Thioredoxin reductase and dihydrolipoamide dehydrogenases of *Plasmodium falciparum*

by

Paul James M^{c}Millan *BSc (Hons)*

Thesis submitted for the degree of Doctor of Philosophy

March 2006

Division of Infection and Immunity,
Institute of Biomedical and Life Sciences,
University of Glasgow

Abstract

Plasmodium falciparum is an obligate intracellular protozoan parasite and is the causative agent of malaria, which infects 270 million people and causes 2 – 3 million human deaths each year. During the erythrocytic stages of their life cycle the parasites are exposed to reactive oxygen species generated by the host and their own metabolism. As a result the parasites have developed efficient antioxidant defence mechanisms, which include the thioredoxin system. A key component of the thioredoxin system is the flavin adenine dinucleotide (FAD)-dependent disulphide oxidoreductase thioredoxin reductase (PfTrxR), which provides the reducing power for the system using NADPH. PfTrxR is a high M_r TrxR and has been genetically and chemically validated as a potential drug target by targeted gene disruption and the identification of several specific inhibitors. PfTrxR contains three redox active centres (FAD, Cys88/Cys93 and Cys535/Cys540) that are in redox communication.

This study examined the catalytic mechanism (i.e. the transfer of reducing equivalents between these redox active centres) of PfTrxR by employing pre-steady state kinetics of the reductive (reduction of PfTrxR by NADPH) and oxidative half reaction (oxidation of reduced PfTrxR by thioredoxin). This identified that the PfTrxR protein conformed to the catalytic mechanism for high M_r TrxR, which was proposed from studies on the *Drosophila melanogaster* TrxR (DmTrxR). However, a few key differences were observed between the proteins from the two organisms. Firstly, the reductive half reaction of PfTrxR was observed to proceed at a faster rate than that of DmTrxR. Secondly, the oxidative half reaction was observed to proceed at a slower rate than that of DmTrxR protein. Finally, PfTrxR was shown not to cycle between a two (EH_2) and four (EH_4) electron reduced species *in vitro* (as demonstrated in DmTrxR) and was observed to be almost fully oxidised following reaction with excess thioredoxin. However, the rate of oxidation from the two electron reduced (EH_2) to oxidised (E_{ox}) forms was extremely slow and would probably not be relevant *in vivo* due to the high NADPH concentration.

The role of conserved PfTrxR active site residues was investigated by producing site directed mutants and analysing them by steady state kinetics, anaerobic static titrations and pre-steady state kinetics. Particular attention was paid to the residues (His509 and Glu514 in PfTrxR) contained in the acid/base catalyst motif (HPXXXE) that has been identified and investigated in other disulphide oxidoreductase proteins. This demonstrated that both of these residues played crucial, yet distinct, roles in the transfer

of reducing equivalents between the redox active centres in PfTrxR. The His509 residue was observed to be involved in the stabilisation of the thiolate flavin charge transfer complex intermediate and Glu514 was involved in the transfer of reducing equivalents between the Cys88/Cys93 and Cys535/Cys540 redox active centres. Mutation of His509 to glutamine and Glu514 to alanine resulted in 0.5 % and 7 % activity (relative to wild-type) with NADPH and thioredoxin as substrates, respectively. Analysis of another residue that had been previously proposed as an acid/base catalyst (His137 in PfTrxR), demonstrated that PfTrxR activity was relatively unaffected by mutation of this residue.

The publishing of the *P. falciparum* genome in 2002 led to the identification of two *dihydrolipoamide dehydrogenase (LipDH)* genes. LipDH is an essential component of α – keto acid dehydrogenase complexes (KADH) and is also a member of disulphide oxidoreductase protein family (along with TrxR and glutathione reductase). Both *LipDH* genes have been recombinantly expressed using the *Escherichia coli* system and the recombinant proteins were biochemically characterised showing that they displayed typical characteristics of LipDH proteins. The major difference between the two proteins was their turnover numbers.

In silico analyses of the *LipDH* genes suggested that the proteins were targeted to different organelles in the parasite (one to the mitochondrion and one to the apicoplast). This suggested that the KADH would also be localised in these organelles and the components of three KADH were identified in the *P. falciparum* genome; an α – ketoglutarate dehydrogenase complex (KGDH), a branched chain α – keto acid dehydrogenase complex (BCKDH) and a pyruvate dehydrogenase complex (PDH).

Targeting of proteins to both the mitochondrion and apicoplast has been previously shown to be due to N-terminal targeting sequences and we experimentally confirmed the *in silico* predictions by fusing N-terminal sections of the components of the three KADH with green fluorescent protein (GFP) and expressing the fusion protein in erythrocytic stage parasites. By analysing the fluorescent staining in these parasites we were able to demonstrate that *P. falciparum* does possess organelle specific LipDH proteins and that components of the KGDH and BCKDH were localised in the mitochondrion and components of the PDH were localised in the apicoplast.

The transcription and protein expression levels of components of the *P. falciparum* KADH were tested by RT-PCR and western blotting, respectively. This confirmed that these complexes were expressed during the erythrocytic stages of the parasite's life

cycle and would therefore be likely to play important roles in the parasites metabolism. The presence of a single apicoplast localised PDH is novel as organisms (such as plants and algae) that contain mitochondrion and plastids usually possess distinct PDH in these organelles. The *P. falciparum* PDH is likely to provide acetyl CoA in the apicoplast for fatty acid biosynthesis, but its absence from the mitochondrion raises further questions as to the function of the tricarboxylic acid cycle (TCA cycle) and the role of the mitochondrion in general. The presence of the KGDH and BCKDH in the mitochondrion is in line with their localisation in other organisms and suggestions for their roles are given.

Author's declaration

I hereby declare that I am the sole author of this thesis and performed all of the work presented, with the following exceptions:

Chapter 3

- CD spectrum analysis was performed by Dr. Sharon Kelly of the Protein Characterisation Facility, University of Glasgow.

- Data analysis was performed in conjunction with Dave Arscott and Dave Ballou from the Department of Biological Chemistry, University of Michigan.

Chapter 4

- MALDI-TOF mass spectrometry analysis was performed by the Fingerprints proteomics facility, University of Dundee.

Chapter 5

- The majority of fluorescence images were taken by Dr. Sylke Müller.

Paul James McMillan

Publications

McMillan, P.J., Stimmler, L.M., Foth, B.J., McFadden, G.I., Müller, S.
The human malaria parasite *Plasmodium falciparum* possesses two distinct dihydrolipoamide dehydrogenases.
Molecular Microbiology (2005) **55:** 27-38.

Günther S, McMillan PJ, Wallace LJ, Müller S.
Plasmodium falciparum possesses organelle-specific alpha-keto acid dehydrogenase complexes and lipoylation pathways.
Biochemical Society Transactions (2005) **33:** 977-80.

McMillan, P.J., Arscott, D.L., Ballou, D.P., Becker, K., Williams, C.H. Jr., Müller S.
Identification of acid-base catalytic residues of high Mr thioredoxin reductase from *Plasmodium falciparum*
Journal of Biological Chemistry (submitted)

Acknowledgement

I'd like to thank my supervisor Dr. Sylke Müller for giving me the opportunity to do a PhD in her laboratory and for her invaluable advice and mentorship. I would like to thank all members of the Müller group (old and new) for the laughs, tears and for putting up with my occasional tantrums. Special thanks go to Susan A, Carsten, Svenja and Lynsey for feeding my transfectants whilst I was working away in the States (your endeavours were greatly appreciated).

Thanks go to the Ann Arbor group (Drs. Charles Williams, Dave Arscott and Dave Ballou) for allowing me to visit their lab and being the most eager of hosts. Especially to Charles and Dave A. for the helpful discussion over my results, and explaining it all to me in plain English. Thanks also go to the members of Charles' and Dave B's groups for making me feel at home, and the staffs at the red roof inn for putting up with that damn Scotsman.

I'd like to thank my family for letting me get on with it all and being happy to only see me every couple of months. I'd especially like to thank my mum, who always said that I should be a doctor. Even though she only based that on my poor penmanship, it was enough to plant the seed. This one is for you mum!!

Finally (and most importantly), I'd like to thank Rachel for her love and support throughout my PhD and for going through the whole experience with me. You know what happens next .....

Table of Contents

List of figures

List of Tables

Abbreviations/Definitions

ε	Extinction coefficient
μ	micro
μg	microgram
μl	microlitre
μM	micromolar
ACP	Acyl carrier protein
ACT	Artemisinin combination therapies
BCKDH	branched chain α-keto acid dehydrogenase complex
bp	base pairs
BSA	Bovine serum albumin
C535A	Cysteine 535 of PfTrxR mutated to alanine
C540A	Cysteine 540 of PfTrxR mutated to alanine
C88A	Cysteine 88 of PfTrxR mutated to alanine
C93A	Cysteine 93 of PfTrxR mutated to alanine
C-terminal	Carboxy terminal
CD	Circular dichroism
cDNA	Complementary DNA
CoA	Coenzyme A
CoA (SH)	Coenzyme A (reduced form)
CTC	Charge transfer complex
CTC1	Flavin thiolate charge transfer complex
CTC2	$NADP^+$ $FADH^-$ charge transfer complex
Da	Daltons
DEPC	Diethyl pyrocarbonate
DMSO	Dimethyl sulphoxide
DmTrxR	*Drosophila melanogaster* thioredoxin reductase
DNA	Deoxyribonucleic acid
DOXP	1-deoxy-D-xylulose-5-phosphate
DTNB	5,5'-dithiobis-(2-nitrobenzoic acid)
E1	Dehydrogenase protein of KADH
E2	Acyl Transferase protein of KADH
E3	Dihydrolipoamide dehydrogenase (LipDH) protein of KADH
E514A	Glutamate 514 of PfTrxR mutated to alanine
E_{ox}	Oxidised form of protein
EH_2	2 electron reduced form of protein
EH_4	4 electron reduced form of protein
EH_6	6 electron reduced form of protein
ECL	Enhanced chemiluminescence
EDTA	Ethylenediamine tetra acetic acid
EMP-1	Erythrocyte membrane protein 1
eq	Molar equivalents
FAD	Flavin adenine dinucleotide
FP IX	Ferriprotoporphyrin IX/Haem
gDNA	Genomic DNA
GFP	Green fluorescent protein
GR	Glutathione reductase
Grx	Glutaredoxin
GSH	Glutathione (reduced form)
GSSG	Glutathione disulphide (oxidised form)
GST	Glutathione transferase
H137N	Histidine 137 of PfTrxR mutated to asparagine
H137Q	Histidine 137 of PfTrxR mutated to glutamine

H137S	Histidine 137 of PfTrxR mutated to serine
H137A	Histidine 137 of PfTrxR mutated to alanine
H_2O_2	Hydrogen peroxide
H509Q	Histidine 509 of PfTrxR mutated to glutamine
hDHFR	Human dihydrofolate reductase
ICAM	Intercellular adhesion molecule 1
IPTG	Isopropyl-ß-D-thiogalactopyranoside
K_m	Michaelis constant
k_{cat}	Catalytic constant
KADH	α-keto acid dehydrogenase complex
kb	kilobases
KGDH	α -ketoglutarate dehydrogenase complex
LB	Luria Bertani media
LB/Amp	LB containing ampicillin
LB/Kan	LN containing kanamycin
LipDH	Dihydrolipoamide dehydrogenase
m	milli
ml	millilitre
mM	millimolar
min	Minute(s)
MQO	Malate quinone oxidoreductase
MR	Mercuric ion reductase
MtPDH	Mitochondrial pyruvate dehydrogenase complex
M_r	Molecular weight
NaCl	Sodium chloride
NAD^+	Nicotinamide adenine dinucleotide, oxidised form
NADH	Nicotinamide adenine dinucleotide, reduced form
$NADP^+$	Nicotinamide adenine dinucleotide phosphate, oxidised form
NADPH	Nicotinamide adenine dinucleotide phosphate, reduced form
NAD(P)H	Either NADH or NADPH
nm	Nanometres
N-terminal	Amino terminal
O_2^-	Superoxide
OD	Optical density
ORF	Open reading frame
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDH	Pyruvate dehydrogenase complex
PfTrx	*Plasmodium falciparum* thioredoxin
PfTrxR	*Plasmodium falciparum* thioredoxin reductase
PMSF	Phenylmethyl Sulphonyl Flouride
PVM	Parasitophorous vacuole
Px	Peroxiredoxin/thioredoxin peroxidase
RBC	Red blood cell/erythrocyte
rpm	Revolutions per minute
ROS	Reactive oxygen species
RT-PCR	Reverse transcriptase PCR
SD	Standard deviation
SDS	Sodium dodecyl sulphate
SOD	Superoxide dismutase
TAE	Tris-acetate containing EDTA
TCA	Tricarboxylic acid
TIC	Translocase of the inner chloroplast membrane

TOC	Translocase of the outer chloroplast membrane
TIM	Translocase of the inner mitochondrial membrane
TOM	Translocase of the outer mitochondrial membrane
TPP	Thiamine pyrophosphate
Tris	Tris [hydroxymethyl] aminomethane
TryR	Trypanothione reductase
V_{max}	Maximum velocity

Amino Acids

Amino acid	Abbreviation	Single letter code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartate	Asp	D
Cysteine	Cys	C
Glutamate	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Chapter 1 Introduction

1.1 *Plasmodium falciparum*

Plasmodium falciparum is one of four species of the genus *Plasmodium* (the others are *P. ovale*, *P. malariae* and *P. vivax*) that infect humans and are the causative agent of malaria (Sherman, 1998). Of the four species it is the infection with *P. falciparum* that causes the most severe clinical symptoms resulting in 2 to 3 million deaths per annum primarily because the parasites have developed resistance against most of the antimalarial drugs in use. Chemotherapy is currently the only way to combat malarial infection, as a vaccine against the parasites is still elusive. Therefore the identification, validation and assessment of potential new targets for the design of antimalarials is urgently required.

Indeed the *P. falciparum* genome project was completed in 2002 (Gardner *et al.*, 2002) and has been widely used by scientists as a powerful tool to further investigate the parasite that causes malaria. *P. falciparum* has 14 chromosomes that range from 0.643 to 3.29 megabases (Mb) and the complete genome size is 23 Mb. The genome wide A/T content is 80.6 %, which was the highest ever sequenced and much higher when compared to that of *Saccharomyces cerevisiae* and *Arabidopsis thaliana* with 61.7 % and 65.1 %, respectively. The over 7,000 intergenic regions were found to have an overall A/T content of 86.5 %. The genome was predicted to contain over 5,000 protein coding genes with approximately 60 % genes that encode hypothetical proteins when it was published (Gardner *et al.*, 2002).

The availability of the entire genome sequence has also enabled scientists to engage in functional studies such as transcriptome and proteome analyses. Transcriptome analyses demonstrated that transcription from the parasite genome is highly regulated with developmental stage specific regulation of pathways (i.e. transcription of genes required for transcription and translation in ring stages, glycolysis in trophozoite stages and merozoite invasion in late schizont stages). Genes that were adjacent in the parasite's genome were rarely co-regulated, although transcription from the plastid genome was found to be highly co-regulated (Bozdech *et al.*, 2003). Proteome analyses further confirmed the stage specific expression of genes and also investigated the roles of expressed proteins in the sexual as well as the asexual stages of the parasites (Florens *et al.*, 2002).

Other studies investigated sequence motifs that are involved in the targeting of proteins into the different organelles within the parasite and also locations outside the parasites such as the Maurer's clefts and trafficking onto the erythrocytic surface (Przyborski *et al.*, 2005). Together these studies are highly informative and provide an excellent basis for future analyses that will help to better understand the metabolic processes governing parasite development and survival and thus offer fantastic opportunities to study these processes in more detail with the aim to find control measures against this disease.

1.1.1 Malaria

Malaria is estimated to infect approximately 270 million people resulting in around 2 - 3 million deaths globally each year (Snow *et al.*, 2005). Ninety percent of this death toll occurs in sub-Saharan Africa (Figure 1.1A), where most of the victims of the disease are young children and pregnant mothers (World health organisation, http://www.who.int/mediacentre/factsheets/fs094/en/). It has been estimated that 40 % of the world's population is at risk from malarial infection (Greenwood *et al.*, 2002). In addition to being a major health threat the disease causes great economic losses – in countries with high levels of malaria transmission the disease can account for up to 40 % of public health expenditure, 30 – 50 % of inpatient admissions, and up to 50 % of outpatient visits (Source: World health organisation, http://www.rbm.who.int/cmc_upload/0/000/015/363/RBMInfosheet_10.htm).

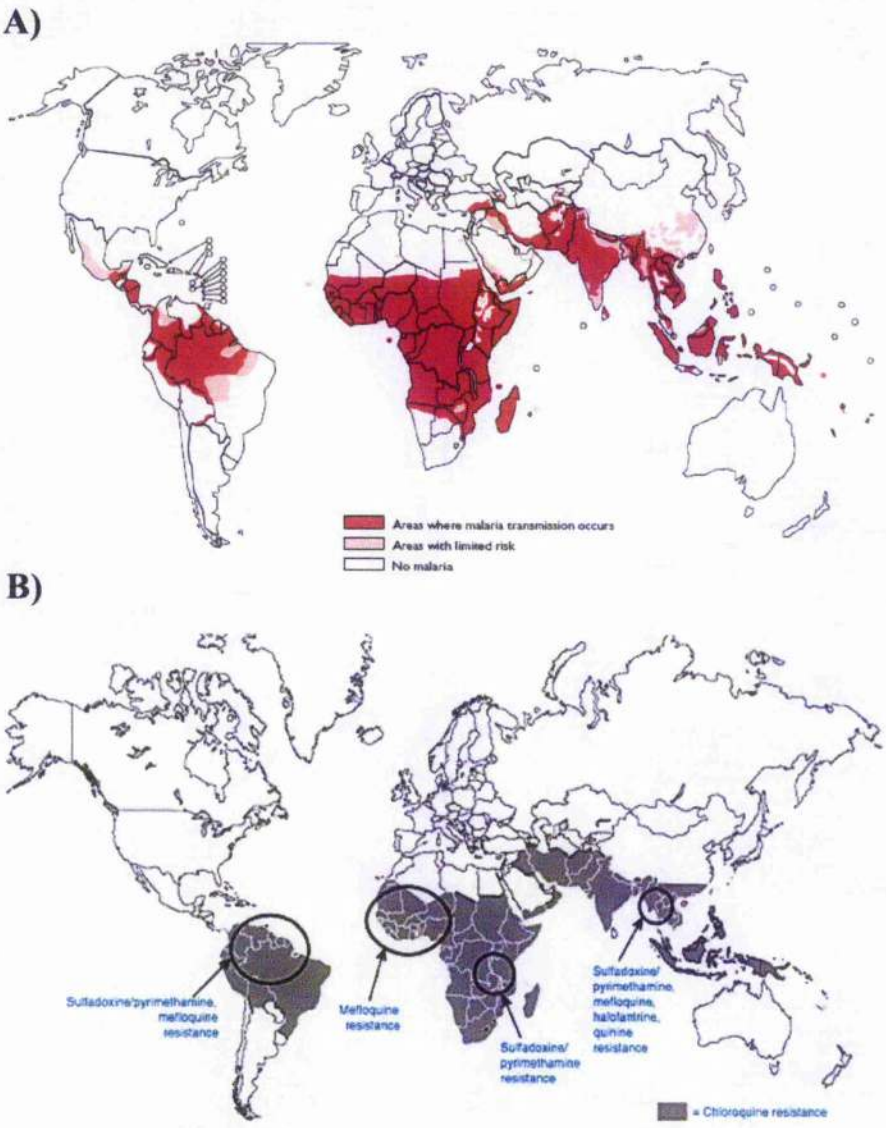


Figure 1.1 – Occurrence of transmission and drug resistance in malaria

The worldwide distribution of malaria transmission and resistance to anti-malarial drugs are displayed. **Panel A** displays the regions where malaria transmission occurs. The dark pink regions are those where transmission occurs, the light pink are regions with limited risk and the white regions have no malaria transmission. **Panel B** displays the regions where drug resistance to anti-malarial drugs occurs. The areas coloured in grey display chloroquine resistance and those circled are areas where resistance to other anti-malarials occurs. This figure was constructed from figures taken from the world health organisation website (Panel A; http://rbm.who.int/rbm/Attachment/20040921/SP-Malaria2004_1.pdf) and the centre for disease control and prevention website (Panel B; http://www.cdc.gov/ncidod/emergplan/box23.htm).

1.1.2 Life cycle

P. falciparum is an obligate intracellular protozoan parasite of the phylum Apicomplexa that are characterised by the presence of apical organelles called micronemes, rhoptries and dense granules (figure 1.2). These organelles are essential for the invasion of their host cells and the formation of the parasitophorous vacuole. The life cycle of *P. falciparum* consists of sexual and asexual stages (figure 1.3).

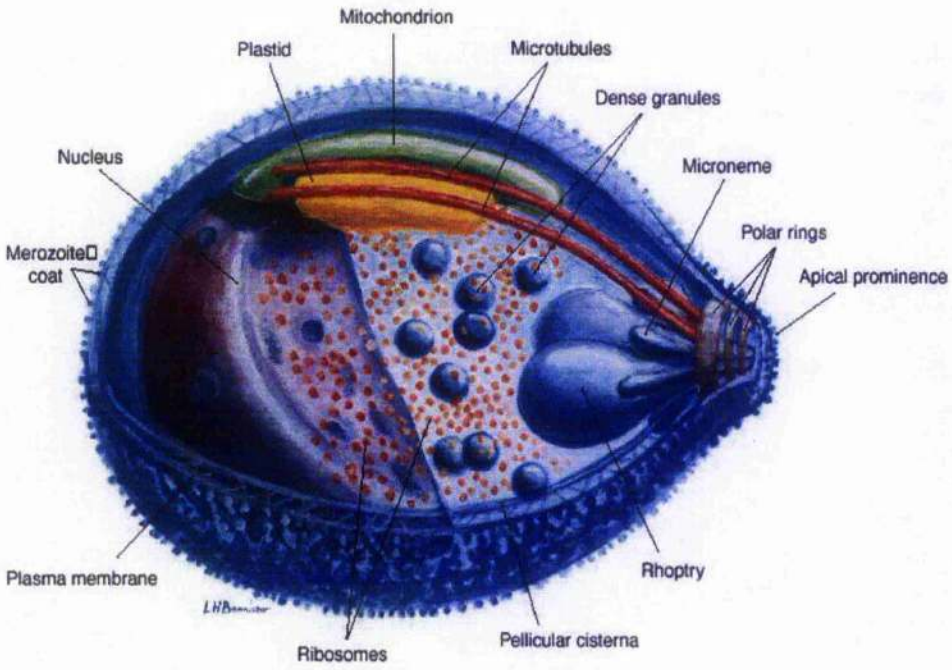


Figure 1.2 – The ultrastructure of a merozoite

This representation of the ultrastructure of a *Plasmodium falciparum* merozoite was based on analysis of serial electron microscope slides. The figure displays the apical organelles (Micronemes, rhoptries and dense granules) that are involved in invasion of the merozoite into the erythrocyte and the single mitochondrion and the plastid (apicoplast), which are usually observed to be closely associated . This figure was taken from Bannister *et al.* 2000.

The asexual stages of the life cycle (figure 1.3A) begin with the injection of infective sporozoites into the human by a female *Anopheles* mosquito as she takes a blood meal (Greenwood *et al.*, 2002). These sporozoites circulate in the bloodstream and may be cleared from the system by the spleen; however sporozoites that avoid the spleen firstly infect liver cells where they multiply to form the exo-erythrocytic schizont, which can contain up to 30,000 merozoites. Once the exo-erythrocytic schizont bursts it releases the merozoites into the bloodstream and these are then able to infect the host's erythrocytes/red blood cells (RBC). Invasion of the RBC by the merozoite is coordinated by the apical organelles of the merozoite (figure 1.2), and results in a parasite that is surrounded by what is termed the parasitophorous vacuole membrane

(PVM) that produces a barrier between the parasite and the host cell. The asexual intra-erythrocytic stages of parasite infection are responsible for the symptoms associated with malaria infection including lactic acidosis and fever. During the 48 hour erythrocytic cycle of *P. falciparum* the parasites roughly undergo three stages of development; the ring stage, the trophozoite stage and the schizont stage. During the ring stage of infection, which occurs 0 – 15 hours following erythrocytic invasion, the parasites are considered to be in G_0 phase and are relatively metabolically inactive. The trophozoite stage of infection occurs between 15 and 28 hours after invasion and it is during this phase that the parasite ingests the host cells cytosol in order to extract the nutrients it requires to develop. During this ingestion the parasite is required to detoxify the haem, which is released upon haemoglobin digestion. This is achieved by a process of haem biomineralisation, which results in an inert crystalline material called 'haemozoin'. This digestion and crystallization occurs in an acidic vacuole called the 'food vacuole' (Egan *et al.*, 2002). In the schizont stage of infection, which occurs between 28 and 48 hours after invasion, the parasites begin to replicate their DNA by mitotic divisions – this process is called 'schizogony'. The parasite replicates to produce up to 16 merozoites that can go on to invade further erythrocytes (figure 1.3). It is this cyclic bursting of erythrocytes that causes the anaemia and cyclical fever associated with malaria infections. The asexual stage can continue for many cycles, although some merozoites will undergo gametocytogenesis which is thought to be triggered by innate and host defence factors and can take 9 – 12 days to be completed (Sinden, 1998). The gametocytes that are formed are infective to the mosquito vector where the later stages of sexual development of the parasites takes place (figure 1.3B). When gametocytes are taken up in a blood meal from an infected individual, they differentiate into male and females gametes. Fertilisation occurs within the gut of the mosquito producing a zygote, which develops into the so-called ookinete. The mobile ookinete traverses the mosquito's midgut and migrates to the basal lamina, where it undergoes rounds of asexual replication to produce sporozoites. The sporozoites are released from the ookinete and enter the salivary glands, ready to be injected into the human host during the next blood meal taken by the insect vector, allowing the cycle of infection to continue between the insect and human host of the malaria parasite.

As it is the erythrocytic stage of infection that causes the primary symptoms of malaria, this is the stage that will be primarily considered in this study. The erythrocytic stage of infection with *P. falciparum* leads to severe clinical manifestations, which is because these infections are capable of producing cerebral malaria, along with the lactic acidosis, fever and anaemia observed in the other human malarial infections. Cerebral

malaria can be caused by two phenomena that are specific to *P. falciparum* infections, which are called sequestration and rosetting (White, 1998; Miller *et al.*, 2002).

Sequestration is caused by the attachment of infected RBC to the vascular endothelium and is caused by the expression of a parasite encoded protein, called erythrocyte membrane protein 1 (EMP-1) on the surface of the RBC (Kriek *et al.*, 2003). It is anchored to the membrane via association with another parasite-encoded protein called knob associated histidine rich protein, which accumulates under the RBC membrane in structures that are called 'knobs'. EMP-1 attaches to the vascular endothelium by interacting with receptor proteins on the cell surface, such as CD36 and intracellular adhesion molecule (ICAM) (Kyes *et al.*, 2001). Sequestration occurs in the later stages of the erythrocytic stage and results in the removal of these later stages from the bloodstream (Miller *et al.*, 2002), and decreases the likelihood of infected RBC being cleared by the spleen. Sequestration can occur in many organs in the body (i.e. heart, liver and kidneys), but it is primarily a problem if it occurs in the capillaries of the brain where it can cause blockage and lead to cerebral malaria. Roseting can also cause cerebral malaria and is produced by the association of an infected RBC with uninfected RBC (Wahlgren *et al.*, 1998).

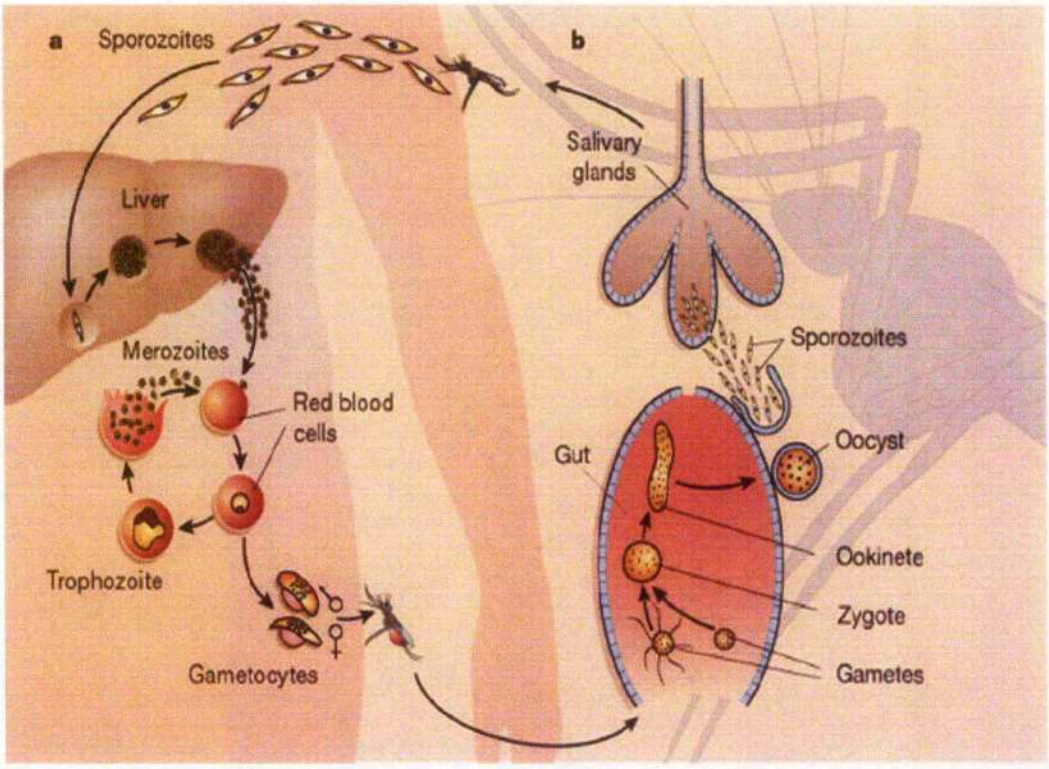


Figure 1.3 – The life cycle of *Plasmodium falciparum*

The life cycle of *Plasmodium falciparum* consists of both sexual and asexual stages. Panel A displays the asexual stage of infection, which occurs in the human host. The sporozoites are injected into human tissue by the *Anopheles* mosquito, where they invade hepatocytes. The parasites replicate to produce up to 30,000 merozoites, which burst from the liver into the bloodstream before infecting erythrocytes. The erythrocytic stage of the infection can cycle many times as parasites replicate and re-infect the red blood cells. Parasites can also undergo gametocytogenesis to produce gametocytes, which are the entry point for the sexual stage of infection. Panel B displays the sexual stage of infection, which occurs in the Anopheline mosquito. The gametes fuse in the mosquitos midgut to produce a zygote which develops into an ookinete. The ookinete traverses the midgut wall and becomes an oocyst in which sporozoites develop before entering the salivary glands ready to infect the human host again. This figure was taken from Wirth, D.F., 2002.

1.1.3 Chemotherapies

As malarial infections are so widespread (highlighted pink in figure 1.1A) and cause so many infections and deaths globally it is necessary to have efficient anti-malarial chemotherapies (in order to keep the effect of the disease as low as possible). Chemotherapies that have been/are being used extensively include chloroquine, mefloquine, antifolate drugs and artemisinin, although drug resistant strains to many of these are now widespread (figure 1.1B).

1.1.3.1 Current chemotherapies

Chloroquine has been the most widespread anti-malarial drug used over the last 50 years and this has led to the emergence of widespread chloroquine resistance in malarial infections (White, 1996). Over the years a number of different mechanisms of action of this drug have been suggested. Currently the most favoured mode of chloroquine action is the blocking of degradation and biomineralisation of haemoglobin and haem, respectively (Egan *et al.*, 1997; Leed *et al.*, 2002). The resulting increase in intracellular levels of toxic haem results in increased oxidative stress and membrane damage, which ultimately causes parasite death (Loria *et al.*, 1999). Chloroquine resistance is associated with lower levels of the drug in the parasite's food vacuole, which is thought to be mediated by an increased efflux of the drug from the parasites via a multi-drug resistance transporter (Volkman *et al.*, 1998). A second mechanism that clearly is associated with chloroquine resistance not only in laboratory strains but also clinical isolates are mutations in the so-called chloroquine resistance transporter gene (Mu *et al.*, 2003). The precise mechanism as to how this transporter confers resistance is not yet conclusively solved. Another theory is that chloroquine resistance correlates with increased levels of the tripeptide glutathione (Ginsburg *et al.*, 2003). The basis of this correlation is that free haem is detoxified by glutathione and that this process is inhibited by chloroquine thus arguing that increased levels of glutathione should lead to a protection of the parasites against the toxic action of free haem (Atamna *et al.*, 1995; Ginsburg *et al.*, 2003).

A second quinine derivative is mefloquine and this drug has been in use since the late 1970s (Davidson *et al.*, 1975; Trenholme *et al.*, 1975). Like chloroquine, mefloquine resistance is rising in areas of Asia and Mauritius and the drug also causes numerous side effects such as nausea, vomiting and dizziness.

Antifolate drugs such as pyrimethamine and sulphonamides have been used successfully against *Plasmodium* species. However, resistances to these compounds are now widespread due to point mutations in both the dihydropterate synthase and dihydrofolate reductase (dehydrogenase) genes (Peterson *et al.*, 1988; Reeder *et al.*, 1996).

Artemisinin is a naturally occurring compound, which can be extracted from the shrub *Artemesia annua* and has been used in Chinese medicine for millennia. Unfortunately, artemisinin is an unstable compound (Dong *et al.*, 2003) and many artemisinin derivatives, such as artemether and artesunate are now being used as treatments for malaria. Artemisinin is thought to act in a number of ways, like chloroquine it can inhibit production of haemozoin by alkylating haem but it also leads to alkylation of parasite proteins (Meshnick, 1998). It has been shown to act on the serco/endoplamic reticulum Ca^{2+} ATPase (Eckstein-Ludwig *et al.*, 2003) and is also redox active and can produce reactive oxygen species that create oxidative stress and damage the cell (Dong *et al.*, 2003). Artemisinin based combination therapies (ACT) have now been introduced by the World Health Organisation for treatment of malaria due to the increasing prevalence of drug resistance to the original drugs of choice.

Other antimalarial treatments include the use of insecticide treated mosquito nets that can reduce transmission of the disease (World health organisation; http://www.rbm.who.int/cmc_upload/0/000/015/368/RBMInfosheet_5.htm). These have been extremely successful, but are preventative cures that are logistically difficult and expensive to implement. It is because of this emerging drug resistance and the relative expense of the ACT treatments that the search for new antimalarial drugs must continue.

1.2 Redox and antioxidant systems

During their intra-erythrocytic development *Plasmodium* parasites have to survive in a pro-oxidant environment that contains both oxygen and iron. Reactive oxygen species are produced by the parasites themselves when they degrade haemoglobin to haem (FP IX) in the food vacuole. The oxidation of haemoglobin results in the formation of superoxide anions ($O_2^{\bullet-}$), which can spontaneously form hydrogen peroxide and oxygen. This reaction generates the substrates for the formation of the highly active and extremely toxic hydroxyl radicals. In addition, the presence of iron favours the Fenton reaction, which can result in the formation of hydroxyl radicals (figure 1.4). FP IX that

is produced from the degradation of haemoglobin can be polymerised into an inert substance called 'haemozoin', although only between 50 % and 90 % of FP IX produced is stored as haemozoin, suggesting that the remaining 10 % to 50 % requires to be degraded in order to stop the production of reactive oxygen species (Atamna *et al.*, 1993; Loria *et al.*, 1999; Zhang *et al.*, 1999; Egan *et al.*, 2002; Becker *et al.*, 2004b).

In order to cope with these reactive oxygen species the parasites have evolved highly efficient antioxidant and redox mechanisms including superoxide dismutases (SOD), the thioredoxin and the glutathione systems (figure 1.4 and figure 1.5) (Krnajski *et al.*, 2001; Becker *et al.*, 2003; Sienkiewicz *et al.*, 2004). However, the parasites lack catalase and a glutathione dependent peroxidase, both of which are important antioxidants in other organisms. Therefore it is currently believed that the thioredoxin system acts as the major antioxidant systems in the parasites whereas glutathione acts as major thiol redox buffer and acts as a co-factor for enzymes like glutathione S-transferase and glyoxalase (Harwaldt *et al.*, 2002; Akoachere *et al.*, 2005).

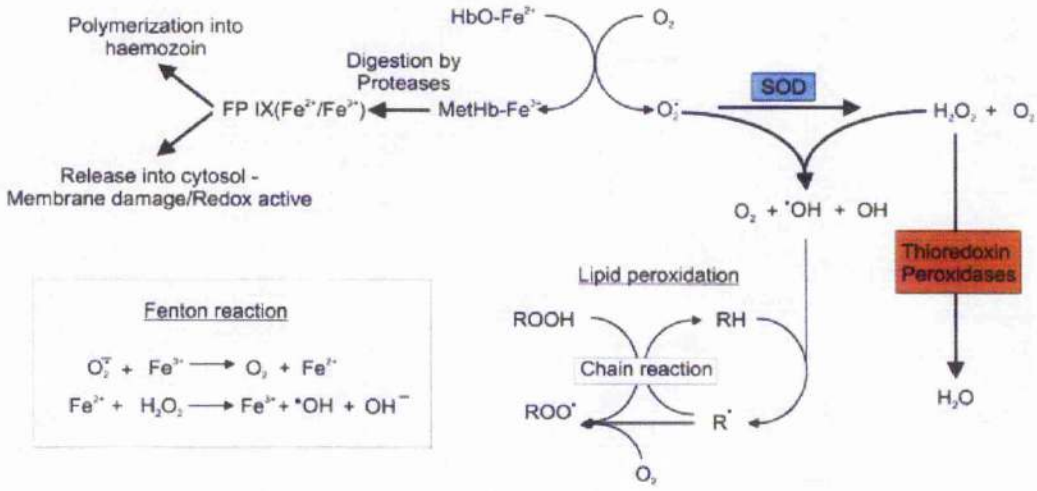


Figure 1.4 – Production of reactive oxygen species in *P. falciparum*

This figure displays the production of reactive oxygen species in *P. falciparum* as a consequence of haemoglobin degradation and biomineralisation in the food vacuole. Haemoglobin (HbO-Fe^{2+}) is degraded by proteases releasing free haem (FP IX), which is then mostly polymerised into haemozoin in the food vacuole. However, up to 50 % of this free haem is released into the cytosol where it can produce reactive oxygen species that cause membrane damage. Hydroxyl radicals can form in the presence of free haem via the Fenton reaction. Superoxide radicals (O_2^-) that are produced by the oxidation of haemoglobin are dismutated to hydrogen peroxide (H_2O_2) by superoxide dismutase (SOD) or can react with H_2O_2 to spontaneously form hydroxyl radicals (OH and ȮH). Hydrogen peroxide is detoxified by thioredoxin peroxidases that form part of the thioredoxin system. Figure was taken from Müller, S., 2004.

1.2.1 Thioredoxin system

As outlined above, it has been suggested that the thioredoxin system is of particular importance in the parasites antioxidant homeostasis (Becker *et al.*, 2004; Becker *et al.*, 2004; Muller, 2004). The thioredoxin system comprises the NADPH-dependent disulphide oxidoreductase thioredoxin reductase (TrxR) and the ~ 12 kDa redox active protein thioredoxin (Trx). Trx acts as a redox messenger in the cell interacting with a wide variety of protein and peptide substrates such as transcription factors, ribonucleotide reductase, methionine sulphoxide reductase, peroxiredoxins and glutathione, and therefore plays an important role in cell proliferation, protection against reactive oxygen species and signal transduction (Powis *et al.*, 2001; Nakamoto *et al.*, 2004; Baier *et al.*, 2005; Kabe *et al.*, 2005).

It is likely that the thioredoxin system has essential roles for *P. falciparum* survival as the knockout of the *TrxR* gene has a lethal effect on the parasites (Krnajski *et al.*, 2002). This supports the finding that there is a single *TrxR* gene present in the parasite genome and supports the suggestion that the functions of the thioredoxin redox cycle cannot be replaced by another system such as the glutathione redox cycle as has been shown in other organisms. The parasites also contain only a single gene encoding a conventional thioredoxin with a number of genes encoding for thioredoxin-like proteins whose function have not yet been established (Muller, 2004). Given that the peroxiredoxin-linked detoxification seems to be the major avenue for the parasites to detoxify ROS, it was expected that these proteins would be of crucial importance for parasite survival. However, there appears to be some redundancy between the different peroxiredoxins identified in the parasites, as the knockout of the cytosolic 2-Cys peroxiredoxin PfTrx-Px-1 merely renders parasites hypersensitive towards oxidative stress but does not have an effect on their viability (Komaki-Yasuda *et al.*, 2003). This emphasises that it is important to further dissect the precise roles of the different peroxiredoxins to be able to assign specific functions of these proteins that are key for parasite survival and thus are potentially suitable for the design of inhibitors that might be developed into new antimalarials.

Both the thioredoxin and glutathione systems play extremely important but distinct roles in *P. falciparum*, with the reducing power for both of these redox cascades provided by their respective disulphide oxidoreductases, TrxR and glutathione reductase (GR) and it is those that I will introduce in the next section.

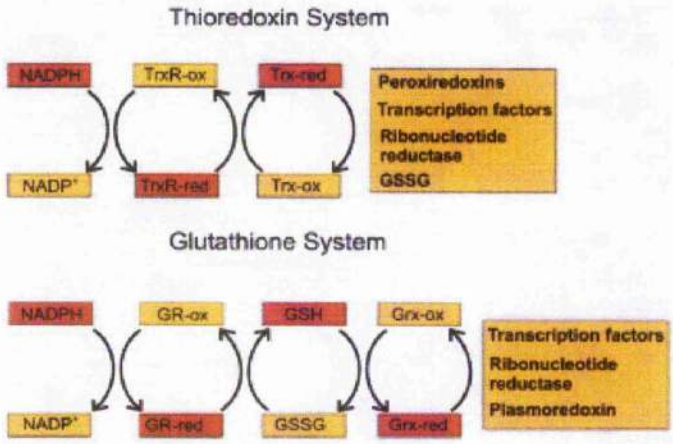


Figure 1.5 – The thioredoxin and glutathione systems of *P. falciparum*

The thioredoxin and glutathione redox systems both consist of a number of redox active proteins/peptides that produce a cascade of electrons from NADPH to acceptor molecules. Both systems contain common receptors such as ribonucleotide reductase and transcription factors, although each system also contains its own specific components as well. This figure displays the thioredoxin and glutathione cascades with the oxidised forms (in yellow) and reduced forms (in orange) of the redox proteins/peptides. The **thioredoxin system** transfers reducing equivalents from NADPH via thioredoxin reductase (TrxR) to thioredoxin (Trx) before finally reducing peroxiredoxins, glutathione (GSSG) and the common acceptor molecules. The **glutathione system** transfers reducing equivalents from glutathione reductase (GR) to glutathione (GSSG in the oxidised form, GSH in the reduced form). Reducing equivalents can then be passed onto glutaredoxin (Grx) before reduction of plasmoredoxin and the common acceptor molecules. Figure was taken from Müller, S., 2004.

1.3 Disulphide oxidoreductases

The family of flavoprotein disulphide oxidoreductases includes glutathione reductase (GR), dihydrolipoamide dehydrogenase (LipDH), thioredoxin reductase (TrxR), trypanothione reductase (TyrR) and mercuric ion reductase (MR) (Argyrou *et al.*, 2004). The proteins of this enzyme family have a high degree of amino acid sequence identity and they have very similar tertiary structures (Williams *et al.*, 1982; Hirt *et al.*, 2002; Argyrou *et al.*, 2004). All members of this protein family are characterised by the presence of a tightly bound flavin co-factor (FAD; figure 1.9), the formation of homodimers and the presence of at least one redox active disulphide. They can however be separated due to the presence of different amino acid domains that have distinct roles in catalysis (figure 1.6). The four most common domains are the FAD binding domain, the NADH/NADPH binding domain, the central domain and the interface domain.

The FAD binding domain and NADH/NADPH (figure 1.6) binding domains contain the heptapeptide binding motif (usually but not always GXGXXG) for binding of their respective co-factor or substrate, with the FAD and the NADH/NADPH being bound to

the proteins in an extended conformation (Pai *et al.*, 1983; Mattevi *et al.*, 1992; Lantwin *et al.*, 1994). The FAD binding domain also usually contains the redox active cysteine motif (CXXXC), although this motif is found in the NADPH domain of the low M_r TrxR (see section 1.3.1.2 and figure 1.6). The central domain contains residues that are important in the binding of the FAD cofactor and also form part of the active site. The interface domain is responsible for dimer formation and contains the acid/base catalyst motif (HPXXXE), which is located in the active site of the homodimeric protein. Two active sites form in the homodimer with each active site containing residues from the FAD binding, NAD(P)H binding and central domains of one subunit with the interface domain of the other. The interface domain of some disulphide oxidoreductase contains a C-terminal extension that possesses a second non-flavin redox centre (figure 1.6B; the extension of high M_r TrxR and the organism specific differences observed will be discussed in section 1.3.1.1).

