- and 3'-*Pf*GK fragments with the native *Pf*GK gene results in integration of the hDHFR cassette, disrupting the gene. pCC-1 contains a negative selection cassette, *S. cerevisiae* cytosine deaminase/uracil phosphoribosyl transferase (CD/UPRT), under the control of a heat-shock protein86 promoter (hsp86) and the *P. berghei* dihydrofolate reductase-thymidylate synthase (PbDT) terminator regions. Cytosine deaminase (CD) converts an innocuous 5-fluorocytosine metabolite into 5-fluoro-uracil, which inhibits RNA synthesis and the enzyme thymidylate synthase (TS). Uracil phosphoribosyl transferase (UPRT) efficiently converts 5-fluoro-uracil into 5-fluoro-uridine monophosphate (5-fluoro-UMP). Host cellular enzymes convert 5-fluoro-UMP into 5-fluoro-UTP and 5-fluoro-dUMP which inhibit RNA synthesis and TS respectively. The CD/UPRT suicide gene codes for a bifunctional chimeric protein that is 1,000-fold more potent than CD itself (Maier *et al.*, 2006).

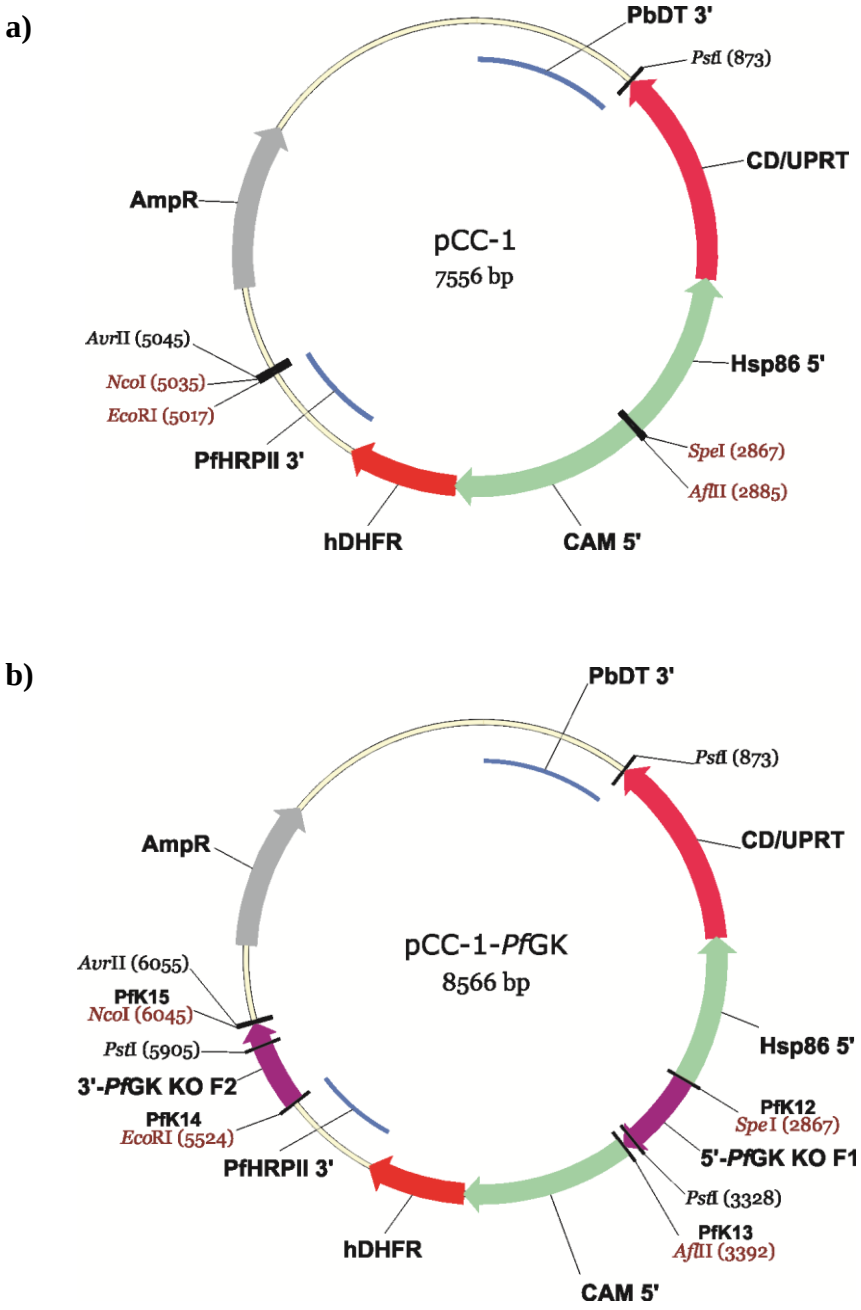



**Figure 64: pCC-1 and pCC-1-*Pf*GK plasmid vector maps.**
Structure and features of the **(a)** pCC-1 and **(b)** pCC-1-*Pf*GK plasmid DNA vectors. The positive selection cassette (hDHFR, orange), with its 5' calmodulin promoter (Cam, green) and histidine rich protein II terminator (HrpII, blue) regions, are indicted. The negative cytosine deaminase/uracil phosphoribosyl transferase (CD/UPRT, red) selection cassette, with its 5' heat-shock protein86 promoter (hsp86, green) and *P. berghei* DHFR-TS terminator (PbDT, blue) regions, are indicted. A 519bp 5'-*Pf*GK region (purple) was cloned into the *Spe*I/*Afl*II restriction sites (maroon), upstream of the hDHFR cassette. A 515bp 3'-*Pf*GK fragment (purple) was cloned into the *Eco*RI/*Nco*I restriction sites (maroon), downstream of the hDHFR cassette. An ampicillin resistance cassette (AmpR, grey) and *Pst*I restriction sites used in Southern hybridization analysis are indicated. Plasmid names and sizes are indicated in each figure. Vector maps were generated using Vector NTi advance, version 10.

**Table 27: Primers used to generate and screen for 3D7Δ*Pf*GK parasites.**

| Primer Name | Primer description | Primer sequence | Restriction Site |
|---|---|---|---|
| PfK03 | pGEX-*Pf*GK-F | gAT ggA TCC ATg AAT gTC ATA TTA AgT | *Bam*HI |
| PfK04 | pGEX-*Pf*GK-R | gAT CTC gAg TTA TAA CTg TAT TAA TgT | *Xho*I |
| PfK12 | pCC1-*Pf*GK_KO1_F | gATC ACT AgT CAA AgT ACA CAA TCA ACC | *Spe*I |
| PfK13 | pCC1-*Pf*GK_KO1_R | gATC CTT AAg ATC TgT ATA ACA ATT CCC | *Afl*II |
| PfK14* | pCC1-*Pf*GK_KO2_F | gATC gAA TTC CAA ACT AAT TgA TgA TCC AA | *Eco*RI |
| PfK15* | pCC1-*Pf*GK_KO2_R | gATC CCA Tgg CAg CTT TAT TCC ATT Cg | *Nco*I |
| Cam-R | pCC-1-Cam-R | CCA ATA gAT AAA ATT TgT AgA | None |
| HrpII | pCC-1-HrpII-F | CTC ATT ATA TAT AAg AAC AT | None |

Restriction sites are underlined
Nucleotides highlighted in blue are additional nucleotides required for efficient restriction enzyme digestion
Nucleotides highlighted in red represent a stop codon
*Primers also used to generate the 3'-DIG-labelled probe for the detection of 3D7Δ*Pf*GK parasites

**pCC-1-EBA**

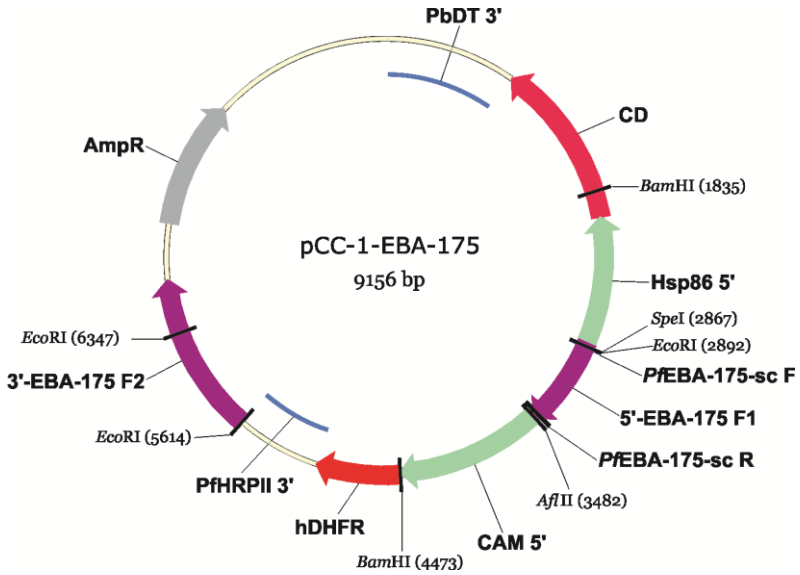



**Figure 65: pCC-1-EBA plasmid vector map.**
Putative structure and features of the EBA-175 knockout plasmid DNA. The vector was donated by Professor A.F Cowman (Walter and Eliza Hall Institute of Medical Research, Victoria, Australia) and Dr S.B Lauterbach performed the partial sequencing analysis of the cloned 5'- and 3'-EBA-175 homologous recombination fragments. The sequenced 579bp 5'-EBA-175 and the partially sequenced 1,091bp 3'EBA-175 regions are illustrated in purple. Primers *Pf*EBA-175-sc_F/*Pf*EBA-175-sc_R amplified an internal 564bp fragment within the 5'-EBA-175 homologous recombination region, which was used as a probe for the Southern hybridization analysis. *Eco*RI and *Bam*HI restriction sites used in Southern hybridization analysis are indicated. The positive selection cassette (hDHFR, orange), with its 5' calmodulin promoter (Cam, green) and histidine rich protein II terminator (HrpII, blue) regions, are indicated. The negative cytosine deaminase/uracil phosphoribosyl transferase (CD/UPRT, red) selection cassette, with its 5' heat-shock protein86 promoter (hsp86, green) and *P. berghei* DHFR-TS terminator (PbDT, blue) regions, are shown. An ampicillin resistance cassette (AmpR, grey) is indicated. Vector maps were generated using Vector NTi advance, version 10.