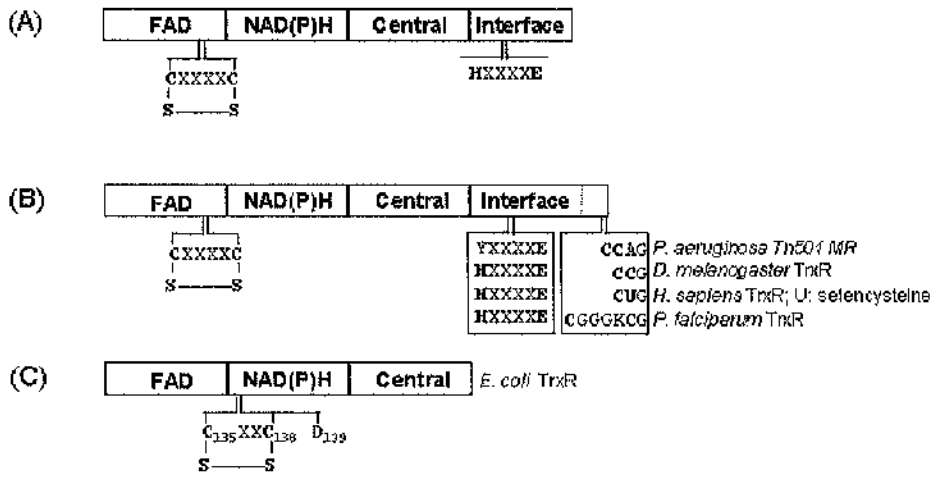


Figure 1.6 – Overview of the domain structure of disulphide oxidoreductases

This figure demonstrates the domain structure of the disulphide oxidoreductases, which can be divided into a flavin cofactor binding domain (FAD), pyridine nucleotide binding domain (NAD(P)H), central domain (Central) and interface domain (Interface). Residues important for catalysis including disulphide forming cysteines and the acid/base catalyst motif are labelled below the domains. **Panel A** displays the domain structure for glutathione reductase and dihydrolipoamide dehydrogenase. **Panel B** displays the domain structure for mercuric ion reductase and high M_r thioredoxin reductase and demonstrates the organism specific differences observed at the C-terminal extension. **Panel C** display the domain structure for low M_r thioredoxin reductase. This figure was modified from Argyrou *et al.* 2004.

Analysis of the tertiary structures of disulphide oxidoreductases also reveals a similarity in the organisation of the active site (except for in the low Mr TrxR; see section 1.3.1.2), with the active site of GR being shown as an example in figure 1.7. The NADH/NADPH binding site is located on the *re* side of the isoalloxazine ring of the bound FAD cofactor. The region in which the nicotinamide ring of NADH/NADPH is located is occupied by an aromatic amino acid (which is usually a tyrosine; Tyr197 in figure 1.7) in the oxidised protein and this is rotated out of position following binding of the pyridine nucleotide (Pai *et al.*, 1983; Mattevi *et al.*, 1992; Biterova *et al.*, 2005). Reducing equivalents are passed from the reduced pyridine nucleotide (on the *re* side) to the flavin before being transferred to the redox active cysteine pair located on the *si* side of the isoalloxazine ring of FAD (Cys58 and Cys63 in figure 1.7). The residues of the acid/base catalyst motif (His467' and Glu472' in figure 1.7) are ideally located close to the N-terminal cysteine pair to promote the transfer of reducing equivalents through the active site to the substrate (or to the C-terminal redox active site). This transfer of reducing equivalents was initially suggested based on the crystal structure of human GR by Pai *et al.* and is now the accepted catalytic mechanism for GR and other disulphide oxidoreductases (Pai *et al.*, 1983). The flow of reducing equivalents in this mechanism and the presence of two distinct substrate binding sites suggest that disulphide oxidoreductases act via a ping-pong mechanism. For disulphide oxidoreductases, this is where one substrate binds, reduces the protein and is released before the second substrate binds and oxidises the protein. Studies into LipDH (Argyrou *et al.*, 2001), GR (Williams, 1992), TryR (Leichus *et al.*, 1992) and MR (Patel *et al.*, 1999) have confirmed this, although GR and LipDH at high substrate concentrations (GSSG for GR and NADH for LipDH) display sequential reactions (Mannervik, 1973; Tsai, 1980; Rakauskiene *et al.*, 1989; Tsai *et al.*, 1992).

Analyses of the spectral intermediates formed during catalysis have confirmed the proposed catalytic mechanism. Using pre-steady state kinetic analyses it is possible to identify the spectral intermediates formed during catalysis; the reaction of *Mycobacterium tuberculosis* LipDH is shown as an example (figure 1.8; Argyrou *et al.*, 2002; Argyrou *et al.*, 2004). Following binding the pyridine nucleotide reduces the flavin and forms a $FADH^-$ NAD^+ charge transfer complex ($FADH_2$ NAD^+ box of figure 1.8) that can be observed spectrally as absorbance at longer wavelengths (600nm – 700 nm in figure 1.8). Reducing equivalents are passed via the flavin (a fully reduced flavin can be observed in the main panel and in the EH_4 box of figure 1.8) to the N-terminal cysteines producing the EH_2 species (EH_2 box of figure 1.8) that forms a charge transfer

complex (CTC) between the flavin interacting thiol and the FAD (Cys63 in figure 1.7), which can be observed spectrally as a band centred around 500nm – 550nm (figure 1.8).

Several studies have investigated the role of the acid/base catalyst motif in the catalysis of disulphide oxidoreductases, although these have been primarily in GR and LipDH. Mutations of the conserved histidine, proline and glutamate residues lead to severe impairment of enzymatic activity (Deonarain *et al.*, 1989; Kim *et al.*, 1992; Gilberger *et al.*, 1997). The conserved histidine residue (His467' in figure 1.7) acts as the acid/base catalyst by acting as a proton donor/acceptor in the dithiol/disulphide transfer reactions that occur during catalysis (Bauer *et al.*, 2003b). The proline residue is in the *cis*-conformation and acts to orientate the histidine residue so that the NE2 of the imidazole ring is positioned near the interchange thiol of the N-terminal cysteine pair (Cys58 in figure 1.7). The conserved glutamate residue (Glu472' in figure 1.7) has been shown to form a hydrogen bond with the conserved histidine residue between the epsilon carboxyl of glutamic acid and the ND1 of histidine (Pai *et al.*, 1983). The role of the conserved glutamate has also been studied by molecular dynamics and density function calculations, which predict its importance in making the histidine residue a better catalyst (Brandt *et al.*, 2005).

The following sections further introduce thioredoxin reductase and dihydrolipoamide dehydrogenase, as these are the members of the disulphide oxidoreductase family that will be investigated in this study.

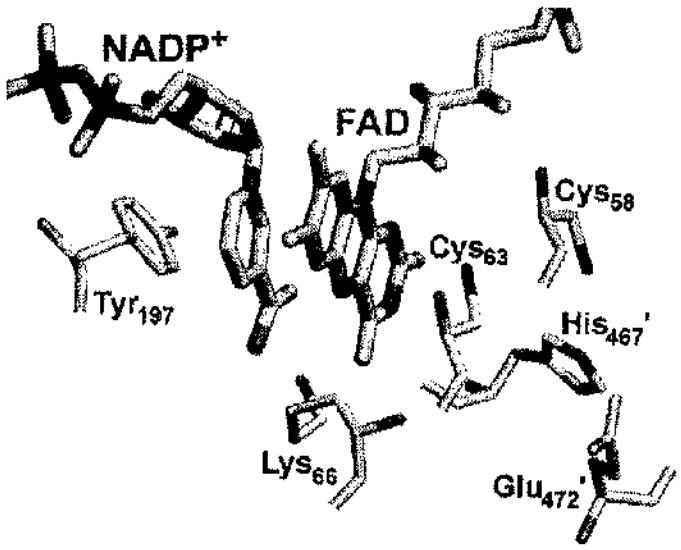


Figure 1.7 – The active site of disulphide oxidoreductases

The active site of human GR is displayed as an example of the organisation of the pyridine nucleotide ($NADP^+$) and FAD-binding and the positioning of the N-terminal redox active cysteine residues (Cys58 and Cys63) and the residues of the acid/base catalyst motif (His467' and Glu472'). The tyrosine residue that shields the *re* side isoalloxazine ring in the oxidised protein is shown rotated outwards due to pyridine nucleotide binding, while the nicotinamide ring and isoalloxazine rings are stacked against each other. This figure was taken from Argyrou *et al.* 2004.

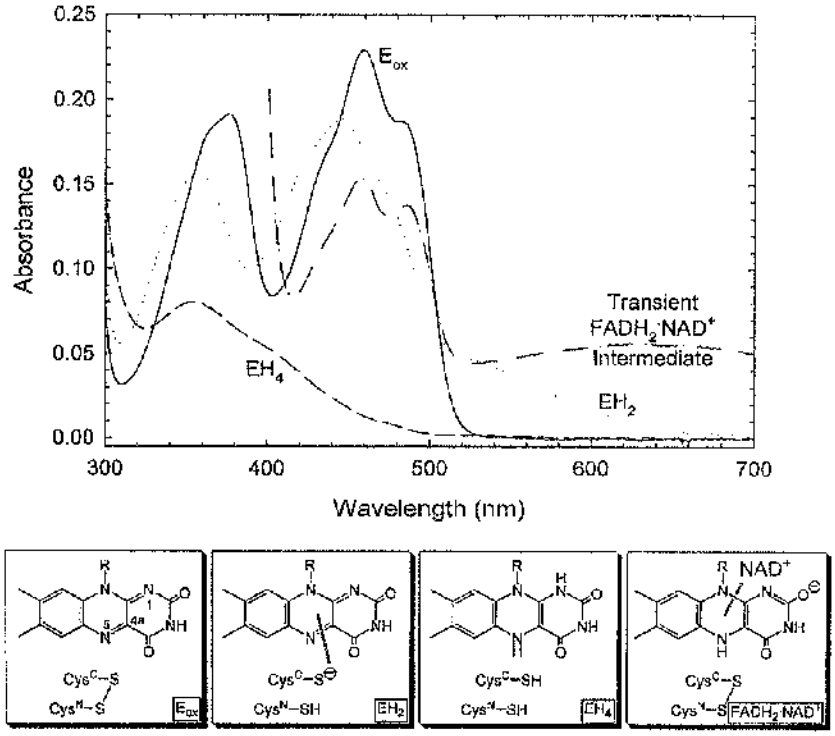


Figure 1.8 – Spectral intermediates of LipDH during catalysis

The intermediates formed during catalysis of *Mycobacterium tuberculosis* LipDH are shown, as an example of those observed in all disulphide oxidoreductases (although high Mr TrxR are complicated by the presence of a second non-flavin redox active centre). The **Main panel** displays the spectra of the catalytic intermediates E_{ox}, $FADH_2^{\cdot}NAD^+$, EH_2 and EH_4. The formation of $FADH_2^{\cdot}NAD^+$ charge transfer complex can be observed as absorbance at longer wavelengths (600nm – 700 nm) and the formation of a flavin thiolate charge transfer complex can be observed as absorbance between 500nm –550 nm. **E_{ox} box** displays the isoalloxazine ring and N-terminal cysteine pair in the oxidised state. **EH_2 box** displays the thiolate flavin charge transfer intermediate during the two electron reduced intermediate. **EH_4 box** displays the full flavin and N-terminal cysteine reduction in the four electron reduced intermediate. **$FADH_2^{\cdot}$ NAD^+ box** displays the charge transfer complex intermediate formed following binding of NADH to the protein. This figure has been taken from Argyrou *et al.* 2004.

1.3.1 Thioredoxin reductase (TrxR)

Thioredoxin reductases catalyse the NADPH-dependent reduction of the 12 kDa thioredoxin protein, and therefore provides the reducing power for the thioredoxin redox cycle (figure 1.5, section 1.2.1). TrxR occur in two distinct forms in nature, the high and low M_r TrxRs that are distinguished by molecular weight (M_r), domain structure and enzyme mechanism as outlined below.

1.3.1.1 High M_r Thioredoxin reductase

The high M_r TrxRs have a subunit molecular size of ~55 kDa. They possess a similar domain structure as GR – they have a protein bound FAD co-factor, an N-terminal redox active pair of cysteines (which is equivalent to GR) and they use NADPH similar to GR (figure 1.6B and figure 1.9). The high Mr TrxR are found in mammals, insects, worms and *Plasmodium falciparum* (Gladyshev *et al.*, 1996; Gilberger *et al.*, 1997; Gladyshev *et al.*, 1999; Williams *et al.*, 2000) and it has been shown by phylogenetic analyses that they are clearly related to GR (Hirt *et al.*, 2002). However, in addition to the two redox active centres that GR has, they contain a third redox active centre which resides on a C-terminal extension of 15 to 20 amino acids (relative to GR; figure 1.6B). The primary structure and nature of this redox active centre is organism specific and can either consist of a disulphide or a selenenylsulphide. Mammalian proteins contain a selenenylsulphide that forms between adjacent selenocysteine and cysteine residues (Gladyshev *et al.*, 1999; Lee *et al.*, 2000), while in insect TrxR a disulphide forms between adjacent cysteine residues (Bauer *et al.*, 2003b) and in *P. falciparum* a disulphide forms between two cysteine residues separated by four amino acids (figure 1.6B) (Gilberger *et al.*, 1998; Wang *et al.*, 1999).

The presence of a selenocysteine residue at the C-terminal redox active site of mammalian TrxR is an interesting observation. Not only is the selenocysteine residue much more reactive than its cysteine homologue, but the extra bond length (15 % greater in comparison to cysteine) allows for the formation of the selenenylsulphide bond between the adjacent selenocysteine and cysteine residues. This additional bond length is in contrast to the situation in insect TrxRs and other proteins, where a disulphide bond that forms from the adjacent cysteine residues is strained (Schulz *et al.*, 1990; Wang *et al.*, 2000). Indeed, replacement of the selenocysteine with cysteine in the human TrxR (producing a GCCG motif) resulted in an almost inactive protein (Lee *et al.*, 2000; Zhong *et al.*, 2000). Replacement of the human TrxR GCUG motif (where U

corresponds to selenocysteine) with that of the *Drosophila melanogaster* TrxR (SCCS), resulted in an even lower activity than that observed for the GCCG motif (L. Johansson, unpublished data), which suggested that the local environment around the C-terminus was important for catalysis.

An in depth study of the *Drosophila melanogaster* TrxR has led to the proposal of the catalytic mechanism of not only DmTrxR but of all high M_r TrxR (Bauer *et al.*, 2003b). By studying the spectral intermediates of the reductive half reaction (anaerobic reduction with NADPH/ sodium borohydride) and the oxidative half reactions (oxidation of the reduced protein by thioredoxin) of DmTrxR, Bauer *et al.* were able to propose the protein's catalytic mechanism based on observations for the related protein GR (Bauer *et al.*, 2003b). Bauer *et al.* proposed that the protein was unable to be fully reduced (i.e. all three redox centres reduced to form an EH_6 species) by natural substrates under physiological conditions, but that the high M_r TrxR cycled in catalysis between a two electron reduced form (EH_2) and a four electron reduced form (EH_4) (Bauer *et al.*, 2003b).

The *P. falciparum* TrxR (PfTrxR) has been studied previously in some detail and these studies have shown that the three redox active centres (FAD, N-terminal disulphide (C88/C93) and C-terminal disulphide (C535/C540) are in redox communication, which guarantees transfer of electrons through the enzymes active site and onto the protein substrate thioredoxin disulphide (Gilberger *et al.*, 1997; Gilberger *et al.*, 1998; Wang *et al.*, 1999). The individual roles of each redox active cysteine residue have been investigated by mutagenesis, and have identified that Cys93 interacts with the flavin while Cys88 is the interchange thiol that forms a disulphide with one of the C-terminal cysteines (Gilberger *et al.*, 1997; Gilberger *et al.*, 1998; Wang *et al.*, 1999). The C-terminal cysteine that forms the disulphide has been suggested to be Cys540, although this has not been categorically proven and both C-terminal cysteines can accept reducing equivalents from the N-terminal thiols (Wang *et al.*, 1999).

In chapter 3 of this study I will further investigate the catalytic mechanism of PfTrxR by studying its reductive and oxidative half reactions and will concentrate primarily on the role that the acid/base catalyst motif plays in the transfer of reducing equivalents between the redox active centres. The further information gained in this study may be an important aid in rational drug design against the PfTrxR protein, which has previously been validated as a potential drug target as gene knockout studies have proved that the protein is essential for parasite survival (Krnajski *et al.*, 2002).

Therefore, any targeted inhibition of the *P. falciparum* thioredoxin reductase (and/or the thioredoxin system) would be a potential candidate for an antimalarial drug. The C-terminal disulphide has been proposed as a possible site of organism specific inactivation due to the organism specific differences observed there. Indeed, this has been shown to be the site of inactivation observed in *Pf*TrxR inhibited with mannich base compounds (Davioud-Charvet *et al.*, 2003) and nitrophenyl compounds (Andricopulo *et al.*, 2006), and is therefore a plausible target for further studies in drug design.

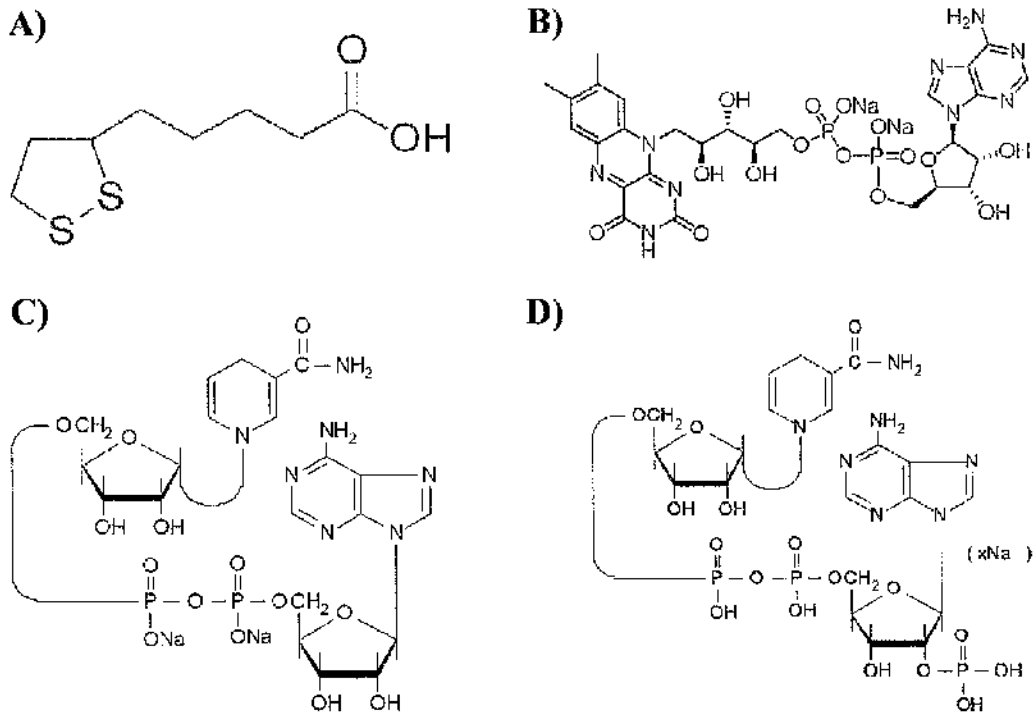


Figure 1.9 - Substrates and cofactors of disulphide oxidoreductases

This figure displays the redox active compounds that are involved in the catalytic reactions of LipDH (Panel A and C) and many other disulphide oxidoreductases (Panels B – D). **Panel A** displays oxidised lipoic acid, with its sulphur groups at the 6 and 8 positions. **Panel B** displays oxidised flavin adenine dinucleotide cofactor of disulphide oxidoreductases, with the picture drawn from the *si* side of the isoalloxazine ring. **Panel C** displays reduced nicotinamide adenine dinucleotide (NADH) substrate that is involved in LipDH catalysis. **Panel D** displays reduced nicotinamide adenine dinucleotide phosphate (NADPH) substrate that is involved in GR, TryR and TrxR catalysis. All structures were obtained from the SigmaAldrich website (www.sigmaaldrich.com).

1.3.1.2 Low M_r Thioredoxin reductase

The low M_r TrxR are ~ 35 kDa and are found in *Escherichia coli*, funghi, plants and the protozoan parasite *Trichomonas vaginalis* (Williams *et al.*, 2000; Hirt *et al.*, 2002; Coombs *et al.*, 2004; Gelhaye *et al.*, 2005). The low M_r TrxR possesses NADPH-binding, FAD-binding and central domains but lack the interface domain and C-terminal extension observed in the high M_r TrxR (Figure 1.6C). The redox active disulphide is present in a CXXC motif that is located in the NADPH binding domain, which is in contrast to other disulphide oxidoreductases that possess a CXXXXC motif in their FAD-binding domain (figure 1.6C). Analysis of the crystal structure from low M_r TrxR demonstrated that both the NADPH-binding domain and the redox active disulphide were on the *re* side of the FAD (Kuriyan *et al.*, 1991), which is unlike the other disulphide oxidoreductases where the NADPH-binding domain is on the *re* side and the redox active disulphide is on the *si* side (Pai *et al.*, 1983; Sandalova *et al.*, 2001). This arrangement requires a large conformational change to occur in the protein for Trx-reduction to occur as the redox active cysteines are buried next to the FAD. Lennon *et al.* demonstrated that a 67° rotation of the NADPH-binding domain relative to the FAD-binding domain resulted in the transfer of the dithiols to a position where it can interact with thioredoxin (Lennon *et al.*, 2000). As they lack the interface domain the low M_r TrxR do not possess the acid-base catalyst motif (HPXXXE) that is found in other disulphide oxidoreductase (Figure 1.6). Instead, it was determined that they utilised an aspartate residue as a acid/base catalyst during catalysis (Mulrooney *et al.*, 1994).

1.3.2 Dihydrolipoamide dehydrogenase (LipDH)

Dihydrolipoamide dehydrogenase (LipDH) is an essential component of the α – keto acid dehydrogenase complexes (KADH; see section 1.4) where it catalyses the oxidation of a dihydrolipoamide moiety (figure 1.10; section 1.4). This dihydrolipoamide moiety is produced from the covalent attachment of lipoic acid to a protein component of the KADH (see section 1.4). In *Plasmodium*, lipoic acid (figure 1.9) is produced by the introduction of two sulphurs at the 6 and 8 positions of a type II fatty acid biosynthesis intermediate called octanoyl-acyl carrier protein. This step is catalysed by a lipoic acid synthase (LipA), which is an iron-sulphur protein that is located in the plastid of *P. falciparum* (Thomsen-Zieger *et al.*, 2003; Wrenger *et al.*, 2004). The lipoic acid produced in the apicoplast can be attached to the KADH by the action of a lipoyl transferase (LipB) that is also located in the apicoplast (Thomsen-

Zieger *et al.*, 2003; Wrenger *et al.*, 2004). Interestingly, *P. falciparum* also contain a mitochondrially localised lipoate protein ligase (LplA) that can catalyse the attachment of salvaged lipoic acid to the KADH (Wrenger *et al.*, 2004). The presence of organelle specific lipoate transferase and ligase proteins suggest that *P. falciparum* may possess KADH in both the mitochondrion and the plastid. This would be reminiscent of the situation in plants and algae, where KADH are located in the mitochondrion and the plastid (Mooney *et al.*, 2002). As a component of the KADH, LipDH proteins are also required to be transported to the mitochondrion and plastids and as a result plastid-containing eukaryotes have developed two different LipDH isoforms (Lutziger *et al.*, 2000). The *P. falciparum* LipDH and components of the KADH will be described in chapter 4 (LipDH) and chapter 5 (KADH components) of this study.

LipDH shares the same domain structure as GR and TyrR, and contains a single redox active cysteine pair in a CXXXXC motif in the FAD domain and the acid-base catalyst motif in the interface domain of the protein (figure 1.6A). However, the protein displays substrate specificity for NADH (figure 1.9) rather than NADPH and although the catalytic mechanism is similar to that of GR, reducing equivalents run in the opposite direction to that of GR in the LipDH physiological reaction (Sahlman *et al.*, 1989). Reducing equivalents are passed from the reduced substrate dihydrolipoamide to the redox active cysteines on the *si* side of the isoalloxazine ring of FAD, then via the bound flavin cofactor to NAD^+ on the *re* side of the isoalloxazine ring of FAD. However, LipDH acts with a ping-pong reaction (similar to GR; see section 1.3) and the transfer of reducing equivalents can also be passed from NADH to lipoamide though the protein. An example of the spectral intermediates formed during catalysis of LipDH is shown in figure 1.8 and is discussed in section 1.3.

In addition to its role as a component of KADH, LipDH can also function as a possible antioxidant by reducing free lipoic acid. Lipoic acid (in free and protein bound forms) has been shown to interact with other redox active compounds, such as thioredoxin, glutathione, nitric oxide and ubiquinone (Jocelyn, 1967; Xia *et al.*, 2001; Bryk *et al.*, 2002; Igamberdiev *et al.*, 2004). LipDH has also been implicated as a potential pro-oxidant (Grinblat *et al.*, 1991; Sreider *et al.*, 1992) and as a result has been extensively studied as a potential drug target, especially in human parasites such as trypanosomatids (Krauth-Siegel *et al.*, 1995).

In humans, a genetic defect in LipDH activity produces a number of disorders due to the fact that the LipDH protein is shared between the mitochondrial KADH, and a

deficiency in the LipDH activity affects all of the complexes. Indeed, targeted gene disruption of the mouse LipDH was found to induce a prenatal mortality phenotype (Johnson *et al.*, 1997). LipDH deficiency can occur as a result of mutations in all of the domains of the LipDH protein, with its clinical manifestations being wide-ranging and including severe and progressive neurological disorders, lactic acidosis, and maple syrup urine disorder (Liu *et al.*, 1993; Shaag *et al.*, 1999; Cerna *et al.*, 2001; Grafakou *et al.*, 2003).

1.4 α–Keto acid dehydrogenase complexes (KADH)

α-Keto acid dehydrogenase complexes (KADH) are large multi-enzyme complexes that can be up to 10 MDa in size. These complexes are formed by multimers of three independent proteins, which are α-keto acid dehydrogenase (E1), acyl transferase (E2) and dihydrolipoamide dehydrogenase (E3/LipDH). There are three KADH that are located in the mitochondria of eukaryotes, the pyruvate dehydrogenase complex (PDH), the α – ketoglutarate dehydrogenase complex (KGDH) and the branched chain α – keto acid dehydrogenase complex (BCKDH) (Perham, 2000).

The α-keto acid dehydrogenase (E1) protein occurs as a homo-dimeric proteins of about 100 - 120 kDa or as hetero-tetrameric proteins that consist of two α and two β subunits ($\alpha_2\beta_2$) of between 35 kDa and 70 kDa. The homodimeric E1 protein displays little sequence similarity to the E1α or E1β subunits except in the thiamine pyrophosphate binding motif (Hawkins *et al.*, 1989; Matuda *et al.*, 1991). The E1α subunit is responsible for the binding of the thiamine pyrophosphate cofactor, while the E1β subunit is involved in protein binding interactions with the E2 protein. The E1 protein confers substrate specificity of the complexes, requires thiamine pyrophosphate and a magnesium ion as cofactors and is the rate-limiting step of the complexes.

The acyl-transferase protein (E2) is organised into three domains, the lipoyl domain, the subunit-binding domain and the catalytic domain that are separated by flexible linker regions. The lipoyl domain contains a lipoic acid prosthetic group that is covalently attached to an ε-amino group of a conserved lysine residue. E2 proteins can contain between 1 and 3 lipoyl domains depending on which complex they belong to (see sections 1.4.1, 1.4.2 and 1.4.3). The lipoyl domain aids in the orientation of the lipoamide moiety into a 20 Å deep pocket that is the site of oxidative decarboxylation in the E1 protein (Aevarsson *et al.*, 1999). The subunit-binding domain of E2 is

responsible for the interaction with E1 and E3 proteins and is the pivot point for the 'swinging arm' of the lipoyl domain(s). The catalytic domain catalyzes the transfer of the acyl group bound to the lipoamide cofactor of the lipoyl domain to coenzyme A, producing acyl CoA and dihydrolipoamide (figure 1.10). This part of the protein is also the region of self-association of E2 and is responsible for the generation of the E2 core – the E2 protein forms homo-trimers that then associate to form 24 mers (cubic organisation) or 60 mers (pentagonal dodecahedron organisation) (Perham, 1991). The flexible linker regions that are found between the three domains are rich in alanine and proline residues and are approximately 20 to 30 amino acids in length (Perham, 1991). This linker region results in a 'swinging arm' that contains the lipoyl domain(s) and is responsible for transferring catalytic intermediates through the complex. This swinging arm produces a substrate channelling effect as the lipoyl domain transfers the acyl group produced by the E1 protein to coenzyme A and then the resulting dihydrolipoamide to the LipDH protein to be oxidised (Aevarsson *et al.*, 1999; Mooney *et al.*, 2002).

The LipDH/E3 protein is a flavoprotein disulphide oxidoreductase belonging to the same protein family as TrxR. The homodimer consists of two identical subunits of approximately 55 kDa and its role within the KADH is the re-oxidation of the dihydrolipoamide after it has reacted with coenzyme A (section 1.3.1.3; figure 1.10). Generally, the complexes share a single LipDH (Bourguignon *et al.*, 1996) but in some organisms it was shown that specific LipDH iso-enzymes exist and are integral to a specific KADH (Lutziger *et al.*, 2000). This protein has been described in more detail in section 1.3.1.3.

Each protein in the KADH plays a distinct role in the catalytic mechanism of the KADH (see figure 1.10 for an overview of what is detailed below). The α-keto acid dehydrogenase (E1) is a decarboxylase that requires thiamine pyrophosphate as a cofactor and catalyses the oxidative decarboxylation of α-keto acids producing a hydroxyl-acyl thiamine diphosphate intermediate and CO_2. The acyl group of the intermediate is then transferred to the lipoamide prosthetic group that is attached to the E2 protein. This acyl group is subsequently transferred to coenzyme A, producing acyl CoA and dihydrolipoamide. Dihydrolipoamide (that is attached to the E2 protein) is then oxidised by the action of the E3 protein, allowing the lipoamide to accept another acyl group from the E1 protein.

KADH activity is regulated in a number of ways, which range from reversible phosphorylation, product inhibition, substrate levels and redox state sensing. The first

three methods will be dealt with in other sections (see sections 1.4.1, 1.4.2 and 1.4.3), but the latter will be discussed here. The redox state of a cell can be thought of being determined by the ratio of redox active compounds such as thiol/disulphides and NAD(P)H/NAD(P)$^+$, both of which are important factors in the activity of KADH. In addition to this the lipoamide that is covalently bound to the E2 protein of all KADH has been shown to be able to interact with other redox active molecules (Holmgren, 1979). This suggests that thioredoxin can interact with KADH and can play a role in regulating the activity of the complexes during periods of low NAD$^+$ concentrations (Bunik *et al.*, 1999). This also suggests that in addition to their physiological functions, the KADH may also be able to play a role in defence against reactive oxygen species (Bryk *et al.*, 2002).

The KADH play a number of important roles in cellular metabolism and I will introduce the biochemical roles and organisation of each of the KADH in the following sections.

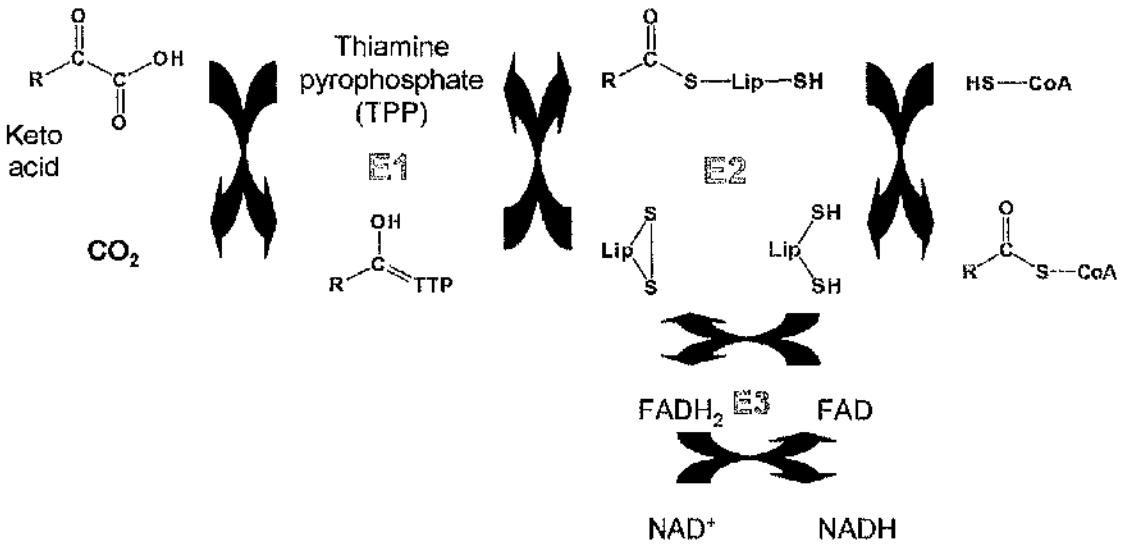


Figure 1.10 – The KADH catalytic mechanism

The catalytic mechanism of KADH involves the three proteins that form this complex; the α – keto acid dehydrogenase (E1), acyl transferase (E2) and dihydrolipoamide dehydrogenase (E3 or LipDH). The E1 protein is responsible for the oxidative decarboxylation of an α – keto acid using thiamine pyrophosphate (TPP) as a cofactor. This reaction produces hydroxyacyl-thiamine pyrophosphate and CO_2. The acyl group from hydroxyacyl-thiamine pyrophosphate is transferred onto the lipoamide moiety that is covalently attached to the E2 protein, before being transferred onto coenzyme A producing and acyl CoA and dihydrolipoamide. The dihydrolipoamide is oxidised to lipoamide by E3, allowing it to accept another acyl group from the E1 protein, with the reaction of E3 requiring a protein bound flavin cofactor (FAD) and producing NADH.

1.4.1 α–Ketoglutarate dehydrogenase complex

The α–ketoglutarate dehydrogenase complex (KGDH) is located in the mitochondria of eukaryotic cells, where it is an integral component of the tricarboxylic acid cycle (TCA cycle). KGDH acts on the α – ketoglutarate that is provided by isocitrate dehydrogenase and produces CO_2, succinyl-CoA and NADH. Succinyl-CoA is either fed into the next reaction of the TCA cycle or may be used for the biosynthesis of haem, whereas NADH is fed into the respiratory chain via complex I (NADH dehydrogenase).

The KGDH shares the multimeric structure of all KADH with some differences. The α – ketoglutarate dehydrogenase (E1 protein; EC 1.2.4.2) is over 100 kDa and forms a homodimer, which is in contrast to the $\alpha_2\beta_2$ hetero-tetramer found in BCKDH and PDH. The succinyl transferase (E2 protein; EC2.1.3.61) forms an octahedral core (24mer) that contains a single lipoyl domain. The dihydrolipoamide dehydrogenase is normally shared with the other mitochondrial KADH (Bourguignon *et al.*, 1996; Lutziger *et al.*, 2001).

The activity of KGDH is not regulated by the reversible phosphorylation of the E1 protein, as is the case for BCKDH and PDH. Any regulation of KGDH activity is therefore likely to occur by transcription/translation control or by competition for substrates such as α ketoglutarate, reduced coenzyme A (CoA SH) and the ratios of NAD^+/NADH and lipoamide/dihydrolipoamide (Dry *et al.*, 1987). The KGDH activity can be inhibited by high levels of products such as NADH, which induces product inhibition of the LipDH protein (Hamada *et al.*, 1975; Lawlis *et al.*, 1981). Indeed, loss of KGDH activity by inhibition/inactivation or by genetic defects can have extreme effects. Genetic defects in KGDH activity in humans has been implicated in the development of many neurodegenerative diseases including Alzheimers disease and Parkinsons disease (Gibson *et al.*, 2000).

1.4.2 Branched chain α–keto acid dehydrogenase complex

The branched chain α–keto acid dehydrogenase complex (BCKDH) is located in the mitochondria of eukaryotes and is involved in the degradation of the branched chain amino acids valine, leucine and isoleucine. The substrates for BCKDH are produced by branched chain amino acid transaminase (EC 2.6.1.42), which converts isoleucine, valine and leucine into α–keto acids (producing α–keto–β-methylvalerate, α-

ketoisovalerate and α–ketoisocaproic acid, respectively). BCKDH acts on these substrates producing acyl CoA products (α–methylbutyryl-CoA, isobutyryl-CoA and isovalyryl–CoA, respectively) that can eventually lead to the production of acetyl CoA and/or succinyl CoA that could be used within the TCA cycle (Ward *et al.*, 1999; Holecek, 2002).

The α–keto acid dehydrogenase protein (E1 protein; EC 1.2.4.4) of the BCKDH are composed of E1α and E1β subunits. The acyl transferase protein (E2; EC 2.3.1.168) forms an octahedral core (24mer) and contains a single lipoyl domain. The dihydrolipoamide dehydrogenase (E3/LipDH; EC 1.8.1.4) is usually shared between the mitochondrial complexes.

In contrast to the KGDH, the activity of the BCKDH is regulated by the reversible phosphorylation of the E1α protein. This is catalysed by branched chain α – keto acid dehydrogenase kinase (EC 2.7.1.115) and phosphatase (EC 3.1.3.52), with two sites of phosphorylation occurring at conserved serine residues (Odessey, 1982). In spite of the fact that it acts on serine residues the branched chain α – keto acid dehydrogenase kinase has been shown to be related to histidine kinases rather than the serine/threonine/tyrosine kinases (Popov *et al.*, 1992).

In humans, genetic defects that result in the loss of BCKDH activity produce an accumulation of branched chain α – keto acids in the body. This can lead to maple syrup urine disease, which produces severe clinical manifestations such as often fatal keto-acidosis and mental retardation (Chuang *et al.*, 2006).

1.4.3 Pyruvate dehydrogenase complex

The pyruvate dehydrogenase complex (PDH) utilises pyruvate, NAD^+ and Coenzyme A as substrates and produces acetyl CoA, NADH and CO_2. The pyruvate dehydrogenase protein (E1 protein; EC 1.2.4.1) can exist as either a homodimeric protein with a subunit molecular mass of 100 kDa or as a hetero-tetramer ($\alpha_2\beta_2$) of α and β subunits with molecular masses of between 35 kDa and 70 kDa. The acetyl transferase (E2 protein; EC 2.3.1.12) primarily forms a 60mer core in an icosahedral arrangement, and can contain between 1 and 3 lipoyl domains at its N-terminus. The dihydrolipoamide dehydrogenase protein (E3/LipDH; EC 1.8.1.4) can be shared between KADH or may be specific to the PDH, depending on the subcellular localisation of the PDH.

There are a number of different forms of PDH, which can be based on their subunit structure, organisation and cellular localisation.

The PDH of bacteria can be found in two forms depending on whether the bacteria are gram-positive or gram-negative. The PDH of gram-negative bacteria possess a pyruvate dehydrogenase protein (E1 protein; EC 1.2.4.1) that forms from a homo-dimeric protein with a molecular mass of 100 kDa (Stephens *et al.*, 1983). The acetyl CoA transferase protein (E2) forms an octahedral core (24mer) and contains three lipoyl domains at its N-terminus (Schulze *et al.*, 1992). The PDH of gram-positive bacteria has a PDH that has a hetero-tetrameric ($\alpha_2\beta_2$) pyruvate dehydrogenase protein (E1 protein; EC 1.2.4.1) and an acetyl transferase protein (E2) that forms an icosahedral core (60mer) with one lipoyl domain (Henderson *et al.*, 1979; Packman *et al.*, 1984; Borges *et al.*, 1990).

In eukaryotes the PDH is located in the mitochondria, where it acts as a link between glycolysis and the TCA cycle. The mitochondrial PDH (mtPDH) is similar in structure and organisation to the PDH of gram-positive bacteria, as its pyruvate dehydrogenase (E1 protein) is composed from E1α and E1β subunits and its acetyl transferase (E2) protein forms an icosahedral (60mer) core. However, the mtPDH E2 protein can contain between one and three lipoyl domains and the complex contains a fourth component called protein X. Protein X has a high degree of sequence homology and domain structure to the E2 protein, but lacks the catalytic activity (Neagle *et al.*, 1989; Harris *et al.*, 1997). It is involved in forming the complexes core together with the E2 protein and is the site of LipDH/E3 association, while the E2 protein binds the E1 protein (Maeng *et al.*, 1994; Sanderson *et al.*, 1996).

The activity of mtPDH is regulated by product inhibition, metabolic effectors and the action of a pyruvate dehydrogenase kinase (EC 2.7.1.99) and phosphatase (EC 3.1.3.43) on the E1α subunit (Gudi *et al.*, 1995; Huang *et al.*, 1998). The pyruvate dehydrogenase kinase is a serine specific kinase, but lack signature domains of the serine/threonine kinase, but instead displays greatest sequence similarity to histidine kinases (Popov *et al.*, 1993). The pyruvate dehydrogenase phosphatase (PDP) is a member of the mammalian PP2C class of phosphatases and occurs in two forms. The PDP1 is an $\alpha\beta$ heterodimer consisting of catalytic and regulatory subunits, while PDP2 is monomeric (Roche *et al.*, 2001).

Genetic defects in the human mtPDH are primarily caused by mutations in the E1α subunit of pyruvate dehydrogenase, although defects in all proteins of the complex

(including protein X) have been implicated (Ho *et al.*, 1989; Neagle *et al.*, 1989; Marsac *et al.*, 1993; Yang *et al.*, 1997). A deficiency in PDH activity is the most common cause of congenital lactic acidosis and other symptoms vary from mild ataxia to severe lesions of the central nervous system (Patel *et al.*, 1992; Robinson, 1995)

In plants there are two distinct PDH that are organelle-specific and are located in the mitochondria and the plastid (Mooney *et al.*, 2002). The plastid located PDH produces acetyl-CoA and NADH for fatty acid biosynthesis (Camp *et al.*, 1985).

The plastid localised PDH has a similar subunit structure as the mitochondrial PDH. The plastid PDH is composed of E1α and E1β subunits (Johnston *et al.*, 2000) and requires a plastid localised E3 protein (see section 1.3.2). No protein X component of the plastid PDH has been identified, although a second E2 gene was identified in *Arabidopsis thaliana*. This predicted protein displayed 60 % sequence identity to the functional E2 protein, although its function has not yet been determined (B.P. Mooney, unpublished data). The plastid PDH is thought to be arranged around an icosahedral core (60mer) as the M_r (4 – 7 MDa) of both the mtPDH and plastid PDH are similar (Camp *et al.*, 1985).

The plastid PDH is not regulated by reversible phosphorylation (as observed in the mitochondrial PDH and BCKDH) (Randall *et al.*, 1989), even though the E1α component of the plastid PDH contains one of the conserved serine residues required (Johnston *et al.*, 1997). Instead, the plastid PDH is thought to be controlled by substrate availability and product inhibition and is more sensitive to the NAD^+/NADH ratio than the acetyl CoA/CoA SH ratio (Camp *et al.*, 1988).

1.5 Organelles in *Plasmodium*

Plasmodium falciparum contain a number of organelles that play distinct and important roles in the life cycle of the parasite (see figure 1.2). The microneme, rhoptries and dense granules are called the apical organelles as they are located at the apical end of the merozoite and are involved in the invasion of the host erythrocyte. Other organelles are involved in the cellular metabolism of the parasite and these include the food vacuole, mitochondrion and apicoplast. The food vacuole has been discussed previously (see section 1.2) and the role of the mitochondrion and the apicoplast will be discussed in more depth in the following sections.

1.5.1 Apicoplast

The apicoplast is a non-photosynthetic plastid that is found in all *Plasmodium* species and most of the Apicomplexan parasites excluding for instance *Cryptosporidium* (Wilson *et al.*, 1994; Kohler *et al.*, 1997). It was first observed by microscopy (McFadden *et al.*, 1997) and has been shown to be surrounded by four membranes (Kohler *et al.*, 1997; McFadden *et al.*, 1999; Cavalier-Smith, 2000). All plastids are derived from the endosymbiosis of cyanobacteria into a eukaryotic cell, which would usually produce two outer membranes and is called primary endosymbiosis (Gray, 1992). However, the four membranes of the apicoplast have formed by a secondary endosymbiosis event, where the eukaryotic cell that was a product of primary endosymbiosis, itself becomes engulfed by the *Plasmodium* cell (van Dooren *et al.*, 2001). The derivation of the apicoplast has been a contentious area for some time, with differing opinions on whether the plastid arose from engulfment of a red or a green algae, however it is now accepted that the apicoplast was most likely derived from a red algal lineage (Williamson *et al.*, 1994; McFadden *et al.*, 1997; Yap *et al.*, 1997; Blanchard *et al.*, 1999).

The apicoplast has been shown to possess its own genome, which has been identified to be circular, 35 kb in length and exists in a copy number of between 1 and 15 (Fichera *et al.*, 1997; Kohler *et al.*, 1997; Matsuzaki *et al.*, 2001; Williamson *et al.*, 2001). The apicoplast genome carries 60 open reading frames, of which only two are not involved in protein synthesis (Wilson *et al.*, 1996). The majority of the genes from the original endosymbiont's genome are thought to have moved to the *Plasmodium* nucleus by lateral gene transfer (Martin *et al.*, 1998; McFadden, 2001; Rujan *et al.*, 2001). This means that the nuclear encoded proteins that are expressed by the parasites require to be transported to, and imported into the plastid.

Due to the plastid lineage of the apicoplast it was assumed that the transport and import of the nuclear encoded proteins would be similar to that observed in plant chloroplasts. The trafficking of nuclear encoded proteins to chloroplasts in plants requires the presence of an N-terminal extension known as a transit peptide (Vothknecht *et al.*, 2000; Jackson-Constan *et al.*, 2001). The transit peptide of plant chloroplasts are rich in hydroxylated amino acids such as serine and threonine and the transfer of the protein across the chloroplast membranes are mediated by import machinery called the TOC and TIC complexes (representing translocon of the outer/inner chloroplast membranes).

The protein to be imported into the chloroplast, which has formed a complex with HSP70 and other chaperone proteins in the cytosol, binds to the TOC complex in a GTP-dependant manner. The protein is then transported across the outer and inner membrane using a mechanism that utilises both the TOC and TIC complexes. Once the transit peptide of the protein has passed the inner chloroplast membrane it is cleaved by a stromal processing peptidase, producing a chloroplast localised mature protein (Richter *et al.*, 1998).

The trafficking of nuclear encoded *Plasmodium* proteins is thought to function by a similar pathway, although the transfer into the plastid is complicated by the presence of the four membranes of the apicoplast. The targeting of protein to the complex plastids (those with 3 or 4 membranes) in many algae provided a clue to the requirements for transport of proteins into the apicoplast (Grossman *et al.*, 1990; Pancic *et al.*, 1993; Apt *et al.*, 1995; Caron *et al.*, 1996). In these complex plastids the transit peptide was preceded by a hydrophobic amino acid sequence known as the signal peptide thus forming a bipartite targeting peptide consisting of a signal peptide and a plastid transit peptide. Prediction programs have been developed to calculate the likelihood of a protein's targeting based on the bipartite nature of proteins targeted to plant and algal plastids (TargetP; http://www.cbs.dtu.dk/services/TargetP/ and SignalP; http://www.cbs.dtu.dk/services/SignalP/ (Nielsen *et al.*, 1997; Emanuelsson O, 2000).

Analysis of the *P. falciparum* genome identified ~ 500 proteins that possessed a potential bipartite targeting peptide that also comprises a signal peptide followed by a possible apicoplast transit peptide (Foth *et al.*, 2003). The nature of this targeting signal was validated and dissected using green fluorescent fusion proteins with the respective parts of the targeting peptide (figure 1.11; Waller *et al.*, 1998; DeRocher *et al.*, 2000; Waller *et al.*, 2000). These studies allowed Foth *et al.* to derive an algorithm that can now be used to predict the possibility of a protein being targeted to the apicoplast (Foth *et al.*, 2003). Many studies have utilised GFP fusion proteins to test the signal transit peptides of potential plastid localised proteins in *P. falciparum* and the related apicomplexan parasite *Toxoplasma gondii.*

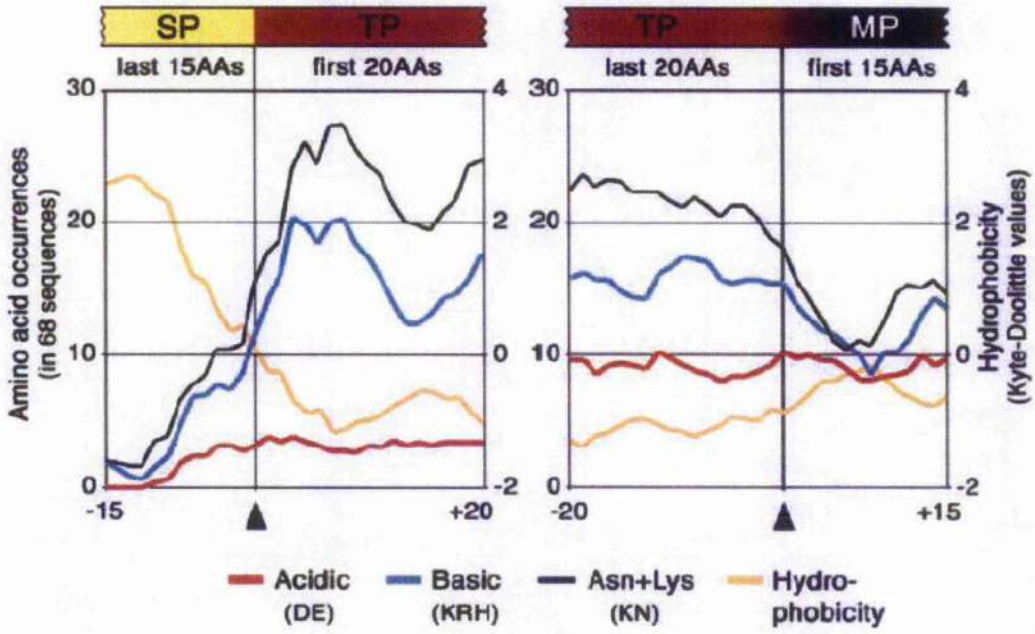


Figure 1.11 – Analysis of apicoplast localisation determinants

This figure displays the results generated from the alignment of 68 putative apicoplast targeted proteins around their estimated signal peptide (SP) cleavage sites (black arrow on the left side) and the estimated boundary between the transit peptide (TP) and the mature protein (MP; black arrow on the right side). This resulted in a clear preference for hydrophobic residues in the signal peptide (yellow line, right hand y axis) and a relatively high occurrence of lysine and asparagines (black line, left hand y axis) and basic residues (blue line, left hand y axis) in the transit peptide. This figure was taken from Foth *et al.*, 2003

The presence of a secondary endosymbiotic plastid obviously was of great interest to the scientific community particularly because it could be expected that this organelle was retained by the parasites to provide some metabolites necessary for their survival. The apicoplast was thought to be a possible drug target as drugs that inhibited the plastids genome replication, transcription and replication led to parasite death (Fichera *et al.*, 1997; Ralph *et al.*, 2001). A transgenic *Toxoplasma gondii* strain that displayed a segregation defect also resulted in the death of parasites lacking the organelle (He *et al.*, 2001). He *et al.* produced a fusion protein containing GFP that was localised to the apicoplast, which was observed to block plastid segregation during endodyogeny (*T. gondii* equivalent of schizogony in *P. falciparum*) and resulted in parasites that displayed a delayed death phenotype (He *et al.*, 2001). This delayed death phenotype is observed in both the transgenic segregation mutant and in the inhibition studies. The parasites seem to develop normally in the host cell during treatment, but the merozoites produced are no longer viable to re-infect host cells. This delayed death phenotype may point towards an essential role for the apicoplast in the invasion of host cells, possibly in the production of the parasitophorous vacuole membrane (PVM).

Of the 500 proteins predicted to be localised in the apicoplast, almost 70 % have no known function. Of the 30 % with known functions a number have been identified to have housekeeping functions such as DNA replication/repair (8.8 %), transcription (2.9 %), translation (25.9 %), transporters (5.3 %) and other roles (15.9 %). Three major pathways have been identified in the apicoplast proteome; fatty acid biosynthesis, isoprenoid precursor synthesis and a partial haem biosynthesis pathway (Gardner *et al.*, 2002; Ralph *et al.*, 2004).

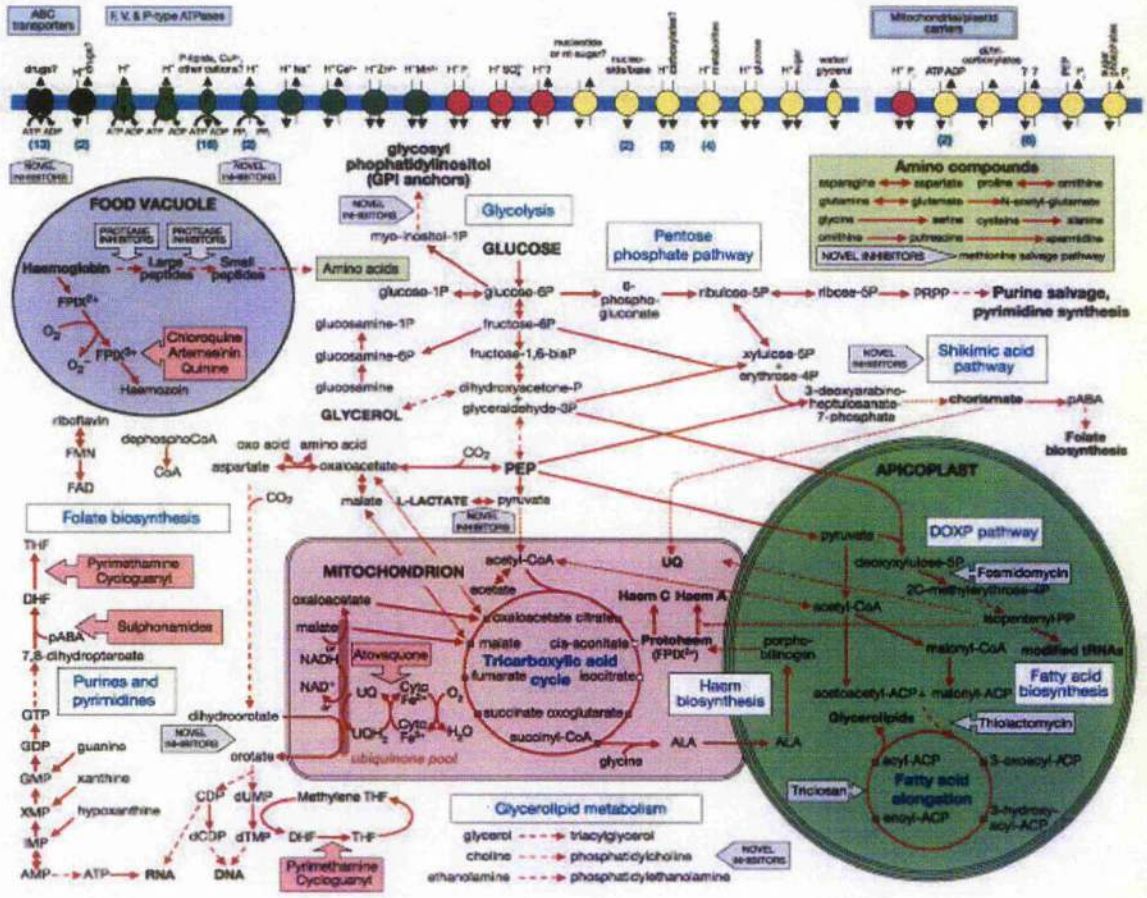


Figure1.12 – Biochemical pathways of *P. falciparum*

This figure gives an overview of the biochemical pathways of *P. falciparum* based on the genes identified in the parasites genome and was taken from Gardner *et al.*, 2002. This study is particularly interested the biochemistry of the mitochondrion and the apicoplast, and this figure displays the pathways that are thought to be active in these organelles. The mitochondrion contains the tricarboxylic acid cycle (TCA cycle), oxidative phosphorylation and a partial haem biosynthesis. The apicoplast contains a type II fatty acid synthesis pathway, a DOXP pathway (non-mevalonate isoprenoid biosynthesis) and a partial haem biosynthesis pathway. The metabolic steps of pathways are indicated with red arrows, while metabolic steps with multiple intervening steps are indicated by a dotted red arrow. Fushia coloured block arrows indicate steps that are inhibited by antimalarials. Abbreviations for mitochondrial and apicoplast pathways: ACP, acyl carrier protein; ALA, aminolevulinic acid; CoA, coenzyme A; DOXP, deoxyxylulose phosphate; PEP, phosphoenolpyruvate; UQ, ubiquinone/coenzyme Q.

The type II fatty acid biosynthesis pathway that has been identified in the parasites is specific for bacteria and plant chloroplasts and requires acetyl CoA as an entry point (Harwood, 1996; Ke *et al.*, 2000). Acetyl CoA is provided from a plastid located pyruvate dehydrogenase complex (PDH; see section 1.4.3) and the presence of this complex will be investigated in chapter 5 of this study. Type II fatty acid biosynthesis is a cyclical pathway which adds two carbons per cycle and usually produces $C_{10} - C_{14}$ fatty acids (Surolia *et al.*, 2001). The fatty acid elongation cycle occurs via four enzymatic steps; condensation, reduction, dehydration and reduction. Triclosan is thought to act specifically against enoyl-ACP reductase (second reduction step of the cycle) of the pathway and has also been shown to have potent antimalarial activity (McLeod *et al.*, 2001; Surolia *et al.*, 2001; Waller *et al.*, 2003).

The presence of the plant-specific non-mevalonate (DOXP) isoprenoid biosynthetic pathway, which uses pyruvate and glyceraldehyde-3-phosphate as precursors rather than mevalonate within the organelle and the absence of an alternative pathway providing isoprenoid units had significance particularly with respect to the design of new antimalarials (Jomaa *et al.*, 1999). Fosmidomycin, an antibiotic that acts against 1-deoxy-D-xylulose-5-phosphate (DOXP) reductoisomerase of the pathway is currently in clinical trials with some success, although recrudescent infections were observed following initial clearing of infection (Jomaa *et al.*, 1999; Lell *et al.*, 2003). Isoprenoids are composed of repeats of isoprene units and form the basis of ubiquinones, dolichols, modified tRNA and prenylated proteins.

P. falciparum appears to have a *de novo* haem biosynthesis pathway, even though it ingests large quantities of haem containing proteins during the erythrocytic stage of its life cycle. The haem biosynthesis pathway in *P. falciparum* contains components in the apicoplast and the mitochondrion (van Dooren *et al.*, 2002; Ralph *et al.*, 2004; Sato *et al.*, 2004; Varadharajan *et al.*, 2004) and some components may also be imported from the host cell (Dhanasekaran *et al.*, 2004). The mitochondrion and the apicoplast have been observed numerous times to be closely association (Hopkins *et al.*, 1999; Bannister *et al.*, 2000; van Dooren *et al.*, 2005). This observation, when coupled with the suggested interchange of substrates between the two organelles (as predicted for haem biosynthesis) has led some researchers to suggest that they are metabolically co-dependent (van Dooren *et al.*, 2005).

1.5.2 Mitochondrion

P. falciparum contains a single mitochondrion and is generally considered to be achristate during the erythrocytic stages of the life cycle (Rudzinska, 1969; Learngaramkul *et al.*, 1999). This is in contrast to the mitochondria observed in gametocytes of *P. falciparum*, which are multiple and christate (Learngaramkul *et al.*, 1999) and the mitochondrion of bird malaria parasites that have also been observed to contain christae (Rudzinska, 1969). The single mitochondrion of *P. falciparum* (like all eukaryotes) was derived from α-protobacterial cells by primary endosymbiosis, which resulted in an organelle that is surrounded by two outer membranes. The mitochondrion also contains a residual genome, which is 6 kb in length and contains fragmented genes encoding large and small ribosomal RNA and three genes that encode cytochrome b, cytochrome oxidase I and cytochrome oxidase III. Therefore, the majority of the genes from the mitochondrial genome must have been passed to the parasites nucleus by lateral gene transfer (Adams *et al.*, 2003). This means that the proteins encoded by the nuclear genes require trafficking back to and into the mitochondrion, from the cytoplasm. The transport of nuclear encoded proteins to the mitochondrion can be mediated by a number of pathways although the most common utilises an N-terminal mitochondrial targeting peptide (Pfanner *et al.*, 2001). The other methods for mitochondrial targeting include using transmembrane domains and internal signals, which have been poorly defined (Pfanner *et al.*, 2001). The N-terminal mitochondrial targeting pre-sequences have been well characterised in other organisms and have led to the development of programs which can predict the likely mitochondrial localisation of a protein based on its N-terminal amino acids (Claros *et al.*, 1996). These N-terminal pre-sequences have the general characteristics at the primary and secondary structure level. The amino acid sequence tends to be high in basic, hydroxylated and hydrophobic residues and an amphipathic α helix with a hydrophobic face and a positive face (von Heijne *et al.*, 1989; Schatz *et al.*, 1996). The mitochondrial-targeted protein is transported into the mitochondria via a group of protein complexes called the TOM and TIM complexes (Truscott *et al.*, 2003). Indeed, homologues of two TOM complex proteins have been identified in the *P. falciparum* genome, which are TOM40 and TOM22 (Macasev *et al.*, 2004). Homologues of two TIM complex associated chaperone proteins (Hsp60 and Hsp10) have also been identified and shown to be localised in the mitochondrion by GFP fusion proteins (Sato *et al.*, 2003; Tonkin *et al.*, 2004; Sato *et al.*, 2005). It is highly likely that *P. falciparum* possess other TIM/TOM complex

proteins although protein homologues have not been reported and require further investigation.

A study that analysed a substantial number of mitochondrial or potential mitochondrial protein sequences identified that the *P. falciparum* mitochondrial targeting pre-sequence was relatively high in isoleucine, leucine, tyrosine and asparagines residues and relatively low in glycine, alanine, tryptophan and arginine residues. The results of this study were used to develop a program that was specific for predicting mitochondrial protein in the *P. falciparum* genome (Bender *et al.*, 2003). Analysis of the *P. falciparum* genome with the Plasmit program predicted between 381 and 1177 (7.1 % and 22 % of total genes) mitochondrial proteins depending on the stringency of the program (Bender *et al.*, 2003).

1.5.2.1 Biochemical pathways in the mitochondrion

Studies into the single mitochondrion in *P. falciparum* have shown that despite being achristate the organelle is still metabolically active (figure 1.12). In mammalian cells the primary source of energy metabolism is via the production of ATP, which is driven in the mitochondria by the tricarboxylic cycle (TCA cycle) and oxidative phosphorylation. Glucose is metabolised into pyruvate by glycolysis and is acted on by a mitochondrial localised pyruvate dehydrogenase complex (PDH; see section 1.4.3 and chapter 5) to produce acetyl CoA, which can then feed into the TCA cycle.

However, energy metabolism in the erythrocytic stages of *P. falciparum* is provided primarily by glycolysis, with up to 90% of the glucose being metabolised into lactic acid by lactate dehydrogenase (Bryant *et al.*, 1964; Scheibel *et al.*, 1970). This release of lactic acid is responsible for causing the lactic acidosis that is a clinical feature of *Plasmodium* infection.

It was therefore initially thought that *Plasmodium* did not therefore contain an active TCA cycle in their erythrocytic stages. In spite of this, all of the genes necessary for an active TCA cycle have now been identified in the *P. falciparum* genome, although there are several key differences in the make up of the *P. falciparum* TCA cycle. Citrate synthase, which is the first step in the TCA cycle has been confirmed to be localised to the mitochondrion of *P. falciparum* by GFP fusion protein (Tonkin *et al.*, 2004). The second protein in the TCA cycle is aconitase and a protein with homology to human iron responsive element protein/aconitase has recently been identified in the *P.*

falciparum genome and has been shown to be (at least partially) mitochondrially localised and active (Hodges *et al.*, 2005). The *P. falciparum* isocitrate dehydrogenase protein has been biochemically characterised and was shown to be active with $NADP^+$ and isocitrate substrates (not NAD^+ specificity as in mammalian cells) and as a result was thought to be involved in mitochondrial redox control rather than in providing electrons for the electron transport chain (Wrenger *et al.*, 2003). The next step in the TCA cycle, the α-ketoglutarate dehydrogenase complex (KGDH; see section 1.4.1) will be investigated in this thesis (Chapters 4 and 5). *P. falciparum* are thought to require succinyl CoA, which is produced by KGDH, for continuing the TCA cycle and also as a precursor for haem biosynthesis. The succinyl CoA is converted to succinate via the succinyl-CoA synthetase /protein, which consists of α and β subunits. Homologues of the succinyl-CoA synthetase of α and β subunits have been identified in the *P. falciparum* genome as have the flavoprotein and iron-sulphur subunits (but not the CybS and CybL subunits) of succinate dehydrogenase (Suraveratum *et al.*, 2000; Takeo *et al.*, 2000). Succinate dehydrogenase plays a role in oxidative phosphorylation (as complex II) and converts succinate into fumarate, which is then acted on by fumarate hydratase. The *P. falciparum* homologue of fumarate hydratase displayed homology to the iron sulphur containing Class I proteins of bacteria and archea rather than the class II proteins from yeast and mammalian cells (Flint *et al.*, 1992). In the mammalian TCA cycle malate is then converted to oxaloacetate by a mitochondrial malate dehydrogenase, although the malate dehydrogenase that has been identified in *P. falciparum* is not thought to be mitochondrial (Lang-Unnasch, 1992; Lang-Unnasch, 1995). However, a possible mitochondrially localised malate quinone oxidoreductase (MQO) homologue has been identified in the *P. falciparum* genome. MQO is a membrane bound flavoprotein, which is found in some bacteria and may contribute electrons via coenzyme Q into oxidative phosphorylation (Uyemura *et al.*, 2004).

The mitochondrion of *Plasmodium* species has been shown to maintain a membrane potential during their erythrocytic stages (Srivastava *et al.*, 1997). This is indicative of the mitochondrion having a functional electron transport chain/oxidative phosphorylation (figure 1.12). The antimalarial drug Atovaquone has been shown to collapse this membrane potential and was later found to interact specifically with the mitochondrial complex III (Srivastava *et al.*, 1997).

It has been proposed that the main role of oxidative phosphorylation in the mitochondrion of *P. falciparum* is to act as an electron sink for the dihydroorotate

dehydrogenase protein that is involved in pyrimidine biosynthesis (Krungkrai, 1995). This is an essential step in parasite biochemistry as *P. falciparum* are unable to scavenge pyrimidines and therefore must rely on *de novo* synthesis. Dihydroorotate dehydrogenase has therefore become an attractive drug target and high throughput drug screens have recently identified some potential *Plasmodium* specific inhibitors (Baldwin *et al.*, 2005).

The mammalian oxidative phosphorylation pathway is composed of a number of membrane associated protein complexes (complex I – IV and an ATP synthase). Protein homologues of components from most of these complexes have been identified in the *P. falciparum* genome (Gardner *et al.*, 2002). Activities of complex I (Krungkrai *et al.*, 2002), complex II (Fry *et al.*, 1991; Uyemura *et al.*, 2000; Takashima *et al.*, 2001; Uyemura *et al.*, 2004), complex III (Fry *et al.*, 1991) and complex IV (Krungkrai *et al.*, 1993) have been shown in a number of studies. Although homologues of the ATP synthase have been identified in the *P. falciparum* genome (Gardner *et al.*, 2002), there have been no reports of any activity. It has been suggested that *P. falciparum* possess and alternate oxidase (Murphy *et al.*, 1997), although no clear homologue has been identified in the genome (G. McFadden, personal communication). Two other proteins have also been suggested to be able to act as a complex III, these being dihydroorotate dehydrogenase (see above) and a malate-quinone oxidoreductase (Fry *et al.*, 1991; Uyemura *et al.*, 2000; Takashima *et al.*, 2001; Uyemura *et al.*, 2004).

Other biochemical functions that occur in the mitochondrion of *P. falciparum* (figure 1.12) are a partial haem biosynthesis pathway (see section 1.5.1), folate metabolism, coenzyme Q synthesis and iron-sulphur cluster biosynthesis. Folates are required in *P. falciparum* as precursors for pyrimidines and methionine (Krungkrai *et al.*, 1990), although the functions of folate metabolism are poorly understood (Hyde, 2005; Nzila *et al.*, 2005a; Nzila *et al.*, 2005b). This pathway is clearly important for parasite biochemistry, as it is a target for a number of antimalarial drugs such as pyrimethamine and sulphonamides. The synthesis of Coenzyme Q (ubiquinone) is required in *P. falciparum* mitochondrion due its essential role as an electron carrier in oxidative phosphorylation. *P. falciparum* require the synthesis of iron-sulphur clusters due to the presence of many iron-sulphur cluster proteins such as ferredoxin, aconitase/iron regulatory protein, fumarate hydratase and proteins involved in isoprenoid biosynthesis and oxidative phosphorylation (Ralph *et al.*, 2004; Johnson *et al.*, 2005; Lill *et al.*, 2005).

1.6 Aims of this study

This thesis can be separated into three distinct research projects. The first (presented in chapter 3) was to investigate the catalytic mechanism of *P. falciparum* thioredoxin reductase, the second (presented in chapter 4) was to investigate the *P. falciparum* dihydrolipoamide dehydrogenases and the last (presented in chapter 5) was to investigate the components of the α–keto acid dehydrogenase complexes. The aims of each of these projects are detailed below:

Chapter 3: *P. falciparum* thioredoxin reductase (PfTrxR)

- Investigate the catalytic mechanism of the PfTrxR using pre-steady state and steady state kinetics of reductive and oxidative half reactions.
- Examine the role of acid/base catalysts in the catalysis of PfTrxR by mutagenesis followed by steady state and pre-steady state kinetics of reductive and oxidative half reactions.

Chapter 4: *P. falciparum* dihydrolipoamide dehydrogenases (LipDH)

- Identify and analyse the genes encoding LipDH in the *P. falciparum* genome.
- Investigate the transcription and protein expression of *LipDH* genes and LipDH proteins in erythrocytic stages of the parasites.
- Analyse the subcellular localisation of LipDH proteins in *P. falciparum*.
- Biochemically characterise recombinant proteins expressed in the *Escherichia coli* system.

Chapter 5: *P. falciparum* α–keto acid dehydrogenase complexes (KADH)

- Identify and analyse the genes encoding for components of the KADH the *P. falciparum* genome.
- Investigate the transcription and protein expression of *KADH* genes and KADH proteins in erythrocytic stages.
- Analyse the subcellular localisation of KADH proteins in *P. falciparum*.
- Biochemically characterise recombinant proteins expressed in the *E. coli* system.

Chapter 2 Materials and Methods

2.1 Materials

2.1.1 Companies from which chemicals, kits and equipment were purchased

Abgene: Reddymix

Amersham biosciences: ÄKTA FPLC, Hyperfilm ECL, HiTrap chelating HP column (1ml and 5ml), hypercassettes (18 x 24 cm), Sephadex S200 16/60 size exclusion column, Superose 6 10/30 size exclusion column, ECL-plus detection kit, protein A sepharose beads.

BDH: Giemsa stain, saponin, formamide, sodium acetate

Beckman: J2-H5 centrifuge with JA-20 rotor

BOC: Malaria culture gas (5% CO_2, 1% O_2 and 94% N_2)

Biorad: Econopac columns, transblot SD semi-dry transfer cell, precision plus all blue protein standards, gel filtration standards, gene pulser electroporation cuvette (0.2 cm), gene pulser Xcell electroporator, protein assay reagent.

Blood transfusion service: Human A^+ erythrocytes

Eppendorf: Eppendorf microcentrifuge with F45-24-11 rotor

Eurogentec: Custom antibody production

Fisher scientific: 0.2 μM filtration units (500 ml and 1000 ml containers)

Hi-Tech scientific: SFA-20 rapid kinetics accessory, SF-61 DX2 double mixing stopped flow spectrophotometer, kinetasyst 3 software.

Invitrogen: Accuprime *Pfx* Supermix, PCR supermix, TOPO TA cloning kit, ZERO BLUNT PCR cloning kit, chemically competent *E.coli* TOP10 cells, RPMI 1640 (with 25 mM HEPES, L-Glutamine, without $NaHCO_3$), Albumax II, gentamycin, NuPAGE 4-

12% Bis/Tris gels, NuPAGE 10 % Bis/Tris gels, 20 x MOPS buffer, Superscript II, DNase I, Random hexamers, dNTP mix, Xcell Surelock™ Mini-cell apparatus.

Jacobus pharmaccuticals: WR99210

Melford: Ampicillin, Isopropyl-beta-D-thiogalactopyranoside (IPTG).

Molecluar probes: MitoTracker CMXRos

Nalgene: 1.0 ml cryotubes

New England Biolabs: All restriction endonucleases, T4 DNA ligase

Novagen: Bugbuster protein extraction solution, benzonase, pET28 expression vector, Chemically competent *E.coli* Novablue

Promega: 1 kb DNA ladder, 6 x DNA loading buffer

Qiagen: Qiaprep spin DNA miniprep columns, Hispeed plasmid maxi kit, Qiaquick gel extraction kit, Ni-NTA agarose, RNase A, pQE30 expression vector.

Roche: Complete mini protease inhibitor tablets, Complete EDTA free protease inhibitor tablets.

Sciquip: Sigma 6K 15 centrifuge with 11150 rotor.

Thermo electron: Custom produced oligonucleotides, Jouan CR3i centrifuge with T40 rotor.

Sartorius: 0.2 μm syringe filters, 0.4 μm syringe filters.

Schleicher & Schuell: Protran nitrocellulose

Shimadzu: UV-2401 and UV-2501 spectrophotometers

Sigma: All chemicals unless otherwise stated.

Stratagene: Chemically competent *E.coli* JM109, chemically competent *E.coli* XL10 GOLD, chemically competent *E.coli* BL21 (DE3), chemically competent *E.coli* BL21 (DE3) RIL

Varian: Cary 300 spectrophotometer

Zeiss: Axioplan 2 microscope

2.1.2 Commonly used reagents

Electrophoresis buffers

1 X TAE: 40 mM Tris.Acetate pH 7.6, 1mM Na_2.EDTA.

6 x SDS-PAGE loading buffer: 62.5 mM Tris-HCl pH 6.8, 2 % (w/v) SDS, 10 % (v/v) glycerol, 0.001 % (w/v) bromophenol blue, 5 % (v/v) 2-mercaptoethanol.

MOPS buffer: 50 mM 3-[N-morpholino] propane sulphonic acid, 50 mM Tris base, 3.5 mM SDS, 1 mM EDTA.

Coomasie stain: 40 % Methanol, 10 % Acetic acid, 0.1 % (w/v) coomasie brilliant blue R – 250.

Destain: 20 % Methanol, 10 % Acetic acid

Western blotting buffers

Towbin buffer: 25mM Tris, 192 mM glycine, 20 % Methanol

PBS: 4.3 mM Sodium phosphate (dibasic), 1.4 mM potassium phosphate (monobasic), 137 mM NaCl, 2.7 mM KCl.

PBS/tween: PBS containing 0.05 % Tween 20

Bacterial culture

LB media: 10 g/l Bacto-Tryptone, 5 g/l Bacto-yeast extract, 10 g/l NaCl

TB media: 12 g/l Bacto-tryptone, 24 g/l yeast extract, 0.4% (v/v) glycerol, 1 X TB salts.

10 x TB salts: 23.1 g/l KH_2PO_4, 125.4 g/l K_2HPO_4 (potassium phosphate, dibasic)

Ampicillin: Stock of 100 mg ml^{-1} in water. Filter sterilised, aliquoted and stored at – 20°C.

Kanamycin: Stock of 50 mg ml^{-1} in water. Filter sterilised, aliquoted and stored at – 20°C.

Protein purification

Qiagen lysis buffer: 50 mM sodium phosphate, 300 mM NaCl, 10 mM imidazole, pH 8.0.

Qiagen wash buffer: 50 mM sodium phosphate, 300 mM NaCl, 20 mM imidazole, pH 8.0.

Qiagen elution buffer: 50 mM sodium phosphate, 300 mM NaCl, 500 mM imidazole, pH 8.0.

Denaturing Ni-NTA purification buffer (Change pH according to wash/elution): 100 mM sodium phosphate, 10 mM Tris-Cl, 8M urea

***Plasmodium* culture**

Complete RPMI medium: Blood wash media supplemented with 5 g litre^{-1} Albumax II.

Blood wash medium: 15.9 g/l RPMI 1640, 2 g litre^{-1} Glucose, 1 g litre^{-1} Sodium bicarbonate, 0.01 mg ml^{-1} gentamycin, 0.03 mg ml^{-1} hypoxanthine, pH 7.4.

Saponin solution: 20 mg ml^{-1} Saponin in 1 x PBS

Sorbitol solution: 5 % (or 27 %) sorbitol (w/v) in 10 mM potassium phosphate buffer pH7.4

Cytomix: 120 mM KCl, 0.15 mM $CaCl_2$, 2 mM EGTA, 5 mM $MgCl_2$, 10 mM potassium phosphate (pH 7.6), 25 mM HEPES (pH 7.6).

Freezing solution: 1 x PBS, 30% glycerol

DNA lysis buffer: 50 mM Tris-HCl pH 8.0, 20 mM EDTA, 2% SDS.

2D lysis buffer: 100 mM Hepes, 5 mM $MgCl_2$, 10 mM EDTA, 0.5% Triton X-100, 5 µg ml^{-1} RNAse A, 1 mM PMSF, 1 mM benzamidine, 2 µg ml^{-1} leupeptin, protease inhibitor cocktail (Roche, 1 tablet per 50 ml of buffer).

TE: 50 mM Tris-HCL pH 7.5, 1 mM EDTA

Immunoprecipitation buffers

IP lysis buffer: 50 mM Tris HCl 7.5, 600 mM KCl, 1% Triton X-100, complete protease inhibitor cocktail (Roche; 1 tablet per 10 ml)

IP wash buffer: 50 mM Tris HCl (pH 7.5), 150 mM NaCl, 1 mM EDTA, 1% (v/v) Triton X-100, 0.25% (w/v) BSA

2.2 Methods

2.2.1 *Plasmodium falciparum* cell culture techniques

2.2.1.1 General culturing

All solutions used for the culturing of *P. falciparum* were filter sterilised using 0.2 µm filters (Fisher scientific) in a category 2 tissue culture hood. *P. falciparum* strains 3D7 (The Netherlands) and D10 (Papua New Guinea) were cultivated in human A^+ erythrocytes (Blood transfusion service, Glasgow) according to Trager and Jensen (Trager *et al.*, 1976). Erythrocytes were washed in RPMI 1640 containing 0.1% sodium bicarbonate, 11 mM glucose, 0.03 mg/ml hypoxanthine and 0.1 mg/ml gentamycin (*Pf* wash medium). Washing was carried out by centrifuging erythrocytes at 1500 rpm (Jouan CR3i centrifuge with T4 rotor), 4°C for 10 min and subsequently aspirating the supernatant and buffy coat/leukocyte layer from the top of the red blood cell pellet. The procedure was repeated 2 to 3 times until the supernatant appeared clear and no white blood cells were observed. Erythrocytes washed in this manner were stored at 4°C and used for culturing for a maximum of 4 weeks following washing.

Parasite strains were cultured in complete RPMI 1640 (*Pf* wash medium supplemented with 0.5% Albumax II) at a 5% haematocrit. Cultures were incubated at 37°C in a humidified atmosphere of 5% CO_2, 1% O_2 and 94% N_2. Medium was exchanged daily and parasitemia was maintained between 0.1 and 10%. Parasitemia was determined by counting the percentage of infected erythrocytes in Giemsa stained blood smears.

2.2.1.2 Synchronisation

Parasite cultures were synchronised as described previously (Hoppe *et al.*, 1991). Cultures were centrifuged at 1500 rpm (Jouan CR3i centrifuge with T4 rotor), 25°C for 10 min and the supernatant was discarded. The resulting pellet was resuspended in 5 pellet volumes of 5% sorbitol (w/v in 10 mM potassium phosphate pH 7.2) at 37°C, incubated at 37°C for 5 min and centrifuged again. The supernatant was removed and the pellet was resuspended in complete RPMI 1640 medium at 37°C and cultured as described above (section 2.2.1.1). Parasite cultures were required to be synchronised for transfection protocols (see section 2.2.1.5).

2.2.1.3 Stabilates

Parasite stabilates were produced from 10 ml cultures with 10% parasitemia in mostly ring stages. Cultures were centrifuged at 1500 rpm (Jouan CR3i centrifuge with T4 rotor), 4°C, 10 mins, before removing the supernatant and resuspending in one pellet volume of cold complete RPMI 1640 media. Two pellet volumes of ice cold stabilate freezing solution (30% glycerol in PBS) was added and 600 µl was transferred into a cryotube (Nalgene) and was incubated on ice for 15 min before storing in liquid nitrogen.

Stabilates were thawed with constant shaking in a 37°C waterbath, and placed on ice immediately following thawing. The contents of the cryotube was transferred to a 15 ml falcon tube (usually 600 µl), and 2 volumes of 27% sorbitol (w/v in 10 mM potassium phosphate buffer pH 7.4) were added drop-wise with constant shaking, before being incubated on ice for 13 mins. A further two volumes of 5% sorbitol (w/v in 10 mM potassium phosphate buffer pH 7.4) was added in a similar way and was incubated on ice for 10 min. The solution was centrifuged at 1500 rpm (Jouan CR3i centrifuge with T4 rotor), 4°C for 5 min, and the supernatant was discarded. The resulting pellet was resuspended in 10 ml of cold complete RPMI 1640 medium, and was centrifuged as before. This wash step was repeated a second time before the pellet was resuspended in complete RPMI 1640 media at 37°C (containing washed RBC at 2.5% haematocrit). Parasites were then cultured as described in section 1.2.1.1, until they reached a parasitemia of 5 % when the appropriate drug selection was applied.

2.2.1.4 Isolation of *P. falciparum* from host erythrocyte

Parasites were extracted from erythrocytes using saponin as described by (Umlas *et al.*, 1971). Parasites were freed from host erythrocytes by incubating cultures with 0.1 volumes of saponin (20 mg/ml in PBS) on ice for 10 min. The solution was then centrifuged at 1500 rpm (Jouan CR3i centrifuge with T4 rotor), 4°C for 10 min before the supernatant was carefully removed; the parasite pellet was washed in 40 ml ice cold PBS to remove residual erythrocytes. The parasite/PBS mixture was centrifuged as described before, the supernatant discarded and the resulting parasite pellet was used for isolation of RNA, genomic DNA (gDNA) or protein and was resuspended in a solution consistent with the isolation necessary. Usually a minimum of 50 ml parasite culture

with a ~10 % parasitemia was necessary to obtain sufficient amounts of RNA, gDNA or protein for further analyses.

RNA isolation was performed following saponin extraction by resuspending the parasite pellet in ten volumes of Trizol (Sigma); this was stored at -80°C until use. RNA was isolated from the Trizol by adding 0.2 volumes chloroform, mixing and centrifugation at 12,000 x g for 30 min at 4°C. The upper aqueous phase was removed and mixed with 0.7 volumes isopropanol, and incubated at 4°C for at least 120 min. The RNA was then centrifuged at 12,000 x g, for 30 min at room temperature. The resulting supernatant was carefully removed and the RNA pellet allowed to air-dry, before being resuspended in 10-20 µl formamide (BDH). The sample was heated to 60°C for 10 min (to aid resuspension) before being cooled on ice and placed at -80°C for long-term storage. The RNA concentration was determined by measuring the absorption at 260nm. The RNA concentration was calculated according to Abs_{260} x 40 = µg of RNA /ml of solution (Sambrook, 1989).

Isolation of genomic DNA from parasite was performed by thoroughly resuspending parasite pellets in 1 ml (per 50 ml of parasite culture) of gDNA lysis buffer (50 mM Tris-HCl pH 8.0, 20 mM EDTA, 2 % SDS). Proteinase K (Sigma) was added to a final concentration of 100 µg/ml and the extraction mixture was incubated at 56°C for 2 hours. The suspension was cooled on ice for 10 min before 0.3 volumes of saturated NaCl was added, mixed and incubated on ice for a further 5 min. The suspension was centrifuged at 2500 rpm, 4°C for 15 min, with the resulting supernatant transferred into fresh tubes and centrifuged at 13000 rpm, 4°C for 15 min. The supernatant was again removed into fresh tubes and RNaseA (Qiagen) was added to a final concentration of 20 µg/ml before being incubated overnight at 4°C. Following overnight incubation, one volume of phenol/chloroform was added and mixed by vortexing. This was centrifuged at room temperature for 10 min at 13000 rpm (Eppendorf microcentrifuge with F45-24-11 rotor), before the aqueous phase was removed and 3 volumes of 100 % ethanol and 0.1 volumes of 3 M sodium acetate pH 5.2 were added, mixed and incubated at -20°C for 60 min. This solution was then centrifuged to pellet the precipitated DNA at 13000 rpm (Eppendorf microcentrifuge with F45-24-11 rotor), 4°C for 15 min with the resulting supernatant being carefully discarded. The pellet was washed in 200 µl 70% ethanol and centrifuged at 13000 rpm (Eppendorf microcentrifuge with F45-24-11 rotor), 4°C for 5 min and the supernatant was again carefully discarded. The DNA pellet was allowed to air-dry and was resuspended in 100-200 µl sterile distilled water. DNA

concentration was determined using the absorbance at 260 nm and visually by agarose gel electrophoresis. The DNA concentration was calculated according to Abs_{260} x 50 = µg DNA/ml of solution (Sambrook et al., 1986).

Protein isolation was performed by resuspending parasite pellets in 10 volumes of 2D lysis buffer, with the resulting cell suspension being stored at -80°C until use. Protein extracts were prepared by five rounds of freeze/thawing (in liquid nitrogen and 37°C waterbath), followed by homogenisation using a micro-pestle and 0.7 mm glass beads (Sigma). Non-soluble proteins and membrane fractions were removed from the soluble proteins by centrifugation at 13000 rpm (Eppendorf microcentrifuge with F45-24-11 rotor), 4°C for 20 min (supernatant fraction), and protein concentrations were determined by Bradford assay (see section 2.2.7.6).

2.2.1.5 Transfection techniques

There are two current methods for the transfection of plasmid DNA into *P. falciparum*; these are "the spontaneous uptake" and "the infected erythrocyte" methods. Each method requires 100 µg of plasmid DNA, tightly synchronised parasite cultures and uses the same Cytomix (25 mM HEPES pH 7.6, 10 mM potassium phosphate pH 7.6, 120 mM KCl, 5 mM $MgCl_2$, 150 µM $CaCl_2$, 2 mM EGTA). 100 µg of plasmid DNA previously prepared using a Hi-speed plasmid maxi kit from Qiagen (see section 2.2.7.2) was ethanol precipitated (see section 2.2.7.3) and the resulting DNA pellet was air-dried in the cell culture cabinet (i.e. under sterile conditions) and resuspended in 30 µl TE buffer (10 mM Tris, 1 mM EDTA pH 7.5).

The spontaneous uptake method was modified from the previously published method (Deitsch *et al.*, 2001). The following details are those that are necessary for a single transfection. 500 µl erythrocytes were washed in 10 volumes cytomix and centrifuged at 1500 rpm (Jouan CR3i centrifuge with T4 rotor), 25°C for 10 min with the supernatant being removed after spinning. 470 µl of erythrocytes and 30 µl of DNA/TE mix were combined and added to a 0.2 cm gene pulser electroporation cuvette (Biorad), and electroporated at 310V and 950 µF using a gene pulser Xcell electroporator (Biorad). This transfection method gave time constant values in the range of 18-20 ms. The electroporated erythrocytes were removed from the cuvette into a culture flask containing 10 ml complete RPMI 1640 medium and erythrocytes infected with late trophozoite stages (sorbitol synchronised) were added to a final parasitemia of 1% and

cultured at 37°C for 4-6 hours after transfection. Following this incubation the culture media was replaced by fresh media containing the necessary drug selection.

The infected erythrocyte transfection method was modified from the protocol published previously (Wu *et al.*, 1995). The following details are those that are necessary for a single transfection. Synchronous parasite cultures with 8-10% ring stages were centrifuged at 1500 rpm (Jouan CR3i centrifuge with T4 rotor), 25°C for 10 mins and the resulting supernatant was discarded. The 30 µl plasmid/TE mix and 200 µl of infected erythrocytes were combined with 370 µl Cytomix and added to a 0.2 cm gene pulser electroporation cuvette (Biorad). The mix was electroporated at 310V and 950 µF using a gene pulser Xcell electroporator (Biorad). This transfection method gave time constant values in the range of 10-15 ms. The electroporated mix was removed from the cuvette into a culture flask containing 10 ml complete RPMI 1640 medium and 300 µl uninfected erythrocytes and cultured for 4-6 hours following transfection. Following this incubation, the culture media was replaced by fresh media containing the appropriate drug selection.

2.2.1.6 Fluorescent microscopy

The subcellular localisation of parasite proteins fused to green fluorescent protein were analysed using an Axioplan 2 microscope (Ziess) equipped with a HB0100 digital camera.

Images were captured at 100x magnification in phase-contrast for live cells. The subcellular localisation of the GFP-fusion proteins were observed using the fluorescent filters for GFP (excitation at 488 nm and emission at 505-530 nm) or FITC (excitation at 494 nm and emission at 518 nm). The parasite's mitochondrion was selectively stained using MitoTracker CMXRos (Molecular Probes) and was visualised using the fluorescent filter for rhodamine (excitation at 570 nm and emission at 590 nm).

Staining of the mitochondrion was performed by incubating 1 ml of culture (10 % parasitemia, 5% haematocrit) with 25 nM MitoTracker CMX ROS (Molecular Probes) for 5 min at 37°C, before washing in 10 ml of warm culture medium (at 37°C). The resulting culture was centrifuged at 1500 rpm (Jouan CR3i centrifuge with T4 rotor), 10 min, 25°C and the supernatant was removed and the erythrocyte pellet was resuspended in 1 ml of warm media (at 37°C) before being analysed as described above.

2.2.2 Bioinformatics

2.2.2.1 Identifying genes in the *P. falciparum* genome

The component genes of the KADH were identified in the *P. falciparum* genome by TBlastN searching the *Plasmodium* genome database PlasmoDb (The *Plasmodium* genome database collaborative (Collaborative, 2001), http://www.plasmodb.org/) with human or yeast protein sequences.

2.2.2.2 Multiple sequence alignments

Multiple protein alignments of the identified sequences were performed using ClustalW (http://www.ebi.ac.uk/clustalw/) with predicted *P. falciparum* protein sequences and those encoding homologous proteins from other organisms, which were identified by EC numbers in the Expasy database (http://us.expasy.org/enzyme/).