**Table 28: Primers used to generate the DIG-labelled probe for detection of 3D7ΔEBA parasites.**

| Primer Name | Primer description | Primer sequence |
|---|---|---|
| *Pf*EBA-175-sc_F | pCC-1-EBA_F | TAA CCg gAT CCT TCA ATg TgC ATA CAA TgA AAT |
| *Pf*EBA-175-sc_R | pCC-1-EBA_R | ggA CCg Cgg CCg CAT CAg gAA TAC ATA CAT AgT |

**Table 29: Primers used to evaluate 3D7Δ*Pf*GK mRNA expression.**

| Primer Name | Primer description | Primer sequence |
|---|---|---|
| PfK05 | pGEX-*Pf*GK-R-Internal | ATC AgC ATA TTT Agg TTT ATC ATT |
| PfK11 | *Pf*GK_Internal_F | gCT ATC TTC gAC gAA ggT gAA g |
| P133 | Rifin-F | CGC TAC ATA TGC TAG TTG CCA GAC |
| P134 | Rifin-R | CGT GTA CCA TAC ATC CTA CCA AC |
| 18s-F | 18s rRNA-F | TgA CTA CGT CCC TgC CCT T |
| 18s-R | 18s rRNA-R | ACA ATT CAT CAT ATC TTT CAA TCg g |

**CytoMix**

120mM KCl, 0.15mM $CaCl_2$ (Saarchem (Pty) Ltd., Wadeville, South Africa), 2mM EGTA (Saarchem (Pty) Ltd., Wadeville, South Africa), 5mM $MgCl_2$ (Saarchem (Pty) Ltd., Wadeville, South Africa), 10mM $K_2HPO_4/KH_2PO_4$ pH7.6 and 25mM HEPES were dissolved in 45ml Milli-Q water. The pH of the solution was adjusted to 7.6. The solution was adjusted to 50ml with Milli-Q, sterilized by filtration through a Millex®-GP 0.22µm Syringe Driven Filter Unit into a sterile tube and stored at 4°C (Cowman *et al.*, 2004).

**WR99210**

1.72mg WR99210 (Jacobus Pharmaceuticals, Princeton, USA) was dissolved in 200µl DMSO to a final concentration of 20mM and stored in 20µl aliquots at -70°C. A 5µM working solution was prepared by diluting WR99210 with incomplete culture medium (Appendix A3). The solution was sterilized by filtration through a Millex®-GP 0.22µm Syringe Driven Filter Unit into a sterile tube and stored at 4°C. The solution is stable for four weeks (Cowman *et al.*, 2004), however, in this study, the drug was used for not more than two weeks.

**5-Fluorocytosine**

A 10mg/ml (77mM) 5-fluorocytosine stock in PBS (Appendix A3) was prepared and stored at 4°C. A 0.77mM working stock was prepared by a 1:100 dilution with PBS and sterilized by filtration through a Millex®-GP 0.22µm Syringe Driven Filter Unit into a sterile tube. The metabolite was stored at 4°C and was used for not more than two weeks (Cowman *et al.*, 2004).

## A16: SOUTHERN BLOT REAGENTS

**20x SSC**

A 3M NaCl and 0.3M sodium citrate (Holpro Analytics (Pty), Ltd., Johannesburg, South Africa) was prepared by dissolving 175.32g NaCl and 88.23g sodium citrate in 900ml Milli-Q water. The pH of the solution was adjusted to 7.0 with 1M acetic acid. The solution was adjusted to 1L with Milli-Q water, autoclaved and stored at room temperature.

**Maleic acid buffer**

A 0.1M Maleic acid (BDH Laboratory Supplies, Poole, UK) and 0.15M NaCl solution was prepared by dissolving 23.21g Maleic acid and 17.53g NaCl in 1,900ml Milli-Q water. The pH of the solution was adjusted to 7.5 with NaOH pellets (Saarchem (Pty) Ltd., Wadeville, South Africa). The solution was adjusted to 2L with Milli-Q water and stored at room temperature.

**Wash buffer**

Three millilitres (0.3% v/v) Tween20 was dissolved in 1L Maleic acid buffer and stored at room temperature.

**Blocking solution**

A 200ml solution was prepared by adding 20ml 10% (w/v) blocking solution (Roche Diagnostics Gmbh, Mannheim, Germany) to 180ml Maleic acid buffer. The solution was autoclaved and stored at 4°C.

## A17: PRIMER BINDING SITES AND SEQUENCES

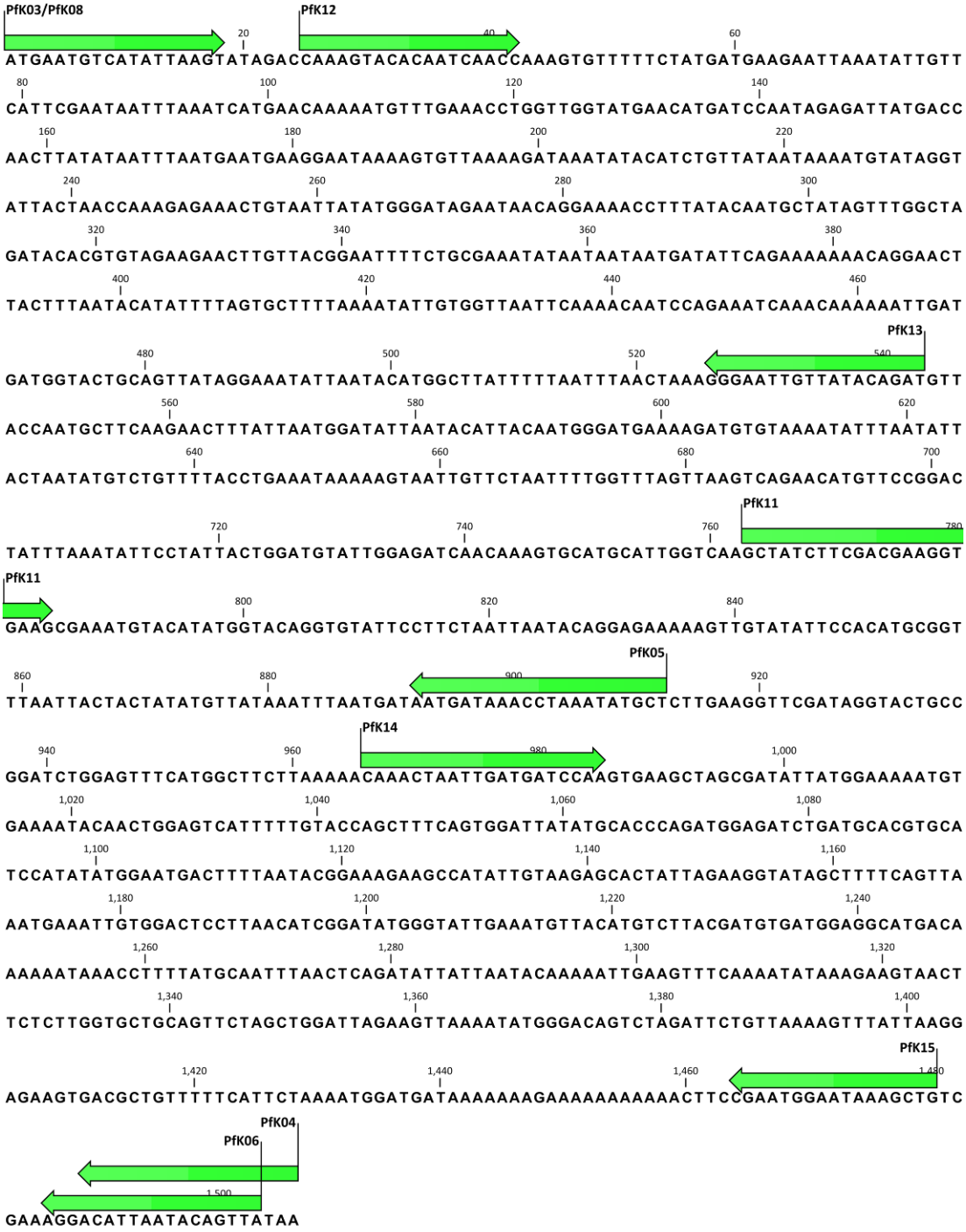



**Figure 66: Primer binding sites on the *Pf*GK gene.**
DNA sequence of the 1,506bp full length *Pf*GK gene. Nucleotide positions are numbered. Oligonucleotide names and binding sites are represented by green arrows.