2.2.2.3 Subcellular localisation predictions

Subcellular localisation predictions using the N-terminal targeting sequences of each protein was performed using SignalP (Nielsen *et al.*, 1997) http://www.cbs.dtu.dk/services/SignalP/, TargetP (Emanuelsson O, 2000) http://www.cbs.dtu.dk/services/TargetP, MitoProt (Claros *et al.*, 1996) http://mips.biochem.mpg.de/cgi-bin/proj/medgen/mitofilter and PlasmoAP (Foth *et al.*, 2003) http://www.PlasmoDB.org/restricted/PlasmoAPcgi.shtml.

2.2.3 Amplification of *Plasmodium falciparum* genes

2.2.3.1 Designing constructs

Constructs designed for recombinant expression in the *E. coli* system were sub-cloned into one of three expression constructs; pQE30, pJC40 and pET28. Recombinant expression of gene products in these vectors is under the control of the T7 promoter and is therefore inducible upon the addition of IPTG. The expression plasmid pQE30 (Qiagen) contains a multiple cloning and ampicillin resistance cassettes, with the *LipDH1* expression construct cloned into *Bam* HI / *Hind* III sites (figure 2.1). Expression with the pQE30 vector produces a recombinant protein with a $(His)_6$ tag attached to it N-terminus. The expression plasmid pJC40 (Clos *et al.*, 1994) contains a

multiple cloning site and ampicillin resistance cassette (figure 2.2). The *LipDH2* expression construct was cloned into *Nde* I / *Bam* HI sites. Expression using the pJC40 vector produces a recombinant protein with a $(His)_{10}$ tag attached to its N-terminus. The expression plasmid pET28 (Novagen) contains a multiple cloning and kanamycin resistance cassette (figure 2.3) and the *KGDH-E2* expression construct was cloned into *Nhe* I / *Hind* III sites which resulted in a $(His)_6$ tag at the N-terminus.

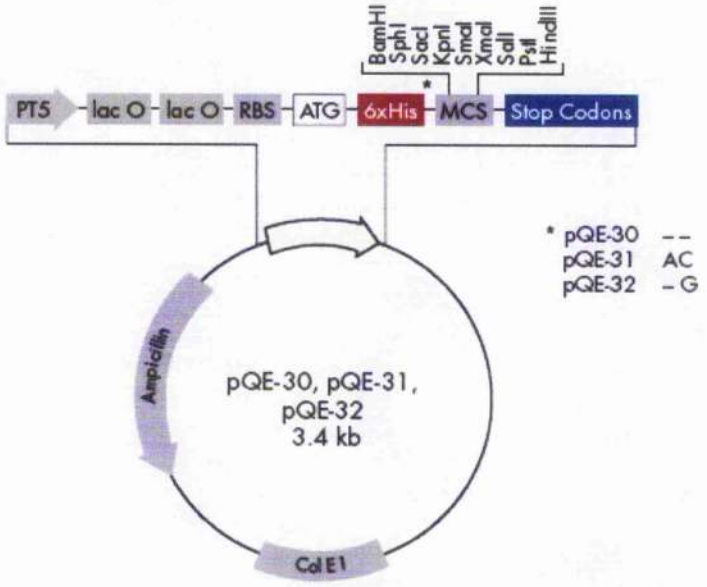


Figure 2.1– pQE30 recombinant expression

This figure displays the important features of the pQE30 recombinant expression plasmid; these include the ampicillin resistance cassette and the multiple cloning site (MCS) that contains the *Bam* HI/*Hind* III sites used for subcloning of the *LipDH1* expression construct. Recombinant expression using this plasmid is under the control of a PT5 promoter and attaches a $His_{(6)}$ tag at the N-terminus of the expressed protein. This figure was taken from the Qiagen website (www.qiagen.com).

ATCC 87114

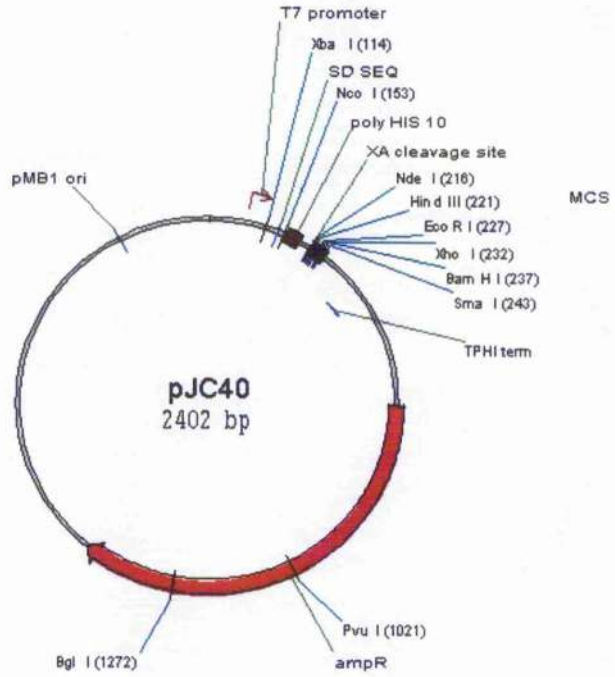


Figure 2.2 – pJC40 recombinant expression plasmid

This figure displays the important features of the pJC40 recombinant expression plasmid; these include the ampicillin resistance cassette (marked in red) and the multiple cloning site that contains the *Nde* I/*Bam* HI sites used for subcloning of the *LipDH2* expression construct. Recombinant expression using this plasmid is under the control of a T7 promoter and attaches a $His_{(10)}$ tag at the N-terminus of the expressed protein.

pET-28a(+) sequence landmarks	
T7 promoter	370-386
T7 transcription start	369
His•Tag coding sequence	270-287
T7•Tag coding sequence	207-239
Multiple cloning sites (*Bam*H I - *Xho* I)	158-203
His•Tag coding sequence	140-157
T7 terminator	26-72
lacI coding sequence	773-1852
pBR322 origin	3286
Kan coding sequence	3995-4807
f1 origin	4903-5358

The maps for pET-28b(+) and pET-28c(+) are the same as pET-28a(+) (shown) with the following exceptions: pET-28b(+) is a 5368bp plasmid; subtract 1bp from each site beyond *Bam*H I at 198. pET-28c(+) is a 5367bp plasmid; subtract 2bp from each site beyond *Bam*H I at 198.

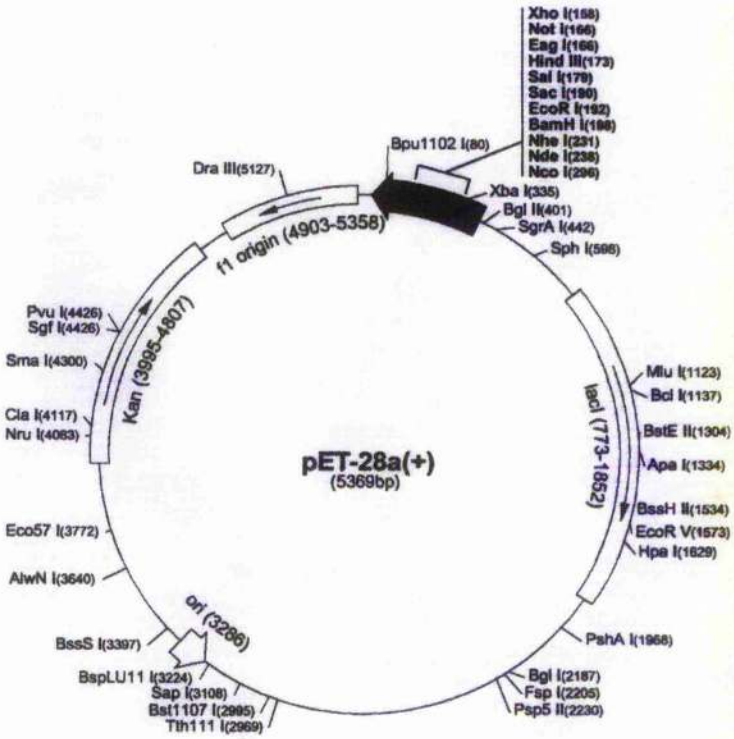


Figure 2.3– pET28 recombinant expression plasmid

This figure displays the important features of the pET28 recombinant expression plasmid; these include the kanamycin resistance cassette and the multiple cloning site. The KGDH-E2 expression construct was subcloned using *Nhe* I/*Hind* III. Recombinant expression using this plasmid is under the control of a T7 promoter and attaches a $His_{(6)}$ tag at the N-terminus of the expressed protein. This figure was taken from the Novagen website (www.novagen.com).

Constructs designed for expression of *P. falciparum* genes in transfected *P. falciparum* cells were sub-cloned into the pHH2 vector (Crabb *et al.*, 2004). This vector employs *Avr* II/ *Bgl* II restriction sites to clone an insert in frame with green fluorescent protein (GFP), producing C-terminally GFP-tagged proteins that are expressed following transfection of the plasmid DNA into *P. falciparum.* Expression of the GFP-fusion protein is under the control of the heat shock protein 86 promoter. pHH2 also contains the calmodulin promoter that controls expression of human dihydrofolate reductase (hDHFR) (figure 2.4), which confers resistance to the antifolate drug WR99210 and was used to select for the presence of the episomal vector following *P. falciparum* transfection (Fidock *et al.*, 1997).

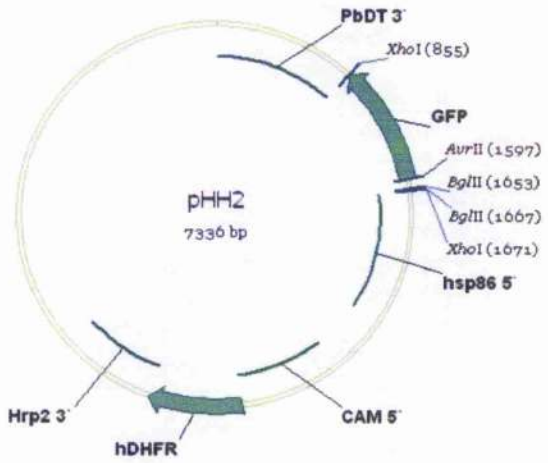


Figure 2.4– pHH2 *Plasmodium* expression plasmid

This figure displays the important features of the pHH2 *Plasmodium* expression plasmid. The human dihydrofolate reductase (hDHFR) gene is under the control of the Calmodulin promoter (CAM 5'), confers resistance to the pyrimethamine derivative WR99210 and is used as a selectable marker in parasite cultures. An ampicillin resistance cassette (not shown) is used during subcloning of GFP fusion constructs into the *Avr* II/*Bgl* II sites. This resulted in proteins that were tagged at C-terminus with green fluorescent protein (GFP), with expression under the control of the heat shock protein 86 promoter (HSP 86 5').

2.2.3.2 Preparation of oligonucleotide primers

Oligonucleotide primers were designed for cloning of expression constructs (full and possible mature lengths), site directed mutagenesis and internal sequencing of constructs. Oligonucleotides designed for amplification of PCR products for recombinant expression/gene knockout studies contained 5' extension that included restriction enzyme recognition sites, to allow cloning into the relevant destination plasmid. Oligonucleotides designed for use in site directed mutagenesis contained nucleotide changes designed to specifically incorporate single amino acid mutations in the open reading frame. The oligonucleotides (Thermo electron) used in this study are shown in table 2.1.

PfTrxR site directed mutagenesis oligonucleotides

PfTrxR H137N For	CAACTGTACAATCT**A**ACATACGTTCATTAAATTTTAGTTATATGAC
PfTrxR H137N Rev	GTCATATAACTAAAATTTAATGAACGTATGT**T**AGATTGTACAGTTG
PfTrxR H137Q For	CAACTGTACAATCTCA**G**ATACGTTCATTAAATTTTAGTTATATGAC
PfTrxR H137Q Rev	GTCATATAACTAAAATTTAATGAACGTAT**C**TGAGATTGTACAGTTG
PfTrxR H137S For	CAACTGTACAATCT**AG**CATACGTTCATTAAATTTTAGTTATATGAC
PfTrxR H137S Rev	GTCATATAACTAAAATTTAATGAACGTATG**CT**AGATTGTACAGTTG
PfTrxR H137A For	CAACTGTACAATCT**GC**CATACGTTCATTAAATTTTAGTTATATGAC
PfTrxR H137A Rev	GTCATATAACTAAAATTTAATGAACGTATG**GC**AGATTGTACAGTTG
PfTrxR E514A For	CCAACAGATGCTG**C**ATCGTTCATG
PfTrxR E514A Rev	CATGAACGAT**G**CAGCATCTGTTGG

Recombinant expression oligonucleotides

Bam HI	LipDH1 For	GCGC **GGATCC** TTA AAA GGA AGT ACA C
Hind III	LipDH1 Rev	CGCG **AAGCTT TTA** GTG AGT TCT TAT TTT TGA TAT AG
Nde I	LipDH2 For	**CATATG** TAT GAT GTT ATA GTC ATT GG
Bam HI	LipDH2 Rev	**GGATCC TTA** CAT GTG TAT AGG TTT ATC
Nhe I	KGDH E2 For	**GCT AGC ATG** ATT AAA GTA CCT AGA CTT GG
Hind III	KGDN E2 Rev	**AAG CTT TTA** ACA ATC AAT TAA CAT TAG

GFP fusion oligonucleotides

Bgl II	LipDH1 GFP For	GCGC **AGATCT** TAGA **ATG** GTC ATA AGG CAA AAT ATT AAA C
Avr II	LipDH1 GFP Rev	CGCG **CCTAGG** GTT CAT AAT ATT TTG TGT ACT TCC
Bgl II	LipDH2A GFP For	GCGC **AGATCT** TAGA **ATG** TAT ATA TTT TTT CCT TTT TTT TGT AGC
Avr II	LipDH2A GFP Rev	CGCG **CCTAGG** AGT ATG TTC ATC TAC TAA ACT CC
Bgl II	LipDH2B GFP For	GCGC **AGATCT** **ATG** AAC AGC GTT ATT TTT AGA GC
Avr II	LipDH2B GFP Rev	CGCG **CCTAGG** TCC ATC AGA TAA ATT ACC CAT GCA C
Bgl II	KGDH E2A GFP For	GCGC **AGATCT** TAGA **ATG** TGT GTG TGT ATT TTT TTT TTT TTT TTT TTG TTA CC
Avr II	KGDH E2A GFP Rev	CGCG **CCTAGG** GAC TCT TCT TTC TGT TCT TTC ATT TAT TG
Bgl II	KGDH E2B GFP For	GCGC **AGATCT** **ATG** ACA AAG AAT CTT GTA TTT CG
Avr II	KGDH E2B GFP Rev	GCGC **CCTAGG** TGC ATC TAC CAA AAC AAC ATC TCC
Bgl II	BCKDH E1β GFP For	GCGC **AGATCT** ATG ATG AGA CTA TTA AGA AAT AAC G
Avr II	BCKDH E1β GFP Rev	GCGC **CCTAGG** TCC ATT TTC AGC TAG ACC TAT AGC AAA TCC
Bgl II	BCKDH E2 GFP For	GCGC **AGATCT** **ATG** TTT GTG AAG AAT GTA CTA AAC GTG CTT AGG
Avr II	BCKDH E2 GFP Rev	GCGC **CCTAGG** CTC ACA GAA ATA TGA TCC AAC TTT TAA C
Bgl II	PDH E1α GFP For	GCGC **AGATCT** **ATG** CTG TTA GTA ATA TGG TTA AAC ATT TTT TGG
Avr II	PDH E1α GFP Rev	CGCG **CCTAGG** GCC GTG TTT TAT GGC ACA CAA TGA TTG TAC ATC TTT TCC
Bgl II	PDH E2 GFP For	GCGC **AGATCT** TAGA **ATG** TTA TAC AAC TTA ATT ATA TTA ATT TTT TAT TTA AG
Avr II	PDH E2 GFP Rev	CGCG **CCTAGG** GTC CCC ATC ATT ATA TTT CTT TTC ATC C

KADH RT PCR oligonucleotides

Hind III	LipDH1 For	GCGC **AAGCTT** ATG GTC ATA AGG CAA AAT ATT AAA C
Bam HI	LipDH1 Rev	GCGC **GGATCC** TAG TGA GTT CTT ATT TTT GAT ATA GAT TTA AAA GC
Nde I	LipDH2 For	GCGC **CATATG** TCA ACT AAG AAA GAC TAT GAT GTT ATA GTC ATT GG
Bam HI	LipDH2 Rev	**GGATCC** TTA CAT GTG TAT AGG TTT ATC
Nhe I	KGDH E1 For	**GCTAGC** ATG CTT GAA AAA GGG AAA ACA G
Hind III	KGDH E1 Rev	**AAGCTT** TTA TAA CGC ATC GGT ATA TTT TTG
Hind III	KGDH E2 For	GCGC **AAGCTT** ATG TCC ATA GAA ACC ATT AAA GTA CC
Bam HI	KGDH E2 Rev	CGCG **GGATCC** TTA ACA ATC AAT TAA CAT TAG ATT AGG
Nde I	BCKDH E1α For	GCGC **CATATG** AGA AAT ATT GTT CAG AAA TAC TTA CAA AGG
Xho I	BCKDH E1α Rev	CGCG **CTCGAG** TCA TCG CTC AAA TTT TGA TGT ATC
Nde I	BCKDH E1β For	GCGC **CATATG** ATG AGA CTA TTA AGA AAT AAC G
Xho I	BCKDH E1β Rev	CGCG **CTCGAG** TTA CTT CAT CAT TTT TTT GAC TTC G
Nde I	BCKDH E2 For	GCGC **CATATG** TTT GTG AAG AAT GTA CTA AAC G
Xho I	BCKDH E12 Rev	GCGC **CTCGAG** TTA TTC CAA TAG TGG TCC TAG GGA TGC
Bam HI	PDH E1α For	GCGC **GGATCC** CGATG TTT AAT TAC GTT TTA TGT GTA GG
Apa I	PDH E1α Rev	CGCG **GGGCCC** TTA ATC TAT TAT TAA GG
Bam HI	PDH E1β For	GCGC **GGATCC** CG ATG GGG AGA AAA AGA AAC
Apa I	PDH E1β Rev	CGCG **GGGCCC** TCA AGA TGA TAG CGA ATG AAG

KADH RT PCR oligonucleotides (Cont.)

PDH E2 For	GTA TTT CTA AGA ACA ATA ATT ACG G
PDH E2 Rev	CTA TAA AAA TAT TTT CAT AAT ATC C

Sequencing oligonucleotides

pHH2 sequencing	CAA GTG TTG GCC ATG GAA C
pHH1 3' sequencing	TTA CAG TTA TAA ATA CAA TCA ATT GG
pQE30 sequencing	CCC GAA AAG TGC CAC CTG

Table 2.1 – Oligonucleotides used during this study

Nucleotides coloured green in are start codons, those highlighted red are stop codons and sites for restriction endonucleases are highlighted in bold. Nucleotides that have been changed in order to manifest site directed mutagenesis are underlined.

2.2.3.3 Reverse transcriptase (RT-PCR)

P. falciparum cDNA was reversed transcribed from RNA isolated from *P. falciparum* strain 3D7 using Superscript II (Invitrogen). All solutions were stored and all additions occurred on ice. 10 µg RNA was incubated at 25°C for 15 min with final concentrations of 5 mM DTT, 1.25x superscript II buffer and 50 units DNase I (Invitrogen) in DEPC treated water with a final volume of 20 µl. The DNase I in the reaction mix was inactivated by incubation at 70°C for 5 min, before the solution was cooled on ice for 10 min. 250 ng of random hexamers (Invitrogen) were added, mixed well and denatured at 70°C for 5 min followed immediately by incubation on ice and addition of dNTP mix (to a final concentration of 0.5 mM for each dNTP) and 200 units of Superscript II (Both Invitrogen). This solution was incubated at 50°C for 45 min and stored at 4°C for second strand synthesis.

Second strand synthesis was carried out using oligonucleotides specific for the components of the *P. falciparum* KADH (see table 2.1 for details of oligonucleotides used for these reactions). 1 µl of the products from the first strand synthesis was used with 50 ng of forward and reverse oligonucleotides, in a 25 µl reaction volume with PCR supermix supplemented with *Pfx* polymerase (Invitrogen). A gradient PCR was set up to enable a range of annealing temperatures between 40 and 65°C, with the PCR reaction run as follows:

Step 1	95 °C, 3 minutes
Step 2	95 °C, 1 minute
Step 3	40-65 °C, 30 seconds
Step 4	60 °C, 2 minutes per kb

Repeat steps 2-4 for 30 cycles

Step 5	60 °C, 4 minutes per kb
Step 6	4 °C, indefinitely

PCR products were analysed by agarose gel electrophoresis, and were cloned into the Zero blunt TOPO PCR kit (Invitrogen) for DNA sequencing using M13F and M13R sequencing primers.

2.2.3.4 PCR from *P. falciparum* genomic DNA

Genes were amplified from genomic DNA (gDNA) isolated from parasites as described in section 2.2.1.4. 100 ng gDNA was mixed with 200 ng of forward and reverse oligonucleotides in *Pfx* supermix or PCR supermix (Invitrogen). The reaction was performed in a gradient PCR as detailed below:

Step 1	95°C, 3 minutes
Step 2	95°C, 30 seconds
Step 3	40 - 60°C, 30 seconds
Step 4	60°C, 2 minutes per kb to be amplified

Repeat steps 2 – 4 for 30 cycles

Step 5	60°C, 4 minutes per kb to be amplified
Step 6	Store at 4°C

The resulting PCR products were analysed by agarose gel electrophoresis before being cloned into the Zero blunt TOPO PCR cloning kit or TOPO TA cloning kit (Invitrogen; figures 2.5 and 2.6). DNA sequences were verified using M13F and M13R sequencing primers, with reactions being performed by The Sequencing Service, University of Dundee (www.dnaseq.co.uk). Clones with correctly verified sequences were sub-cloned into their relevant destination vector as described in section 2.2.4.2.

2.2.3.5 Site directed mutagenesis

Site directed mutagenesis was carried out according to the method of Papworth *et al.* (Papworth, 1996). The method requires both forward and reverse oligonucleotides (containing the required mutation) of the same stretch of sequence; this allows the PCR of the entire template plasmid in both directions, producing complete plasmids containing the desired mutation. The template used in the mutagenesis PCR was the pJC40 *Pf*TrxR wild type construct described previously (Gilberger *et al.*, 1997).

10 ng of double stranded plasmid DNA (pJC40 *Pf*TrxR) was used with 50 ng of forward and reverse oligonucleotides, in a 25 µl reaction volume with Accuprime *Pfx* supermix (Invitrogen). A gradient PCR was set up to enable a range of annealing temperatures between 40 and 65°C. The mutagenesis reaction was run in the gradient PCR block as follows:

Step 1	95 °C, 5 minutes
Step 2	95 °C, 30 seconds
Step 3	40-65 °C, 15 seconds
Step 4	68 °C, 5 minutes

Repeat steps 2-4 for 15 cycles

Step 5	68 °C, 10 minutes
Step 6	4 °C, indefinitely

The PCR products were subsequently treated with 20 units *Dpn* I at 37°C overnight, to digest any template DNA. *Dpn* I is a restriction endonuclease that is specific to the sequence GATC, but only when adenine is methylated. This methylation occurs in the bacterial host of the template plasmid, but not in the products of the PCR; therefore digestion with *Dpn* I degrades only the template DNA and not the PCR products thus resulting in a positive selection for the mutated plasmid.

The *Dpn* I-digested PCR products were then transformed into *E. coli* strain TOP 10 cells (Invitrogen), selected on Luria-Bertani plates containing 100 µg/ml Ampicillin (LB/Amp) and incubated overnight at 37°C. Single colonies were selected and DNA was extracted using the Qiagen miniprep kit (Qiagen), with mutants determined by DNA sequence analyses using relevant sequencing primers (table 2.1).

2.2.3.6 Colony PCR

Colony PCR was necessary for the identification of correctly ligated clones from the ligation of small inserts (up to 300 bp) from TOPO clones into the 7.3 kb destination vector pHH2. Therefore, transformation plates that contained more than one hundred colonies were analysed by colony PCR to identify possible positive clones. Colony PCR uses a single bacterial colony as the template for the polymerase reaction that is

performed using Reddymix (ABgene). Reddymix is a *Taq*-based polymerase that contains a DNA loading dye that allows direct loading of agarose gels after completion of the PCR. An extended first denaturation step assists in release of DNA from the colony allowing the PCR reaction to proceed using oligonucleotides that are specific for insert (the forward oligonucleotide) and vector (the reverse oligonucleotide). By using this combination of oligonucleotides only a colony containing a correctly ligated insert will produce a PCR product.

The PCR was set up using a stock solution for the PCR reaction mix, the total volume of this depended on the number of colonies to be screened, with 10 µl of the stock solution used per colony:

	1 colony	20 colonies	60 colonies
Oligonucleotide F (0.1 mg/ml)	0.5	10	30
Oligonucleotide R (0.1 mg/ml)	0.5	10	30
2 x Reddymix	4.5	90	270
H_20	4.5	90	270
Total Volume	10 µl	200 µl	600 µl

Once the mix was made it was aliquoted into numbered PCR tubes, and a fresh LB/Amp plate was labelled with a numbered grid corresponding to the numbered tubes. A single colony was selected from the transformation plate using a sterile 2 µl tip, and the colony inoculated onto the gridded plate at the relevant position (i.e. position 1) before being placed into the relevant tube (i.e. tube 1). This was repeated for all colonies that were to be screened, with particular attention being paid to matching the grid position with the PCR tube. The tips were removed from the tubes and discarded after 5 mins of incubation at room temperature. Appropriate negative controls were necessary to disregard unspecific binding, for instance by using a colony from an unrelated cloning experiment. The gridded plate was incubated overnight at 37^0C and the PCR reaction was run as follows;

Step 1 94 °C, 7 minutes

Step 2 94 °C, 60 seconds

Step 3 45 °C, 60 seconds

Step 4 60 °C, 2 minutes per kb to be amplifed

Repeat steps 2-4 for 29 cycles

Step 5 60 °C, 3 minutes per kb to be amplified

Step 6 4 °C, indefinitely

PCR products were analysed by agarose gel electrophoresis, with possible positive clones identified by the presence of a PCR product at the correct size. Positive clones were subsequently analysed by DNA miniprep, restriction digestion and DNA sequencing.

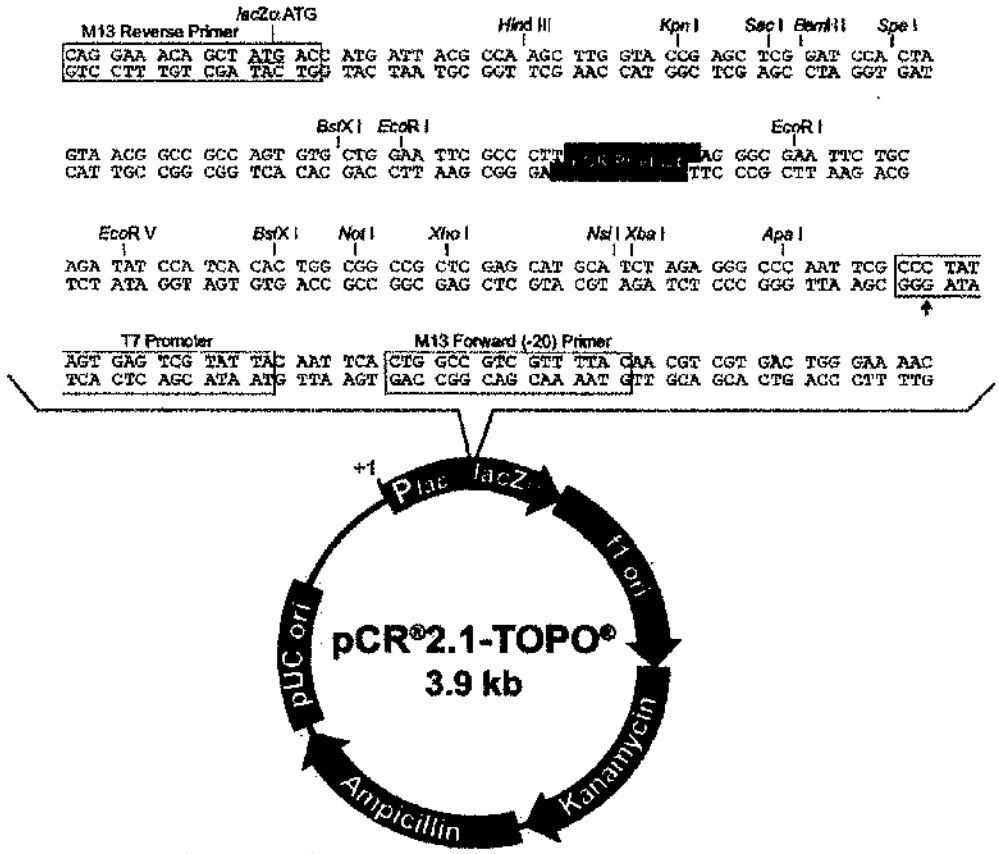
2.2.4 Cloning techniques

2.2.4.1 TOPO cloning of PCR products

All PCR products were initially cloned using a TOPO PCR cloning kit (Invitrogen); two types of vectors were used depending on the polymerase used. PCR products amplified using *Taq* based polymerases produce adenine overhangs on the 3' ends of PCR products and required the TOPO TA® PCR cloning kit (Invitrogen; figure 2.5). While products generated using *Pfx/Pfu* based polymerases amplify as blunt ended products and required the Zero Blunt® TOPO® PCR cloning kit (Invitrogen; figure 2.6). Both kits provide linearised vectors with topoisomerase I covalently bound to the 3' end of the vector. This topoisomerase catalyses the ligation of the PCR product into the vector at room temperature in 5-10 mins, and disrupts the *ccdB* lethal gene, allowing for further selection of correctly ligated products. The ligation reaction was performed as per manufacturer's instructions and was transformed into chemically competent *E. coli* TOP10 cells (Invitrogen) and selected on LB plates containing 50 µg kanamycin/ml.

A number of single bacterial colonies were selected from the transformation plates and plasmid DNA was isolated using the Qiaprep spin miniprep kit (Qiagen) following the manufacturer's guidelines. 500 ng of isolated DNA was digested using 4 units *Eco*RI at 37°C for 2-3 hours before being analysed by agarose gel electrophoresis. Possible

positive clones were determined by the presence of the correct banding pattern and verified by DNA sequencing using M13 F, M13 R or internal sequencing primers.

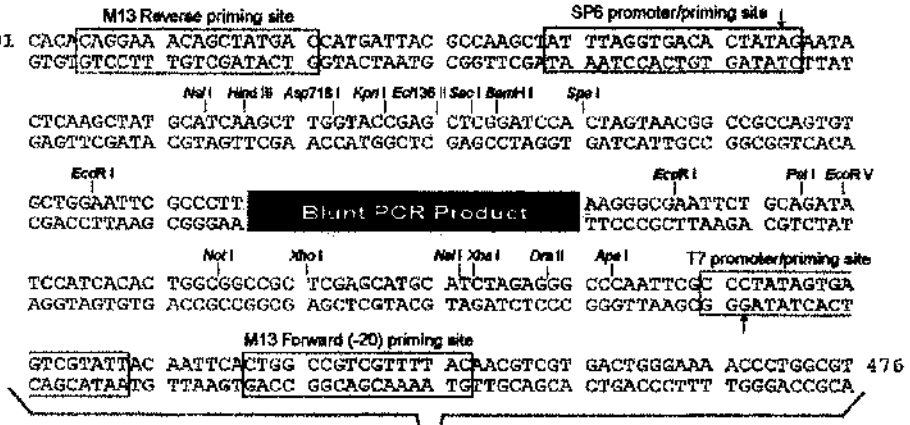


Comments for pCR®2.1-TOPO®
3931 nucleotides

LacZα fragment: bases 1-547
M13 reverse priming site: bases 205-221
Multiple cloning site: bases 234-357
T7 promoter/priming site: bases 364-383
M13 Forward (-20) priming site: bases 391-406
f1 origin: bases 548-985
Kanamycin resistance ORF: bases 1319-2113
Ampicillin resistance ORF: bases 2131-2991
pUC origin: bases 3136-3809

Figure 2.5 – pCR2.1- TOPO plasmid (TOPO TA cloning)

This figure displays the important features of the pCR2.1 TOPO plasmid, which was used to clone PCR products amplified with Taq based polymerases. Key features of this plasmid include the ampicillin and kanamycin resistance cassettes, the multiple cloning site and the position of M13 forward and reverse primers that were used for sequence verification by DNA sequencing.



Comments for pCR®-Blunt II-TOPO®
3519 nucleotides

lac promoter/operator region: bases 95-216
M13 Reverse priming site: bases 205-221
LacZ-alpha ORF: bases 217-576
SP6 promoter priming site: bases 239-256
Multiple Cloning Site: bases 269-399
TOPO®-Cloning site: bases 336-337
T7 promoter priming site: bases 406-425
M13 (-20) Forward priming site: bases 433-448
Fusion joint: bases 577-585
ccdB lethal gene ORF: bases 586-888
kan gene: bases 1099-2031
kan promoter: bases 1099-1236
Kanamycin resistance gene ORF: bases 1237-2031
Zeocin resistance ORF: bases 2238-2612
pUC origin: bases 2724-3397

Figure 2.6 – pCR-Blunt II- TOPO plasmid (TOPO cloning)

This figure displays the important features of the pCR-Blunt II TOPO plasmid, which was used to clone PCR products amplified with polymerases with proof reading abilities (e.g. Pfx supermix). Key features of this plasmid include the kanamycin resistance cassette, the multiple cloning site and the position of M13 forward and reverse primers that were used for sequence verification by DNA sequencing.

2.2.4.2 Sub-cloning into destination vectors

TOPO clones which have had their sequences verified by DNA sequencing were then sub-cloned into their respective destination plasmid. The inserts were digested from the TOPO vector, gel purified and ligated together. The resulting ligation reactions were transformed into either *E. coli* strains JM109, TOP10 or XL-10 gold (see section 2.2.7.1).

Restriction digests of 3 µg of isolated DNA (from both destination vector and TOPO clone plasmids) were performed overnight using 10 units each of the relevant restriction enzymes (table 2.1). The digests were separated by agarose gel electrophoresis and digested products (inserts or plasmids) were isolated by gel purification.

Gel purification of digested DNA was performed on both destination vector and insert bands excised following agarose gel electrophoresis using the Qiaquick gel extraction kit (Qiagen) according to manufacturers guidelines. Samples were eluted from the gel purification columns in 30 µl sterile distilled water, with 3 µl being analysed for purity and visually quantified by agarose gel electrophoresis.

Ligation of isolated vector and insert DNA were performed using T4 DNA ligase (NEB) with molar vector to insert ratios ranging from 1:3 to 1:10. Ligation reactions were incubated in a 14°C waterbath overnight, and were subsequently transformed into *E. coli* TOP10 cells and selected on LB/Amp plates. Possible positive clones were identified by DNA miniprep and digested with the relevant restrictions enzymes (table 2.1). Clones with the correct destination vector and insert band size were verified by DNA sequencing.

2.2.5 Recombinant protein expression and purification

In order to maximise the parameters for recombinant protein expression in *E. coli*, test expressions were set up and analysed using the Bugbuster protein extraction reagent (Novagen). Single colonies were selected and grown in 2 ml LB (containing the relevant antibiotic) cultures overnight at 37 °C with shaking at 200 rpm. Overnight cultures were diluted 1:50 into 10 ml LB/Amp (in 50 ml falcon tubes) and were incubated at 37 °C with shaking at 200 rpm. Once the cultures reached mid-log phase (OD_{600} = 0.4 - 0.8), the expression of the proteins was induced by adding IPTG (Melford) to a final concentration of 1mM. The cultures were then incubated overnight

at various temperatures (37 °C, 25 °C and 15 °C) with 200 rpm shaking. 1 ml aliquots of cultures were removed immediately after induction, 4 h after induction and after overnight expression. These 1 ml aliquots were centrifuged at 13,000 rpm (Eppendorf microcentrifuge with F45-24-11 rotor), 4 °C for 4 mins and the supernatant was discarded. The resulting cell pellets were stored at – 20 °C until further analyses was performed.

Bacterial pellets were resuspended in 50 µl Bugbuster protein extraction reagent containing 0.5 µl benzonase (Novagen). The Bugbuster solution contains a mixture of non-ionic detergents to aid the lysis of the bacterial cells, while the benzonase degrades the DNA and RNA released following this lysis. The cell lysates were mixed at room temperature for 15 min before being centrifuged at 13,000 rpm (Eppendorf microcentrifuge with F45-24-11 rotor), 4 °C for 20 min. The soluble fraction (supernatant) was transferred to a fresh tube and 10 µl of 6 x SDS loading dye was added. The pellet was resuspended in 50 µl 1 x SDS loading dye. The soluble and pellet mixes were heated for 5 min at 100 °C, and then separated by SDS-PAGE followed by Coomassie staining or western blotting (see sections 2.2.7.7 and 2.2.7.9).

2.2.5.1 Thioredoxin reductase (PfTrxR)

PfTrxR (wild type and mutants) was expressed as described previously (Gilberger *et al.*, 1997) with modifications. pJC40 cloned expression constructs were transformed into *E. coli* BL21 (DE3) chemically competent cells (Stratagene) for expression. Starter cultures were inoculated from a glycerol stock and were incubated for 18 h at 37°C with constant shaking at 200 rpm. The starter cultures were diluted 1:50 into 1L LB media containing 50 µg ml^{-1} ampicillin in a 2L Erlenmeyer flask and incubated at 37°C with shaking at 200 rpm. The recombinant expression was then induced by the addition of 1 mM IPTG when the cultures had reached an OD_{600} of 0.4 – 0.8, before being incubated for 18 h at 37°C with constant shaking. Bacteria were pelleted by centrifugation at 4,000 rpm (Sigma 6K15 centrifuge with 12500 rotor), 15 min, 4°C and pellets were resuspended in Qiagen lysis buffer and were stored at –80°C for purification. Recombinant PfTrxR was purified using Ni-NTA batch purification (see section 2.2.6.1) and was quantified using an extinction coefficient of 11,300 mM^{-1} cm^{-1} at 460 nm (except for PfTrxR C88A and C93A mutants, which used an extinction coefficient of 12,800 mM^{-1} cm^{-1} at 451 nm and 12,620 mM^{-1} cm^{-1} 456 nm, respectively).

2.2.5.2 Thioredoxin (PfTrx)

PfTrx was expressed and purified as described previously (Krnajski *et al.*, 2001) with some alterations. pJC40 cloned PfTrx was expressed in *E. coli* BL21 (DE3) by inoculating a LB/Amp starter culture from a glycerol stock and incubating it at 37°C overnight with shaking at 200 rpm. The starter culture was diluted 50 fold in a 1L LB/Amp culture (in 5L Erlenmeyer flask) and incubated at 37°C with shaking at 200 rpm until the culture reached an OD_{600} of between 0.4 and 0.8. The expression of recombinant PfTrx was induced by the addition of 1mM IPTG and the culture was incubated for a further 14 hours at 28°C, with shaking at 200 rpm. The bacterial culture was harvested at 3000 rpm (Sigma 6K15 centrifuge with 12500 rotor), 4°C for 20 mins and the resultant bacterial pellet was resuspended in Qiagen lysis buffer (10 ml of buffer per L of culture) and stored at –80°C until purification. Recombinant PfTrx was purified by Ni-NTA batch purification (see section 2.2.6.1), with purified PfTrx concentration being measured spectrally at 280 nm using the extinction coefficient 11,700 mM^{-1} cm^{-1} (Krnajski *et al.*, 2001) and was confirmed using the Bradford assay.

2.2.5.3 LipDH1

LipDH1 was recombinantly expressed in the pQE30 expression vector producing a recombinant protein with a $His_{(6)}$-tag at the N-terminus. The expression construct was transformed into chemically competent *E. coli* Novablue cells, and single colonies were used to produce a starter culture in LB/Amp media, which was incubated at 37°C overnight with shaking at 200 rpm. The starter culture was diluted 50 fold into 1L LB/amp medium (5L Erlenmeyer flasks), and incubated at 37°C with shaking at 200 rpm until OD_{600} reached 0.4 – 0.8. Expression of the recombinant protein was induced with the addition of 1mM IPTG and the cultures were incubated for a further 14 hours at 30°C. Bacteria were pelleted by centrifugation at 4,000 rpm (Sigma 6K15 centrifuge with 12500 rotor), 15 min, 4°C and pellets were resuspended in Qiagen lysis buffer and were stored at –80°C for purification. Recombinant LipDH1 was purified by Ni-NTA batch purification (see section 2.2.6.1) and purified LipDH1 was quantified spectrally using an extinction coefficient of 11,300 mM^{-1} cm^{-1} at 460 nm.

2.2.5.4 LipDH2

LipDH2 expression was carried out in *E. coli* BL21 (DE3) RIL cells transformed with the pJC40 cloned LipDH2 expression construct. A single colony was inoculated into a starter culture containing LB/Amp media and was incubated overnight at 37°C with shaking at 200 rpm. The starter culture was diluted 50 fold into 1L LB/Amp media (in 5L Erlenmeyer flasks) and incubated at 37°C with shaking at 200 rpm until the culture reaches an OD_{600} of 0.4 – 0.8. Expression of recombinant protein was induced by the addition of 1mM IPTG and the cultures were incubated for a further 14 hours at 37°C with shaking at 200 rpm. The bacterial culture was harvested at 3000 rpm (Sigma 6K15 centrifuge with 12500 rotor), 4°C for 20 min and the resultant bacterial pellet was resuspended in Qiagen lysis buffer (10 ml of buffer per L of culture) and stored at –80°C until purification. Recombinant LipDH2 was purified in a two-step method using Ni-NTA affinity chromatography. The first step utilised Ni-NTA batch purification (see section 2.2.6.1), with the elution sample being extensively dialysed into Qiagen lysis buffer. The second step utilised a Ni-NTA purification with a linear imidazole gradient performed on a FPLC (see section 2.2.6.2) to obtain purified LipDH2. The purified protein was quantified spectrally using an extinction coefficient of 11,300 mM^{-1} cm^{-1} at 460 nm.

2.2.5.5 α–Ketoglutarate succinyltransferase (KGDH-E2)

The recombinant expression of KGDH-E2 was performed in *E. coli* BL21 (DE3) cells (Novagen) transformed with the pET28 KGDH-E2 expression construct. A single colony was used to inoculate a LB/Kan starter culture that was incubated overnight at 37°C with shaking at 200 rpm. The starter culture was diluted 50 fold in a 1L LB/Amp culture that was incubated at 37°C for 24 hours, with the addition of 1mM IPTG when the culture reached an OD_{600} of between 0.4 and 0.8. The bacterial culture was harvested at 3000 rpm (Sigma 6K15 centrifuge with 12500 rotor), 4°C for 20 mins and the resulting pellet was resuspended in Qiagen lysis buffer (10 ml per 1L of bacterial culture) and stored at –80°C for purification. Recombinant KGDH-E2 proteins were purified using Ni-NTA batch purification (see section 2.2.6.1) and purified proteins were quantified using Bradford assay (see section 2.2.7.6).

2.2.6 Purification techniques

Bacterial pellets were flash thawed in running warm water to lyse bacterial cells before being incubated on ice for 30 min with 50 μg ml^{-1} lysozyme. 1mM of the protease inhibitor PMSF (and 100μM FAD for purifications of disulphide oxidoreductases) was added to the bacterial lysate, before centrifuging at 20,000rpm (Beckman J2-H5 centrifuge with a JA-20 rotor), 4°C for 60 min. The resulting bacterial supernatant was removed and filtered through a 0.4 μm syringe filter before the recombinant protein was purified by one of the methods below.

2.2.6.1 Batch Nickel affinity chromatography

Batch Ni-NTA purification was performed using Ni-NTA agarose (Qiagen) according to manufacturers instructions with some modifications. Briefly, 1 ml Ni-NTA agarose per litre of bacterial culture was mixed with the filtered bacterial supernatant (see previous section) at 4°C for 60 min using a blood wheel. The lysate/Ni-NTA mix was poured through an econopac column (Biorad) and the flow-through fraction was collected by gravity flow. The resulting Ni-NTA pellet was washed twice in 5 pellet volumes of Qiagen lysis buffer (which contains 10 mM Imidazole). A subsequent wash was performed with 5 pellet volumes of Qiagen wash buffer (containing 20 mM Imidazole) and a third wash (only for proteins expressed by pJC40) was performed in 50 mM Imidazole. The recombinant proteins were eluted from the Ni-NTA beads by Qiagen elution buffer (containing 500 mM Imidazole) and collected by gravity flow. Eluted proteins were quantified spectrally as described in section 2.2.5. All fractions (bacterial pellet, bacterial lysate, Ni-NTA flow-through, washes and elutions) were stored at 4°C for analysis by SDS-PAGE or western blot following completion of purification. LipDH2 was further purified by FPLC nickel chromatography.