**Table 30: Summary of primer descriptions and sequences used in the study.**

| Primer Name | Primer description | 5' - - - - 3' Primer sequence | Restriction Site |
|---|---|---|---|
| PfK03 | pGEX-*Pf*GK-F | gAT ggA TCC ATg AAT gTC ATA TTA AgT | *Bam*HI |
| PfK04 | pGEX-*Pf*GK-R | gAT CTC gAg TTA TAA CTg TAT TAA TgT | *Xho*I |
| PfK05 | pGEX-*Pf*GK-R-Internal | AgC ATA TTT Agg TTT ATC ATT | None |
| PfK06 | pTriEx-*Pf*GK-R | gAT CTC gAg TAA CTg TAT TAA TgT CCT | *Xho*I |
| PfK08 | pTriEx-*Pf*GK-F | gAT ggA TCC **G** ATg AAT gTC ATA TTA AgT | *Bam*HI |
| PfK11 | *Pf*GK_ Internal_F | gCT ATC TTC gAC gAA ggT gAA g | None |
| PfK12 | pCC1-*Pf*GK_KO1_F | gATC ACT AgT CAA AgT ACA CAA TCA ACC | *Spe*I |
| PfK13 | pCC1-*Pf*GK_KO1_R | gATC CTT AAg ATC TgT ATA ACA ATT CCC | *Afl*II |
| PfK14 | pCC1-*Pf*GK_KO2_F | gATC gAA TTC CAA ACT AAT TgA TgA TCC AA | *Eco*RI |
| PfK15 | pCC1-*Pf*GK_KO2_R | gATC CCA Tgg CAg CTT TAT TCC ATT Cg | *Nco*I |
| PfK16 | *Pf*GK_Insect_ Internal_F | gCT ATC TTC gAC gAg ggC gAg | None |
| PfK17 | *Pf*GK_Insect_ Internal_R | AgC gTA CTT ggg CTT gTC gTT | None |
| Cam-R | pCC-1-Cam-R | CCA ATA gAT AAA ATT TgT AgA | None |
| HrpII | pCC-1-HrpII-F | CTC ATT ATA TAT AAg AAC AT | None |
| P133 | Rifin-F | CgC TAC ATA TgC TAg TTg CCA gAC | None |
| P134 | Rifin-R | CgT gTA CCA TAC ATC CTA CCA AC | None |
| 18s-F | 18s rRNA-F | TgA CTA CgT CCC TgC CCT T | None |
| 18s-R | 18s rRNA-R | ACA ATT CAT CAT ATC TTT CAA TCg g | None |
| SE-pGEX-F | Forward | ggg CTg gCA AgC CAC gTT Tgg Tg | None |
| SE-pGEX-R | Reverse | CCg ggA gCT gCA TgT gTC AgA gg | None |
| SE-pTriEx3-F | Forward | gTT ATT gTg CTg TCT CAT CA | None |
| SE-pTriEx3-R | Reverse | CAC AAA TAC CAC TgA gAT CgA | None |
| *Pf*EBA-175-sc_F | pCC-1-EBA_F | TAA CCg gAT CCT TCA ATg TgC ATA CAA TgA AAT | None |
| *Pf*EBA-175-sc_R | pCC-1-EBA_R | ggA CCg Cgg CCg CAT CAg gAA TAC ATA CAT AgT | None |

Restriction sites are underlined
Nucleotides highlighted in blue are additional nucleotides required for efficient restriction enzyme digestion
Nucleotides highlighted in red represent a stop codon
The added **G** in PfK08 was necessary for in-frame translation of *Pf*GK in the pTriEx-3 expression system

# REFERENCES


Agnandji, S.T., Lell, B., Soulanoudjingar, S.S., *et al.* (2011) First results of phase 3 trial of RTS,S/AS01 malaria vaccine in African children. *N Engl J Med*, **365**, 1863-75.

Agre, P., King, L.S., Yasui, M., Guggino, W.B., Ottersen, O.P., Fujiyoshi, Y., Engel, A. and Nielsen, S. (2002) Aquaporin water channels-from atomic structure to clinical medicine. *J Physiol*, **542**, 3-16.

Alano, P. and Billker, O. (2005) Gametocytes and gametes. In *Molecular Approaches to Malaria*. (Ed, Sherman, I.W.) ASM Press, Washington, D.C.

Alberge, B., Gannoun-Zaki, L., Bascunana, C., Tran van Ba, C., Vial, H. and Cerdan, R. (2010) Comparison of the cellular and biochemical properties of *Plasmodium falciparum* choline and ethanolamine kinases. *Biochem J*, **425**, 149-58.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*, **25**, 3389-402.

Ancelin, M.L., Calas, M., Vidal-Sailhan, V., Herbute, S., Ringwald, P. and Vial, H.J. (2003) Potent inhibitors of *Plasmodium* phospholipid metabolism with a broad spectrum of *in vitro* antimalarial activities. *Antimicrob Agents Chemother*, **47**, 2590-7.

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

Arechaga, I., Miroux, B., Runswick, M.J. and Walker, J.E. (2003) Over-expression of *Escherichia coli* F1F(o)-ATPase subunit a is inhibited by instability of the uncB gene transcript. *FEBS Lett*, **547**, 97-100.

Aurrecoechea, C., Brestelli, J., Brunk, B.P., Dommer, J., Fischer, S., Gajria, B., Gao, X., Gingle, A., Grant, G., Harb, O.S., Heiges, M., Innamorato, F., Iodice, J., Kissinger, J.C., Kraemer, E., Li, W., Miller, J.A., Nayak, V., Pennington, C., Pinney, D.F., Roos, D.S., Ross, C., Stoeckert, C.J., Jr., Treatman, C. and Wang, H. (2009) PlasmoDB: a functional genomic database for malaria parasites. *Nucleic Acids Res*, **37**, D539-43.

Bahl, A., Brunk, B., Coppel, R.L., Crabtree, J., Diskin, S.J., Fraunholz, M.J., Grant, G.R., Gupta, D., Huestis, R.L., Kissinger, J.C., Labo, P., Li, L., McWeeney, S.K., Milgram, A.J., Roos, D.S., Schug, J. and Stoeckert, C.J., Jr. (2002) PlasmoDB: the *Plasmodium* genome resource. An integrated database providing tools for accessing, analyzing and mapping expression and sequence data (both finished and unfinished). *Nucleic Acids Res*, **30**, 87-90.

Beitz, E., Pavlovic-Djuranovic, S., Yasui, M., Agre, P. and Schultz, J.E. (2004) Molecular dissection of water and glycerol permeability of the aquaglyceroporin from *Plasmodium falciparum* by mutational analysis. *Proc Natl Acad Sci U S A*, **101**, 1153-8.

Berman, H.M., Bhat, T.N., Bourne, P.E., Feng, Z., Gilliland, G., Weissig, H. and Westbrook, J. (2000) The Protein Data Bank and the challenge of structural genomics. *Nat Struct Biol*, **7 Suppl**, 957-9.

Bhattacharyya, M.K. and Kumar, N. (2008) Culturing of sexual, oocyst, and sporozoite stages. In *Methods in Malaria Research*. 5th edition (Eds, Moll, K., Ljungström, I., Perlmann, H., Scherf, A. and Wahlgren, M.) MR4 / ATCC, Manassas, Virginia.

Biagini, G.A., Pasini, E.M., Hughes, R., De Koning, H.P., Vial, H.J., O'Neill, P.M., Ward, S.A. and Bray, P.G. (2004) Characterization of the choline carrier of *Plasmodium falciparum*: a route for the selective delivery of novel antimalarial drugs. *Blood*, **104**, 3372-7.

Biagini, G.A., Richier, E., Bray, P.G., Calas, M., Vial, H. and Ward, S.A. (2003) Heme binding contributes to antimalarial activity of bis-quaternary ammoniums. *Antimicrob Agents Chemother*, **47**, 2584-9.

Bietz, S., Montilla, I., Kulzer, S., Przyborski, J.M. and Lingelbach, K. (2009) Recruitment of human aquaporin 3 to internal membranes in the *Plasmodium falciparum* infected erythrocyte. *Mol Biochem Parasitol*, **167**, 48-53.

Billker, O., Lindo, V., Panico, M., Etienne, A.E., Paxton, T., Dell, A., Rogers, M., Sinden, R.E. and Morris, H.R. (1998) Identification of xanthurenic acid as the putative inducer of malaria development in the mosquito. *Nature*, **392**, 289-92.

Billker, O., Shaw, M.K., Margos, G. and Sinden, R.E. (1997) The roles of temperature, pH and mosquito factors as triggers of male and female gametogenesis of *Plasmodium berghei in vitro*. *Parasitology*, **115 (Pt 1)**, 1-7.

Birkholtz, L.M., Blatch, G., Coetzer, T.L., Hoppe, H.C., Human, E., Morris, E.J., Ngcete, Z., Oldfield, L., Roth, R., Shonhai, A., Stephens, L. and Louw, A.I. (2008) Heterologous expression of plasmodial proteins for structural studies and functional annotation. *Malar J*, **7**, 197.

Birnboim, H. (1983) A rapid alkaline extraction method for the isolation of plasmid DNA. *Methods in Enzymology*, **100**, 243-255.

Bozdech, Z. and Ginsburg, H. (2005) Data mining of the transcriptome of *Plasmodium falciparum*: the pentose phosphate pathway and ancillary processes. *Malar J*, **4**, 17.

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

Brenner, S.E. (1999) Errors in genome annotation. *Trends Genet*, **15**, 132-3.

Brisson, D., Vohl, M.C., St-Pierre, J., Hudson, T.J. and Gaudet, D. (2001) Glycerol: a neglected variable in metabolic processes? *Bioessays*, **23**, 534-42.

Bunnage, M.E. (2011) Getting pharmaceutical R&D back on target. *Nat Chem Biol*, **7**, 335-9.

Carter, R., Ranford-Cartwright, L. and Alano, P. (1993) The culture and preparation of gametocytes of *Plasmodium falciparum* for immunochemical, molecular, and mosquito infectivity studies. *Methods Mol Biol*, **21**, 67-88.

Cheng, A.C., Coleman, R.G., Smyth, K.T., Cao, Q., Soulard, P., Caffrey, D.R., Salzberg, A.C. and Huang, E.S. (2007) Structure-based maximal affinity model predicts small-molecule druggability. *Nat Biotechnol*, **25**, 71-5.