2.2.6.2 Denaturing Nickel affinity chromatography

Denaturing Ni-NTA purification was performed using Ni-NTA (Qiagen) according to manufacturers guidelines with some modifications. Recombinant expression cultures were harvested as described in section 2.2.6.1, but were resuspended in denaturing Ni-NTA purification buffer (at pH 8.0) rather than Qiagen lysis buffer. This suspension was mixed for between 15 – 60 min at room temperature with occasional vortexing. The lysate was then centrifuged at 20,000rpm (Beckman J2-H5 centrifuge with a JA-20 rotor), 4°C for 30 min to remove cellular debris. Ni-NTA (1 ml per litre of culture) was

added to the resulting supernatant and incubated at room temperature of constant shaking (an a rotary shaker) for 60 min. The Ni-NTA/lyaste mix was then applied to an econopac column (Biorad) and the flow-through was collected by gravity flow. The resulting Ni-NTA pellet was washed twice with 2 pellet volumes of denaturing Ni-NTA purification buffer at pH 6.3 and once with 2 pellet volumes of denaturing Ni-NTA purification buffer at pH 5.9. The recombinant protein was eluted in three times in 2 pellet volumes of denaturing Ni-NTA purification buffer at pH 4.4. All fractions were collected to be analysed by SDS-PAGE.

2.2.6.3 FPLC Nickel affinity chromatography

FPLC nickel affinity chromatography used HiTrap chelating HP columns (either 1 ml or 5ml; Amersham biosciences). HiTrap columns were washed sequentially with sterile distilled water, 100 mM $NiSO_4$, sterile distilled water and Qiagen lysis buffer before protein samples were pre-loaded onto a HiTrap chelating column (Amersham biosciences). The column was then attached to an ÄKTA FPLC (Amersham biosciences), washed with five column volumes of Qiagen lysis buffer before the sample was eluted using a linear gradient of 10 mM – 500 mM imidazole. The linear gradient of imidazole was prepared by the ÄKTA FPLC (Amersham biosciences), by mixing Qiagen lysis buffer (10mM imidazole) with Qiagen Elution buffer (500 mM imidazole). The elution profile of the proteins loaded onto the HiTrap chelating HP column was followed spectrally at 280 nm. Samples from OD_{280} peaks were analysed by SDS-PAGE and western blotting.

2.2.6.4 Gel filtration chromatography

Gel filtration chromatography was employed in order to investigate the oligomeric state of the recombinantly expressed protein. A Sephadex S200 16/60 or Superose 6 10/30 column was attached to an ÄKTA FPLC (All Amersham biosciences) and pre-equilibrated with 50 mM potassium phosphate pH 7.6, 150 mM NaCl with a flow rate of 1 ml min^{-1}. The protein elution profile of samples was followed at 280 nm. Before samples were separated the columns were calibrated in the same buffer with gel filtration standards (Biorad), with elution profiles of the standards used to produce a calibration curve. The calibration curve was produced by plotting the log_{10} of the proteins molecular mass (on the y-axis) against the Ve/Vo (on the x-axis), where Ve is the elution volume of the protein and Vo is the void volume of the column. This

resulted in the calibration curves of y = 7.9844 + (-1.8492x) for the Sephadex S200 16/60 and y = 7.8235 + (-1.5140x) for the Superose 6 10/30 column.

1.4 mg of the purified LipDH1 protein was loaded onto the Superose 6 10/30 column (Amersham biosciences) and 300 µg of the purified LipDH2 protein sample was loaded onto the Sephadex S200 16/60 column (Amersham biosciences) and the resulting elution profiles were used to calculate the apparent molecular size of the proteins using the calibration curves.

2.2.7 Molecular biology techniques

2.2.7.1 *E. coli* transformation

Plasmid DNA was transformed into chemically competent *E. coli* strains JM109 (Stratagene), XL10 gold (Stratgene) and TOP10 (Invitrogen) for general cloning and sub-cloning; BL21(DE3) and BL21(DE3) RIL (both Stratagene) and Novablue cells (Novagen) were used for recombinant expression. The genotypes of the *E. coli* strains used are displayed in table 2.2. Transformations were performed to manufacturers guidelines with modifications. Transformations were performed with either whole ligation reactions (maximum of 30 µl) or 100 ng of destination plasmid DNA, with transformations being plated onto LB agar plates with the appropriate antibiotic selection and incubated overnight at 37°C. Possible positive clones were identified by DNA miniprep and digested with the relevant restrictions enzymes (table 2.1). Clones with the correct destination vector and insert band size were verified by DNA sequencing.

BL21 (DE3) strain	*E. coli*B F- *dcm ompT hsd*S(r_B – m_B-) gal λ (DE3)
BL21 (DE3) RIL strain	*E. coli* B F– *ompT hsdS*(r_B– m_B–) *dcm*+ Tet^r *gal λ* (DE3) *endA* Hte *[argU ileY leuW* Cam^r]
XL10 Gold strain	Tet^r *Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* Hte [F' *proAB* $lacI^qZ\Delta M15$ *Tn10* (Tet^r) Amy Cam^r][a]
JM109 strain	*e14-(McrA-) recA1 endA1 gyrA96 thi-1 hsdR17*(rK- mK+) *supE44 relA1 Δ(lac-proAB)* [F' *traD36 proAB* $lacI^qZ\Delta M15$]
Novablue (DE3)	*endA1 ksdR17*(r_{K12}^- m_{K12}^+) *supE44 tki-1recA1 gyrA96 relA1 lac*^{c} (DE3) F' [$proA^+B^+$ *lacI ZΔM15*::Tn*10*] (Tet^R)

Table 2.2 – *E. coli* strains used during this study

The genotypes of all the *E. coli* strains used during this study are given above.

2.2.7.2 Plasmid DNA isolation from *E. coli*

Plasmid DNA was isolated from transformed *E.coli* cultures using two methods that depended on the levels of plasmid DNA required.

Small-scale isolations were carried out using the Qiaprep spin miniprep kit (Qiagen). DNA was isolated from single colonies on transformation plates by inoculating them into 2 ml LB media containing either 50 μg/ml kanamycin (for TOPO clones) or 100 μg/ml ampicillin (for destination clones) antibiotics; these cultures were incubated at 37°C overnight with shaking at 200 rpm. DNA was isolated from the resulting cultures using the Qiaprep spin miniprep kit (Qiagen) by following the manufacturers guidelines, with the DNA eluted in 50 μl sterile distilled water. Miniprep plasmid DNA was analysed by either restriction digests followed by agarose gel electrophoresis or by DNA sequencing using the appropriate primers.

Large-scale isolations were carried out using the Hi-speed plasmid maxi kit (Qiagen) and were performed following the manufacturers guidelines. It was necessary to perform large-scale isolations on constructs, which were due to be transfected into parasites (pHH2 constructs; see section 2.2.3.1). Isolated plasmid DNA was quantified by spectrophotometric analyses, using the absorbance at 260 nm to calculate DNA concentration as described in section 2.2.1.4. Using this kit, 250 ml of culture usually yielded 200 - 400 μg of purified DNA, which provided enough DNA for 2 – 4 transfections (see section 2.2.1.5).

2.2.7.3 Ethanol precipitation of DNA

The precipitation of DNA was required as part of the transfection protocol (to concentrate and buffer exchange the DNA solution) and was also utilised during sub-cloning as a buffer exchange method for restriction enzyme digests with two non-compatible enzymes.

Ethanol precipitation was performed by mixing the DNA solution with 0.1 volumes of 3M NaAc pH 5.2 and 3 volumes of 100% ethanol. The resulting mix was incubated at –20°C or –80°C for at least an hour before being centrifuged at 13,000 rpm (Eppendorf microcentrifuge with F45-24-11 rotor), 4°C for 30 min. The supernatant was carefully removed and the pellet was washed in 70% ethanol before being centrifuged at 13,000 rpm (Eppendorf microcentrifuge with F45-24-11 rotor), 4°C for 10 min and the

supernatant was again carefully removed. The pellet was allowed to air-dry before being resuspended in the appropriate buffer by gently pipetting and storing at 4°C overnight to aid resuspension.

2.2.7.4 DNA sequencing

DNA sequencing was performed by the sequencing service, School of Life Sciences, University of Dundee (http://www.dnaseq.co.uk/home.html) using Applied Biosystems big-dye version 3.1 chemistry on an Applied Biosystems 3730 automated capillary DNA sequencer. This reaction required 200-300 ng of plasmid and 3.2 picomoles of sequencing primer per reaction. All relevant sequencing primers, except for sequencing primers for pQE30 and pHH2 (table 2.1), were provided by the sequencing service.

2.2.7.5 Agarose gel electrophoresis

DNA was routinely examined by electrophoresis in 1.0% (w/v) agarose gel that contained TAE buffer with 0.5 μg/ml ethidium bromide. The DNA sample of interest was mixed with 6 x DNA loading buffer (Promega) and electrophoresis was performed at 100 V for 30-60 min in TAE buffer. 1 kb ladders (Promega) were routinely run alongside samples to allow size determination of the bands visualised by UV illumination at 302 nm or 365 nm. DNA was visualised at 365 nm when subsequent gel extraction was required because under this condition DNA damage was minimised.

2.2.7.6 Determining protein concentration

The protein concentration of purified and isolated proteins was analysed by the Bradford assay (Bradford, 1976), using the Biorad protein assay reagent (Biorad). Protein concentration was observed spectrally as the absorbance at 595 nm and was determined in reference to a standard curve produced by the addition of 1 – 20 μg ml^{-1} bovine serum albumin (BSA) to a 1 x Bradford solution (Biorad).

2.2.7.7 Sodium dodecyl sulphate polyacrylamide gel electrophoresis

Protein samples were separated by sodium dodecyl sulphate polyacrlamide gel electrophoresis (SDS-PAGE) using the NuPAGE Bis/Tris electrophoresis system (Invitrogen). Novex Bis/Tris pre-cast gels (either 10 % or 4 –12 % polyacrylamide) were used in Xcell Surelock™ Mini-cell apparatus with 1 x MOPS running buffer (all

Invitrogen). Proteins to be run on SDS-PAGE were commonly quantified by Bradford assay and gels were loaded with 10 μg protein in 1 x SDS loading buffer following heating at 100°C for 5 min (see section 2.2.7.6). Electrophoresis was performed according to manufacturers guidelines and SDS-PAGE gels were stained by incubation (with shaking) in Coomasie stain for 10 mins and washed in destain.

2.2.7.8 Generation of polyclonal antibodies

Polyclonal antibodies were generated from purified recombinantly expressed proteins by Eurogentec (http://uk.eurogentec.com/code/en/hp.asp). 400 μg of recombinant LipDH2, BCKDH E2 and KGDH E2 proteins were separated by SDS-PAGE, gel excised and sent to Eurogentec using their standard immunisation protocol (Eurogentec). Each bleed was analysed on recombinant protein and *P. falciparum* extract to test for specificity.

2.2.7.9 Western blot analyses

Following protein separation by SDS-PAGE (see section 2.2.7.8) the proteins were transferred to Protran nitrocellulose (Schleicher & Schuell) using a Transblot semi-dry blotting system (Biorad) according to manufacturers instructions. Transfer was in a downward direction in a sandwich of Towbin buffer soaked blotting paper, SDS-PAGE gel, nitrocellulose and blotting paper (from top to bottom in the sandwich). The transfer was carried out at a fixed voltage of 20V for 30 min per gel. The nitrocellulose membrane was subsequently stained using Ponceau S stain (Sigma) to allow analyses of transfer and marking of the gel lanes with a blunt pencil. Ponceau S was removed by washing the blot in PBS with 0.05% Tween 20. Following washing the blot was blocked by incubating (with shaking) in PBS containing 5 % Marvel milk powder for either 60 min at room temperature or overnight at 4°C. Primary antibodies were diluted to the necessary concentration (table 2.3) in 1% (w/v) Marvel in PBS, and incubated on the blot at room temp for 60 min with constant shaking. The blots were then washed in PBS containing 0.05% (v/v) Tween 20 for three washed of 10 min with constant shaking. Secondary antibodies were then incubated at the necessary dilution at room temperature for 60 min in 1% (w/v) Marvel in PBS, before the wash step was repeated as before. The HRP conjugated secondary antibodies were detected using the ECL-plus detection kit (Amersham biosciences), and the blots were exposed to ECL film

(Amersham biosciences) for the necessary time in a hyperfilm cassette (Amersham biosciences).

Primary antibody	Dilution	Secondary antibody	Dilution
anti - $His_{(6)}$ Tag	1 : 5,000	anti – Mouse	1 : 5,000
anti – GFP	1 : 5,000	anti – Rabbit	1 : 5,000
anti - LipDH2	1 : 500	anti – Rabbit	1 : 1,000
anti - KGDH-E2	1 : 100	anti – Rat	1 : 10,000
anti - BCKDH-E2	1 : 500	anti – Rabbit	1 : 1,000

Table 2.3 – Antibody dilutions for western blots

This table displays the dilutions factors of the commonly used primary and secondary antibodies.

2.2.7.10 Immunoprecipitation

Parasites were isolated from IRBC by Saponin lysis (see section 2.2.1.4), before being resuspended in 5 pellet volumes of IP lysis buffer. This was flash frozen at -80°C before being thawed at room temperature, and then incubated on ice for 30 min. The lysed solution was then cleared by centrifugation at 12,000g for 30 min at 4°C. A 50% slurry (v/v) of protein A sepharose beads (Amersham biosciences) in IP wash buffer was then added to the lysate in order to pre-clear the solution of any proteins that un-specifically bind to the beads. Simultaneously, 5 µg of antibody were added to a 50% slurry (v/v) of Protein A sepharose beads in IP wash buffer in a separate tube. Both the pre-cleared and antibody/bead mix were mixed on a blood wheel at 4°C for 60 min. The beads were pelleted at 10,000g, 4°C for 4 min and the supernatant from the antibody/beads incubation was removed and replaced with the supernatant from the pre-clear incubation. This was incubated on a blood wheel at 4°C for 2 hours. The beads were pelleted again by centrifugation at 10,000g and washed in IP wash buffer, this step was then repeated a further 3 times. Finally the beads were washed a further 2 times in PBS and eluted from the beads in 2x SDS loading buffer. Samples were then separated on SDS-PAGE and detected by either Coomassie staining or western blotting.

2.2.8 Spectral analyses of disulphide oxidoreductases

The spectral features of the recombinantly expressed PfTrxR (wild-type and mutants) and LipDH proteins were analysed by reduction with either NADH, NADPH or sodium dithionite (all Sigma).

2.2.8.1 Reductive half-reaction of dihydrolipoamide dehydrogenase

Reduction of LipDH was measured under aerobic conditions by the addition of excess NADH. Spectra between 300 and 700 nm were measured before and immediately after addition of 20 molar excess of NADH to a LipDH solution (11 - 22 μM) on a UV-2401 spectrophotometer (Shimadzu). The LipDH concentration was calculated using the extinction coefficient for protein bound FAD at 460 nm of 11,300 mM cm^{-1} and the NADH concentration was calculated using the extinction coefficient of 6,200 mM cm^{-1} at 340 nm.

2.2.8.2 Reductive half-reaction of thioredoxin reductase

The reductive half-reaction in thioredoxin reductase was analysed under anaerobic conditions in order to counteract the enzyme re-oxidation observed under aerobic conditions. *Pf*TrxR solutions were made anaerobic in an anaerobic cuvette by 8 – 10 cycles of vacuum and ultrapure nitrogen. The solution was left to equilibrate in ultrapure nitrogen before an air tight Hamilton syringe, containing the reductant, was attached to the cuvette. The anaerobic *Pf*Trx R solution was titrated with the reductants NADPH or sodium dithionite at 25°C under anaerobic conditions; each addition was analysed between 350 and 700 nm in a Cary 300 spectrophotometer (Varian). *Pf*TrxR solutions were quantified using the extinction coefficient for protein bound FAD at 460 nm of 11,300 mM^{-1} cm^{-1}, except for PfTrxR C88A and C93A which used the extinction coefficients of 12,800 M^{-1} cm^{-1} at 451 nm and 12,620 M^{-1} cm^{-1} at 456 nm respectively (Gilberger *et al.*, 1997).

2.2.8.2.1 NADPH reduction

Freshly prepared 1.5 mM NADPH was quantified using an extinction coefficient at 340 nm of 6,220 mM^{-1} cm^{-1}, and degassed by bubbling in ultrapure nitrogen for 10 min in a glass syringe. The anaerobic NADPH solution was drawn into an airtight Hamilton

syringe and titrated (in 2 µl additions) into an anaerobic cuvette containing an anaerobic 10 - 15 µM *Pf*TrxR solution.

2.2.8.2.2 Sodium dithionite reduction

Sodium dithionite was freshly prepared and quantified by reduction of lumiflavin-3-acetic acid, before *Pf*TrxR reduction was performed. A 1.5 mM stock solution of sodium dithionite (Sigma) was prepared in anaerobic 50 mM pyrophosphate buffer pH 9.0. Anaerobisis was achieved by 10 cycles of purified nitrogen and vacuum, the buffer was equilibrated with ultrapure nitrogen before being mixed with the sodium dithionite powder.

The dithionite solution was then quantified by anaerobic static titration with lumiflavin-3-acetic acid (LF-3-AA). 1 ml of a 20 µM LF-3-AA solution was prepared in 50 mM pyrophosphate pH 9.0, and was made anaerobic (as in section 2.2.8.2) in an anaerobic cuvette. The dithionite solution was titrated in 5 µl steps using an airtight Hamilton syringe; the dithionite solution was quantified using the reduction of LF-3-AA at 443 nm with an extinction coefficient of 12,400 mM^{-1} cm^{-1}.

The anaerobic dithionite solution was drawn into an airtight Hamilton syringe and titrated into an anaerobic 15 µM *Pf*TrxR solution in an anaerobic cuvette. Dithionite titrations were performed in the presence of 1.5 µM methyl viologen to act as an electron mediator between the dithionite and TrxR.

2.2.9 Steady state kinetic analyses of LipDH

The kinetic parameters of recombinant LipDH proteins were analysed for both forward and reverse reactions by varying the substrates used in the assay. The reactions were measured using a stopped flow fast kinetic system (SFA-20, Rapid kinetics accessory, Hi-Tech scientific) in order to measure initial rates as described previously (Akerman *et al.*, 2003). The reactions were measured using the consumption/production of NADH at 340 nm with the extinction coefficient of 6,220 mM^{-1} cm^{-1}.

2.2.9.1 Forward reaction

The forward reaction of LipDH uses the substrates dihydrolipoamide and NAD^+. Dihydrolipoamide was produced from lipoamide (Sigma) by reduction with 100 fold

molar excess of sodium borohydride, as modified from previous studies (Reed *et al.*, 1958). A stock solution of 60 mM lipoamide was prepared in 100% ethanol, before sodium borohydride was added to the solution to a final concentration of 600 mM and nitrogen was bubbled into the solution until it was clear. The solution was diluted to the necessary concentration in 50 mM potassium phosphate, 1 mM EDTA, pH 8.0, and was used for assays following complete hydrolysis of the sodium borohydride.

NAD^+ was freshly prepared in reaction buffer and quantified at 340 nm using the extinction co-efficient of 6,200 mM^{-1} cm^{-1}. The reaction was performed with 1 µg of protein at 25°C in 50 mM potassium phosphate, 1 mM EDTA, pH 8.0 in order to prevent spontaneous oxidation of dihydrolipoamide. The $k_m^{dihydrolipoamide}$ was determined by holding the NAD^+ concentration constant at 2mM and varying dihydrolipoamide between 50 µM and 2 mM. The k_m^{NAD+} was determined by holding the dihydrolipoamide concentration constant at 2 mM and varying NAD^+ between 62.5 mM and 2mM.

2.2.9.2 Reverse reaction

The reverse reaction of LipDH uses the substrates lipoamide and NADH. Stock solutions of 60 mM lipoamide were prepared in 100% ethanol and it was necessary to heat to 37°C and vortex to make the lipoamide soluble. NADH was freshly prepared in reaction buffer and quantified at 340 nm using the extinction co-efficient of 6,200 mM^{-1} cm^{-1}. The reaction was performed at 25°C in 50 mM potassium phosphate, 1 mM EDTA, pH 7.0. The $k_m^{lipoamide}$ was determined by holding the NADH concentration constant at 200 µM and varying lipoamide between 50 µM and 4 mM. The k_m^{NADH} was determined by holding the lipoamide concentration constant at 2 mM and varying NADH between 10 µM and 200 µM.

2.2.9.3 Analysing pH optimum

The pH optimum of both the forward and reverse reactions were measured using the assays detailed above over a range of pH values. Buffers were prepared between pH 5.5 and 10.0 using a mixture of 50 mM bicine/Bis Tris Propane/MES buffers.

2.2.9.4 Determining catalytic mechanism

The reverse reaction of LipDH was utilised to investigate the catalytic mechanism of LipDH2. Reactions were performed as described previously and measured over a range of lipoamide concentrations (0.5 mM to 2 mM) at three different NADH concentrations (10 μM, 20 μM, 40 μM).

2.2.10 Determination of thioredoxin reductase steady state kinetic parameters

The kinetic parameters of PfTrxR can be analysed using NADPH as reductant and either PfTrx (the natural substrate) or DTNB as electron acceptors. The activity assays were followed spectrophotometrically (Shimadzu UVPC 2501) at diagnostic wavelengths.

2.2.10.1 Thioredoxin/insulin assay

The TrxR activity with the natural substrates NADPH and thioredoxin was determined as described previously (Luthman *et al.*, 1982). The reaction was performed at 25°C in 50 mM potassium phosphate, 1 mM EDTA, 0.2 mg ml^{-1} human insulin (Sigma), pH 7.6, with 1 – 20 μg PfTrxR (depending on mutant), 200 μM NADPH and 50 μM PfTrx. The reaction was followed at 340 nm for the oxidation of NADPH using the extinction coefficient of 6220 M^{-1} cm^{-1}. A 10 mM NADPH stock solution was prepared in 2 mM Tris pH 10.0 and stored at 4°C. The K_m^{PfTrx} was determined by holding the NADPH concentration at 200 μM and varying PfTrx between 2 μM and 50 μM.

2.2.10.2 DTNB assay

The DTNB assay follows the release of TNB anions at 412 nm using the extinction coefficient of 13,600 M^{-1} cm^{-1} per TNB released as described previously (Holmgren, 1977). The reaction was performed at 25°C in 50 mM potassium phosphate, 1 mM EDTA, pH 7.6, with 1 –20 μg PfTrxR (depending on variant assayed), 200 μM NADPH and 3 mM DTNB. The K_m^{DTNB} was determined by holding the NADPH concentration constant at 200 μM and varying DTNB between 25 μM and 3 mM. The K_m^{NADPH} was determined by holding the DTNB concentration constant at 3 mM and varying NADPH between 0.5 μM and 200 μM. A stock solution of 100 mM DTNB was prepared in

DMSO, with the final DMSO concentration in the assay being held constant at 3% (v/v). A 10 mM NADPH stock solution was prepared in 2 mM Tris pH 10.0 and stored at 4°C.

2.2.11 Pre-steady state kinetic analyses of *P. falciparum* TrxR

Pre-steady state kinetics of the *Pf* TrxR reductive and half-reactions were analysed using a SF-61 DX2 double mixing stopped flow spectrophotometer (Hi-tech scientific). The stopped flow spectrophotometer was operated in either photodiode array (Spectra) or photomultiplier (Single wavelength) modes. The kinetic traces collected during the reactions were analysed and quantified using the kinetasyst 3 program (Hi-tech scientific), with data fitted using the Marquadt algorithm for solving differential equations (Press, 1992).

2.2.11.1 Reductive half-reaction

*Pf*TrxR reduction with NADPH was observed in the stopped flow spectrophotometer in both photodiode array (Spectra) or photomultiplier (Single wavelength) modes. The reaction was followed at the wavelengths 355, 440, 540 and 670 nm in photomultiplier mode. All solutions were made anaerobic as described previously (see section 2.2.8.2) in tonometers (anaerobic vessels). The tonometers were attached to the stopped flow spectrophotometer, and the reactions were followed until completion after mixing of 10 - 15 μM *Pf*TrxR (concentration after mixing) with 6 molar equivalents NADPH at 5°C.

2.2.11.2 Oxidative half-reaction

The oxidative half-reaction of *Pf* TrxR was followed in double mixing mode (4 syringes) in the stopped flow spectrophotometer at 5 °C or 25 °C depending on the rate of reaction. All solutions were made anaerobic as described previously in anaerobic tonometers (see section 2.2.8.2), and attached to the stopped flow spectrophotometer. 40 μM *Pf*TrxR was pre-reduced with 2.2 molar equivalents NADPH and was incubated until completion of reductive half-reaction. The reduced *Pf* TrxR (EH_4) was then mixed with between 0.5 and 20 molar equivalents *Pf*Trx, with data being acquired in photomultiplier mode at 490 and 540 nm until the reaction was complete.

Chapter 3 The role of acid-base catalysts in the catalytic mechanism of *Plasmodium falciparum* thioredoxin reductase.

3.1 Introduction

3.1.1 *Plasmodium falciparum* thioredoxin reductase (PfTrxR)

Thioredoxin reductase (TrxR) is a component of the thioredoxin redox system and a member of the disulphide oxidoreductases family that includes glutathione reductase, dihydrolipoamide dehydrogenase, trypanothione reductase and mercuric ion reductase (Williams, 1992). TrxR acts as an electron donor in the thioredoxin redox system and therefore plays an important role in cell proliferation, defence against reactive oxygen species and signal transduction (Powis *et al.*, 2001; Nakamoto *et al.*, 2004; Baier *et al.*, 2005; Kabe *et al.*, 2005). Indeed, gene knockout studies of the *TrxR* gene of *Plasmodium falciparum* demonstrated this protein to be essential for parasite survival during erythrocytic stages (Krnajski *et al.*, 2002), and therefore genetically validated it as a potential drug target.

TrxR from *P. falciparum* (PfTrxR) belongs to the high M_r TrxRs and contains three redox-active centres that are involved in catalysis; a protein bound FAD cofactor, an N-terminal disulphide (Cys88 and Cys93) and a C-terminal disulphide (Cys535 and Cys540). It is at the C-terminal disulphide that the most notable differences can be observed between high Mr TrxRs. In *P. falciparum* TrxR the redox active cysteines are separated by four amino acids, while the *Drosophila melanogaster* TrxR disulphide is formed by two adjacent cysteine residues and the mammalian and worm TrxRs form a selenenylsulphide from adjacent cysteine and selenocysteine residues (Williams *et al.*, 2000).

The catalytic mechanism of high M_r TrxR has been investigated previously with the *D. melanogaster* TrxR (DmTrxR). This resulted in a proposed reaction mechanism for the high M_r TrxRs, and demonstrated that the enzyme cycled between the two (EH_2) and four (EH_4) electron reduced states in catalysis (Bauer *et al.*, 2003b).

The transfer of reducing equivalents between the redox active centres in the related proteins glutathione reductase and dihydrolipoamide dehydrogenase has been shown to require an acid-base catalyst, which is in a conserved motif (HPXXXE) in all disulphide oxidoreductase of this protein family (Benen *et al.*, 1991; Benen *et al.*, 1992; Rietveld *et*

al., 1994). The role of the potential acid-base catalyst has however not been investigated in high M_r TrxR. A recent study by Gromer *et al.* suggested an additional acid-base catalyst present in high M_r TrxR (Gromer *et al.*, 2003). This appeared to be an appealing hypothesis because during catalysis three instead of two dithiol-disulphide interchange reactions take place, which need to be supported by acid-base catalysis.

3.1.2 Aims

The aims of this part of the study were to investigate the catalytic mechanism of the *P. falciparum* TrxR and examine the role of the active site residues and those located in the acid-base catalyst motif in catalysis. A site directed mutagenesis approach was taken to investigate the roles of the chosen residues, whereby analyses of the characteristics of these mutated residues would allow investigation of their role in catalysis. The mutants were to be analysed by steady state kinetics, static anaerobic titrations with reductants and by stopped-flow rapid reaction kinetics of their reductive and oxidative half-reactions.

3.2 Results

3.2.1 Identification of conserved active site residues

In order to analyse the conserved active site residues I performed a ClustalW alignment of the PfTrxR amino acid sequence with those of other high M_r TrxR (figure 3.1). The *P. falciparum* TrxR amino acid sequence displayed 45% sequence identity to the other proteins in the alignment. The amino acid sequences of high M_r TrxR can be separated into three domains; the FAD binding, NADPH binding and interface domains (Sandalova *et al.*, 2001).

The FAD binding domain, which is located between residues 1 – 193 and 320 - 388 contains the conserved FAD binding motif (GGGXXG; highlighted yellow in figure 3.1) for binding of the pyrophosphate moiety of FAD in part of a Rossman-like fold ($\beta\alpha\beta\alpha\beta\alpha$ units that form parallel β sheets; Schierbeek *et al.*, 1989; Mattevi *et al.*, 1991). The N-terminal cysteine pair that is conserved in all the high M_r TrxR (highlighted red in figure 3.1) is also located within this domain. This domain also contains a proposed acid-base catalyst (Gromer *et al.*, 2003), which corresponds to His137 in the *P. falciparum* TrxR and is shown to be conserved in the alignment with high M_r TrxR (highlighted blue in figure 3.1). Alignment of the *P. falciparum* TrxR with *P. falciparum* glutathione reductase and both dihydrolipoamide dehydrogenase (see chapter 4) shows that this residue is not conserved between these related disulphide oxidoreductase and suggests that it may play a role only in the high M_r TrxR (data not shown).

The NADPH domain forms a Rossman-like fold, is located between residues 194 and 319 and is the site of association with the NADPH substrate. Two arginine residues, corresponding to Arg253 and Arg258 in PfTrxR (highlighted pink in figure 3.1), are involved in binding of the 2' phosphate moiety of $NADP^+$ and are therefore imparting substrate specificity (Sandalova *et al.*, 2001). The NADPH domain also contains a conserved tyrosine residue (Tyr145 in PfTrxR; highlighted orange in figure 3.1) that is responsible for shielding the isoalloxazine ring of FAD in the oxidised protein; this residue has been shown to be displaced by the nicotinamide ring of the bound NADPH substrate in the reduced protein (Biterova *et al.*, 2005).

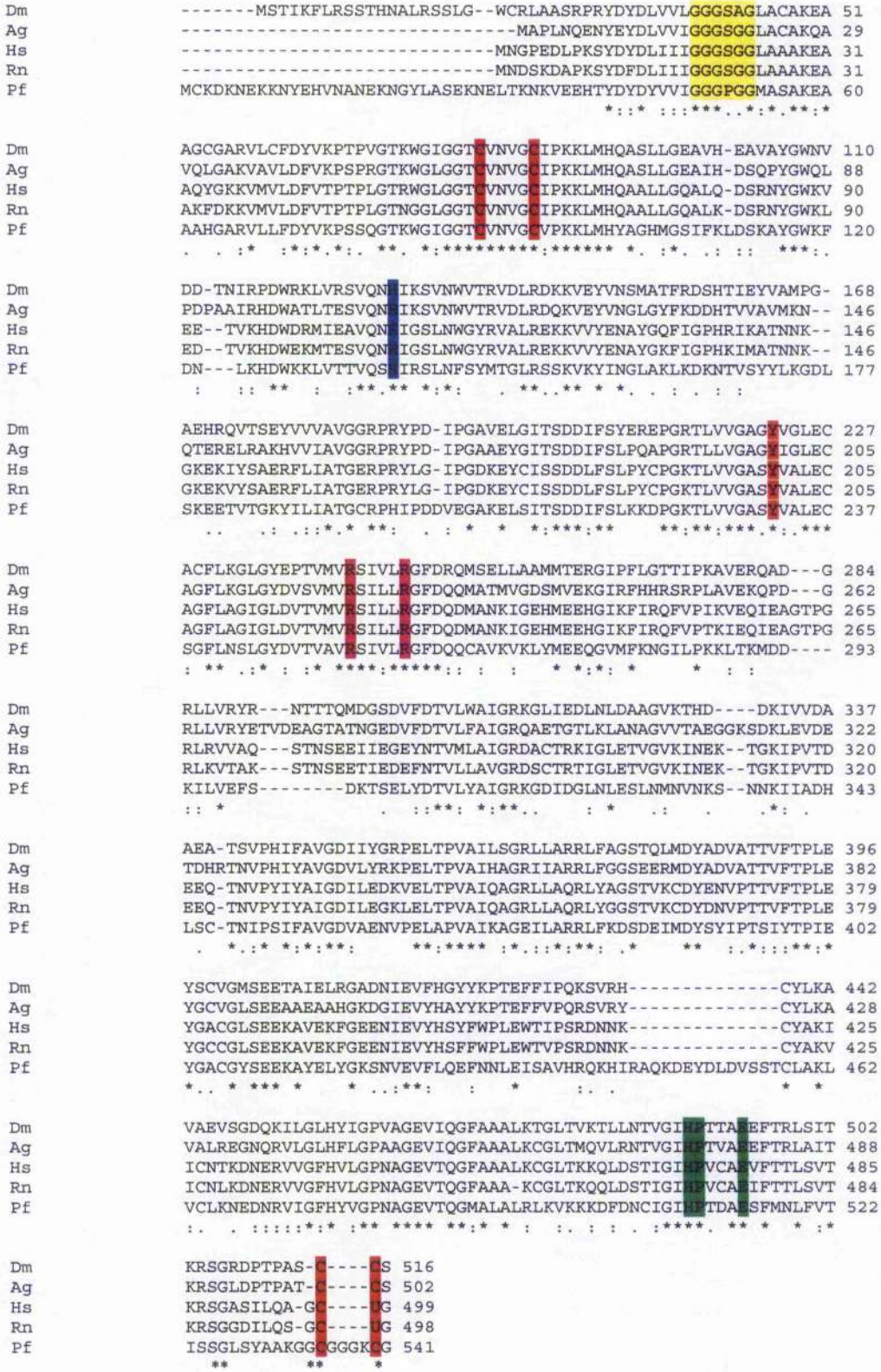

Dm -------MSTIKFLRSSTHNALRSSLG--WCRLAASRPRYDYDLVVLGGGSAGLACAKEA 51
Ag ------------------------------MAPLNQENYEYDLVVIGGGSGGLACAKQA 29
Hs ------------------------------MNGPEDLPKSYDYDLIIIGGGSGGLAAAKEA 31
Rn ------------------------------MNDSKDAPKSYDFDLIIIGGGSGGLAAAKEA 31
Pf MCKDKNEKKNYEHVNANEKNGYLASEKNELTKNKVEEHTYDYDYVVIGGGPGGMASAKEA 60
Dm AGCGARVLCFDYVKPTPVGTKWGIGGTCVNVGCIPKKLMHQASLLGEAVH-EAVAYGWNV 110
Ag VQLGAKVAVLDFVKPSPRGTKWGLGGTCVNVGCIPKKLMHQASLLGEAIH-DSQPYGWQL 88
Hs AQYGKKVMVLDFVTPTPLGTRWGLGGTCVNVGCIPKKLMHQAALLGQALQ-DSRNYGWKV 90
Rn AKFDKKVMVLDFVTPTPLGTNGGLGGTCVNVGCIPKKLMHQAALLGQALK-DSRNYGWKL 90
Pf AAHGARVLLFDYVKPSSQGTKWGIGGTCVNVGCVPKKLMHYAGHMGSIFKLDSKAYGWKF 120
Dm DD-TNIRPDWRKLVRSVQNEIKSVNWVTRVDLRDKKVEYVNSMATFRDSHTIEYVAMPG- 168
Ag PDPAAIRHDWATLTESVQNHIKSVNWVTRVDLRDQKVEYVNGLGYFKDDHTVVAVMKN-- 146
Hs EE--TVKHDWDRMIEAVQNHIGSLNWGYRVALREKKVVYENAYGQFIGPHRIKATNNK-- 146
Rn ED--TVKHDWEKMTESVQNHIGSLNWGYRVALREKKVVYENAYGKFIGPHKIMATNNK-- 146
Pf DN---LKHDWKKLVTTVQSEIRSLNFSYMTGLRSSKVKYINGLAKLKDKNTVSYYLKGDL 177
Dm AEHRQVTSEYVVVAVGGRPRYPD-IPGAVELGITSDDIFSYEREPGRTLVVGAGYVGLEC 227
Ag QTERELRAKHVVIAVGGRPRYPD-IPGAAEYGITSDDIFSLPQAPGRTLLVGAGYIGLEC 205
Hs GKEKIYSAERFLIATGERPRYLG-IPGDKEYCISSDDLFSLPYCPGKTLVVGASYVALEC 205
Rn GKEKVYSAERFLIATGERPRYLG-IPGDKEYCISSDDLFSLPYCPGKTLVVGASYVALEC 205
Pf SKEETVTGKYILIATGCRPHIPDDVEGAKELSITSDDIFSLKKDPGKTLVVGASYVALEC 237
Dm ACFLKGLGYEPTVMVRSIVLRGFDRQMSELLAAMMTERGIPFLGTTIPKAVERQAD---G 284
Ag AGFLKGLGYDVSVMVRSILLRGFDQQMATMVGDSMVEKGIRFHHRSRPLAVEKQPD---G 262
Hs AGFLAGIGLDVTVMVRSILLRGFDQDMANKIGEHMEEHGIKFIRQFVPIKVEQIEAGTPG 265
Rn AGFLAGIGLDVTVMVRSILLRGFDQDMANKIGEHMEEHGIKFIRQFVPTKIEQIEAGTPG 265
Pf SGFLNSLGYDVTVAVRSIVLRGFDQQCAVKVKLYMEEQGVMFKNGILPKKLTKMDD---- 293
Dm RLLVRYR---NTTTQMDGSDVFDTVLWAIGRKGLIEDLNLDAAGVKTHD----DKIVVDA 337
Ag RLLVRYETVDEAGTATNGEDVFDTVLFAIGRQAETGTLKLANAGVVTAEGGKSDKLEVDE 322
Hs RLRVVAQ---STNSEEIIEGEYNTVMLAIGRDACTRKIGLETVGVKINEK--TGKIPVTD 320
Rn RLKVTAK---STNSEETIEDEFNTVLLAVGRDSCTRTIGLETVGVKINEK--TGKIPVTD 320
Pf KILVEFS--------DKTSELYDTVLYAIGRKGDIDGLNLESLNMNVNKS--NNKIIADH 343
Dm AEA-TSVPHIFAVGDIIYGRPELTPVAILSGRLLARRLFAGSTQLMDYADVATTVFTPLE 396
Ag TDHRTNVPHIYAVGDVLYRKPELTPVAIHAGRIIARRLFGGSEERMDYADVATTVFTPLE 382
Hs EEQ-TNVPYIYAIGDILEDKVELTPVAIQAGRLLAQRLYAGSTVKCDYENVPTTVFTPLE 379
Rn EEQ-TNVPYIYAIGDILEGKLELTPVAIQAGRLLAQRLYGGSTVKCDYDNVPTTVFTPLE 379
Pf LSC-TNIPSIFAVGDVAENVPELAPVAIKAGEILARRLFKDSDEIMDYSYIPTSIYTPIE 402
Dm YSCVGMSEETAIELRGADNIEVFHGYYKPTEFFIPQKSVRH----------CYLKA 442
Ag YGCVGLSEEAAEAAHGKDGIEVYHAYYKPTEFFVPQRSVRY----------CYLKA 428
Hs YGACGLSEEKAVEKFGEENIEVYHSYFWPLEWTIPSRDNNK----------CYAKI 425
Rn YGCCGLSEEKAVEKFGEENIEVYHSFFWPLEWTVPSRDNNK----------CYAKV 425
Pf YGACGYSEEKAYELYGKSNVEVFLQEFNNLEISAVHRQKHIRAQKDEYDLDVSSTCLAKL 462
Dm VAEVSGDQKILGLHYIGPVAGEVIQGFAAALKTGLTVKTLLNTVGIHPTTAEEFTRLSIT 502
Ag VALREGNQRVLGLHFLGPAAGEVIQGFAAALKCGLTMQVLRNTVGIHPTVAEEFTRLAIT 488
Hs ICNTKDNERVVGFHVLGPNAGEVTQGFAAALKCGLTKKQLDSTIGIHPVCAEVFTTLSVT 485
Rn ICNLKDNERVVGFHVLGPNAGEVTQGFAAA-KCGLTKQQLDSTIGIHPVCAEIFTTLSVT 484
Pf VCLKNEDNRVIGFHYVGPNAGEVTQGMALALRLKVKKKDFDNCIGIHPTDAESFMNLFVT 522
Dm KRSGRDPTPAS-C----CS 516
Ag KRSGLDPTPAT-C----CS 502
Hs KRSGASILQA-GC----UG 499
Rn KRSGGDILQS-GC----UG 498
Pf ISSGLSYAAKGGCGGGKCG 541

Figure 3.1 - ClustalW alignment of high M_r TrxR

ClustalW alignment of amino acid sequences from the high M_r TrxR from *Plasmodium falciparum* (Pf; NC_004330), *Homo sapiens* (Hs; Q16881), *Rattus norvegicus* (Rn; O89049), *Anopheles gambiae* (Ag; AJ459821) and *Drosophila melanogaster* (Dm; AAM51940). The ClustalW program aligns identical (*), conserved (:) and semi-conserved (.) residues and displays conserved residues and motifs that are associated with the disulphide oxidoreductases. The FAD binding motif is highlighted in yellow; active site cysteines are highlighted in red (selenocysteines are represented by the letter U); acid-base catalyst motif is highlighted in green; a possible second acid-base catalyst residue is highlighted in blue; a conserved tyrosine residue involved in NADPH binding is highlighted orange and two arginine residues that confer substrate specificity for NADPH are highlighted in pink.

The interface domain is located between residues 389 and 541 and is the site of subunit interactions. This domain contains the C-terminal disulphide and the acid-base catalyst motif. The ClustalW alignment displays the differences observed between members of the high M_r TrxRs at the C-terminal disulphide, with PfTrxR containing two cysteine residues separated by four amino acids, whereas the other proteins contain adjacent cysteines or adjacent cysteine/selenocysteines (highlighted red in figure 3.1). The acid-base catalyst motif (HPXXXE) is conserved in all high M_r TrxR proteins and corresponds to amino acids 509 to 514 in the *P. falciparum* sequence (highlighted green in figure 3.1).

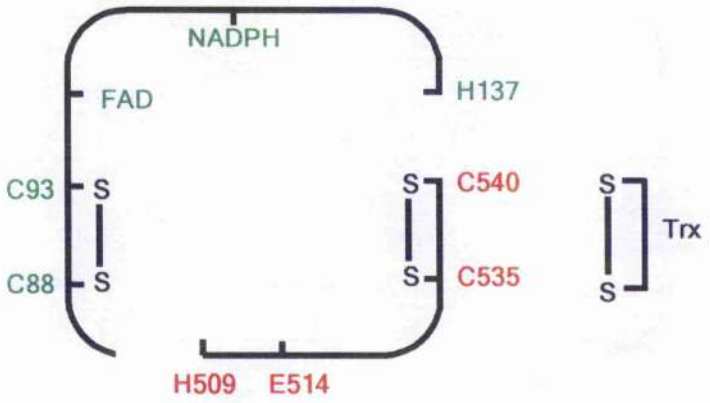


Figure 3.2 – Overview of conserved active site residues

The *Pf*TrxR active site consists of many conserved residues and binding sites. *Pf*TrxR contains substrate-binding domains (NADPH and *Pf*Trx), along with the redox active centres that consist of a bound FAD cofactor, N-terminal cysteine pair (Cys88 & Cys93) and C-terminal cysteine pair (Cys535 & Cys540). The catalytic mechanism is proposed to also involve acid-base catalysts (His 137, His509 & Glu514). The active site consists of residues from both subunits of the homodimeric protein; residues from each monomer are labelled with a different colour (green for one and red for the other monomer).

Figure 3.2 displays an overview of the residues that were to be investigated in this study, and demonstrates that the active site is generated by residues from both subunits of the homodimer. Residues are coloured in red or in green according to the subunit they belong to. The residues coloured in green include those responsible for NADPH and FAD-binding, along with the N-terminal cysteine pair, while the other subunit (residues coloured in red) provides the C-terminal cysteine pair and acid-base catalyst motif and is responsible for passing electrons to the thioredoxin substrate.

The PfTrxR active site cysteine residues (Cys 88, Cys 93, Cys 535 & Cys 540) and the histidine residue in the acid-base catalyst motif (His 509) had been mutated previously to either alanine for cysteine residues or glutamine for histidine residue and these mutants were further analysed in this study (Gilberger *et al.*, 1997; Gilberger *et al.*, 1998). Site directed mutants were also generated for the glutamate residue in the acid-base catalyst motif (Glu 514) and the proposed acid-base catalyst (His 137). Glutamate 514 was mutated to alanine, while histidine 137 was mutated to asparagine (N), glutamine (Q), serine (S) and alanine (A).

3.2.2 Recombinant expression and purification

The recombinant expression of both PfTrxR and PfTrx were carried out as described previously (Gilberger *et al.*, 1997; Krnajski *et al.*, 2001) and resulted in 2 – 5 mg of purified proteins per litre of bacterial culture. PfTrxR was purified in the oxidised form, which was unlike the insect forms of TrxR (*Anopheles gambiae* and *Drosophila melanogaster*) that were purified in a reduced form (Bauer *et al.*, 2003a; Bauer *et al.*, 2003b). An example of the purified samples used for these analyses is given in figure 3.3 and multimers that could not be resolved by reducing SDS-PAGE can be observed in all lanes. The PfTrx sample (Lane 1 in figure 3.3) displays a monomeric band at 14 kDa and a dimeric band at 28 kDa, which is consistent with samples observed in previous purifications (Krnajski *et al.*, 2001). The PfTrxR wild-type and mutant proteins all display a major band at 63 kDa and also higher M_r bands that correspond to homodimers at 126 kDa (Lanes 2 to 8 in figure 3.3).