Chow, M.K., Amin, A.A., Fulton, K.F., Fernando, T., Kamau, L., Batty, C., Louca, M., Ho, S., Whisstock, J.C., Bottomley, S.P. and Buckle, A.M. (2006a) The REFOLD database: a tool for the optimization of protein expression and refolding. *Nucleic Acids Res*, **34**, D207-12.

Chow, M.K., Amin, A.A., Fulton, K.F., Whisstock, J.C., Buckle, A.M. and Bottomley, S.P. (2006b) REFOLD: an analytical database of protein refolding methods. *Protein Expr Purif*, **46**, 166-71.

Chowdhury, D.R., Angov, E., Kariuki, T. and Kumar, N. (2009) A potent malaria transmission blocking vaccine based on codon harmonized full length Pfs48/45 expressed in *Escherichia coli*. *PLoS One*, **4**, e6352.

Clyde, D.F. (1972) The problem of drug-resistant malaria. *Am J Trop Med Hyg*, **21**, 736-43.

Cohen, P. (2002) Protein kinases--the major drug targets of the twenty-first century? *Nat Rev Drug Discov*, **1**, 309-15.

Cowman, A., Crabb, B., Maier, A., Tonkin, C., Healer, J., Gibson, P. and De-Koning Ward, T. (2004) Transfection of *Plasmodium falciparum*. In *Methods in Malaria Research*. 4th edition (Eds, Ljunstrom, I., Perlman, H., Schlichtherle, M., Scherf, A. and Wahlgren, M.) MR4/ATCC, Manassas, Virginia.

Crabb, B.S. and Cowman, A.F. (1996) Characterization of promoters and stable transfection by homologous and nonhomologous recombination in *Plasmodium falciparum*. *Proc Natl Acad Sci U S A*, **93**, 7289-94.

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

Daily, J.P., Scanfeld, D., Pochet, N., Le Roch, K., Plouffe, D., Kamal, M., Sarr, O., Mboup, S., Ndir, O., Wypij, D., Levasseur, K., Thomas, E., Tamayo, P., Dong, C., Zhou, Y., Lander, E.S., Ndiaye, D., Wirth, D., Winzeler, E.A., Mesirov, J.P. and Regev, A. (2007) Distinct physiological states of *Plasmodium falciparum* in malaria-infected patients. *Nature*, **450**, 1091-5.

Dechamps, S., Maynadier, M., Wein, S., Gannoun-Zaki, L., Marechal, E. and Vial, H.J. (2010a) Rodent and nonrodent malaria parasites differ in their phospholipid metabolic pathways. *J Lipid Res*, **51**, 81-96.

Dechamps, S., Shastri, S., Wengelnik, K. and Vial, H.J. (2010b) Glycerophospholipid acquisition in *Plasmodium* - a puzzling assembly of biosynthetic pathways. *Int J Parasitol*, **40**, 1347-65.

Dechamps, S., Wengelnik, K., Berry-Sterkers, L., Cerdan, R., Vial, H.J. and Gannoun-Zaki, L. (2010c) The Kennedy phospholipid biosynthesis pathways are refractory to genetic disruption in *Plasmodium berghei* and therefore appear essential in blood stages. *Mol Biochem Parasitol*, **173**, 69-80.

DePristo, M.A., Zilversmit, M.M. and Hartl, D.L. (2006) On the abundance, amino acid composition, and evolutionary dynamics of low-complexity regions in proteins. *Gene*, **378**, 19-30.

Doerig, C., Abdi, A., Bland, N., Eschenlauer, S., Dorin-Semblat, D., Fennell, C., Halbert, J., Holland, Z., Nivez, M.P., Semblat, J.P., Sicard, A. and Reininger, L. (2010) Malaria: targeting parasite and host cell kinomes. *Biochim Biophys Acta*, **1804**, 604-12.

Doerig, C., Billker, O., Haystead, T., Sharma, P., Tobin, A.B. and Waters, N.C. (2008) Protein kinases of malaria parasites: an update. *Trends Parasitol*, **24**, 570-7.

Dondorp, A.M., Nosten, F., Yi, P., Das, D., Phyo, A.P., Tarning, J., Lwin, K.M., Ariey, F., Hanpithakpong, W., Lee, S.J., Ringwald, P., Silamut, K., Imwong, M., Chotivanich, K., Lim, P., Herdman, T., An, S.S., Yeung, S., Singhasivanon, P., Day, N.P., Lindegardh, N., Socheat, D. and White, N.J. (2009) Artemisinin resistance in *Plasmodium falciparum* malaria. *N Engl J Med*, **361**, 455-67.

Dorin-Semblat, D., Sicard, A., Doerig, C., Ranford-Cartwright, L. and Doerig, C. (2008) Disruption of the PfPK7 gene impairs schizogony and sporogony in the human malaria parasite *Plasmodium falciparum*. *Eukaryot Cell*, **7**, 279-85.

Dowhan, W. (1997) Molecular basis for membrane phospholipid diversity: why are there so many lipids? *Annu Rev Biochem*, **66**, 199-232.

Durand, P.M., Naidoo, K. and Coetzer, T.L. (2008) Evolutionary patterning: a novel approach to the identification of potential drug target sites in *Plasmodium falciparum*. *PLoS ONE*, **3**, e3685.

Dvorin, J.D., Martyn, D.C., Patel, S.D., Grimley, J.S., Collins, C.R., Hopp, C.S., Bright, A.T., Westenberger, S., Winzeler, E., Blackman, M.J., Baker, D.A., Wandless, T.J. and Duraisingh, M.T. (2010) A plant-like kinase in *Plasmodium falciparum* regulates parasite egress from erythrocytes. *Science*, **328**, 910-2.

Eda, S. and Sherman, I.W. (2002) Cytoadherence of malaria-infected red blood cells involves exposure of phosphatidylserine. *Cell Physiol Biochem*, **12**, 373-84.

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*, **32**, 1792-7.

Egan, T.J. (2008) Haemozoin formation. *Mol Biochem Parasitol*, **157**, 127-36.

Elabbadi, N., Ancelin, M.L. and Vial, H.J. (1994) Characterization of phosphatidylinositol synthase and evidence of a polyphosphoinositide cycle in *Plasmodium*-infected erythrocytes. *Mol Biochem Parasitol*, **63**, 179-92.

Elabbadi, N., Ancelin, M.L. and Vial, H.J. (1997) Phospholipid metabolism of serine in *Plasmodium*-infected erythrocytes involves phosphatidylserine and direct serine decarboxylation. *Biochem J*, **324 (Pt 2)**, 435-45.

Farris, J.S. (1989) The retention index and the rescaled consistency index. *Cladistics*, **5**, 417-19.

Feese, M.D., Faber, H.R., Bystrom, C.E., Pettigrew, D.W. and Remington, S.J. (1998) Glycerol kinase from *Escherichia coli* and an Ala65-->Thr mutant: the crystal structures reveal conformational changes with implications for allosteric regulation. *Structure*, **6**, 1407-18.

Felsenstein, J. (1985) Confidence of limits on phylogenies: an approach using bootstrap. *Evolution*, **39**, 793-91.

Fernandez, V. (2004) Sorbitol-synchronization of *Plasmodium falciparum*-infected erythrocytes. In *Methods in Malaria Research*. 4th edition (Eds, Ljungström, I., Perlmann, H., Schlichtherle, M., Scherf, A. and Wahlgren, M.) MR4 / ATCC, Manassas, Virginia.

Fichera, M.E. and Roos, D.S. (1997) A plastid organelle as a drug target in apicomplexan parasites. *Nature*, **390**, 407-9.

Fidock, D.A. (2011) Drug discovery: Priming the antimalarial pipeline. *Nature*, **465**, 297-8.

Fidock, D.A. and Wellems, T.E. (1997) Transformation with human dihydrofolate reductase renders malaria parasites insensitive to WR99210 but does not affect the intrinsic activity of proguanil. *Proc Natl Acad Sci U S A*, **94**, 10931-6.

Gamo, F.J., Sanz, L.M., Vidal, J., de Cozar, C., Alvarez, E., Lavandera, J.L., Vanderwall, D.E., Green, D.V., Kumar, V., Hasan, S., Brown, J.R., Peishoff, C.E., Cardon, L.R. and Garcia-Bustos, J.F. (2011) Thousands of chemical starting points for antimalarial lead identification. *Nature*, **465**, 305-10.

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

Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D. and A, B. (2005) Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook, Humana Press*. (Ed, Walker, J.M.) Humana Press, New Jersey, USA.

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

Gerold, P. and Schwarz, R.T. (2001) Biosynthesis of glycosphingolipids *de-novo* by the human malaria parasite *Plasmodium falciparum*. *Mol Biochem Parasitol*, **112**, 29-37.

Ginsburg, H. (2011) Metabolism: Malaria parasite stands out. *Nature*, **466**, 702-3.

Gowda, D.C. and Davidson, E.A. (1999) Protein glycosylation in the malaria parasite. *Parasitol Today*, **15**, 147-52.

Greenwood, B.M., Fidock, D.A., Kyle, D.E., Kappe, S.H., Alonso, P.L., Collins, F.H. and Duffy, P.E. (2008) Malaria: progress, perils, and prospects for eradication. *J Clin Invest*, **118**, 1266-76.

Guiguemde, W.A., Shelat, A.A., Bouck, D., Duffy, S., Crowther, G.J., Davis, P.H., Smithson, D.C., Connelly, M., Clark, J., Zhu, F., Jimenez-Diaz, M.B., Martinez, M.S., Wilson, E.B., Tripathi, A.K., Gut, J., Sharlow, E.R., Bathurst, I., El Mazouni, F., Fowble, J.W., Forquer, I., McGinley, P.L., Castro, S., Angulo-Barturen, I., Ferrer, S., Rosenthal, P.J., Derisi, J.L., Sullivan, D.J., Lazo, J.S., Roos, D.S., Riscoe, M.K., Phillips, M.A., Rathod, P.K., Van Voorhis, W.C., Avery, V.M. and Guy, R.K. (2011) Chemical genetics of *Plasmodium falciparum*. *Nature*, **465**, 311-5.

Haeggström, M. and Schlichtherle, M. (2004) Freezing and thawing of asexual *Plasmodium* spp. In *Methods in Malaria Research*. 4th edition (Eds, Ljungström, I., Perlmann, H., Schlichtherle, M., Scherf, A. and Wahlgren, M.) MR4 / ATCC, Manassas, Virginia.

Haeggström, M., Schlichtherle, M. and Bolad, A. (2004) Culturing of erythrocytic asexual stages of *Plasmodium falciparum* and *P. vivax*. In *Methods in malaria research*. 4th edition (Eds, Ljungström, I., Perlmann, H., Schlichtherle, M., Scherf, A.a. and Wahlgren, M.) MR4/ATCC, Manassas, Virginia.

Hajduk, P.J., Huth, J.R. and Tse, C. (2005a) Predicting protein druggability. *Drug Discov Today*, **10**, 1675-82.

Hall, T. (1999) BioEdit: A user friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Series*, **41**, 95-98.

Hanada, K., Mitamura, T., Fukasawa, M., Magistrado, P.A., Horii, T. and Nishijima, M. (2000) Neutral sphingomyelinase activity dependent on $Mg^{2+}$ and anionic phospholipids in the intraerythrocytic malaria parasite *Plasmodium falciparum*. *Biochem J*, **346 Pt 3**, 671-7.

Hanada, K., Palacpac, N.M., Magistrado, P.A., Kurokawa, K., Rai, G., Sakata, D., Hara, T., Horii, T., Nishijima, M. and Mitamura, T. (2002) *Plasmodium falciparum* phospholipase C hydrolyzing sphingomyelin and lysocholinephospholipids is a possible target for malaria chemotherapy. *J Exp Med*, **195**, 23-34.

Hannun, Y.A. and Obeid, L.M. (2008) Principles of bioactive lipid signalling: lessons from sphingolipids. *Nat Rev Mol Cell Biol*, **9**, 139-50.

Hansen, M., Kun, J.F., Schultz, J.E. and Beitz, E. (2002) A single, bi-functional aquaglyceroporin in blood-stage *Plasmodium falciparum* malaria parasites. *J Biol Chem*, **277**, 4874-82.

Hayakawa, T., Culleton, R., Otani, H., Horii, T. and Tanabe, K. (2008) Big bang in the evolution of extant malaria parasites. *Mol Biol Evol*, **25**, 2233-9.

Helenius, A. and Aebi, M. (2004) Roles of N-linked glycans in the endoplasmic reticulum. *Annu Rev Biochem*, **73**, 1019-49.

Hill, D.R., Baird, J.K., Parise, M.E., Lewis, L.S., Ryan, E.T. and Magill, A.J. (2006) Primaquine: report from CDC expert meeting on malaria chemoprophylaxis I. *Am J Trop Med Hyg*, **75**, 402-15.

Hillis, D.M. and Huelsenbeck, J.P. (1992) Signal, noise, and reliability in molecular phylogenetic analyses. *J Hered*, **83**, 189-95.

Hjelmeland, L., Allenmark, S., An der Lan, B., Jackiw, B.A., Nguyen, N.Y. and Chrambach, A. (1981) Electrophoresis in gels containing zwitterionic groups. *Electrophoresis*, **2**, 82-90.

Holms, H. (1996) Flux analysis and control of the central metabolic pathways in *Escherichia coli*. *FEMS Microbiol Rev*, **19**, 85-116.

Holz, G.G., Jr. (1977) Lipids and the malarial parasite. *Bull World Health Organ*, **55**, 237-48.

Hughes, A.L. (2004) The evolution of amino acid repeat arrays in *Plasmodium* and other organisms. *J Mol Evol*, **59**, 528-35.

Hurley, J.H. (1996) The sugar kinase/heat shock protein 70/actin superfamily: implications of conserved structure for mechanism. *Annu Rev Biophys Biomol Struct*, **25**, 137-62.

Hurley, J.H., Faber, H.R., Worthylake, D., Meadow, N.D., Roseman, S., Pettigrew, D.W. and Remington, S.J. (1993) Structure of the regulatory complex of *Escherichia coli* $III^{Glc}$ with glycerol kinase. *Science*, **259**, 673-7.

Hurtley, S., Ash, C. and Roberts, L. (2010) Tuberculosis & malaria. Landscapes of infection. Introduction. *Science*, **328**, 841.

Hyde, J.E. (2005) Drug-resistant malaria. *Trends Parasitol*, **21**, 494-8.

Jomaa, H., Wiesner, J., Sanderbrand, S., Altincicek, B., Weidemeyer, C., Hintz, M., Turbachova, I., Eberl, M., Zeidler, J., Lichtenthaler, H.K., Soldati, D. and Beck, E. (1999) Inhibitors of the nonmevalonate pathway of isoprenoid biosynthesis as antimalarial drugs. *Science*, **285**, 1573-6.

Kabsch, W. and Holmes, K.C. (1995) The actin fold. *Faseb J*, **9**, 167-74.

Kabsch, W. and Sander, C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**, 2577-637.

Kapust, R.B. and Waugh, D.S. (1999) *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci*, **8**, 1668-74.

Karlin, S., Brocchieri, L., Bergman, A., Mrazek, J. and Gentles, A.J. (2002) Amino acid runs in eukaryotic proteomes and disease associations. *Proc Natl Acad Sci U S A*, **99**, 333-8.

Kawamoto, F., Kido, N., Hanaichi, T., Djamgoz, M.B. and Sinden, R.E. (1992) Gamete development in *Plasmodium berghei* regulated by ionic exchange mechanisms. *Parasitol Res*, **78**, 277-84.

Kent, C. (1995) Eukaryotic phospholipid biosynthesis. *Annu Rev Biochem*, **64**, 315-43.

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

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

Kool, M., Voncken, J.W., van Lier, F.L., Tramper, J. and Vlak, J.M. (1991) Detection and analysis of *Autographa californica* nuclear polyhedrosis virus mutants with defective interfering properties. *Virology*, **183**, 739-46.

Kost, T.A., Condreay, J.P. and Jarvis, D.L. (2005) Baculovirus as versatile vectors for protein expression in insect and mammalian cells. *Nat Biotechnol*, **23**, 567-75.

Kralova, I., Rigden, D.J., Opperdoes, F.R. and Michels, P.A. (2000) Glycerol kinase of *Trypanosoma brucei*. Cloning, molecular characterization and mutagenesis. *Eur J Biochem*, **267**, 2323-33.

Krishnegowda, G. and Gowda, D.C. (2003) Intraerythrocytic *Plasmodium falciparum* incorporates extraneous fatty acids to its lipids without any structural modification. *Mol Biochem Parasitol*, **132**, 55-8.

Kumar, N. and Zheng, H. (1990) Stage-specific gametocytocidal effect *in vitro* of the antimalaria drug qinghaosu on *Plasmodium falciparum*. *Parasitol Res*, **76**, 214-8.

Kuntz, I.D., Chen, K., Sharp, K.A. and Kollman, P.A. (1999) The maximal affinity of ligands. *Proc Natl Acad Sci U S A*, **96**, 9997-10002.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680-5.

Lambros, C. and Vanderberg, J.P. (1979) Synchronization of *Plasmodium falciparum* erythrocytic stages in culture. *J Parasitol*, **65**, 418-20.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J. and Higgins, D.G. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics*, **23**, 2947-8.

Lasonder, E., Janse, C.J., van Gemert, G.J., Mair, G.R., Vermunt, A.M., Douradinha, B.G., van Noort, V., Huynen, M.A., Luty, A.J., Kroeze, H., Khan, S.M., Sauerwein, R.W., Waters, A.P., Mann, M. and Stunnenberg, H.G. (2008) Proteomic profiling of *Plasmodium* sporozoite maturation identifies new proteins essential for parasite development and infectivity. *PLoS Pathog*, **4**, e1000195.

Lauer, S.A., Chatterjee, S. and Haldar, K. (2001) Uptake and hydrolysis of sphingomyelin analogues in *Plasmodium falciparum*-infected red cells. *Mol Biochem Parasitol*, **115**, 275-81.

Laufer, M.K., Thesing, P.C., Eddington, N.D., Masonga, R., Dzinjalamala, F.K., Takala, S.L., Taylor, T.E. and Plowe, C.V. (2006) Return of chloroquine antimalarial efficacy in Malawi. *N Engl J Med*, **355**, 1959-66.

Le Roch, K.G., Johnson, J.R., Ahiboh, H., Chung, D.W., Prudhomme, J., Plouffe, D., Henson, K., Zhou, Y., Witola, W., Yates, J.R., Mamoun, C.B., Winzeler, E.A. and Vial, H. (2008) A systematic approach to understand the mechanism of action of the bisthiazolium compound T4 on the human malaria parasite, *Plasmodium falciparum*. *BMC Genomics*, **9**, 513.

Le Roch, K.G., Johnson, J.R., Florens, L., Zhou, Y., Santrosyan, A., Grainger, M., Yan, S.F., Williamson, K.C., Holder, A.A., Carucci, D.J., Yates, J.R., 3rd and Winzeler, E.A. (2004) Global analysis of transcript and protein levels across the *Plasmodium falciparum* life cycle. *Genome Res*, **14**, 2308-18.