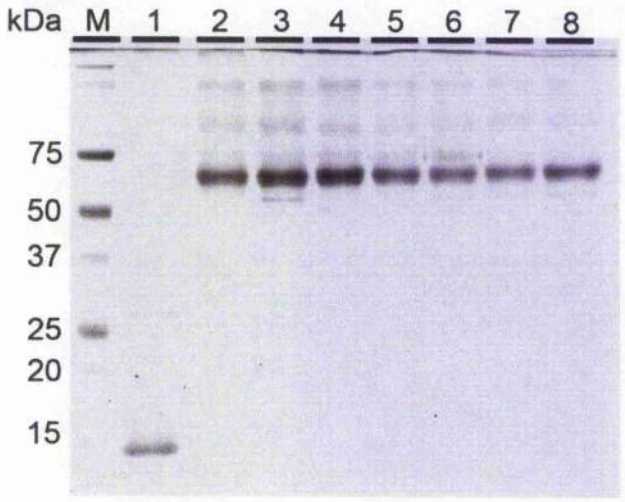


Figure 3.3 – Purification of PfTrxR & PfTrx

An example of the purified samples used in these studies were analysed on 4 – 12 % SDS – PAGE and stained with Coomassie blue. The gel was loaded with 2 μg of each of the following recombinant proteins; PfTrx (lane 1), PfTrxR (lane 2), PfTrxR H509Q (lane 3), PfTrxR E514A (lane 4), PfTrxR H137N (lane 5), PfTrxR H137Q (lane 6), PfTrxR H137S (lane 7) and PfTrxR H137A (lane 8). Lane M contains precision protein markers (Biorad). Lane 1 displays a major band at 14.2 kDa that corresponds to monomeric PfTrx and a band at 28.4 kDa that corresponds to dimeric PfTrx. Lanes 2 – 8 display a major band at 63 kDa that corresponds to monomeric PfTrxR and a number of minor bands of higher molecular mass, that correspond to dimers and other multimers of PfTrxR.

3.2.3 Effect of mutations on tertiary structure

Before in-depth characterisation of the mutant proteins could occur it was necessary to analyse if the mutations had affected the folding of the purified proteins. Although all proteins were purified with FAD bound and displayed spectral properties of oxidised flavoprotein disulphide oxidoreductase, their tertiary structure and protein folding was analysed by circular dichroism spectrometry (CD). The CD analyses of the wild-type and mutant forms of PfTrxR were carried out by Dr. Sharon Kelly of the Protein Characterisation Facility, Institute of Biomedical and Life Sciences, University of Glasgow.

Analyses of the CD spectra observed in the far UV (panel A in figure 3.4) and near UV/visible ranges (panels B and C in figure 3.4) demonstrated that there have been no gross structural changes of the mutants in comparison to the wild-type protein. Further analysis of the far UV spectra demonstrates that most of the spectra are unchanged, however the spectra of PfTrxR E514A (red line in panel A of figure 3.4) and H137S (blue line in panel A of figure 3.4) display increases in CD intensity at 222 nm (compared to wild-type; black line in panel A of figure 3.4) of 18.7% and 12.3%, respectively. The near UV/visible spectra (panel B and C in figure 3.4) display that there may be small re-arrangements in the environments around the FAD and disulphide chromophores. This was observed between 300 and 500 nm for the FAD chromophore and at 260 nm for the disulphides, and demonstrates that all mutants display greater intensities at these wavelengths than the wild-type protein. This suggests that the protein structure around these chromophores are more rigid in the mutants than in the wild-type protein.

Overall the CD analyses displays that the mutated proteins do not produce any major structural changes, but that small changes in the micro-environment of the active site can be observed. This would be expected, as the residues being analysed are conserved active site residues.

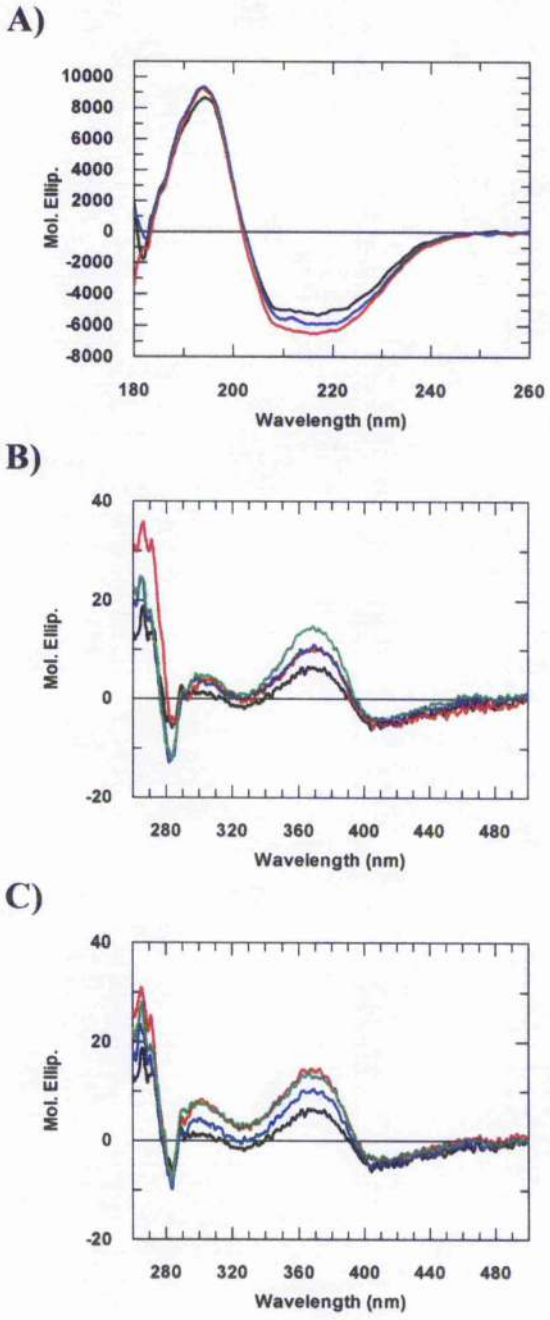


Figure 3.4 – Circular dichroism (CD) analyses of PfTrxR variants

The circular dichroism spectra of the wild-type and mutant forms of PfTrxR were analysed at far UV and near UV/visible wavelengths in order to investigate the effect of the mutations on the tertiary structure of the proteins. These experiments were carried out by Dr. Sharon Kelly of the Protein Characterisation Facility, Institute of Biological and Life Sciences, University of Glasgow. Data were collected in a JASCO J-180 instrument with protein samples at 0.5 mg/ml. **Panel A** displays the far UV CD spectra of wild-type and mutant proteins. Most spectra overlay with wild-type (black line), although two outliers corresponding to E514A (red line) and H137S (blue line) are shown. **Panel B** displays a comparison of visible/near UV CD spectra of wild-type (black line), H509Q (red line), E514A (blue line), H137N (green line). **Panel C** displays a comparison of visible/near UV CD spectra of wild-type (black line), H137Q (red line), H137S (blue line) and H137A (green line) proteins.

3.2.4 Steady state kinetic analyses

PfTrxR enzymatic activities were determined for wild-type and mutant forms of the enzyme using either DTNB or thioredoxin disulphide as electron acceptors. The steady state kinetic parameters of the wild-type and mutant proteins were analysed using both of these assay systems and are summarised in table 3.1.

The catalytic activities observed for the wild-type proteins were similar to those reported previously (Gilberger *et al.*, 1997). Mutation of His509 to glutamine resulted in the largest decreases in activity of the mutants studied. The H509Q mutant retained only 0.06 % (with *Pf*Trx) and 8 % (with DTNB) activities, when compared to those of wild-type *Pf*TrxR; this was consistent with the previously published activities for PfTrxR H509Q (Gilberger *et al.*, 1997). Replacement of Glu514 with alanine resulted in activities of 7.5 % and 48 % for *Pf*Trx and DTNB, respectively. These data confirm that glutamate 514 plays an important role in the catalysis of PfTrxR. Mutation of the His 137 residue had the smallest effect on catalytic activity, with mutants retaining between 102 and 60 % activity in the *Pf*Trx assay. Analyses of these mutants with DTNB displayed that H137N retained the most activity (145 %) and H137A the least (81 %). The greater specific activity observed in the H137N (also observed in H137Q and H137S) mutants may be due to a change in the extinction coefficient of these proteins as subsequent analysis has shown that H137N acts essentially as wild-type PfTrxR (see sections 3.2.6.4 and 3.2.7.4).

The Michaelis-Menten graphs and Lineweaver-Burk plots of the data collected for the wild-type and mutant proteins with DTNB, NADPH and PfTrx substrates are displayed in figures 3.5, 3.6 and 3.7, respectively. The K_m for DTNB was altered from 212 μM in the wild-type to 50 μM in H509Q and 80 μM in E514A, but was not drastically altered in the His137 mutants (Figure 3.5; Table 3.1). This change in K_m may arise as a consequence of the interaction of DTNB with both the N and C-terminal cysteines, unlike PfTrx that only accepts reducing equivalents from the C-terminal cysteines (Fujiwara *et al.*, 2001; Bauer *et al.*, 2003b). The K_m for NADPH was determined to be 2.0 μM for wild-type PfTrxR using the DTNB assay system and did not change drastically with mutation of the active site residues studied (Figure 3.6; Table 3.1). The K_m for PfTrx was determined to be 3.4 μM for the wild-type protein and was again not altered by the mutation of the active site residues (Figure 3.7; Table 3.1).

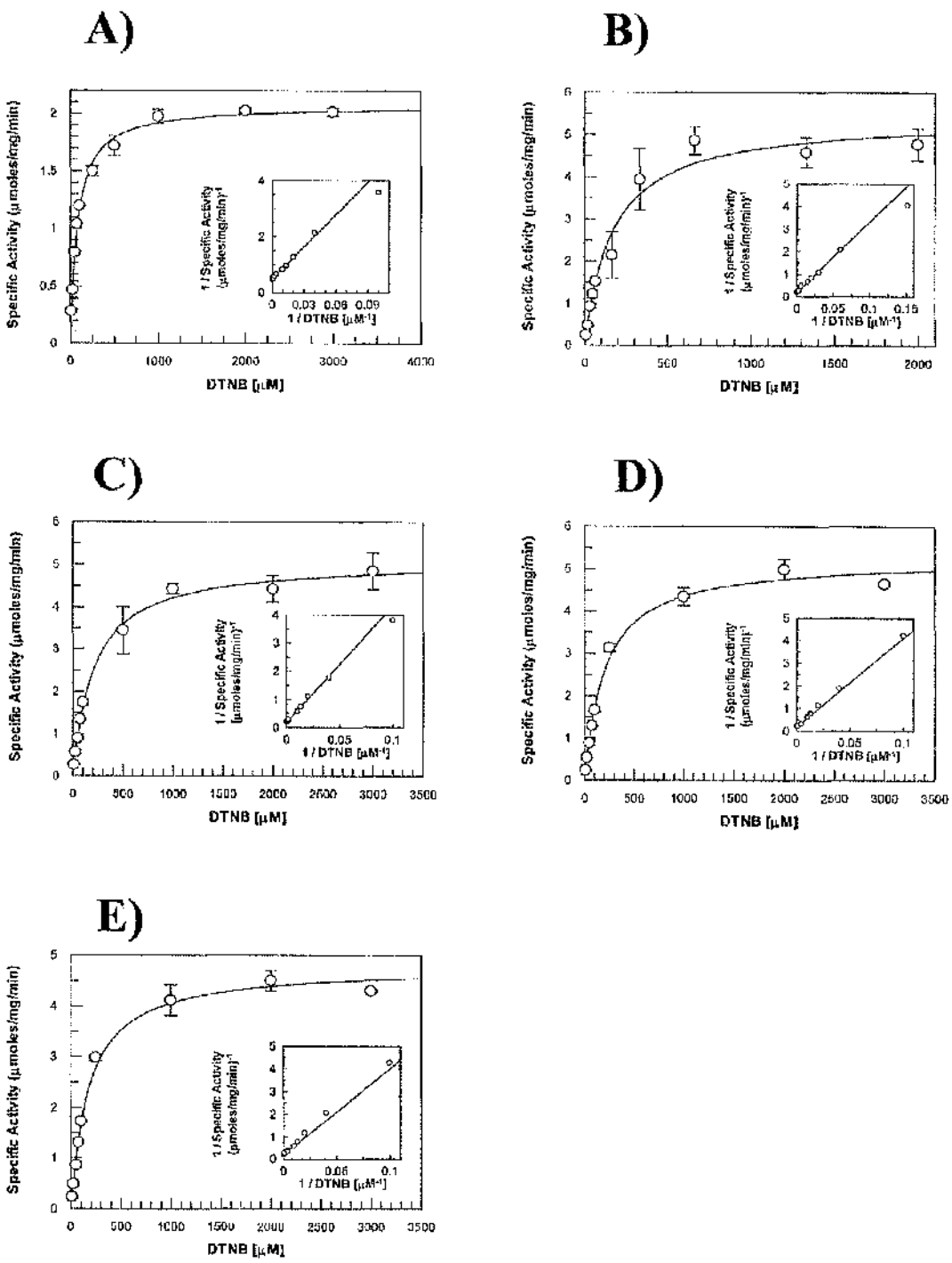


Figure 3.5 - The steady state kinetic parameters of PfTrxR with DTNB

The kinetic parameters for DTNB were determined using the DTNB assay, which was performed at 25°C in 50mM potassium phosphate buffer pH 7.6, 1 mM EDTA. The reaction mix contained 1 – 20 µg of protein (depending on the mutant), 100 µM NADPH and varying concentrations of DTNB (25 µM to 3 mM) in a final volume of 1 ml. The reaction was followed at 412 nm for the production of TNB^- with specific activities calculated using an $\varepsilon_{412\,nm}$ of 13,600 $M^{-1}\,cm^{-1}$ per TNB^- released. Michaelis-Menten curves and Lineweaver-Burk are displayed with each point being the mean ± standard deviation of 2 – 4 independent measurements. Kinetic parameters of PfTrxR E514A (**Panel A**), H137N (**panel B**), H137Q (**panel C**), H137S (**panel D**), H137A (**panel E**) are displayed, with K_m and V_{max} values (Table 3.1) determined using the Grafit 5.0 software (Erithacus software).

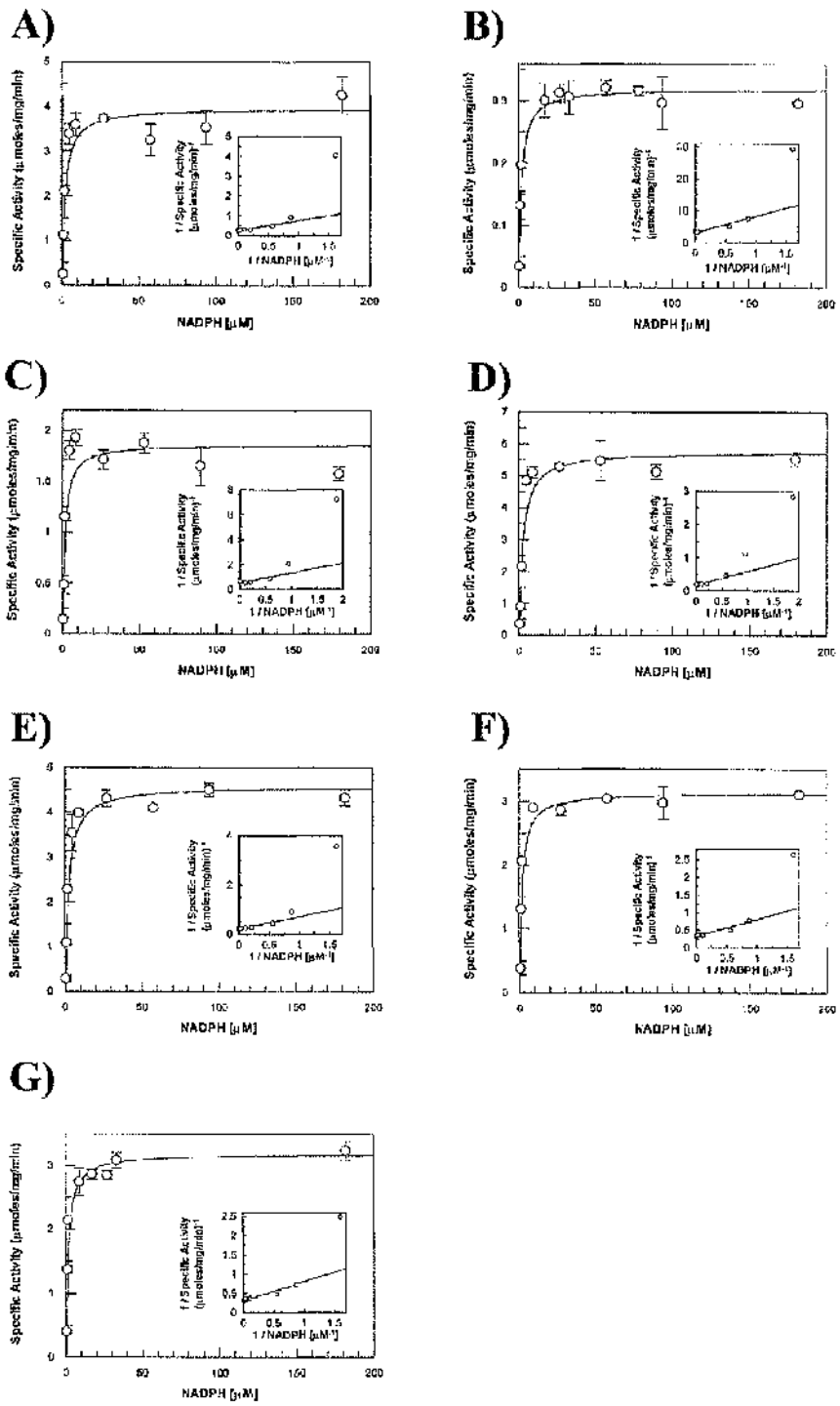


Figure 3.6 – The steady state kinetic parameters of PfTrxR with NADPH

The kinetic parameters for NADPH were determined using the DTNB assay, which was performed at 25°C in 50mM potassium phosphate buffer pH 7.6, 1 mM EDTA. The reaction mix contained 1 – 20 µg of protein (depending on the mutant), 3 mM DTNB and varying concentrations of NADPH (0.5 µM to 200 µM) in a final volume of 1 ml. The reaction was followed at 412 nm for the production of TNB^- with specific activities calculated using an $\varepsilon_{412\,nm}$ of 13,600 M^{-1} cm^{-1} per TNB^- released. Michaelis-Menten curves and Lineweaver-Burk plots are displayed with each point the mean +- standard deviation of 2 – 4 independent measurements. The kinetic parameters for PfTrxR wild-type (**panel A**), H509Q (**panel B**), E514A (**panel C**), H137N (**panel D**), H137Q (**panel E**), H137S (**panel F**), H137A (**panel G**) are displayed with apparent K_m and V_{max} values (Table 3.1) determined using the Grafit 5.0 software (Erithacus software).

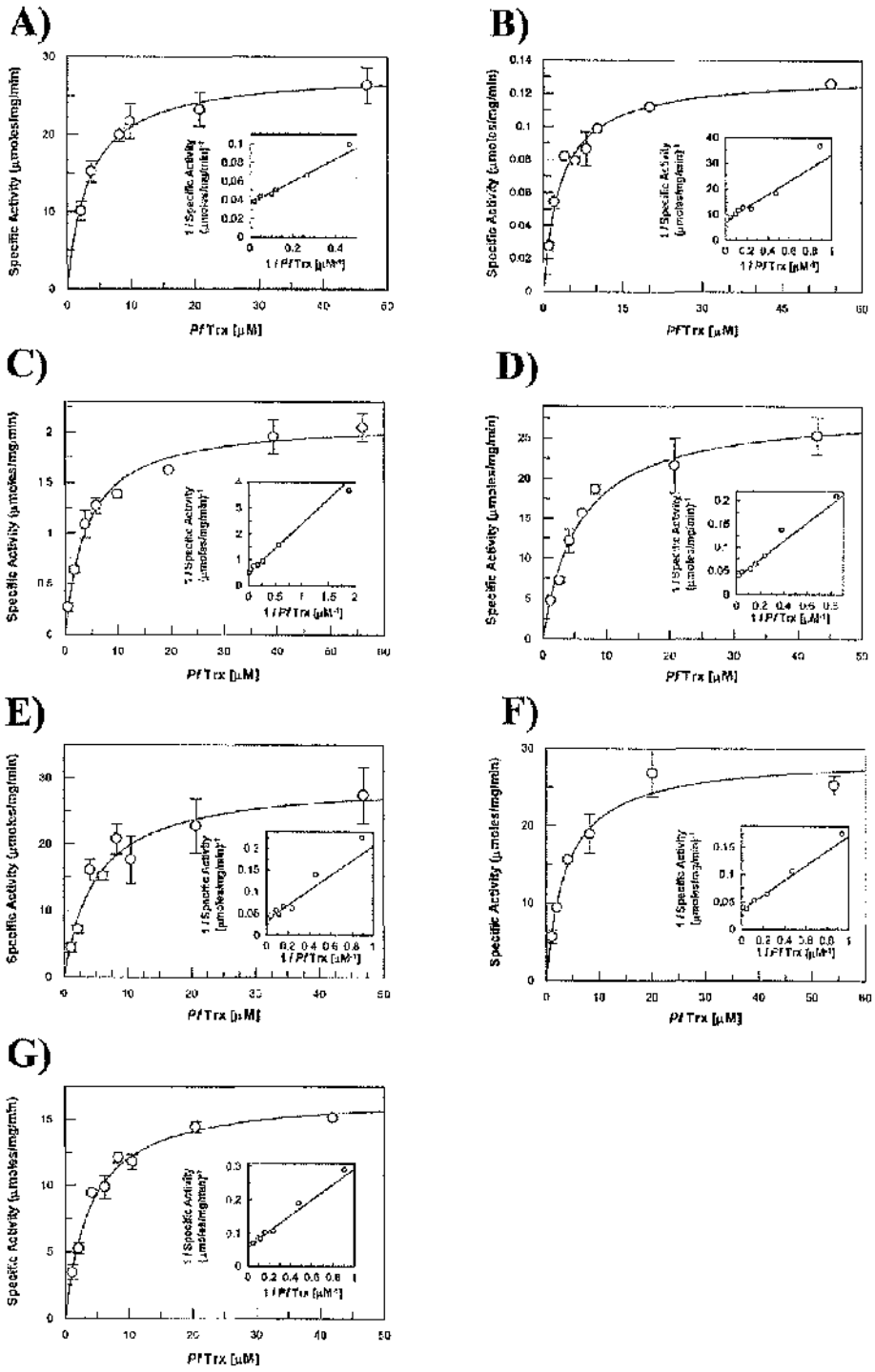


Figure 3.7 - The steady state kinetic parameters of PfTrxR with PfTrx

The kinetic parameters for reduction of PfTrx were determined using the insulin assay, which was performed at 25°C in 50mM potassium phosphate buffer pH 7.6, 1 mM EDTA, 0.2 mg ml^{-1} insulin. The reaction mix contained 0.5 – 20 μg of protein (depending on the mutant), 200 μM NADPH and varying concentrations of PfTrx (2 μM to 60 μM) in a final volume of 1 ml. The reaction was followed at 340 nm for the consumption of NADPH with specific activities calculated using an $\varepsilon_{340\ nm}$ of 6,200 $M^{-1}\ cm^{-1}$. Michaelis-Menten curves and Lineweaver-Burk plots are displayed with each point the mean ± standard deviation of 2 – 4 independent measurements. Kinetic parameters for PfTrxR wild-type (**panel A**), H509Q (**panel B**), E514A (**panel C**), H137N (**panel D**), H137Q (**panel E**), H137S (**panel F**), H137A (**panel G**) are displayed with apparent K_m and V_{max} values (Table 3.1) determined using the Grafit 5.0 software (Erithacus software).

Protein	K_m^{DTNB} (μM)	k_{cat}^{DTNB} (min^{-1})	K_m^{NADPH} (μM)	k_{cat}^{NADPH} (min^{-1})	K_m^{Trx} (μM)	k_{cat}^{Trx} (min^{-1})
WT	212[a]	275[a]	2.0 ± 0.6	249	3.4 ± 0.3	1774
H509Q	50[a]	14[a]	1.7 ± 0.4	20	3.3 ± 0.5	8
E514A	80 ± 5	130	1.5 ± 0.6	118	4.0 ± 0.5	133
H137N	169 ± 31	341	2.4 ± 0.8	362	5.6 ± 0.8	1805
H137Q	209 ± 18	321	2.3 ± 0.6	289	5.0 ± 1.1	1868
H137S	204 ± 20	330	1.5 ± 0.4	197	3.9 ± 0.8	1831
H137A	181 ± 16	301	1.6 ± 0.4	201	4.0 ± 0.8	1066

Table 3.1 – Kinetic parameters of PfTrxR wild-type and mutants

The kinetic parameters of DTNB and NADPH were determined using the DTNB assay and those of PfTrx were determined using the insulin assay. The apparent K_m and V_{max} data was determined using the Grafit 5.0 software (figures 3.5, 3.6 and 3.7; Erithacus software). [a] Data taken from Gilberger et al., 1997.

3.2.5 Spectral features of oxidised and reduced proteins

The wild-type and mutant proteins were anaerobically titrated with either NADPH or sodium dithionite, to analyse how the mutations affected the distribution of electrons within the redox active centres and the formation of charge transfer complexes.

Spectral analysis of the oxidised *Pf*TrxR wild-type protein demonstrated that it exhibited the characteristic shape of a disulphide oxidoreductase (Gilberger *et al.*, 1997; Farber *et al.*, 1998), with peaks at around 460 and 380 nm (black line in figure 3.8A). Oxidised spectra of the mutant forms display the same characteristic shape with minor changes in the position of the flavin peak around 460 nm (Gilberger *et al.*, 1997).

Reduction of wild-type *Pf*TrxR with 6 eq NADPH (blue line in figure 3.8A) produced several spectral differences when compared to the oxidised protein (black line in figure 3.8A). The absorbance of the flavin peak at 462 nm decreases and shifts to shorter wavelengths (445 nm) as the flavin becomes reduced. Two charge transfer complexes were observed in the spectral features; the flavin thiolate charge transfer complex band (CTC1) was centred around 540 nm and a small absorbance at longer wavelengths (670 – 700 nm) associated with the $NADP^+$ $FADH^-$ charge transfer complex (CTC2) was also observed in the reduced protein. This spectral pattern was consistent with those features described in previous studies (Gilberger *et al.*, 1997). Reduction with sodium dithionite, which reduces the protein via the FAD cofactor, removes the spectral features associated with NADPH binding and reaction (red line in figure 3.8A). The spectrum of the sodium dithionite reduced protein displays the reduced flavin peak at 444 nm and the CTC1 peak at 540 nm, but lacks the absorbance at longer wavelengths associated with CTC2 and the peak produced at 340 nm which is a result of excess NADPH; an isobestic point is also observed at approximately 355 nm, which will be used for calculations in the pre-steady state kinetic studies.

Reduction of the H509Q mutant with NADPH (blue line in figure 3.8B) resulted in a spectrum that displays several different properties to those observed in the wild-type protein (blue line in figure 3.8A). The CTC1 absorbance had decreased to ε_{540nm} 2.09 mM^{-1} cm^{-1} in the mutant from ε_{540nm} 3.42 mM^{-1} cm^{-1} in the wild-type. The flavin peak appears more reduced at 5.81 mM^{-1} cm^{-1} in the mutant when compared with 6.86 mM^{-1} cm^{-1} in the wild-type (at 440 nm), although the peak retains a partially oxidised shape (presence of two clear shoulders on the flavin peak). The CTC2 absorbance (at 670 nm) displayed no significant changes with an absorbance of 1.58 mM^{-1} cm^{-1} in comparison

to 1.55 mM^{-1} cm^{-1} in the wild-type. Reduction of H509Q with sodium dithionite (red line in figure 3.8B) showed remarkable differences to the wild-type protein reduced with sodium dithionite (red line in figure 3.8A); the charge transfer absorbances (CTC1 and 2) between 540 and 700 nm were almost completely lost (ε_{540nm} = 1.08 mM^{-1} cm^{-1} & ε_{670nm} = 0.42 mM^{-1} cm^{-1}) in the H509Q protein upon reduction. The flavin peak was more reduced (ε_{440nm} of 6.85 mM^{-1} cm^{-1}), although it retained a more oxidised shape than that in the dithionite reduced wild-type protein (compare red lines in figure 3.8A and B). A small quantity of flavin anion semi-quinone can also be observed at 405 nm in the dithionite reduced H509Q spectrum (red line in figure 3.8B). Analysis of the effects of the H509Q mutation suggests that His509 has a crucial role in stabilising CTC1, and is therefore also required for electron transfer between the redox active centres.

Reduction of the E514A mutant with NADPH (blue line in figure 3.8C) led to an increase in the CTC1 absorbance (at 540 nm) formed, from 3.42 mM^{-1} cm^{-1} in wild-type to 3.78 mM^{-1} cm^{-1} in the E514A mutant. The CTC2 absorbance (at 670 nm) was slightly increased in the mutant to 1.79 mM^{-1} cm^{-1} compared with 1.55 mM^{-1} cm^{-1} in the wild-type protein. The flavin peak resembles that of the NADPH reduced wild-type protein and the absorption coefficients are also very similar (6.91 mM^{-1} cm^{-1} in the mutant compared with 6.86 mM^{-1} cm^{-1} in the wild-type (at 440 nm). Reduction of the E514A protein with sodium dithionite displays a spectrum (red line in figure 3.8C) with major changes to the corresponding wild-type spectrum (red line in figure 3.8A). The flavin peak (440 nm) is reduced in shape and yet has a higher absorbance (ε_{440nm} of 10.1 mM^{-1} cm^{-1} compared to an ε_{444nm} of 8.3 mM^{-1} cm^{-1} in the wild-type protein). The CTC1 absorbance is also increased in the mutant with a ε_{540nm} of 4.7 mM^{-1} cm^{-1} compared to an ε_{540nm} of 3.0 mM^{-1} cm^{-1} in the wild-type protein. The mutation of glutamate 514 therefore seems to enhance the level of CTC1 observed in the static titrations. This is presumably due to the mutation favouring the formation of CTC1 (being stabilised by His509), but also conferring an inhibition of the transfer of reducing equivalents beyond CTC1 to the C-terminal cysteines.

Replacement of the second proposed acid-base catalyst residue His137 to an asparagine residue produced only minor changes when compared with the wild-type protein. Reduction of H137N with NADPH (blue line in figure 3.8D) produced minor increases in both CTC1 (ε_{540nm} from 3.42 mM^{-1} cm^{-1} in the wild-type to 4.06 mM^{-1} cm^{-1} in the mutant), CTC2 (ε_{670nm} from 1.55 mM^{-1} cm^{-1} in the wild-type to 2.2 mM^{-1} cm^{-1} in the

mutant), and flavin peak (ε_{440nm} from 6.86 mM^{-1} cm^{-1} in the wild-type to 7.45 mM^{-1} cm^{-1} in the mutant) absorbances (blue line in figure 3.8D), although the increases in absorbances for each diagnostic wavelength may be due to a fine precipitate that formed during titration with NADPH. The reduction of the H137N mutant with sodium dithionite (red line figure 3.8D) again yielded a spectrum that was similar to that observed in wild-type PfTrxR. The CTC1 absorbance was ε_{540nm} 2.9 mM^{-1} cm^{-1} in the mutant compared to 3.0 mM^{-1} cm^{-1} in the wild-type protein, the CTC2 absorbance was ε_{670nm} 0.8 mM^{-1} cm^{-1} in the mutant (compared to ε_{670nm} 0.7 mM^{-1} cm^{-1} in the wild-type) and the flavin peak was 8.6 mM^{-1} cm^{-1}at 444 nm in the mutant (compared with 8.3 mM^{-1} cm^{-1}at 444 nm in the wild-type enzyme. The mutation of His137 with three other amino acids (glutamine, serine and alanine) led to only minor spectral changes, with the NADPH reduced spectra of all His137 mutants being essentially identical to wild-type (figure 3.9).

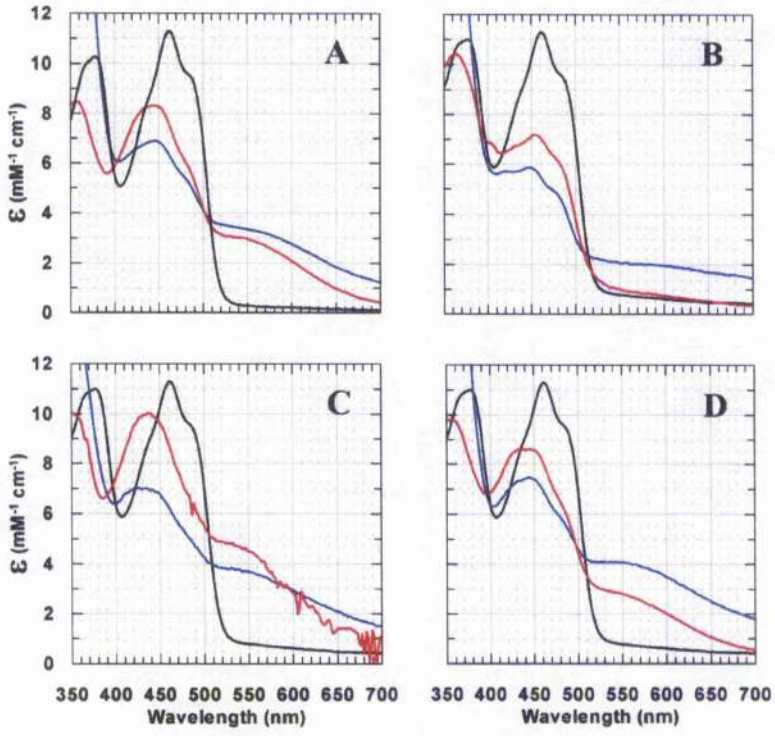


Figure 3.8 – Static anaerobic reductive titrations of Pf TrxR

Visible absorption spectra of 15 μM *Pf*TrxR wild-type and mutant forms were measured at 25°C under anaerobic conditions in 50 mM potassium phosphate, 1 mM EDTA pH 7.6. Spectra were recorded for the wild-type oxidised (black line) and either NADPH (blue line) or dithionite (red line) reduced forms of the proteins. NADPH reduction was achieved by the addition of 6 eq NADPH; dithionite spectra were selected for maximum absorbance at 540 nm during titration of dithionite between 0 – 8 eq. **Panel A** wild-type *Pf*TrxR; **Panel B** *Pf*TrxR H509Q mutant; **Panel C** *Pf*TrxR E514A mutant; **Panel D** *Pf*TrxR H137N mutant.

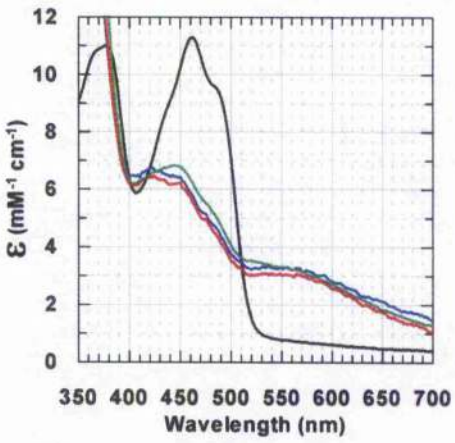


Figure 3.9 – Anaerobic static titrations of H137 mutants with NADPH

Visible absorption spectra of 15 μM *Pf*TrxR His137 mutants were measured between 350 nm and 700 nm at 25°C under anaerobic conditions in 50 mM potassium phosphate, 1 mM EDTA pH 7.6. Spectra were recorded between for the oxidised (black line) and NADPH reduced H137Q (green line), H137S (blue line), and H137A (red line). NADPH reduction was achieved by the addition of 6 eq NADPH.

Mutation of the N-terminal cysteine pair affected the NADPH reduced spectra of the proteins drastically. Replacement of the charge transfer thiol (Cys93) with an alanine residue produced spectra (blue line in figure 3.10A) with a highly reduced flavin peak (ε_{440nm} = 4.36 mM^{-1} cm^{-1}), a long wavelength peak associated with CTC2 (ε_{670nm} = 1.79 mM^{-1} cm^{-1}) and very little absorbance for the CTC1 peak (ε_{540nm} = 1.48 mM^{-1} cm^{-1}). Replacement of the interchange thiol (Cys88) with an alanine residue resulted in a spectrum (red line in figure 3.10A) displaying a reduced flavin peak with higher absorbance than observed in the wild-type protein (ε_{440nm} of 7.77 mM^{-1} cm^{-1} in the mutant, compared with 6.86 mM^{-1} cm^{-1} in the wild-type spectra), while the CTC1 (ε_{540nm} = 1.35 mM^{-1} cm^{-1}) absorbance was similar to that observed in the C93A spectra. The CTC2 absorbance (ε_{670nm} = 1.42 mM^{-1} cm^{-1}) was essentially similar to wild-type spectra. The flavin peaks of both N-terminal cysteine mutants maintained partially oxidised shape, unlike that observed in the wild-type spectra. This is in agreement with previous reports, which investigated the static titrations of the C88A and C93A mutants (Gilberger *et al.*, 1997).

Mutation of the C-terminal cysteines to alanine residues resulted in only minor changes to the reduced spectra of the proteins. The NADPH reduced C535A spectrum (blue line in figure 3.10B) displayed a more reduced flavin peak (ε_{440nm} = 5.89 mM^{-1} cm^{-1}), less CTC1 absorbance (ε_{540nm} of 3.1 mM^{-1} cm^{-1}) and less CTC2 absorbance (ε_{670nm} = 1.28

mM^{-1} cm^{-1}) than was observed in the wild-type protein. The NADPH reduced spectra of the C540A mutant (red line in figure 3.10B) displayed almost identical features to the C535A mutant, with the values observed for C540A of flavin reduction (ε_{440nm} = 6.09 mM^{-1} cm^{-1}), CTC1 (ε_{540nm} = 3.15 mM^{-1} cm^{-1}) and CTC2 (ε_{670nm} = 1.36 mM^{-1} cm^{-1}) being extremely similar. This is in agreement with previous reports, which investigated the C535A and C540A mutants (Wang *et al.*, 1999).

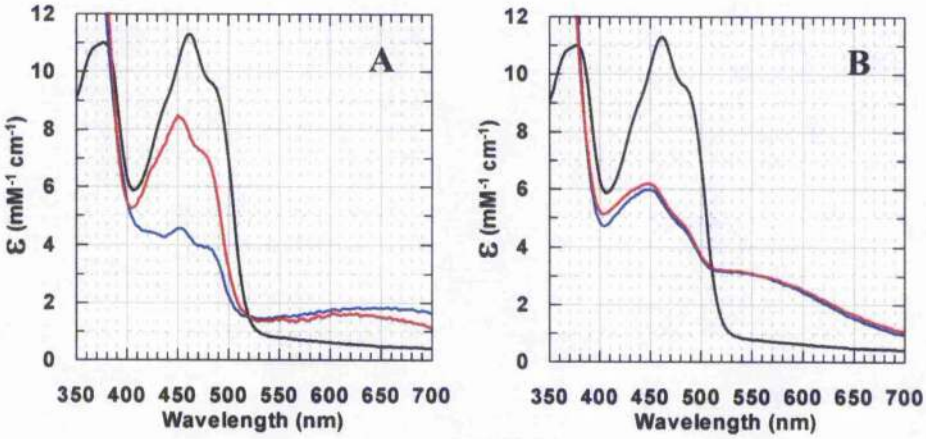


Figure 3.10 – Anaerobic static titrations of PfTrxR cysteine mutants with NADPH

Visible absorption spectra of 15 μM *Pf*TrxR cysteine mutants were measured between 350 nm and 700 nm at 25°C under anaerobic conditions in 50 mM potassium phosphate, 1 mM EDTA pH 7.6. Spectra were recorded for the oxidised (black line) and NADPH reduced proteins (achieved by the addition of 6 Eq NADPH). **Panel A** displays the reduced spectrum of PfTrxR C93A (blue line) and C88A (red line). **Panel B** displays the reduced spectrum of PfTrxR C535A (blue line) and C540A (red line).

3.2.6 The pre-steady state reductive half-reaction

The reduction of PfTrxR wild-type and mutant proteins with six molar equivalents of NADPH was studied under anaerobic conditions using a stopped-flow spectrophotometer; this allowed me to study the spectral intermediates formed during catalysis from PfTrxR E_{ox} to PfTrxR EH_4 and therefore the effect that the mutations have on catalysis.

3.2.6.1 Reductive half-reaction of PfTrxR wild-type

The reductive half-reaction of the wild-type protein with NADPH displayed similar kinetics to the *D. melanogaster* TrxR (DmTrxR) that has been studied previously; this was not surprising as the reaction mechanism proposed for DmTrxR is thought to be relevant for all high M_r TrxR (Bauer *et al.*, 2003b). The reductive half-reaction of PfTrxR wild-type protein was fitted to three phases with the first phase being dependent on NADPH concentration and reaching a saturating rate of 700 – 800 s^{-1} (table 3.2). The first phase corresponds primarily to the formation of $FADH^-$ - $NADP^+$ charge transfer complex (CTC2; observed at 440 nm and 670 nm), which can be observed in the spectrum taken 5 ms following reduction (red line in figure 3.11). This phase can also be observed in the kinetic traces as an increase of absorbance at 670 nm (blue line in the inset of figure 3.11) and a decrease in absorbance at 440 nm (red line in the inset of figure 3.11) in the inset. This phase also demonstrates the consumption of NADPH, which is observed as a decrease in absorbance at 355 nm (black line in the inset of figure 3.11).

The phases following the first phase proved difficult to resolve when PfTrxR was reacted with more than 2 molar equivalents of NADPH. During the second phase (which was determined to proceed at an observed rate constant of 360 s^{-1}; Table 3.2) the reducing equivalents are passed from CTC2 to the N-terminal disulphide, producing a thiolate flavin charge transfer complex between Cys93 and the bound FAD cofactor (CTC1; observed at 540 nm). This formation of CTC1 was observed by the increase in absorbance at 540 nm in the spectrum taken 36 ms following reduction (green line in figure 3.11) and in the 540 nm trace in the inset (green line in the inset of figure 3.11). The formation of CTC1 can also be observed by an increase in absorbance at 440 nm and a shifting of the flavin peak to a shorter wavelength (green line in figure 3.11 and the red line in the inset of figure 3.11).

In the third phase the reducing equivalents are passed from the N-terminal cysteines to the C-terminal cysteines. This transfer occurs by the interchange thiol (Cys88) reacting with the C-terminal disulphide to form a transient mixed disulphide probably between Cys88 and C540, which is followed by the transfer of reducing equivalents to the C-terminal cysteine pair. Phase three was associated with an increase in absorbance at 440 nm and a further shifting of the FAD peak to a shorter wavelength (blue line in figure 3.11 and red line in the inset of figure 3.11). The CTC1 absorbance (observed at 540 nm) was also enhanced during this phase due to the displacement of $NADP^+$ by a second molecule of NADPH producing a decrease in CTC2, and therefore a shift of the equilibrium to CTC1 (blue line in figure 3.11 and green line in the inset of figure 3.11). The third phase was determined to proceed at an observed rate constant of 20 s^{-1} and produced an endpoint spectrum (blue line in figure 3.11) that was similar to the spectrum observed in the static NADPH titrations (blue line in figure 3.8A).

Comparison of the NADPH reduced spectra from the static titration and stopped-flow analysis identified an increase in the CTC1 absorbance in the latter when compared to the former (ε_{540nm} of 3.0 mM^{-1} cm^{-1} in the static titration compared with 4.2 mM^{-1} cm^{-1} in the stopped-flow analysis). This was probably due to the different temperatures that these experiments were performed at modulating the pKa of active site residues and therefore the equilibria that were formed. Comparison of static titrations performed at 25°C and 4°C confirmed that this was a possible explanation, as spectra from the 4°C experiment displayed higher CTC1 absorbances than those collected at 25°C using the same equipment (data not shown).

The NADPH consumption was analysed using the data collected for 355 nm (black line in figure 3.11). This wavelength was used as it is an isobestic point between the PfTrxR E_{ox} and EH_4 species (red line in figure 3.8A). The consumption of NADPH can be observed by a decrease in absorbance at 355 nm and seems to proceed in a biphasic manner (although it was fitted to 3 phases), which would fit well with the priming of the oxidised protein to EH_2 (in the first phase) followed by a subsequent production of EH_4 (in phase 3), which is consistent with the proposed catalytic mechanism.