Le Roch, K.G., Zhou, Y., Blair, P.L., Grainger, M., Moch, J.K., Haynes, J.D., De La Vega, P., Holder, A.A., Batalov, S., Carucci, D.J. and Winzeler, E.A. (2003) Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science*, **301**, 1503-8.

Lee, P.J., Bhonsle, J.B., Gaona, H.W., Huddler, D.P., Heady, T.N., Kreishman-Deitrick, M., Bhattacharjee, A., McCalmont, W.F., Gerena, L., Lopez-Sanchez, M., Roncal, N.E., Hudson, T.H., Johnson, J.D., Prigge, S.T. and Waters, N.C. (2009) Targeting the fatty acid biosynthesis enzyme, beta-ketoacyl-acyl carrier protein synthase III (PfKASIII), in the identification of novel antimalarial agents. *J Med Chem*, **52**, 952-63.

Lehane, A.M., Saliba, K.J., Allen, R.J. and Kirk, K. (2004) Choline uptake into the malaria parasite is energized by the membrane potential. *Biochem Biophys Res Commun*, **320**, 311-7.

Lemieux, J.E., Gomez-Escobar, N., Feller, A., Carret, C., Amambua-Ngwa, A., Pinches, R., Day, F., Kyes, S.A., Conway, D.J., Holmes, C.C. and Newbold, C.I. (2009) Statistical estimation of cell-cycle progression and lineage commitment in *Plasmodium falciparum* reveals a homogeneous pattern of transcription in ex vivo culture. *Proc Natl Acad Sci U S A*, **106**, 7559-64.

LeRoux, M., Lakshmanan, V. and Daily, J.P. (2009) *Plasmodium falciparum* biology: analysis of *in vitro* versus *in vivo* growth conditions. *Trends Parasitol*, **25**, 474-81.

Li, S.F., Wang, H.L., Hu, Z.H. and Deng, F. (2012) Genetic modification of baculovirus expression vectors. *Virol Sin*, **27**, 71-82.

Lian, L.Y., Al-Helal, M., Roslaini, A.M., Fisher, N., Bray, P.G., Ward, S.A. and Biagini, G.A. (2009) Glycerol: an unexpected major metabolite of energy metabolism by the human malaria parasite. *Malar J*, **8**, 38.

Liang, H., Narum, D.L., Fuhrmann, S.R., Luu, T. and Sim, B.K. (2000) A recombinant baculovirus-expressed *Plasmodium falciparum* receptor-binding domain of erythrocyte binding protein EBA-175 biologically mimics native protein. *Infect Immun*, **68**, 3564-8.

Lichtarge, O., Bourne, H.R. and Cohen, F.E. (1996) An evolutionary trace method defines binding surfaces common to protein families. *J Mol Biol*, **257**, 342-58.

Lichtarge, O. and Sowa, M.E. (2002) Evolutionary predictions of binding surfaces and interactions. *Curr Opin Struct Biol*, **12**, 21-7.

Lim, L. and McFadden, G.I. (2010) The evolution, metabolism and functions of the apicoplast. *Philos Trans R Soc Lond B Biol Sci*, **365**, 749-63.

Lin, E.C. (1977) Glycerol utilization and its regulation in mammals. *Annu Rev Biochem*, **46**, 765-95.

Liu, J., Istvan, E.S., Gluzman, I.Y., Gross, J. and Goldberg, D.E. (2006) *Plasmodium falciparum* ensures its amino acid supply with multiple acquisition pathways and redundant proteolytic enzyme systems. *Proc Natl Acad Sci U S A*, **103**, 8840-5.

Liu, Y., Promeneur, D., Rojek, A., Kumar, N., Frokiaer, J., Nielsen, S., King, L.S., Agre, P. and Carbrey, J.M. (2007) Aquaporin 9 is the major pathway for glycerol uptake by mouse erythrocytes, with implications for malarial virulence. *Proc Natl Acad Sci U S A*, **104**, 12560-4.

Llinas, M., Bozdech, Z., Wong, E.D., Adai, A.T. and DeRisi, J.L. (2006) Comparative whole genome transcriptome analysis of three *Plasmodium falciparum* strains. *Nucleic Acids Res*, **34**, 1166-73.

Lu, J.Z., Lee, P.J., Waters, N.C. and Prigge, S.T. (2005) Fatty Acid synthesis as a target for antimalarial drug discovery. *Comb Chem High Throughput Screen*, **8**, 15-26.

Maguire, P.A. and Sherman, I.W. (1990) Phospholipid composition, cholesterol content and cholesterol exchange in *Plasmodium falciparum*-infected red cells. *Mol Biochem Parasitol*, **38**, 105-12.

Maier, A.G., Braks, J.A., Waters, A.P. and Cowman, A.F. (2006) Negative selection using yeast cytosine deaminase/uracil phosphoribosyl transferase in *Plasmodium falciparum* for targeted gene deletion by double crossover recombination. *Mol Biochem Parasitol*, **150**, 118-21.

Makler, M.T., Lee, L.G. and Recktenwald, D. (1987) Thiazole orange: a new dye for *Plasmodium* species analysis. *Cytometry*, **8**, 568-70.

Makrides, S.C. (1996) Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiol Rev*, **60**, 512-38.

Mamoun, B.C., Prigge, S.T. and Vial, H. (2010) Targeting the Lipid Metabolic Pathways for the Treatment of Malaria. *Drug Dev Res*, **71**, 44-55.

Martin, R.E., Ginsburg, H. and Kirk, K. (2009) Membrane transport proteins of the malaria parasite. *Mol Microbiol*, **74**, 519-28.

Martinez Agosto, J.A. and McCabe, E.R. (2006) Conserved family of glycerol kinase loci in *Drosophila melanogaster*. *Mol Genet Metab*, **88**, 334-45.

Massingham, T. and Goldman, N. (2005) Detecting amino acid sites under positive selection and purifying selection. *Genetics*, **169**, 1753-62.

McRobert, L., Taylor, C.J., Deng, W., Fivelman, Q.L., Cummings, R.M., Polley, S.D., Billker, O. and Baker, D.A. (2008) Gametogenesis in malaria parasites is mediated by the cGMP-dependent protein kinase. *PLoS Biol*, **6**, e139.

Mehlin, C., Boni, E., Buckner, F.S., Engel, L., Feist, T., Gelb, M.H., Haji, L., Kim, D., Liu, C., Mueller, N., Myler, P.J., Reddy, J.T., Sampson, J.N., Subramanian, E., Van Voorhis, W.C., Worthey, E., Zucker, F. and Hol, W.G. (2006) Heterologous expression of proteins from *Plasmodium falciparum*: results from 1000 genes. *Mol Biochem Parasitol*, **148**, 144-60.

Mi-Ichi, F., Kita, K. and Mitamura, T. (2006) Intraerythrocytic *Plasmodium falciparum* utilize a broad range of serum-derived fatty acids with limited modification for their growth. *Parasitology*, **133**, 399-410.

Mita, T., Tanabe, K. and Kita, K. (2009) Spread and evolution of *Plasmodium falciparum* drug resistance. *Parasitol Int*, **58**, 201-9.

Mlambo, G., Kumar, N. and Yoshida, S. (2010) Functional immunogenicity of baculovirus expressing Pfs25, a human malaria transmission-blocking vaccine candidate antigen. *Vaccine*, **28**, 7025-9.

Mok, S., Imwong, M., Mackinnon, M.J., Sim, J., Ramadoss, R., Yi, P., Mayxay, M., Chotivanich, K., Liong, K.Y., Russell, B., Socheat, D., Newton, P.N., Day, N.P., White, N.J., Preiser, P.R., Nosten, F., Dondorp, A.M. and Bozdech, Z. (2011) Artemisinin resistance in *Plasmodium falciparum* is associated with an altered temporal pattern of transcription. *BMC Genomics*, **12**, 391.

Moore, R.B., Obornik, M., Janouskovec, J., Chrudimsky, T., Vancova, M., Green, D.H., Wright, S.W., Davies, N.W., Bolch, C.J., Heimann, K., Slapeta, J., Hoegh-Guldberg, O., Logsdon, J.M. and Carter, D.A. (2008) A photosynthetic alveolate closely related to apicomplexan parasites. *Nature*, **451**, 959-63.

Moreira, C.K., Marrelli, M.T. and Jacobs-Lorena, M. (2004) Gene expression in *Plasmodium*: from gametocytes to sporozoites. *Int J Parasitol*, **34**, 1431-40.

Nakamura, Y., Gojobori, T. and Ikemura, T. (2000) Codon usage tabulated from international DNA sequence databases: status for the year 2000. *Nucleic Acids Res*, **28**, 292.

Needleman, S.B. and Wunsch, C.D. (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J Mol Biol*, **48**, 443-53.

Nevoigt, E. and Stahl, U. (1997) Osmoregulation and glycerol metabolism in the yeast *Saccharomyces cerevisiae*. *FEMS Microbiol Rev*, **21**, 231-41.

Novagen (2004) *BugBuster Protein Extraction Reagent*. EMD Biosciences Inc, Darmstadt, Germany.

Novagen (2005) *Overnight Express™ Autoinduction Systems*. EMD Biosciences Inc, Darmstadt, Germany.

Novagen (2007) *BacMagic™ DNA Kit*. EMD Biosciences Inc, Darmstadt, Germany.

Novagen (2008a) *Sf9 Cells*. EMD Chemicals Inc, Darmstadt, Germany.

Novagen (2008b) *TriEx Sf9 Cells*. EMD Chemicals Inc, Darmstadt, Germany.

Novagen (2009a) *Competent Cells*. EMD Chemicals Inc, Darmstadt, Germany.

Novagen (2009b) *Overnight ExpressTM Instant LB and TB Media*. EMD Chemicals Inc, Darmstadt, Germany.

Olszewski, K.L. and Llinas, M. (2010) Central carbon metabolism of *Plasmodium* parasites. *Mol Biochem Parasitol*, **175**, 95-103.

Olszewski, K.L., Mather, M.W., Morrisey, J.M., Garcia, B.A., Vaidya, A.B., Rabinowitz, J.D. and Llinas, M. (2010) Branched tricarboxylic acid metabolism in *Plasmodium falciparum*. *Nature*, **466**, 774-8.

Orjih, A.U., Ryerse, J.S. and Fitch, C.D. (1994) Hemoglobin catabolism and the killing of intraerythrocytic *Plasmodium falciparum* by chloroquine. *Experientia*, **50**, 34-9.

Ormo, M., Bystrom, C.E. and Remington, S.J. (1998) Crystal structure of a complex of *Escherichia coli* glycerol kinase and an allosteric effector fructose-1,6-bisphosphate. *Biochemistry*, **37**, 16565-72.

Page, R.D.M. (1996) TreeView: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci*, **12**, 357-8.

Palacpac, N.M., Hiramine, Y., Mi-ichi, F., Torii, M., Kita, K., Hiramatsu, R., Horii, T. and Mitamura, T. (2004) Developmental-stage-specific triacylglycerol biosynthesis, degradation and trafficking as lipid bodies in *Plasmodium falciparum*-infected erythrocytes. *J Cell Sci*, **117**, 1469-80.

Pasvol, G. (2010) Protective hemoglobinopathies and *Plasmodium falciparum* transmission. *Nat Genet*, **42**, 284-5.

Paul, S.M., Mytelka, D.S., Dunwiddie, C.T., Persinger, C.C., Munos, B.H., Lindborg, S.R. and Schacht, A.L. (2010) How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nat Rev Drug Discov*, **9**, 203-14.

Pessi, G., Kociubinski, G. and Mamoun, C.B. (2004) A pathway for phosphatidylcholine biosynthesis in *Plasmodium falciparum* involving phosphoethanolamine methylation. *Proc Natl Acad Sci U S A*, **101**, 6206-11.

Petersen, I., Eastman, R. and Lanzer, M. (2011) Drug-resistant malaria: molecular mechanisms and implications for public health. *FEBS Lett*, **585**, 1551-62.

Phan, I.Q., Pilbout, S.F., Fleischmann, W. and Bairoch, A. (2003) NEWT, a new taxonomy portal. *Nucleic Acids Res*, **31**, 3822-3.

Phyo, A.P., Nkhoma, S., Stepniewska, K., Ashley, E.A., Nair, S., McGready, R., ler Moo, C., Al-Saai, S., Dondorp, A.M., Lwin, K.M., Singhasivanon, P., Day, N.P., White, N.J., Anderson, T.J. and Nosten, F. (2012) Emergence of artemisinin-resistant malaria on the western border of Thailand: a longitudinal study. *Lancet*, **379**, 1960-6.

Pierce (2009) *Protein stability and storage*. Thermo Fisher Scientific Inc, Rockford, USA.

Pizzi, E. and Frontali, C. (2001) Low-complexity regions in *Plasmodium falciparum* proteins. *Genome Res*, **11**, 218-29.

Promeneur, D., Liu, Y., Maciel, J., Agre, P., King, L.S. and Kumar, N. (2007) Aquaglyceroporin PbAQP during intraerythrocytic development of the malaria parasite *Plasmodium berghei*. *Proc Natl Acad Sci U S A*, **104**, 2211-6.

Pukrittayakamee, S., White, N.J., Davis, T.M., Supanaranond, W., Crawley, J., Nagachinta, B. and Williamson, D.H. (1994) Glycerol metabolism in severe *falciparum* malaria. *Metabolism*, **43**, 887-92.

Raabe, A.C., Wengelnik, K., Billker, O. and Vial, H.J. (2011) Multiple roles for *Plasmodium berghei* phosphoinositide-specific phospholipase C in regulating gametocyte activation and differentiation. *Cell Microbiol*, **13**, 955-66.

Rich, S.M. and Ayala, F.J. (2006) Evolutionary Origins of Human Malaria Parasites. In *Emerging Infectious Diseases of the 21st Century: Malaria - Genetic and Evolutionary Aspects*. (Eds, Krishna, R. and Dronamraju, P.A.) Springer Science and Business Media, Inc, New York.

Rock, R.C. (1971) Incorporation of 14 C-labelled non-lipid precursors into lipid of *Plasmodium knowlesi in vitro*. *Comp Biochem Physiol B*, **40**, 657-69.

Rontein, D., Nishida, I., Tashiro, G., Yoshioka, K., Wu, W.I., Voelker, D.R., Basset, G. and Hanson, A.D. (2001) Plants synthesize ethanolamine by direct decarboxylation of serine using a pyridoxal phosphate enzyme. *J Biol Chem*, **276**, 35523-9.

Roth, E., Jr. (1990) *Plasmodium falciparum* carbohydrate metabolism: a connection between host cell and parasite. *Blood Cells*, **16**, 453-60; discussion 461-6.

Safran, M., Chalifa-Caspi, V., Shmueli, O., Olender, T., Lapidot, M., Rosen, N., Shmoish, M., Peter, Y., Glusman, G., Feldmesser, E., Adato, A., Peter, I., Khen, M., Atarot, T., Groner, Y. and Lancet, D. (2003) Human Gene-Centric Databases at the Weizmann Institute of Science: GeneCards, UDB, CroW 21 and HORDE. *Nucleic Acids Res*, **31**, 142-6.

Santiago, T.C., Zufferey, R., Mehra, R.S., Coleman, R.A. and Mamoun, C.B. (2004) The *Plasmodium falciparum* PfGatp is an endoplasmic reticulum membrane protein important for the initial step of malarial glycerolipid synthesis. *J Biol Chem*, **279**, 9222-32.

Schlichtherle, M. and Wahlgren, M. (2004) Small-scale genomic DNA isolation from *Plasmodium falciparum*. In *Methods in Malaria Research*. 4th edition (Eds, Ljungström, I., Perlmann, H., Schlichtherle, M., Scherf, A. and Wahlgren, M.) MR4 / ATCC, Manassas, Virginia.

Schnick, C., Polley, S.D., Fivelman, Q.L., Ranford-Cartwright, L.C., Wilkinson, S.R., Brannigan, J.A., Wilkinson, A.J. and Baker, D.A. (2009) Structure and non-essential function of glycerol kinase in *Plasmodium falciparum* blood stages. *Mol Microbiol*, **71**, 533-45.

Schumann, W. and Ferreira, L.C.S. (2004) Production of recombinant proteins in *Escherichia coli*. *Genetics and Molecular Biology*, **27**, 442-453.

Sijwali, P.S., Brinen, L.S. and Rosenthal, P.J. (2001) Systematic optimization of expression and refolding of the *Plasmodium falciparum* cysteine protease falcipain-2. *Protein Expr Purif*, **22**, 128-34.

Silvestrini, F., Lasonder, E., Olivieri, A., Camarda, G., van Schaijk, B., Sanchez, M., Younis Younis, S., Sauerwein, R. and Alano, P. (2010) Protein export marks the early phase of gametocytogenesis of the human malaria parasite *Plasmodium falciparum*. *Mol Cell Proteomics*, **9**, 1437-48.

Singh, B., Kim Sung, L., Matusop, A., Radhakrishnan, A., Shamsul, S.S., Cox-Singh, J., Thomas, A. and Conway, D.J. (2004) A large focus of naturally acquired *Plasmodium knowlesi* infections in human beings. *Lancet*, **363**, 1017-24.

Singh, S.M. and Panda, A.K. (2005) Solubilization and refolding of bacterial inclusion body proteins. *J Biosci Bioeng*, **99**, 303-10.

Slavic, K., Straschil, U., Reininger, L., Doerig, C., Morin, C., Tewari, R. and Krishna, S. (2010) Life cycle studies of the hexose transporter of *Plasmodium* species and genetic validation of their essentiality. *Mol Microbiol*, **75**, 1402-13.

Slep, K.C., Kercher, M.A., He, W., Cowan, C.W., Wensel, T.G. and Sigler, P.B. (2001) Structural determinants for regulation of phosphodiesterase by a G protein at 2.0 A. *Nature*, **409**, 1071-7.

Sorensen, H.P. and Mortensen, K.K. (2005a) Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *J Biotechnol*, **115**, 113-28.

Sorensen, H.P. and Mortensen, K.K. (2005b) Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*. *Microb Cell Fact*, **4**, 1.

Sowa, M.E., He, W., Slep, K.C., Kercher, M.A., Lichtarge, O. and Wensel, T.G. (2001) Prediction and confirmation of a site critical for effector regulation of RGS domain activity. *Nat Struct Biol*, **8**, 234-7.

Swofford, D.L. (2002) *PAUP* Phylogenetic analysis using parsimony (*and Other Methods)*. Sinauer Associates, Sunderland, Massachusetts.

Tarun, A.S., Peng, X., Dumpit, R.F., Ogata, Y., Silva-Rivera, H., Camargo, N., Daly, T.M., Bergman, L.W. and Kappe, S.H. (2008) A combined transcriptome and proteome survey of malaria parasite liver stages. *Proc Natl Acad Sci U S A*, **105**, 305-10.

Thorner, J.W. and Paulus, H. (1971) Composition and subunit structure of glycerol kinase from *Escherichia coli*. *J Biol Chem*, **246**, 3885-94.

Trager, W. and Jensen, J.B. (1976) Human malaria parasites in continuous culture. *Science*, **193**, 673-5.