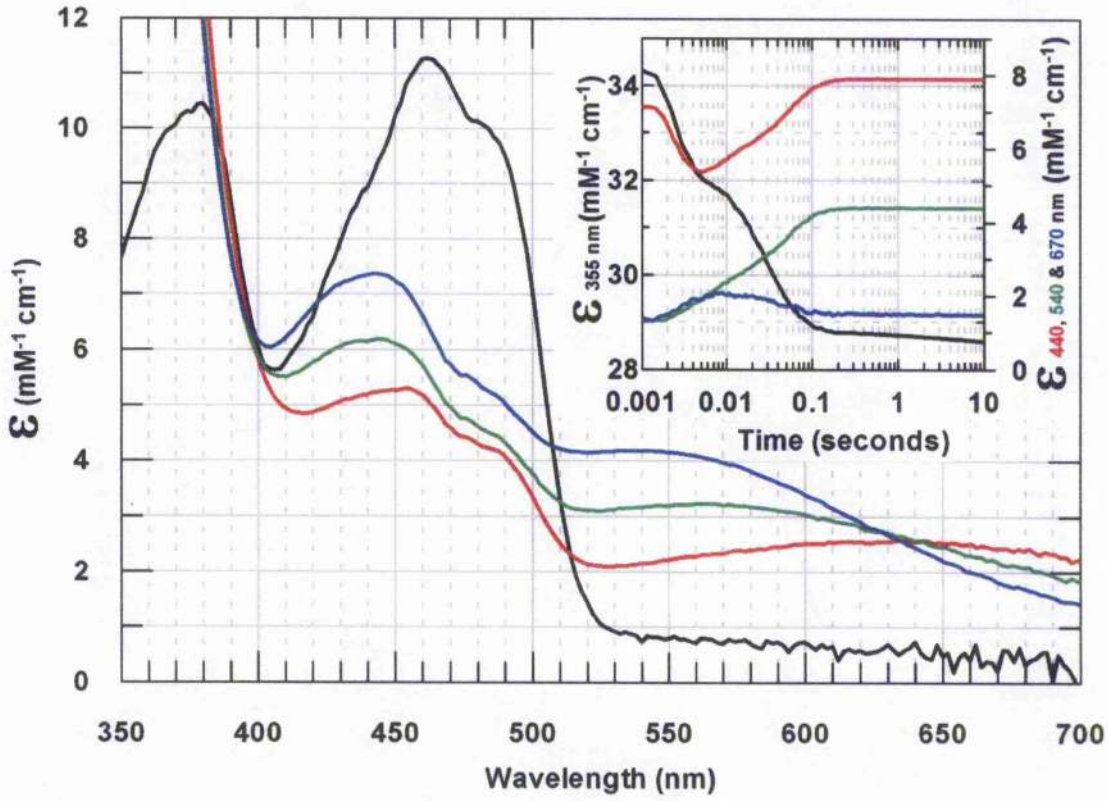


Figure 3.11 – Reductive half-reaction of wild-type PfTrxR

Reduction of 15.6 μM PfTrxR wild-type with 6 eq of NADPH was performed under anaerobic conditions at 5°C. Spectra are shown before (Eox, black line) and 5 ms (red line), 36 ms (green line) and 10 s (blue line) following mixing with NADPH. The 5 ms spectrum (red line) displays the $FADH^-$-$NADP^+$ charge transfer complex (CTC2) at 670 nm and 460 nm. The 10 s spectrum displays the thiolate-flavin charge transfer complex (CTC1) at 540 nm and 440 nm. **Inset** displays kinetic traces of the reaction at diagnostic wavelengths 355 nm (black line), 440 nm (red line), 540 nm (green line), and 670 nm (blue line). The kinetic traces were fitted to three distinct phases using the KinetAsyst 3 software (Hi-Tech scientific) (Table 2).

3.2.6.2 Reductive half-reaction of PfTrxR H509Q

The reductive half-reaction of PfTrxR H509Q was fitted to three phases, which was similar to the situation in the wild-type, although the data collected were considerably different. This would be expected as the steady state kinetic and static titration data for this mutant were very different to those observed for the wild-type protein. These data had suggested that His509 would be important in the formation and stabilisation of the thiolate flavin charge transfer complex (CTC1), and the reductive half-reaction of PfTrxR H509Q was analysed in order to confirm this.

The first phase seen in the reductive half-reaction was assignable to production of CTC2 and was essentially the same as that observed in the wild-type protein. The first phase was also dependent of NADPH concentration (as observed for the wild-type protein) and was determined to proceed at 567 s^{-1}, which was slightly slower than the maximum rate in the first phase of the wild-type protein (table 3.2; 700 – 800 s^{-1}). The spectrum taken 20 ms following reduction displayed a decrease in flavin absorbance (440 nm) and an absorbance band at longer wavelengths (650-700 nm), which was indicative of CTC2 production (green line in figure 3.12). However, this spectrum also displayed much less absorbance at 440 nm (ε_{440} of 4.6 mM^{-1} cm^{-1}in H509Q compared with 5.2 mM^{-1} cm^{-1} in wild-type) and the flavin peak had not shifted to shorter wavelengths and also maintained an oxidised-like shape, which suggests that the primary species is CTC2 after reaction with NADPH (green line in figure 3.12).

The second phase in the reductive half-reaction proceeded much slower (by 1000 fold) than in wild-type reductions (figure 3.12 inset). This second phase is normally assigned to transfer of electrons from the flavin to the N-terminal cysteines and the subsequent formation of CTC1 (Bauer *et al* 2003). In the wild-type protein this phase proceeds at an observed rate constant of 360 s^{-1}, while the H509Q proteins reaction was slowed to 0.37 s^{-1} (table 3.2). This phase is characterised by an increase in absorbance at 440 nm (figure 3.12 inset) and a decrease in CTC2 and is postulated to involve the transfer of reducing equivalents from the flavin to the N-terminal cysteines without the formation of CTC1.

The effect of pH on the reductive half-reaction of H509Q was investigated as the pH optimum of this mutant has been shifted from pH 7.6 (the wild-type optimum) to pH 8.2 (data not shown). The second phase was observed to proceed faster at pH 8.2 (blue line in figure 3.13) than at pH 7.6 (red line in figure 3.13; table 3.2), presumably as the

solvent partially compensated for the loss of the base (His509). This did not however result in an increase in the CTC1 formed (observed at 540 nm) at higher pH (up to pH 8.2, data not shown) confirming that the reducing equivalents were passing between the flavin and the N-terminal cysteines without the formation of a CTC1 intermediate. The higher rate observed in the second phase of the reaction performed at pH 7.0 (table 3.2) may be due to the formation of a different set of intermediates.

The third phase of the PfTrxR H509Q reductive half-reaction produced an endpoint spectrum (blue line in figure 3.12) that was similar to the spectrum observed in the static titration of H509Q with NADPH (blue line in figure 3.8B). This phase proceeds at 0.071 s^{-1} (table 3.2) and was not affected by altering the pH of the reaction. The production of absorbance at 540 nm (blue line in figure 3.12) that occurs during this phase may be due to displacement of $NADP^+$ with NADPH and also the generation of a small amount of CTC1. This small degree of CTC1 formation was also observed in the sodium dithionite titration of H509Q (red line in figure 3.8B).

Analysis of the NADPH consumption appeared biphasic (as observed in the wild-type) with each phase consuming 0.9 equivalents, which is consistent with the reduction of the protein to the EH_2 and then EH_4 species. Utilising the rate of the first phase at different NADPH concentrations allowed me to analyse the $K_{d\ app}$ for NADPH, with H509Q yielding a higher $K_{d\ app}$ (59 μM; table 3.2) than wild-type protein (25 μM; table 3.2). This is in contrast to the $K_{m\ app}$ data determined for H509Q, which was not different from that of the wild-type protein (table 3.1) and may be a function of the effect of the mutation on the observed rate constants altering the dissociation of the substrate.

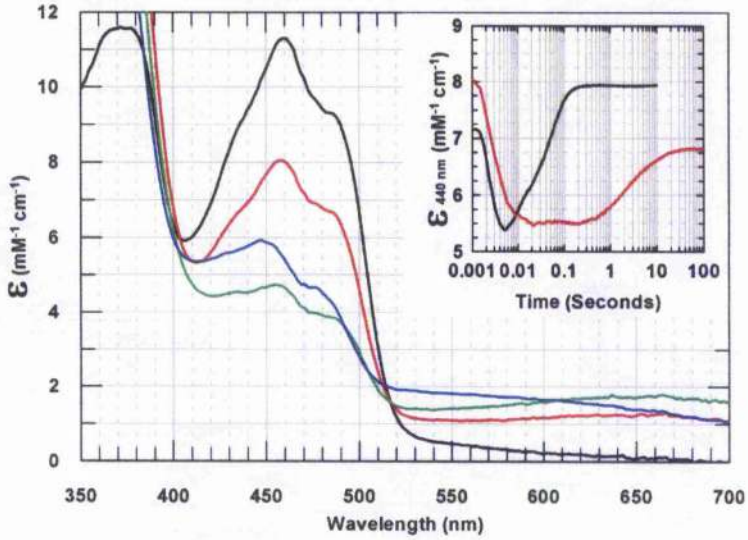


Figure 3.12 – Reductive half-reaction of PfTrxR H509Q

Reduction of 17.9 μM PfTrxR H509Q with 6 eq of NADPH was performed in 50 mM potassium phosphate pH 7.6, 1 mM EDTA under anaerobic conditions at 5°C. Spectra are shown before (Eox, black line) and 1ms (red line), 20ms (green line) and 140s (blue line) following mixing with NADPH. **Inset** displays rapid reaction kinetic traces of 440 nm for 15.6 μM PfTrxR wild-type (black line) and 17.9 μM PfTrxR H509Q (red line) reduced with 6 eq of NADPH. The kinetic traces of the H509Q reductive half-reaction were fitted to three exponential phases, using the KinetAsyst 3 software (Hi-Tech scientific), with the second and third phases of the reductive half-reaction being much slower than with wild-type enzyme (table 3.2).

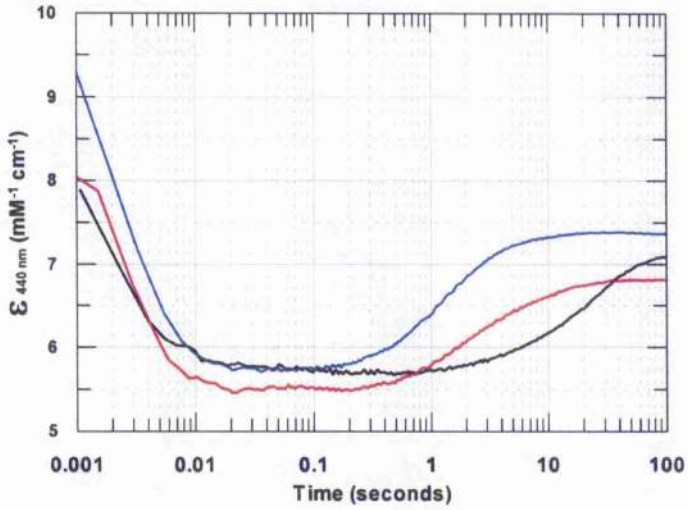


Figure 3.13 – Effect of pH on the reductive half-reaction of PfTrxR H509Q

PfTrxR H509Q was reduced at 5°C with 6 eq of NADPH. The reaction was followed at 440 nm in 50 mM potassium phosphate, 1 mM EDTA at pH 7.0 (black line; 9.7 μM PfTrxR H509Q), pH 7.6 (red line; 13.5 μM PfTrxR H509Q) and pH 8.2 (blue line; 12.1 μM PfTrxR H509Q). The reductive half-reaction proceeds much faster at higher pH, suggesting that the protonation state of the residue at position 509 is crucial for the formation and stabilization of the thiolate-flavin CTC and for the dithiol-disulfide interchange reactions. Rates observed with wild-type enzyme are much less sensitive to pH in this range (table 3.2).

3.2.6.3 Reductive half-reaction of PfTrxR E514A

The reductive half-reaction of the PfTrxR E514A mutant was fitted to three phases, as observed for both wild-type and H509Q forms. The first phase was again associated with formation of $FADH^-$ $NADP^+$ charge transfer complex (CTC2) and was determined to proceed at an observed rate constant of 803 s^{-1} (essentially the same as wild-type; table 3.2). However, the flavin reduction observed in this phase proceeds much further than those of either wild-type or H509Q proteins, with the absorbance at 440 nm reaching 4.5 mM^{-1} cm^{-1} (green line in figure 3.14 and inset of figure 3.14) when compared with an ε_{440nm} of 5.2 mM^{-1} cm^{-1} in wild-type (figure 3.14 inset). The resulting 11.2 ms spectrum resembles that of reduced flavin ($FADH_2$)as it was observed in the titration of PfTrxR wild-type with 6 eq of sodium dithionite (data not shown; see EH_4 spectrum in figure 1.8 for an example). The second phase was slowed by 32 fold relative to wild-type (table 3.2), with the slower formation of the thiolate flavin charge transfer complex (CTC1) allowing for more of the CTC2 species to be observed (green line in figure 3.14). The third phase of the reductive half-reaction was slowed by 20 fold relative to the wild-type reaction (table 3.2) and resulted in an endpoint spectrum that displayed a much more pronounced CTC1 absorbance at 540 nm ($\varepsilon_{540\ nm}$ of 4.9 mM^{-1} cm^{-1} in E514A compared with 4.2 mM^{-1} cm^{-1} in the wild-type protein). This is consistent with the spectrum obtained in the static titrations although more CTC1 absorbance is observed in the stopped-flow experiment due to the lower temperature that the experiment was performed at (see section 3.2.6.1 for an explanation).

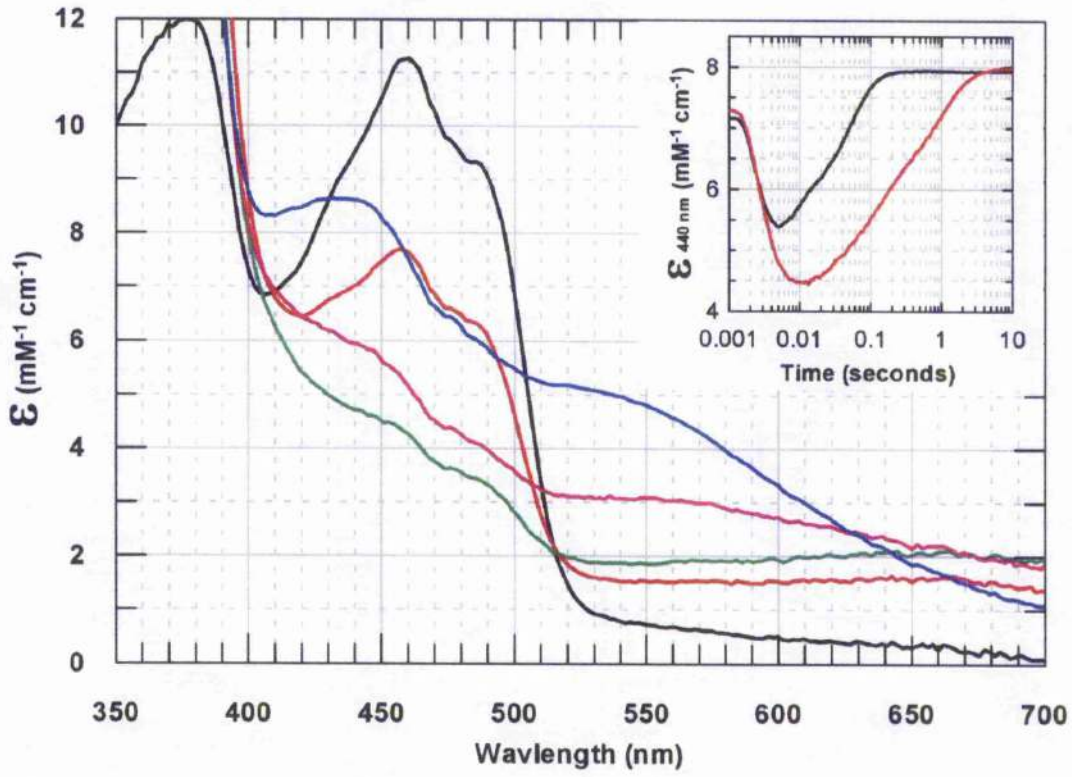


Figure 3.14 – Reductive half-reaction of PfTrxR E514A

Reduction of 12.4 μM PfTrxR E514A with 6 eq of NADPH was performed under anaerobic conditions at 5°C in 50 mM potassium phosphate pH 7.6, 1 mM EDTA. Spectra shown are taken before (Eox, black line) and 3.7 ms (red line), 11.2 ms (green line), 101.2 ms (fuchsia line) and 7.12 s (blue line) following mixing with NADPH. **Inset** displays rapid reaction kinetic traces of 440 nm for 15.6 μM PfTrxR wild-type (black line) and 12.4 μM PfTrxR E514A (red line) reduced with 6 eq of NADPH. The kinetic trace of the E514A reductive half-reaction was fitted to three exponential phases, using the KinetAsyst 3 software (Hi-Tech scientific), with the second and third phases of the reductive half-reaction being much slower than in the case of wild-type enzyme (table 3.2).

3.2.6.4 Reductive half-reaction of PfTrxR H137N

The reductive half-reaction of the PfTrxR H137N mutant was essentially the same as that for the wild-type protein (inset of figure 3.15), and could be fitted with the same three phases. Analysis of the different phases demonstrated the same features observed in the wild-type protein, with phase 1 being the formation of $FADH^-$ $NADP^+$ charge transfer complex (CTC2; red line in figure 3.15), phase 2 corresponding to the formation of the thiolate flavin charge transfer (CTC1; green line in figure 3.15) and phase 3 being the transfer of reducing equivalents to the C-terminal cysteine pair and an enhancement of CTC1 (blue line in figure 3.15).

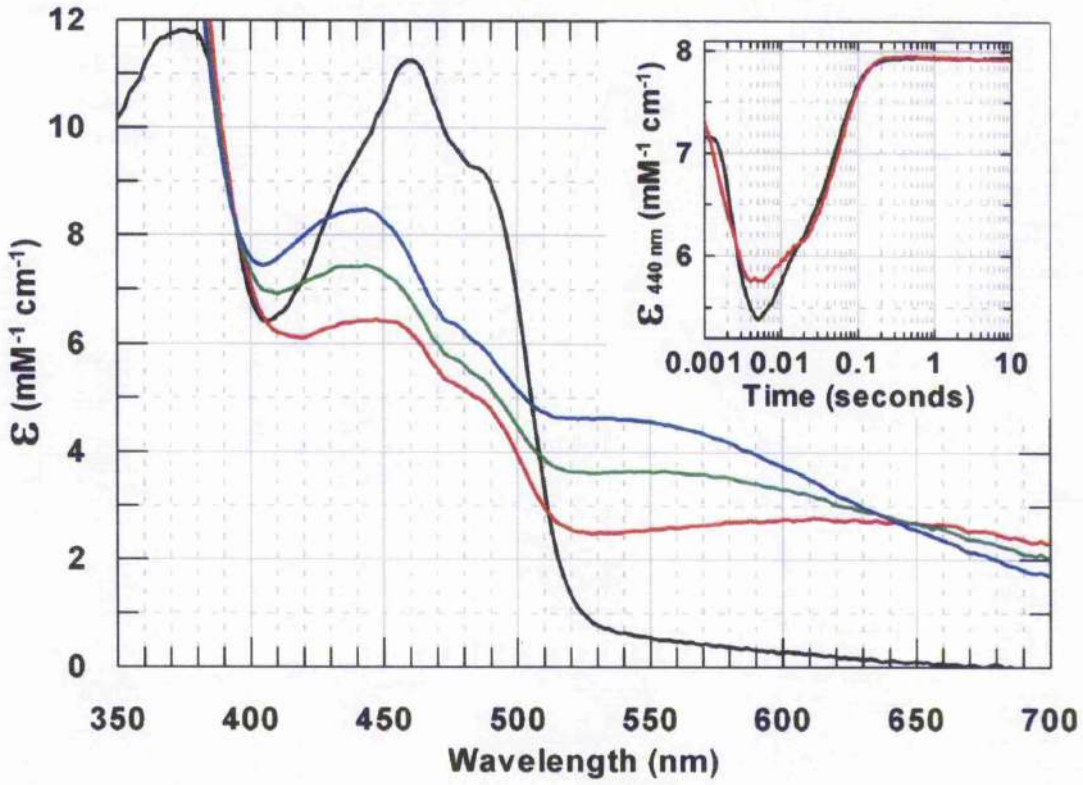


Figure 3.15 – Reductive half-reaction of PfTrxR H137N

Reduction of 14.7 μM PfTrxR H137N with 6 eq of NADPH was performed under anaerobic conditions at 5°C in 50 mM potassium phosphate pH 7.6, 1 mM EDTA. Spectra shown are taken before (Eox, black line) and 3.7 ms (red line), 30.7 ms (green line) and 1.43 s (blue line) following mixing with NADPH. **Inset** displays rapid reaction kinetic traces of 440 nm for 15.6 μM PfTrxR wild-type (black line) and 14.7 μM PfTrxR H137N (red line) reduced with 6 eq of NADPH. The kinetic trace of the H137N reductive half-reaction were essentially the same as wild-type.

3.2.6.5 Reductive half-reaction of PfTrxR C93A

Analysis of the reductive half-reaction of PfTrxR C93A mutant served as a control for the analyses of the acid base motif mutants, as mutation of Cys93 (the charge transfer thiol) means that no thiolate flavin charge transfer (CTC1) can be formed (blue line in figure 3.10A). Comparison of the flavin reduction (at 440 nm) in the C93A and wild-type reductive half-reactions shows that no further reduction occurs after the first phase (figure 3.16). This suggests that the C93A mutant can form $FADH^-$ $NADP^+$ charge transfer complex (CTC2) which can be observed at longer wavelengths (650 – 700 nm), by the decrease in flavin absorbance at 440 nm and the oxidised like shape of the flavin peak (green and blue lines and inset of figure 3.16). This also confirms that there is a small amount of absorbance displayed at 540 nm, even in a protein that cannot form the CTC1, which is also the case (although to a greater degree) in the H509Q reductive half-reaction.

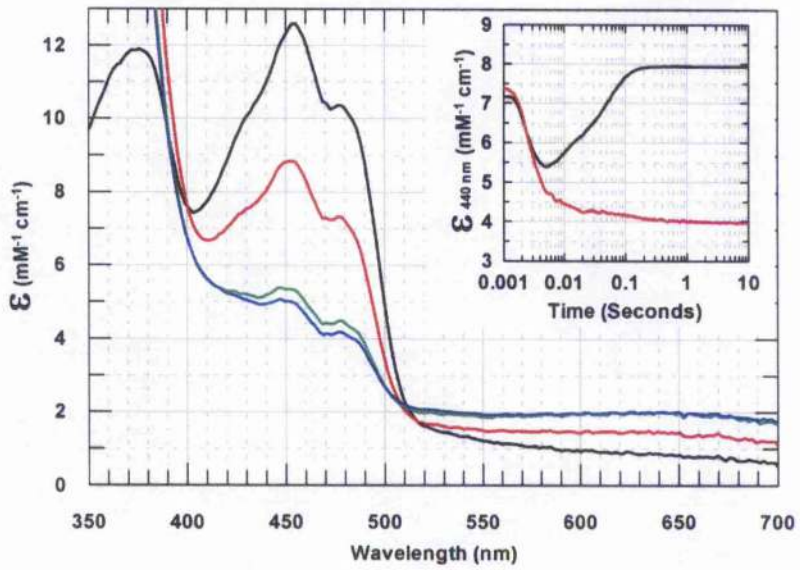


Figure 3.16 – Reductive half-reaction of PfTrxR C93A

Reduction of 19.8 μM PfTrxR C93A with 6 eq of NADPH was performed under anaerobic conditions at 5°C in 50 mM potassium phosphate pH 7.6, 1 mM EDTA. Spectra shown are taken before (Eox, black line) and 7.5 ms (red line), 22.5 ms (green line) and 14.25 s (blue line) following mixing with NADPH. **Inset** displays rapid reaction kinetic traces of 440 nm for 15.6 μM PfTrxR wild-type (black line) and 19.8 μM PfTrxR C93A (red line) reduced with 6 eq of NADPH. The spectra and kinetic traces of the C93A reductive half-reaction demonstrated that this mutant could form the $FADH^-$ $NADP^+$ charge transfer complex (CTC2; observed at 670 nm and 460 nm), but was unable to form the thiolate flavin charge transfer complex (CTC1; observed at 540 nm and 440 nm).

Protein	pH	$k_{1\_obs\ (max)}$ (s^{-1})	$k_{2\ obs}$ (s^{-1})	$k_{3\ obs}$ (s^{-1})	$K_{d\ app}$ NADPH (μM)
Wt	7.0	720 (± 134)	485[a]	20[b]	14.2 (± 3)
Wt	7.6	793 (+ 104)	360[a]	20[b]	25.5 (± 2.5)
Wt	8.2	815 (± 10)	168 (± 11)	~30[b]	32 (± 6)
H509Q	7.0	425 (± 39)	5[c]	0.08 (± 0.02)	ND[c]
H509Q	7.6	567 (± 72)	0.37 (± 0.04)	0.07 (± 0.03)	59 (± 18)
H509Q	8.2	663 (± 75)	0.63 (± 0.05)	0.07 (± 0.04)	77 (± 18)
E514A	7.6	803 (± 56)	11.2 (± 1.5)	1.02 (± 0.08)	47 (± 8)

Table 3.2 – Observed rate constants of the reductive half-reaction

The reductive half-reaction of PfTrxR with 6 equivalents of NADPH was performed at 5°C in 50 mM potassium phosphate buffer (at different pH), 1 mM EDTA. The observed rate constants for PfTrxR wild-type and mutants were fitted to three exponential phases, using the KinetAsyst 3 software (Hi-Tech scientific). [a] This kinetic phase is only observed with 2 or more equivalents. [b] A decreasing hyperbolic looking rate is observed vs μM NADPH for 440 nm traces between 0.5 Eq to over 9 Eq NADPH. [c] Not determined as saturation had already occurred, therefore k_d must be below 30 μM. Therefore the $k_{1\ obs\ max}$ is an average of all rates observed between 3 eq and 34 eq NADPH.

3.2.7 The pre-steady state oxidative half-reaction

The oxidative half-reaction of PfTrxR was studied by pre-reducing the enzyme to EH_4 with 2.2 equivalents of reductant (either sodium dithionite or NADPH), before adding between 1 and 20 equivalents of PfTrx and following the reaction in a stopped-flow spectrophotometer. These reactions were performed under anaerobic conditions at 5°C, and the data collected have been corrected to remove the baseline shift caused by the addition of excess PfTrx. Traces were collected at 490 nm and 540 nm as these wavelengths were where the largest changes occur for flavin redox and thiolate flavin charge transfer (CTC1), respectively.

3.2.7.1 Oxidative half-reaction of PfTrxR wild-type

The oxidative half-reaction of PfTrxR wild-type was initially performed by reducing the protein with 2.2 eq of sodium dithionite before analysing the oxidation of the enzyme after addition of 6 eq PfTrx. Spectra were taken of the oxidised protein (E_{ox}; black line in figure 3.17A), the sodium dithionite reduced protein (EH_4; red line in Figure 3.17A) and following oxidation of the reduced protein by PfTrx (blue line in Figure 3.17A). The sodium dithionite reduced spectrum (red line in Figure 3.17A) was characteristic of the spectrum observed during static titration of the protein with the reductant (data not shown). The spectrum taken 16 minutes after addition of PfTrx (to the reduced protein), displayed that the protein had almost returned to its oxidised state. This was shown by the similarity of the E_{ox} spectrum (black line in figure 3.17A) and PfTrx oxidised spectrum (blue line in figure 3.17A) between 350 nm and 530 nm, although some long wavelength absorbance (530 – 700 nm) remained in the PfTrx reduced spectrum. The oxidative half-reaction of the sodium dithionite reduced protein was further analysed by rapid reaction kinetics and the data collected (at 490 nm and 540 nm; figure 3.17B) were fitted to three exponential phases, with the observed rate constants of 40 s^{-1}, 0.016 s^{-1} and 0.003 s^{-1} (table 3.3). The first phase displayed no dependence on PfTrx concentration, which suggested that PfTrx is bound rapidly and that we are only observing the re-oxidation of the enzyme due to reduction of the bound PfTrx. The second and third phases are thought to be due to the slow reaction of the resultant EH_2 species with a second molecule of PfTrx disulphide.

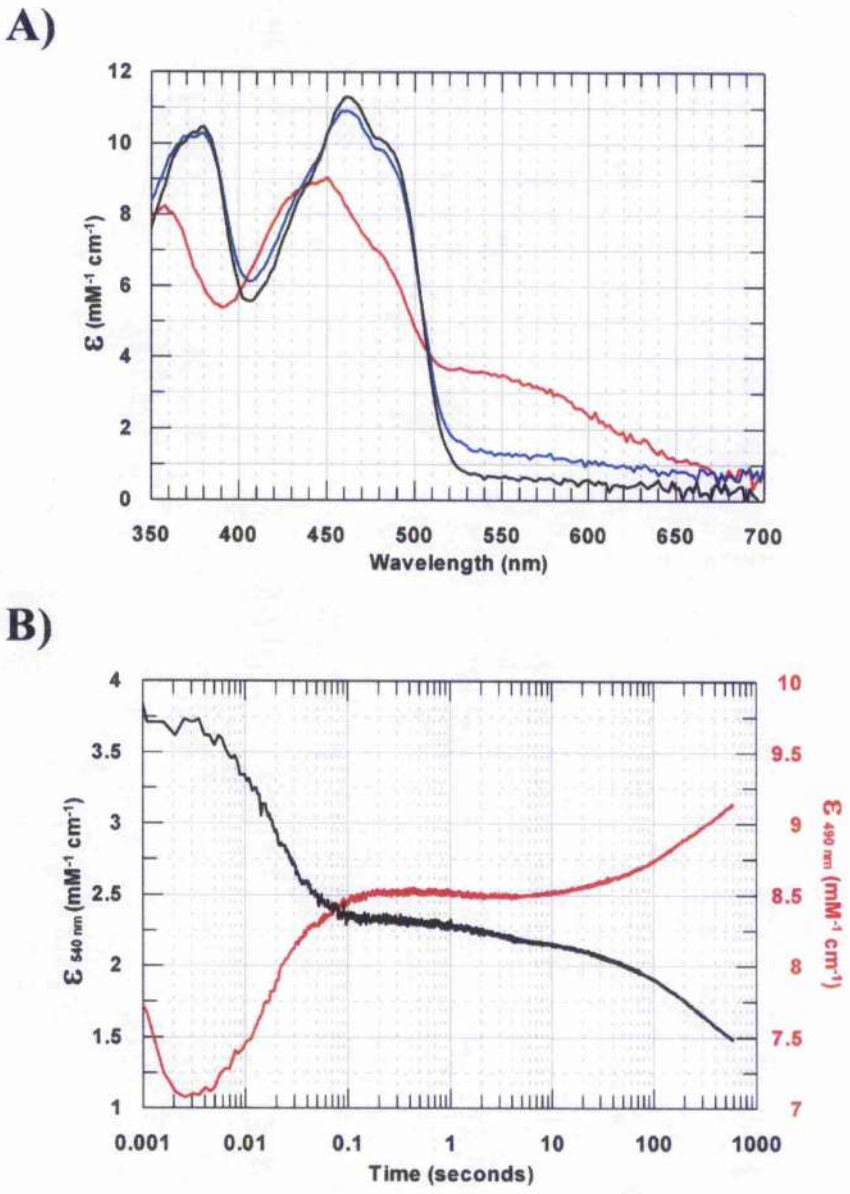


Figure 3.17 – Oxidative half-reaction of PfTrxR wild-type (dithionite reduced)

Wild-type *Pf*TrxR (16.1 μM) reduced with 2.2 eq of dithionite was reacted with *Pf*Trx under anaerobic conditions at 5 °C in 50 mM potassium phosphate pH 7.6, 1 mM EDTA. **Panel A** displays spectra were recorded before (E_{ox}; black line) and after reduction (EH_4; red line), and 16 minutes after the addition of 6 eq of *Pf*Trx (blue line). **Panel B** displays kinetic traces from the oxidative half-reaction using 6 eq of *Pf*Trx. Traces for 490 nm (red line, right hand y-axis) and 540 (black line, left hand y-axis) were measured for 10 minutes. The kinetic traces were fitted with three exponential phases with observed rates of 40 s^{-1}, 0.016 s^{-1} and 0.003 s^{-1}. Note: the apparent decrease in absorbance from 1 to 3 ms of the 490 nm trace is due to a flow artefact in this experiment.

The oxidative half-reaction of PfTrxR wild-type was also analysed following reduction with 2.2 eq NADPH and subsequent reaction with 10 eq PfTrx. Again spectra were taken of the oxidised protein (Eox; black line in figure 3.18A), after reduction with 2.2 eq NADPH (EH_4; red line in figure 3.18A) and 16 min following addition of PfTrx (blue line in figure 3.18A). The NADPH reduced spectrum was similar to those observed at 2.2 eq NADPH during the static titration of the protein (data not shown), and the PfTrx oxidised protein displayed a spectrum very similar to that of the E_{ox} protein (compare black line with blue line in figure 3.18A). The major differences observed between the E_{ox} and PfTrx reacted spectra are observed at longer wavelengths (530 nm – 700 nm), which was likely to be due to the presence of protein bound $NADP^+$ or NADPH. This spectral feature was also similar to that observed in the static titration of PfTrxR H509Q with NADPH (Figure 3.8). The oxidative half-reaction of the NADPH reduced protein was analysed by rapid reaction kinetics with the data collected (at 490 nm and 540 nm; Figure 3.18B) being fitted to three exponential phases with observed rate constants of 20-28 s^{-1}, 3.1 s^{-1} and 0.05 s^{-1}. The first phase of the NADPH reduced protein was slower than those determined for the sodium dithionite reduced protein (table 3.2), which was probably due to the presence of bound $NADP^+$ in the EH_4 reduced protein stabilising the reduced flavin (via $FADH^-$ $NADP^+$ charge transfer complex/CTC2).

The return of both sodium dithionite and NADPH reduced protein to near E_{ox} species suggests that PfTrxR may not cycle between the EH_2 and EH_4 species, as is the situation in the *D. melanogaster* TrxR (Bauer *et al.*, 2003b). From these experiments it can be deduced that PfTrxR is rapidly reduced to a PfTrxR EH_2 species, which can be observed after 200 ms in the sodium dithionite experiment and after 2 seconds in the NADPH experiment. The oxidation of this EH_2 species to an E_{ox} species proceeds at a much slower rate, which may not be catalytically relevant due to the high NADPH concentration in the cytosol (50 - 100 μM in humans; H. Schirmer, personal communication).

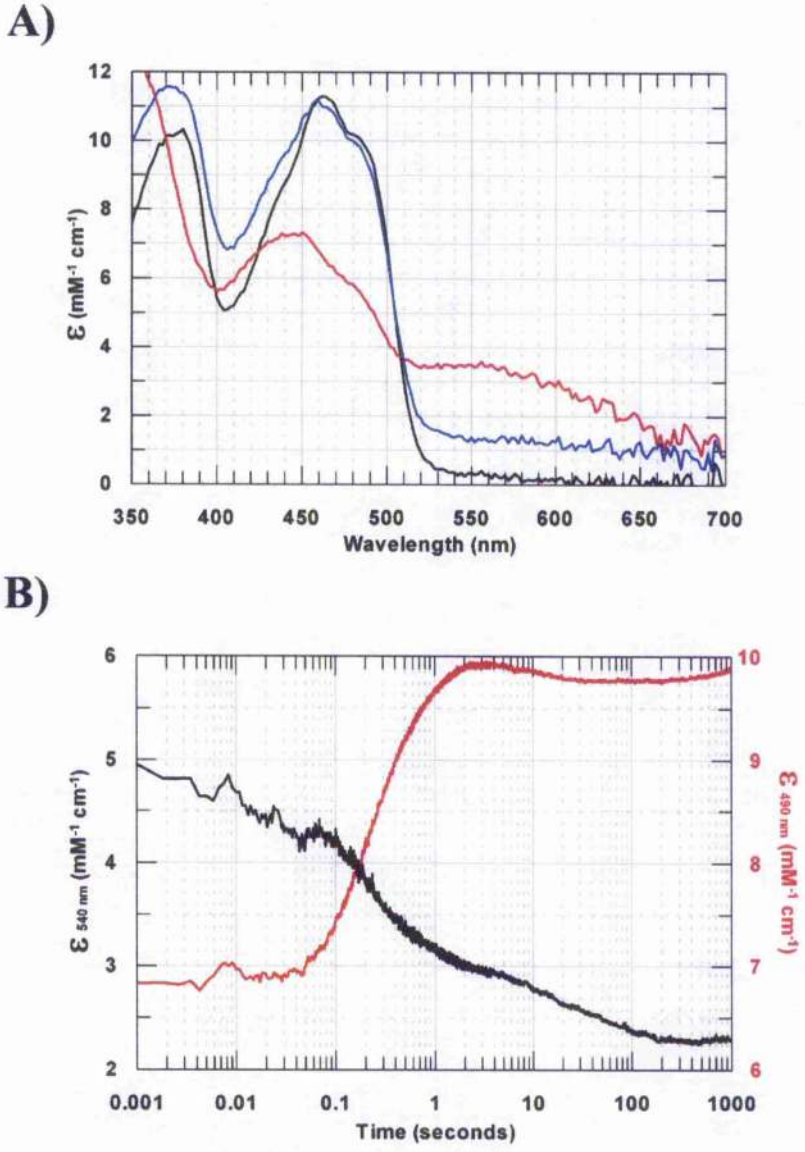


Figure 3.18 – Oxidative half-reaction of PfTrxR wild-type (NADPH reduced)

Wild-type *Pf*TrxR (10 μM) pre-reduced with 2.2 eq of NADPH was reacted with *Pf*Trx under anaerobic conditions at 5 °C in 50 mM potassium phosphate pH 7.6, 1 mM EDTA. **Panel A** displays spectra recorded before (E_{ox}; black line) and after reduction (EH_4; red line), and 16 minutes after the addition of 10 equivalents of *Pf*Trx (blue line). **Panel B** displays kinetic traces from the oxidative half-reaction using 10 eq of *Pf*Trx. Traces for 490 nm (red line, right hand y-axis) and 540 (black line, left hand y-axis) were measured for 16 minutes. The kinetic traces were fitted with three exponential phases with observed rates of 20-28 s^{-1}, 3.1 s^{-1} and 0.05 s^{-1}.

3.2.7.2 Oxidative half-reaction of PfTrxR H509Q

The oxidative half-reaction of PfTrxR H509Q was analysed following reduction with 2.2 eq NADPH and subsequent reaction with 19 eq PfTrx. Spectra were taken of the oxidised protein (E_{ox}; black line in figure 3.19A), the NADPH reduced protein (EH_4; red line in figure 3.19A) and the protein following reaction with PfTrx (blue line in figure 3.19A). The NADPH reduced spectrum was similar to those observed during the static titration of the protein with NADPH (data not shown). The spectrum following reaction with PfTrx displays similar features to those observed in the E_{ox} protein.

The oxidative half-reaction of PfTrxR H509Q was further analysed by rapid reaction kinetics with the data collected at 490 nm (red line in figure 3.19B) being fitted to three exponential phases; 11.1 s^{-1}, 0.0181 s^{-1} and 0.00013s^{-1} (table 3.3). The first phase of the PfTrxR H509Q oxidative half-reaction represented only 10% of the total extinction change at 490 nm. This suggests that flavin oxidation does not occur during this phase, but that it is more likely to be due to $NADP^+$ dissociation and/or reduction of the first equivalent of thioredoxin. Phases two and three, which are slower and are responsible for 90% of the extinction change at 490 nm, are more likely responsible for flavin oxidation as reducing equivalents pass to the C-terminal redox centre to react with the second equivalent of Trx (figure 3.19A).

The oxidative half-reaction of PfTrxR H509Q was 100 fold slower than that of the wild-type protein (compare black line and red lines in figure 3.19A). This confirms that the mutation of His509 to glutamine severely affects not only the reductive half reaction but also the oxidative half-reaction of PfTrxR, by perturbing formation of CTC1 and therefore the transfer of reducing equivalents between the redox active sites.

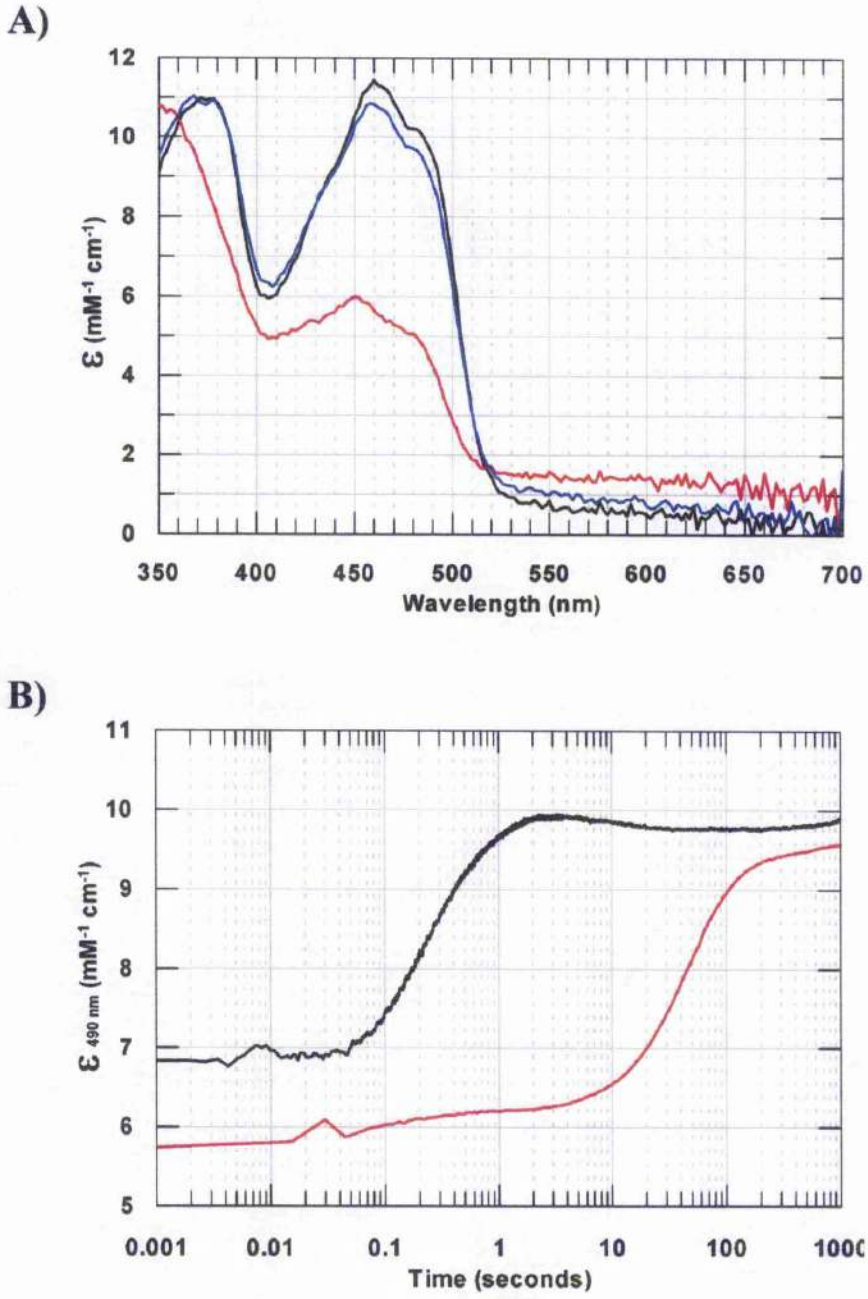


Figure 3.19 – Oxidative half-reaction of PfTrxR H509Q (NADPH reduced)

The oxidative half-reaction of 10.1 μM PfTrxR H509Q, reduced with 2.2 eq of NADPH, was mixed with 19 eq *Pf*Trx under anaerobic conditions at 5 °C in 50 mM potassium phosphate pH 7.6, 1 mM EDTA. **Panel A** displays spectra recorded before (E_{ox}; black line) and after reduction (EH_4; blue line) and 16 minutes after the addition of 19 equivalents *Pf*Trx (red line). **Panel B** displays kinetic traces from the oxidative half-reaction with *Pf*Trx. Traces at 490 nm are shown for *Pf*TrxR wild-type reacted with 10 eq of *Pf*Trx (black line) and *Pf*TrxR H509Q reacted with 19 equivalents *Pf*Trx (red line). The PfTrxR H509Q 490 nm trace was fitted with three exponential phases with observed rates of 11.1 s^{-1}, 0.0181 s^{-1} and 0.00013 s^{-1} (table 3.3).

3.2.7.3 Oxidative half-reaction of PfTrxR E514A

The oxidative half-reaction of PfTrxR E514A was analysed following reduction with 2.2 eq NADPH and subsequent addition of 18 eq PfTrx. Spectra were taken for the oxidised protein (black line in figure 3.20A), after reduction with NADPH (red line in figure 3.20A) and 16 min after reaction with PfTrx (blue line in figure 3.20A). The NADPH reduced spectrum was similar to those observed in the static NADPH titrations with this mutant (data not shown). The spectrum of the PfTrx reacted sample did not display features of the E_{ox} protein, as had been observed in the wild-type protein reacted with PfTrx (figure 3.20A). The E514A PfTrx oxidised spectrum was consistent with the presence of the thiolate flavin charge transfer (CTC1), with a high absorbance at 540 nm (2.2 mM^{-1} cm^{-1}) and a partially reduced flavin peak with a maximum at 450 nm. This suggested that, similar to the spectra observed in the reductive half-reaction of E514A (figure 3.14A), the mutant formed a stable CTC1. This prevented the production of an E_{ox} species during the oxidative half-reaction (as observed in the wild-type and H509Q proteins) and resulted in a spectrum that appeared closer to an EH_2 species than an E_{ox} species.