Tripathi, T., Rahlfs, S., Becker, K. and Bhakuni, V. (2007) Glutathione mediated regulation of oligomeric structure and functional activity of *Plasmodium falciparum* glutathione S-transferase. *BMC Struct Biol*, **7**, 67.

UniProt (2012) Reorganizing the protein space at the Universal Protein Resource (UniProt). *Nucleic Acids Res*, **40**, D71-5.

Vaid, A., Ranjan, R., Smythe, W.A., Hoppe, H.C. and Sharma, P. (2010) PfPI3K, a phosphatidylinositol-3 kinase from Plasmodium falciparum, is exported to the host erythrocyte and is involved in hemoglobin trafficking. *Blood*, **115**, 2500-7.

Vallejo, L.F. and Rinas, U. (2004) Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins. *Microb Cell Fact*, **3**, 11.

van Oers, M.M. (2011) Opportunities and challenges for the baculovirus expression system. *J Invertebr Pathol*, **107 Suppl**, S3-15.

Vassilev, L.T., Vu, B.T., Graves, B., Carvajal, D., Podlaski, F., Filipovic, Z., Kong, N., Kammlott, U., Lukacs, C., Klein, C., Fotouhi, N. and Liu, E.A. (2004) *In vivo* activation of the p53 pathway by small-molecule antagonists of MDM2. *Science*, **303**, 844-8.

Vaughan, A.M., O'Neill, M.T., Tarun, A.S., Camargo, N., Phuong, T.M., Aly, A.S., Cowman, A.F. and Kappe, S.H. (2009) Type II fatty acid synthesis is essential only for malaria parasite late liver stage development. *Cell Microbiol*, **11**, 506-20.

Vedadi, M., Lew, J., Artz, J., Amani, M., Zhao, Y., Dong, A., Wasney, G.A., Gao, M., Hills, T., Brokx, S., Qiu, W., Sharma, S., Diassiti, A., Alam, Z., Melone, M., Mulichak, A., Wernimont, A., Bray, J., Loppnau, P., Plotnikova, O., Newberry, K., Sundararajan, E., Houston, S., Walker, J., Tempel, W., Bochkarev, A., Kozieradzki, I., Edwards, A., Arrowsmith, C., Roos, D., Kain, K. and Hui, R. (2007) Genome-scale protein expression and structural biology of *Plasmodium falciparum* and related Apicomplexan organisms. *Mol Biochem Parasitol*, **151**, 100-10.

Vial, H.J. and Ancelin, M.L. (1992) Malarial lipids. An overview. *Subcell Biochem*, **18**, 259-306.

Vial, H.J. and Ancelin, M.L. (1998) Malarial lipids. In *Malaria: Parasite biology, pathogenesis and protection*. (Ed, Sherman, I.) ASM Press, Washington DC.

Vial, H.J., Eldin, P., Tielens, A.G. and van Hellemond, J.J. (2003) Phospholipids in parasitic protozoa. *Mol Biochem Parasitol*, **126**, 143-54.

Vial, H.J. and Mamoun, C.B. (2005) *Plasmodium* lipids: metabolism and function. In *Molecular approaches to malaria*. (Ed, In: Sherman, I.W.) ASM Press, Washington DC.

Victor, M.E., Bengtsson, A., Andersen, G., Bengtsson, D., Lusingu, J.P., Vestergaard, L.S., Arnot, D.E., Theander, T.G., Joergensen, L. and Jensen, A.T. (2010) Insect cells are superior to *Escherichia coli* in producing malaria proteins inducing IgG targeting PfEMP1 on infected erythrocytes. *Malar J*, **9**, 325.

Vielemeyer, O., McIntosh, M.T., Joiner, K.A. and Coppens, I. (2004) Neutral lipid synthesis and storage in the intraerythrocytic stages of *Plasmodium falciparum*. *Mol Biochem Parasitol*, **135**, 197-209.

Vlak, J.M., Klinkenberg, F.A., Zaal, K.J., Usmany, M., Klinge-Roode, E.C., Geervliet, J.B., Roosien, J. and van Lent, J.W. (1988) Functional studies on the p10 gene of *Autographa californica* nuclear polyhedrosis virus using a recombinant expressing a p10-beta-galactosidase fusion gene. *J Gen Virol*, **69 (Pt 4)**, 765-76.

Vogel, G. (2010) The 'do unto others' malaria vaccine. *Science*, **328**, 847-8.

von Itzstein, M., Plebanski, M., Cooke, B.M. and Coppel, R.L. (2008) Hot, sweet and sticky: the glycobiology of *Plasmodium falciparum*. *Trends Parasitol*, **24**, 210-8.

Ward, J.J., McGuffin, L.J., Bryson, K., Buxton, B.F. and Jones, D.T. (2004) The DISOPRED server for the prediction of protein disorder. *Bioinformatics*, **20**, 2138-9.

Wengelnik, K., Vidal, V., Ancelin, M.L., Cathiard, A.M., Morgat, J.L., Kocken, C.H., Calas, M., Herrera, S., Thomas, A.W. and Vial, H.J. (2002) A class of potent antimalarials and their specific accumulation in infected erythrocytes. *Science*, **295**, 1311-4.

WHO (2007) *World Malaria Report: 2007*. Geneva.

WHO (2010a) *Global report on antimalarial drug efficacy and drug resistance: 2000-2010*. Geneva.

WHO (2010b) *World malaria report: 2010*. Geneva.

WHO (2011a) Malaria Vaccine Rainbow Tables. http://www.who.int/vaccine_research/links/Rainbow/en/index.html.

WHO (2011b) *World Malaria Report: 2011*. Geneva.

Wirth, D., Daily, J., Winzeler, E., Mesirov, J.P. and Regev, A. (2009) *In vivo* profiles in malaria are consistent with a novel physiological state. *Proc Natl Acad Sci U S A*, **106**, E70; author reply E71-2.

Witola, W.H., El Bissati, K., Pessi, G., Xie, C., Roepe, P.D. and Mamoun, C.B. (2008) Disruption of the *Plasmodium falciparum* PfPMT gene results in a complete loss of phosphatidylcholine biosynthesis via the serine-decarboxylase-phosphoethanolamine-methyltransferase pathway and severe growth and survival defects. *J Biol Chem*, **283**, 27636-43.

Woodrow, C.J., Burchmore, R.J. and Krishna, S. (2000) Hexose permeation pathways in *Plasmodium falciparum*-infected erythrocytes. *Proc Natl Acad Sci U S A*, **97**, 9931-6.

Wootton, J.C. (1994) Non-globular domains in protein sequences: automated segmentation using complexity measures. *Comput Chem*, **18**, 269-85.

Worthington, K. and Worthington, V. (Eds) (1993) *Glycerol kinase*. Worthington Biochemical Corporation, New Jersey, USA.

Xia, X. and Xie, Z. (2001) DAMBE: software package for data analysis in molecular biology and evolution. *J Hered*, **92**, 371-3.

Xia, X., Xie, Z., Salemi, M., Chen, L. and Wang, Y. (2003) An index of substitution saturation and its application. *Mol Phylogenet Evol*, **26**, 1-7.

Xue, H.Y. and Forsdyke, D.R. (2003) Low-complexity segments in *Plasmodium falciparum* proteins are primarily nucleic acid level adaptations. *Mol Biochem Parasitol*, **128**, 21-32.

Yang, Z. (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput Appl Biosci*, **13**, 555-6.

Young, C.L., Britton, Z.T. and Robinson, A.S. (2012) Recombinant protein expression and purification: a comprehensive review of affinity tags and microbial applications. *Biotechnol J*, **7**, 620-34.

Young, J.A., Fivelman, Q.L., Blair, P.L., de la Vega, P., Le Roch, K.G., Zhou, Y., Carucci, D.J., Baker, D.A. and Winzeler, E.A. (2005) The *Plasmodium falciparum* sexual development transcriptome: a microarray analysis using ontology-based pattern identification. *Mol Biochem Parasitol*, **143**, 67-79.

Yu, M., Kumar, T.R., Nkrumah, L.J., Coppi, A., Retzlaff, S., Li, C.D., Kelly, B.J., Moura, P.A., Lakshmanan, V., Freundlich, J.S., Valderramos, J.C., Vilcheze, C., Siedner, M., Tsai, J.H., Falkard, B., Sidhu, A.B., Purcell, L.A., Gratraud, P., Kremer, L., Waters, A.P., Schiehser, G., Jacobus, D.P., Janse, C.J., Ager, A., Jacobs, W.R., Jr., Sacchettini, J.C., Heussler, V., Sinnis, P. and Fidock, D.A. (2008) The fatty acid biosynthesis enzyme FabI plays a key role in the development of liver-stage malarial parasites. *Cell Host Microbe*, **4**, 567-78.

Yuvaniyama, J., Chitnumsub, P., Kamchonwongpaisan, S., Vanichtanankul, J., Sirawaraporn, W., Taylor, P., Walkinshaw, M.D. and Yuthavong, Y. (2003) Insights into antifolate resistance from malarial DHFR-TS structures. *Nat Struct Biol*, **10**, 357-65.

Zhou, Z., Schnake, P., Xiao, L. and Lal, A.A. (2004) Enhanced expression of a recombinant malaria candidate vaccine in *Escherichia coli* by codon optimization. *Protein Expr Purif*, **34**, 87-94.

Zilversmit, M.M., Volkman, S.K., DePristo, M.A., Wirth, D.F., Awadalla, P. and Hartl, D.L. (2010) Low-complexity regions in *Plasmodium falciparum*: missing links in the evolution of an extreme genome. *Mol Biol Evol*, **27**, 2198-209.

Zwaig, N., Kistler, W.S. and Lin, E.C. (1970) Glycerol kinase, the pacemaker for the dissimilation of glycerol in *Escherichia coli*. *J Bacteriol*, **102**, 753-9.