The oxidative half-reaction of PfTrxR E514A was further analysed by rapid reaction kinetics with the data collected at 490 nm (red line in figure 3.20B) being fitted to three exponential phases; 10 s^{-1}, 0.14 s^{-1} and 0.004 s^{-1} (table 3.3). The oxidative half-reaction of PfTrxR E514A was affected in a similar manner as that of the H509Q mutant, in that the first phase was representing flavin re-oxidation, but was most likely associated with $NADP^+$ dissociation, with phases two and three being responsible for the majority of the flavin re-oxidation. The oxidative half-reaction of PfTrxR E514A was 20 fold slower than the wild-type reaction (compare black line with red line in figure 3.20B), which confirms that Glu514 plays an important role in the transfer of reducing equivalents between the redox active centres; although it plays a clearly distinct role to that of the His509 residue.

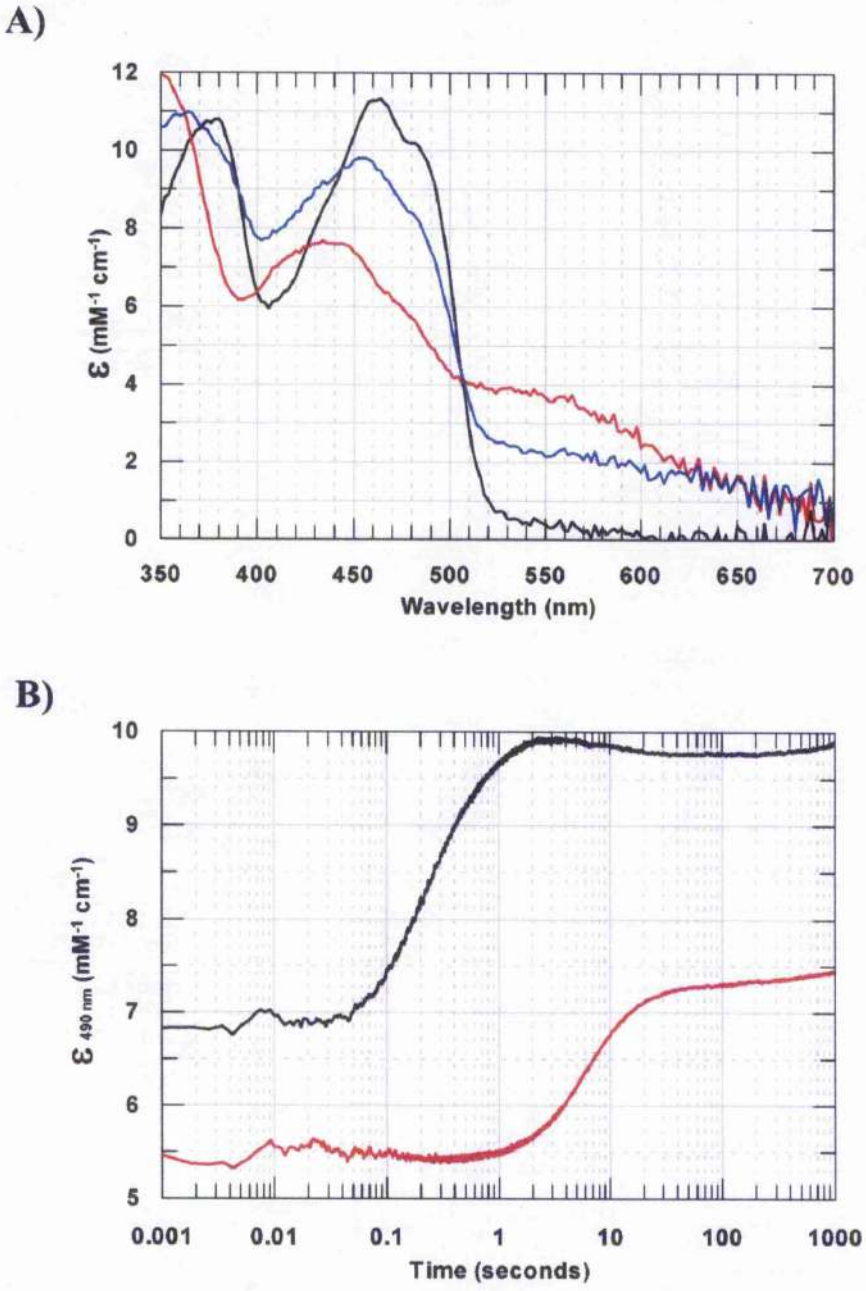


Figure 3.20 – Oxidative half-reaction of PfTrxR E514A (NADPH reduced)

8.6 μM PfTrxR E514A, reduced with 2.2 eq of NADPH, was mixed with 18 eq *Pf*Trx under anaerobic conditions at 5 °C in 50 mM potassium phosphate pH 7.6, 1 mM EDTA. **Panel A** displays spectra recorded before (E_{ox}; black line) and after reduction (EH_4; red line) and 16 minutes after the addition of 19 equivalents *Pf*Trx (blue line). **Panel B** displays kinetic traces at 490 nm for reduced *Pf*TrxR wild-type reacted with 10 eq of *Pf*Trx (black line) and *Pf*TrxR E514A reacted with 18 eq of *Pf*Trx (red line). The PfTrxR E514A 490 nm trace was fitted with three exponential phases with observed rates of 10 s^{-1}, 0.14 s^{-1} and 0.004 s^{-1} (table 3.3).

3.2.7.4 Oxidative half-reaction of PfTrxR H137N

The oxidative half-reaction of PfTrxR H137N was analysed following reduction with 2.2 eq NADPH and subsequent addition of 10 eq PfTrx. The reaction was analysed by rapid reaction kinetics at 490 nm and the observed trace (red line in figure 3.21) displayed the same features as those observed in the corresponding wild-type trace (red line in figure 3.21). The increase in absorbance 10 seconds after addition of the PfTrx was due to the formation of a fine precipitate in the reaction mix; this precipitation made analysis of the spectra collected very difficult (data not shown). From this experiment, as also observed in the static titration and reductive half-reaction, it was concluded that mutation of His137 to glutamine resulted in no change to the proteins catalytic activity. Although the kinetics of the H137N mutant were essentially the same as that of the wild-type protein, the mutant was prone to precipitation as observed in both the NADPH static titration and oxidative half-reaction.

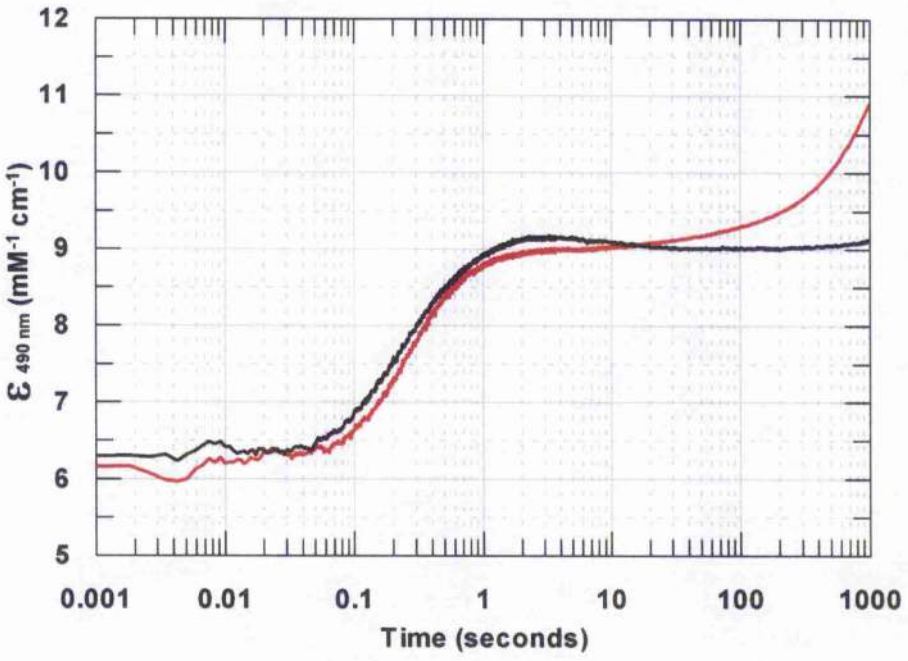


Figure 3.21 – Oxidative half-reaction of PfTrxR H137N (NADPH reduced)

9.5 μM PfTrxR H137N, reduced with 2.2 equivalents NADPH, was mixed with 10 eq *Pf*Trx under anaerobic conditions at 5 °C. The 490 nm kinetic traces from the oxidative half-reaction with *Pf*Trx are shown for *Pf*TrxR wild-type reacted with 10 equivalents *Pf*Trx (black line) and *Pf*TrxR H137N reacted with 10 eq of *Pf*Trx (red line). The kinetic trace was essentially the same as that of wild-type PfTrxR. Note: the increase in absorbance observed after 10 seconds was due to the formation of a precipitate.

Protein	Reductant	$k_{1\ obs}$ (s^{-1})	$k_{2\ obs}$ (s^{-1})	$k_{3\ obs}$ (s^{-1})
Wt	Sodium dithionite	40 (± 6)	0.016 (± 0.14)	0.003 (± 0.001)
Wt	NADPH	20 -28	3.1 (± 0.6)	0.05 (± 0.03)
H509Q	NADPH	11.1 (± 0.1)	0.0181 (± 0.0002)	0.00013 (± 1e-5)
E514A	NADPH	10 (± 2)	0.14 (± 0.02)	0.004 (± 0.003)

Table 3.3 – Observed rate constants of the oxidative half-reaction

PfTrxR pre-reduced with 2.2 equivalents of NADPH, was reacted with varying equivalents of PfTrx at 5°C in 50 mM potassium phosphate buffer pH 7.6, 1 mM EDTA. The observed rate constants for PfTrxR wild-type and mutants were fitted to three exponential phases, using the KinetAsyst 3 software (Hi-Tech scientific).

3.3 Discussion

3.3.1 Crystal structure of high M_r TrxR

In this study I have investigated the role of a number of the key active site residues involved in catalysis of the PfTrxR. Although there is no crystal structure of the PfTrxR protein, structural studies on rat TrxR and mouse mitochondrial TrxR identified that the bound NADPH, FAD and active site residues were all well positioned to aid transfer of the reducing equivalents between the redox active centres (Sandalova *et al.*, 2001; Biterova *et al.*, 2005). For example, in the rat TrxR structure the charge transfer thiol in the N-terminal cysteine pair (Cys93 in PfTrxR) is within 3.6 Å of the isoalloxazine ring of FAD (Sandalova *et al.*, 2001) allowing it to form the thiolate flavin charge transfer complex (CTC1).

The C-terminal disulphide is located on a 16-21 amino acids extension (21 amino acids in PfTrxR) when compared to glutathione reductase. Studies in the rat TrxR identified several residues involved in glutathione binding as being conserved in the TrxR structure and postulated that the presence of the C-terminal tail conferred substrate specificity for thioredoxin in the TrxR protein (Kanzok *et al.*, 2000; Sandalova *et al.*, 2001). In both high M_r TrxR structural studies the C-terminal tail had to be modelled into the structure as this region was not ordered in either structure. This flexibility of the C-terminal tail was also suggested by Gromer *et al.* who demonstrated that the C-terminal tail could be cleaved (by proteases) in the reduced but not the oxidised human TrxR (Gromer *et al.*, 1998). These findings supported the idea that the flexibility of the C-terminal tail in high M_r TrxR may play some role in catalysis, although this needs to be investigated further. This C-terminal flexibility also raises questions to the validity of thioredoxin docking models that have been proposed (Sandalova *et al.*, 2001; Brandt *et al.*, 2005). In order to elucidate the interaction of thioredoxin with TrxR it may be necessary to form a stable mixed disulphide between the two proteins, and try to solve the structure of the protein-protein complex.

3.3.2 The acid-base catalyst diad is essential for catalysis

The crystal structures of both mammalian TrxRs also confirmed the acid-base catalyst residues (equivalent to His509 and Glu514 in PfTrxR) to be suitably located to perform their postulated role within the active site. This was reminiscent to the studies on human glutathione reductase that initially proposed these residues as being important in catalysis (Pai *et al.*, 1983). Pai *et al* identified that the histidine and glutamate residues of the acid-

base catalyst were closely associated and formed a hydrogen bond in the oxidised protein. They also demonstrated that the imidazole ring of the histidine residue was closely associated with the bound FAD and the distal cysteine (equivalent to Cys88 in PfTrxR) was therefore ideally placed to act in catalysis. Mutagenesis studies on the histidine residue from glutathione reductase (from various organisms) confirmed that this residue was essential for catalysis, as mutants displayed highly reduced enzymatic activities (below 1% of wild-type) and demonstrated altered K_m values for both substrates (Berry *et al.*, 1989; Deonarain *et al.*, 1989; Scrutton *et al.*, 1990; Deonarain *et al.*, 1992).

Studies on dihydrolipoamide dehydrogenase have concentrated on the histidine and the glutamate residues with mutation studies confirming their crucial role by showing the loss of enzymatic activity, alteration of K_m values for the substrates and a change in catalytic mechanism (Williams *et al.*, 1989; Benen *et al.*, 1991; Benen *et al.*, 1992; Kim *et al.*, 1992). The change in catalytic mechanism, from ping-pong to sequential, was observed in a protein with the glutamate residue mutated to glutamine and was probably due to the alteration of K_m producing a slower release of substrates than that observed in the wild-type protein (Kim *et al.*, 1992). Mutants of the glutamate residue in *Azotobacter vinelandii* LipDH produced a shift in equilibrium in favour of the thiolate flavin charge transfer complex (Benen *et al.*, 1991), which was also observed in this study on the analogous residue (E514A) of PfTrxR.

These results reported on GR and LipDH were similar to those presented in this thesis and to studies previously carried out on PfTrxR, where mutation of His509 to glutamine resulted in 95 % loss of enzymatic activity relative to wild-type (Gilberger *et al.*, 1997). Overall the previous studies strongly suggest that the residues of the acid-base catalyst had a key role in catalysis with the histidine residue acting to stabilise catalytic intermediates and the glutamate residue acted to modulate the pKa value of the histidine residue during catalysis.

The predicted second acid-base catalyst (His137 in PfTrxR), which was proposed to stabilise the C-terminal cysteines during thiol-disulphide exchange (Gromer *et al.*, 2003), has been shown in this study to not act as an acid-base catalyst. This was also confirmed by Jacob *et al.*, who mutated the corresponding residue in the *Drosophila melanogaster* TrxR into asparagines, glutamine and phenylalanine (Jacob *et al.*, 2005). These mutants displayed a marked decrease in activity with DmTrx (90% reduction for the glutamine mutant), although the substrate K_m and pH optima of the mutants were not altered from that of wild-type (Jacob *et al.*, 2005). Jacob *et al.* postulated that the residue was more

likely to have a structural role in the protein than act as an active site acid-base catalyst. A closely associated tyrosine residue (equivalent to Tyr145 in PfTrxR) has also been studied as a possible acid-base catalyst in *E. coli* glutathione reductase, although the residue was not observed to play any such function in catalysis (Deonarain *et al.*, 1989).

3.3.3 The catalytic mechanism of PfTrxR

High M_r thioredoxin reductases contain three redox active centres; a bound FAD cofactor, an N-terminal disulphide and a C-terminal disulphide. Previous studies on the *P. falciparum* TrxR have demonstrated that these centres are in redox communication (Wang *et al.*, 1999; Williams *et al.*, 2000). The catalytic mechanism of the high M_r TrxR has been proposed based on studies carried out on the *D. melanogaster* TrxR protein (Bauer *et al.*, 2003b) and I have modified this model to explain the data observed during my studies on the PfTrxR (figure 3.22).

In the proposed catalytic mechanism the oxidised protein is reduced to an EH_2 species by 1 molar equivalent of NADPH. This reduction initially forms a $FADH^-$ $NADP^+$ charge transfer complex (CTC2) on the *re* side of the flavin that is visible as a spectral intermediate in the reductive half reaction. This was attributable to the first phase in the reductive half reaction of all PfTrxR forms and proceeded at a faster rate than that observed for the *D. melanogaster* TrxR (Bauer *et al.*, 2003b). The reductive half-reaction of both proteins were complete after 200 ms, however if one takes into account the temperature that the experiment was performed at (5^0C for PfTrxR and 10^0C for DmTrxR), the PfTrxR has a much faster reductive half-reaction than DmTrxR.

Reducing equivalents are then passed to the *si* side of the flavin and form a thiolate flavin charge transfer complex (CTC1; EH_2 (B) in figure 3.22) with Cys93. Analysis of the C93A mutant's reductive half reaction displayed that it was able to form CTC2 but not CTC1, confirming this residue as the charge transfer residue (figure 3.16). Another residue important in the formation of EH_2 (B) is His509 of the acid/base catalyst motif, as mutation of this residue to glutamine resulted in a loss of the CTC1 spectral intermediate during the reductive half reaction. This was most dramatically demonstrated in the static reduction with sodium dithionite as the resulting spectrum lacked the spectral properties of the bound NADP(H) (figure 3.8). The His509 residue is therefore postulated to stabilise the CTC1 by de-protonating the Cys93 thiol formed in this charge transfer complex. Mutation of the Glu514 residue of the acid/base catalyst motif resulted in an enhancement of the EH_2 (B) spectral intermediate during the reductive half reaction. This may be due to the loss of the

negative charge of Glu514 favouring the de-protonation of the Cys93 thiol, which results in an accumulation of the EH_2 (B) species (figure 3.22). Reduction of PfTrxR wild-type by one equivalent of NADPH results in an equilibrium of EH_2 species (A – D), although the major species is EH_2 (B). Formation of EH_2 (B) was attributable to the second observed phase in the reductive half reactions of PfTrxR proteins and this phase proceeded much slower in H509Q and E514A forms than in PfTrxR wild-type, confirming the crucial roles that both these residues play in this part of the catalytic mechanism.

Reaction of a second equivalent of NADPH results in the further enhancement of CTC1 spectral intermediate and produces a four electron reduced form (EH_4 A – B in figure 3.22). This was attributed to the third observed phase of the reductive half reaction and this phase also proceeded slower in the H509Q and E514A mutants. This was postulated to be due to the inhibition of CTC1 formation (observed in EH_2 B, EH_2 C and EH_4 B in figure 3.22) in H509Q or the accumulation of CTC1 in E514A resulting in a slower transfer of reducing equivalents through the active site of the mutant forms. The faster reaction of the PfTrxR H509Q mutant at higher pH (coupled with the lack of wild-type pH dependency) was further evidence that this residue acted as an acid/base catalyst.

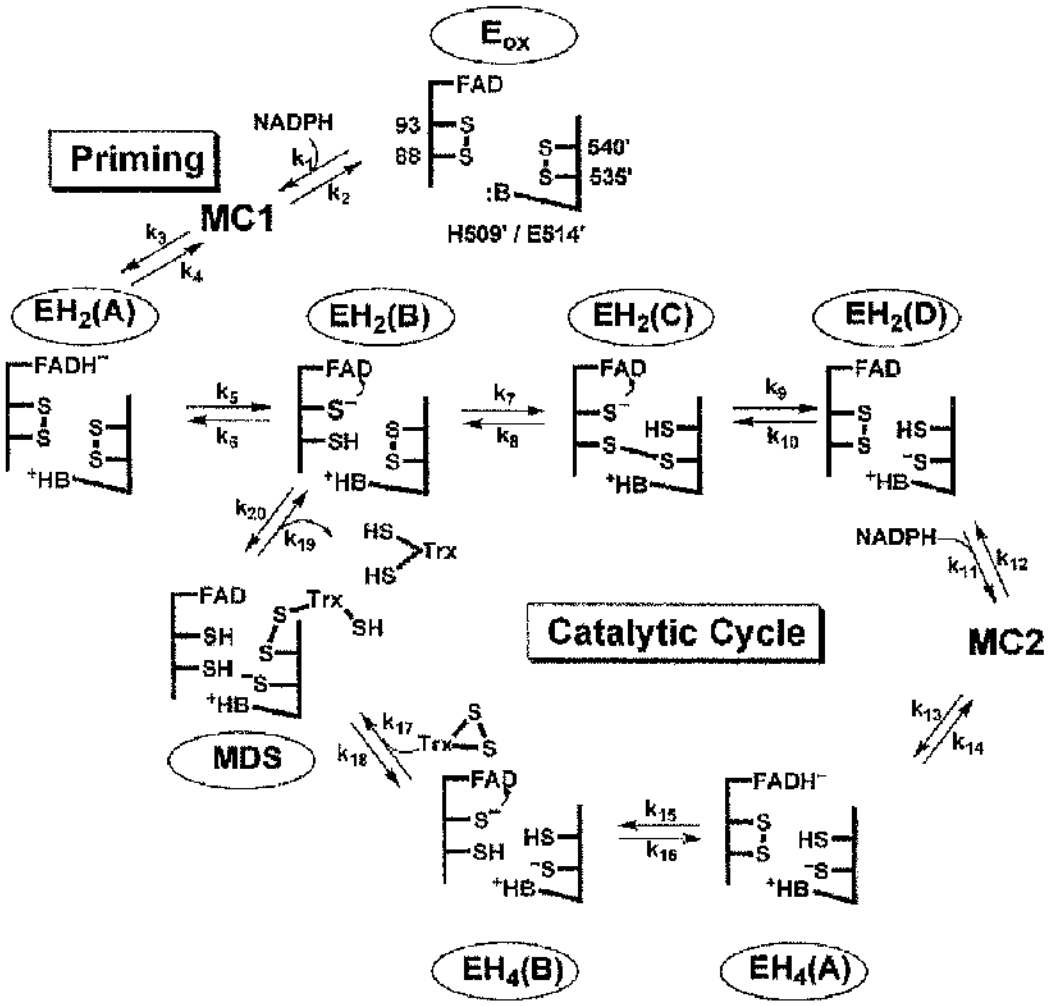


Figure 3.22 – The proposed catalytic mechanism of PfTrxR

This model of the catalytic cycle of *Pf*TrxR has been adapted from Bauer *et al*. The oxidised protein is 'primed' by NADPH forming a Michaelis complex (MC) which leads to a two electron reduced species (EH_2 A-D). EH_2 B displays the thiolate flavin charge transfer complex (CTC1; dashed arrows indicate charge-transfer interactions), and reducing equivalents are passed from the N-terminal cysteines (EH_2 B) to the C-terminal cysteines (EH_2 D) via the formation of a disulphide bond between Cys88 and Cys540 (EH_2 C). A second NADPH then forms a second Michaelis complex (MC2) which results in the four electron reduced form of the protein (EH_4 A – B). The thioredoxin substrate then binds to the protein and oxidises the C-terminal cysteines via the formation of a mixed disulphide (MDS), before the reduced thioredoxin is released and the PfTrxR returns to the two electron reduced state (EH_2B). The reductive half-reaction involves the steps designated by rate constants k_1-k_{16}, and the oxidative half-reaction by the rate constants k_{17}-k_{20}.

Analysis of the reaction of PfTrxR EH_4 species with Trx (the oxidative half reaction) demonstrated three observable phases. The first phase was attributable to the reduction of Trx (via a mixed disulphide; MDS in figure 3.22) and subsequent production of a PfTrxR EH_2 species (equilibrium of EH_2 A – D). The steps involved in oxidation of the EH_4 species (such as Trx SH_2 dissociation and formation of EH_2 equilibrium) are first order reactions and this can therefore explain the independence of the rate (of the first phase) on Trx concentration. The second and third phases proceeded much slower than the first and represented the oxidation of PfTrxR EH_2 to E_{ox} by Trx (represented by k_{10}, k_8, k_6, k_4 and k_2 in figure 3.22). The slower rate of EH_2 oxidation to E_{ox} suggests that PfTrxR, like DmTrxR, cycles in catalysis between EH_2 and EH_4 (Bauer *et al.*, 2003b). This is further likely to be true due to the high cytosolic NADPH concentration that would readily reduce any E_{ox} that does form back to EH_2. The oxidative half reactions of wild-type PfTrxR reduced with sodium dithionite and NADPH display major differences. The sodium dithionite reduced EH_4 species reacts much faster with Trx than the NADPH reduced EH_4 species (compare figure 3.17B with figure 3.18B). This is thought to be due to the different equilibria that form during reduction of the proteins (i.e. the formation of a $FADH^-$ $NADP^+$ intermediate in the NADPH reduced and not in the sodium dithionite reduced). The oxidative half reactions of the H509Q and E514A mutants reduced with NADPH are also slower than that of the wild-type PfTrxR (compare figures 3.19 and 3.20 with figure 3.18). The oxidative half reaction of PfTrxR H509Q is substantially slowed but oxidises to near the same extent as the wild type reaction (compare wild-type and H509Q extinctions at 1000s in figure 3.19). This is likely to be due to impaired transfer of reducing equivalents from the N-terminal to the C-terminal redox centres. The E514A mutant does not oxidise to the extent that the wild-type and H509Q forms do (compare figure 3.20 with figures 3.18 and 3.19). This is probably due to the loss of the negative charge of Glu514 encouraging the formation of CTC1 and therefore 'trapping' reducing equivalents in this intermediate. This results in an inhibition of transfer of reducing equivalents between N-terminal and C-terminal redox centres and can be observed in the oxidative half reaction as well as the reductive half reaction.

From these experiments it can be concluded that the residues of the acid/base catalyst motif play crucial roles in the transfer of reducing equivalents between the redox active centres and that the PfTrxR protein in likely to cycle between EH_2 and EH_4 *in vivo*.

3.3.4 Importance of the C-terminal tail in high Mr TrxR

The C-terminal tail of the high Mr TrxR is located in the interface domain and is the site of Trx binding and reduction. The flexibility observed in this tail has been suggested to be important for the catalytic mechanism (Gromer *et al.*, 1998). Mutation of residues (outwith the active site cysteine/selenocysteine residues) within this C-terminal tail have also been shown to alter the enzymatic activity of the protein, and have identified that the distribution of residues in this tail is tailored to the active site cysteine/selenocysteine residues involved in catalysis (Gromer *et al.*, 2003).

It may therefore be possible to utilise the major differences observed at the C-terminal redox centres to specifically inhibit a TrxR from one organism, while not inhibiting those from other organisms. The development of an inhibitor that would work specifically against the PfTrxR (but not the human TrxR), may result in a potent antimalarial. Indeed, the C-terminal disulphide in *P. falciparum* has already been shown to be a site of inhibition by Mannich bases (Davioud-Charvet *et al.*, 2003). Mannich bases are intermediates of the Mannich reaction, and is the reaction product of carbonyl and an amine compound that engage in a nucleophilic addition reaction followed by elimination to the imine salt. These Mannich bases were initially identified in high throughput screens of the Pfizer compound library as having possible anti-thioredoxin reductase activity and are dependent on the enzyme mechanism. The inactivation at the C-terminal disulphide was proven by the loss of enzymatic activity with PfTrx as a substrate and is a result of chemical modification of the C-terminal cysteines by the inhibitor (Davioud-Charvet *et al.*, 2003).

3.4 Conclusions

In this study I have investigated the reaction mechanism and role within that mechanism of an acid-base catalyst motif in the high M_r TrxR from *Plasmodium falciparum*.

My studies have shown that the PfTrxR reaction mechanism is similar to that proposed for other high M_r TrxR based on previous studies on the *D. melanogaster* protein. Although some major differences in the catalysis of the DmTrxR and PfTrxR proteins were observed:

1. PfTrxR appears to have a faster reductive half-reaction

2. PfTrxR has a slower oxidative half-reaction

3. PfTrxR appears to be able to return to E_{ox} following oxidation with PfTrx (*in vitro*).

Investigations into the proposed acid-base catalyst residues have shown that H509 and E514 of the conserved motif are crucial for the efficient transfer of electrons through the redox active centres and also display distinct roles.

1. The H509Q mutant inhibited formation of the thiolate flavin charge transfer complex (CTC1) and this confirmed that the histidine residue was vital in stabilising the formation of this reaction intermediate.

2. The E514A mutant enhanced the level of CTC1 observed and is thought to be involved in modulating the pKa of the histidine residue during catalysis.

Analysis of a proposed acid-base catalyst (His137) demonstrated that this residue does not form a function important in catalysis as observed by the wild-type like parameters of His137 mutants.

Chapter 4 *Plasmodium falciparum* possesses two distinct dihydrolipoamide dehydrogenases

4.1 Introduction

4.1.1 Dihydrolipoamide dehydrogenase

Dihydrolipoamide dehydrogenase (LipDH) is a member of the disulphide oxidoreductase family that includes thioredoxin reductase, glutathione reductase, trypanathione reductase and mercuric ion reductase (Williams, 1976). LipDH is an essential component of α-keto acid dehydrogenase multi-enzyme complexes (Perham *et al.*, 1996; Douce *et al.*, 2001; Mooney *et al.*, 2002), where it is responsible for the re-oxidation of dihydrolipoamide covalently bound to the complexes (figure 1.10; figure 4.1).

LipDH is a homodimeric protein that uses a bound flavin (FAD) cofactor and a redox active cysteine pair to transfer reducing equivalents from dihydrolipoamide to NAD^+ producing NADH and lipoamide (figure 4.1).

4.1.2 Aims

The aims of this part of the study were to identify and characterise the genes that encode dihydrolipoamide dehydrogenase in the *Plasmodium falciparum* genome. The genes were then to be amplified and cloned for recombinant expression, biochemical characterisation and subcellular localisation studies.

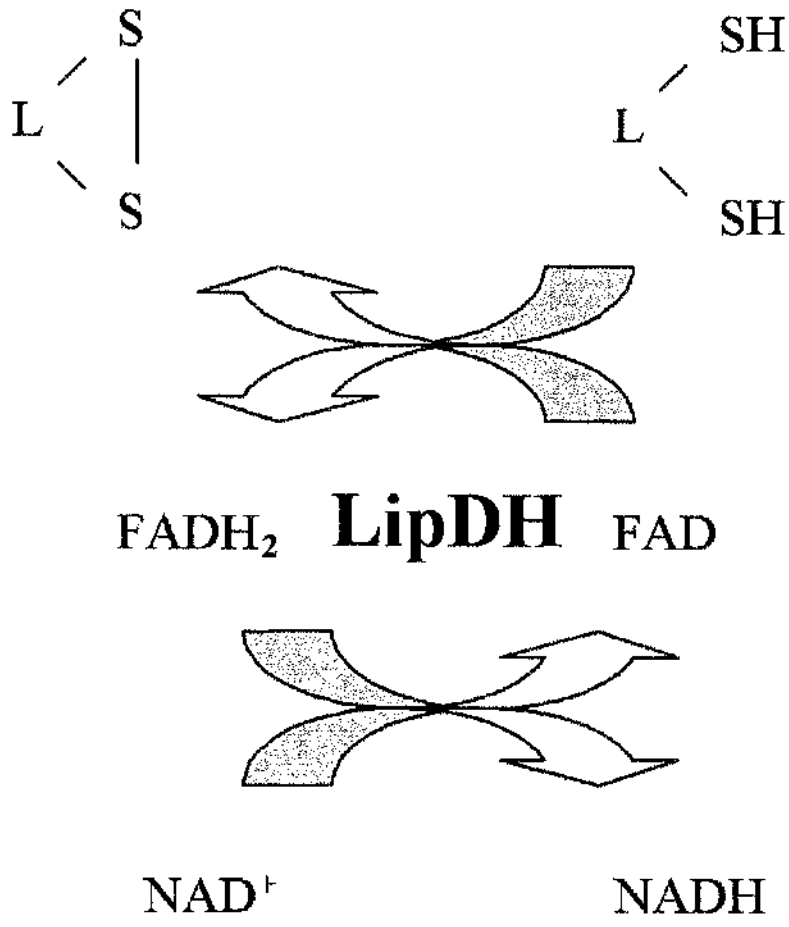


Figure 4.1 – The catalytic mechanism of dihydrolipoamide dehydrogenase (LipDH)

The physiological reaction of LipDH catalyses the oxidation of dihydrolipoamide, which results in the production of lipoamide and NADH. The LipDH protein can also catalyse the reduction of lipoamide by NADH in the reverse reaction. LipDH is a member of the disulphide oxidoreductase family and requires a protein bound flavin cofactor (FAD) for transfer of reducing equivalents from one substrate to the other.

4.2 Results

4.2.1 Identification of two distinct dihydrolipoamide dehydrogenase genes.

The *Plasmodium falciparum* genome database (www.plasmodb.org) was searched by TBLASTN using the human (P09622) and yeast (P09624) LipDH amino acid sequences as queries. These searches identified two LipDH open reading frames (ORFs) located on chromosomes 8 and 12, which I named *LipDH1* and *LipDH2*, respectively. LipDH1 had BLAST scores of 211 and 191 with the human and yeast proteins, respectively. LipDH2 had BLAST scores of 316 and 332 with the human and yeast proteins, respectively. TBLASTN searches identified the related disulphide oxidoreductases glutathione reductase (with BLAST scores of 135 and 139 for the human and yeast LipDH proteins, respectively) and thioredoxin reductase proteins (with BLAST scores of 150 and 132 for the human and yeast LipDH proteins, respectively) from *P. falciparum* as the next best hits.

4.2.1.1 Sequence analyses of LipDH1

The *LipDH1* gene identified on chromosome 8 (which was later annotated PF08_0066 in PlasmoDB) was 2001 bp in length and displayed 21 % and 23.3 % amino acid identity to human and yeast *LipDH* genes respectively. The gene encodes a 667 amino acid protein with a predicted molecular mass of 75.6 kDa (figure 4.2). Comparison of *P. falciparum* LipDH1 with predicted homologues from other *Plasmodium* species confirmed the predicted start site (data not shown). Homologues identified in *P. gallinecium*, *P. yoelii yoelii*, *P. berghei*, *P. knowlesi* and *P. vivax* had 74 %, 60 %, 59 %, 56 % and 53 % amino acid identity, respectively.

When the *LipDH1* deduced amino acid sequence was subjected to BLASTP analysis, it displayed greatest amino acid identity to the plastid isoforms of *Oryza sativa* (Japanese rice; 36 % identity), *Arabidopsis thaliana* (Thale cress; 35 % identity), *Trichodesmium erythraceum* (Cyanobacteria; 38 % identity) and *Nostoc punctiforme* (Cyanobacteria; 37 % identity). This observed homology to plastid isoforms of LipDH suggests that LipDH1 may be localised to the apicoplast in *P. falciparum*. The observation of a large N-terminal extension in the LipDH1 deduced amino acid sequence (when compared with the human and yeast proteins; data not shown) suggests the presence of the bipartite signal/transit peptide that is responsible for apicoplast targeting in *P.*

falciparum (Foth *et al.*, 2003). In order to further investigate the possible subcellular localisation of LipDH1, I analysed the amino acid sequence using sub-cellular predictions programs (see section 2.2.2.3). The results obtained from these programs confirmed the possible apicoplast localisation for LipDH1 with good confidence (95%) and a predicted signal peptide cleavage site after amino acid 33 (table 4.1; highlighted blue in figure 4.1).

```
ATGGTCATAAGGCAAAATATTAAACATATCGTTAAACTTAACGTCGTTACTCTAATTTGG  60
M  V  I  R  Q  N  I  K  H  I  V  K  L  N  V  V  T  L  I  W   20

TTATCTTATCTTTTTCTTCTTAAACCTCATGGAACTTTGAAAAATATGATGGTGTGTAAT 120
L  S  Y  L  F  L  L  K  P  H  G  T  ▮  K  N  M  M  V  C  N   40

GCTGTCCTTCTTCCATTTAATGAGAAAAATAAAGGCATAAATAATTTTGTGTATATTAAC 180
A  V  L  L  P  F  N  E  K  N  K  G  I  N  N  F  V  Y  I  N   60

CCAAAGAATATTATTTTGAATAAAATAAAAAAAGATGTCATAAAACTAGAAAAAGATAAT 240
P  K  N  I  I  L  N  K  I  K  K  D  V  I  K  L  E  K  D  N   80

ATCATCCTCTGCCAACACAATAGAAAAAGAGATAACTATATTAAACAACAAAAACGAAAA 300
I  I  L  C  Q  H  N  R  K  R  D  N  Y  I  K  Q  Q  K  R  K  100

GAAAAAAACGCGAATAATTTTACATTTATGTTAAAAGGAAGTACACAAAATATTATGAAC 360
E  K  N  A  N  N  F  T  F  ▮  L  K  G  S  T  Q  N  I  M  ▮  120

ATTAATGAAAAAGAATATGATCTTGCTATAATCGGTTGTGGTGTTGGAGGACATGCGGCC 420
I  N  E  K  E  Y  D  L  A  I  I  G  C  G  V  G  G  H  A  A  140

GCAATTAATGCTATGGAAAGAAATTTAAAAGTCATTATATTTGCAGGGGATGAAAATTGT 480
A  I  N  A  M  E  R  N  L  K  V  I  I  F  A  G  D  E  N  C  160

ATTGGAGGAACATGTGTTAATGTTGGCTGTATACCAAGCAAAGCGCTATTATACGCTACA 540
I  G  G  T  C  V  N  V  G  C  I  P  S  K  A  L  L  Y  A  T  180

AATAAATATAGAGAATTAAAAAATTTAGATAAATTATATTATTATGGTATCCATAGCAAT 600
N  K  Y  R  E  L  K  N  L  D  K  L  Y  Y  Y  G  I  H  S  N  200

ATTTTTCAAAATAATAAAAATACAGAAATTGAAAATAATCAACTCGTTTCAAACAGCTTC 660
I  F  Q  N  N  K  N  T  E  I  E  N  N  Q  L  V  S  N  S  F  220

CAAATTAATATTACGAAATTGAAAGAATACACACAAAGTGTTATTGACAAATTAAGAAAT 720
Q  I  N  I  T  K  L  K  E  Y  T  Q  S  V  I  D  K  L  R  N  240

GGAATTTCACATGGATTTAAAACATTAAAATTTAATAAAAATTCTGAACATGTTCAGGTA 780
G  I  S  H  G  F  K  T  L  K  F  N  K  N  S  E  H  V  Q  V  260

ATTTATGAACATGGTCAGCTATTAGATAAAAATACTATAAAAAGTAAAAAAAGTGGTAAT 840
I  Y  E  H  G  Q  L  L  D  K  N  T  I  K  S  K  K  S  G  N  280

ACATATAAAGTAAAAAATATCATTATAGCAACAGGATCTGTACCTAATATTCCAAATAAC 900
T  Y  K  V  K  N  I  I  I  A  T  G  S  V  P  N  I  P  N  N  300

GTTGAGATAGATGACAAGAGTGTTTTTACAAGTGATATGGCAGTAAAATTAGTTGGTTTA 960
V  E  I  D  D  K  S  V  F  T  S  D  M  A  V  K  L  V  G  L  320

AAAAATTATATGAGTATAATTGGTATGGGAATAATTGGTTTAGAATTTGCTGATATATAT 1020
K  N  Y  M  S  I  I  G  M  G  I  I  G  L  E  F  A  D  I  Y  340

ACAGCCTTAGGTTCAGAAATAACATTTTTGGAATATTCTTCTGAATTATTACCAATAATT 1080
T  A  L  G  S  E  I  T  F  L  E  Y  S  S  E  L  L  P  I  I  360

GATAATGATGTTGCAAAATATTTTGAAAGGGTATTTTTAAAAAATAAACCTGTAAATTAC 1140
D  N  D  V  A  K  Y  F  E  R  V  F  L  K  N  K  P  V  N  Y  380

CATTTAAATACAGAGGTTAAATATATAAAGGCTTCTAAAAATAATAATCCGGTTATTATT 1200
H  L  N  T  E  V  K  Y  I  K  A  S  K  N  N  N  P  V  I  I  400

GGATATTCACATAGAACTGGTAATGATGATAATGAAAAAAAAAATATGACAGATGTTAAA 1260
G  Y  S  H  R  T  G  N  D  D  N  E  K  K  N  M  T  D  V  K  420

GAATTATATGTTGATAGTTGTTTAGTTGCTACTGGACGAAATCCAAATACACAAAACTTA 1320
E  L  Y  V  D  S  C  L  V  A  T  G  R  N  P  N  T  Q  N  L  440
```

```
GGTTTGGAAAAATTAAAAATACAAATGAATAGAGGATATGTTTCAGTAAATGACAATTTA 1380
G  L  E  K  L  K  I  Q  M  N  R  G  Y  V  S  V  N  D  N  L   460

CAAGTAAAAATGGAAAATAATGAAATTTATGATAATATATTTTGTATAGGAGATGCAAAT 1440
Q  V  K  M  E  N  N  E  I  Y  D  N  I  F  C  I  G  D  A  N   480

GGGAAACAAATGTTAGCACACACAGCATCATATCAAGCATTAAAAGTCATTGATTTTATT 1500
G  K  Q  M  L  A  H  T  A  S  Y  Q  A  L  K  V  I  D  F  I   500

GAAAAGAAAGAAAAAAAGAATGTAAACATAAATGTTGAAAATAATTTGAGTAAACCCATA 1560
E  K  K  E  K  K  N  V  N  I  N  V  E  N  N  L  S  K  P  I   520

TTATATAAAAATATACCTTCTGTTTGTTATACAAATCCTGAATTAGCTTTTATCGGATTA 1620
L  Y  K  N  I  P  S  V  C  Y  T  N  P  E  L  A  F  I  G  L   540

ACAGAAAAAGAAGCAAAAGTATTATATCCAGATAATGTCGGTGTTGAAATTTCATATTAT 1680
T  E  K  E  A  K  V  L  Y  P  D  N  V  G  V  E  I  S  Y  Y   560

AAATCAAATTCAAAAATATTGTGTGAAAATAATATTTCCTTAAATAATAATAAAAAAAAT 1740
K  S  N  S  K  I  L  C  E  N  N  I  S  L  N  N  N  K  K  N   580

AATTCATATAATAAAGGACAATATAACATAAATGATAATACAAATGGTATGGTAAAAATA 1800
N  S  Y  N  K  G  Q  Y  N  I  N  D  N  T  N  G  M  V  K  I   600

ATATATAAAGAAGACACAAAGGAAATATTAGGAATGTTCATAGTTGGAAATTATGCTTCC 1860
I  Y  K  E  D  T  K  E  I  L  G  M  F  I  V  G  N  Y  A  S   620

GTTTTAATTCATGAAGCAGTATTGGCAATTAATCTGAAATTATCAGCATTCGATTTGGCA 1920
V  L  I  H  E  A  V  L  A  I  N  L  K  L  S  A  F  D  L  A   640

TATATGGTTCATTCTCATCCAACAGTAAGTGAAGTTCTGGACACAGCTTTTAAATCTATA 1980
Y  M  V  H  S  H  P  T  V  S  E  V  L  D  T  A  F  K  S  I   660

TCAAAAATAAGAACTCACTAA                                        2001
S  K  I  R  T  H  *                                           667
```

Figure 4.2 – LipDH1 nucleotide and amino acid sequence

The cDNA sequence of Pf*LipDH1*, with the deduced amino acid sequence displayed underneath. Amino acid numbering includes the initiator methionine; the termination codon is marked with an asterisk. This figure also highlights the amino acids of the predicted signal peptide cleavage site (highlighted in blue), the predicted start site of the mature protein used for recombinant expression (highlighted in red) and the last amino acid before attachment to GFP in the GFP-fusion construct (highlighted in green).

4.2.1.2 Sequence analyses of LipDH2

The *LipDH2* gene identified on chromosome 12 (which was later annotated PFL1550w in PlasmoDB) was initially identified to be 1497 base pairs in length. This gene (named *LipDH2A*) encoded a protein comprising of 499 amino acids with a predicted molecular mass of 55.6 kDa. However, ClustalW alignment of this deduced amino acid sequence with the predicted open reading frames of homologues from other *Plasmodium* species suggested the presence of an intron at the N-terminus of the gene. This prediction was due to the un-conserved nature of the LipDH2A translation initiation site (Met) and the apparent loss of identity in the N-terminal 10 – 20 amino acids (figure 4.3A). Analyses of the genomic DNA sequence from *P. falciparum LipDH2* identified possible intron/exon boundaries in the 5' end of the gene. The predicted intron produced a shift of frames between the two exons, from frame +1 in exon 1 to frame +3 in exon 2 (figure 4.3B) and displayed a much higher A/T content (86.6 %) when compared to that of 64.8 % and 72.1 % for exons 1 and 2, respectively. The positions of the intron/exon boundaries were conserved in the amino acid sequences between the *Plasmodium* species (figure 4.3C), although the length of the predicted intron did vary between the species; with the introns being 216 bp for *P. falciparum*, 182 bp for *P. berghei*, 185 bp for *P. chabaudi*, 239 bp for *P. vivax* and 187 for *P. yoelii yoelii*.

I named the *P. falciparum* gene that contained the intron, *LipDH2B*. The genomic sequence of the *LipDH2B* gene was 1763 bp in length, while the complementary DNA (following removal of the intron from mRNA) was 1539 bp and encoded for a protein containing 513 amino acids with a predicted molecular mass of 57.2 kDa (figure 4.4). The deduced amino acid sequence of LipDH2B displayed high levels of amino acid sequence identity with homologues from other *Plasmodium* species *P. yoelii yoelii* (79 % identity), *P. vivax* (73 % identity), *P. chabaudi* (79 % identity) and *P. berghei*(78 % identity).

Subsequent BLASTP analyses of the predicted amino acid sequence displayed that LipDH2B had greatest amino acid identity to mitochondrial isoforms of LipDH from *Euglena gracilis* (Protist; 40 % identity), *Dictyostelium discoideum* (Slime mould; 38 % identity), *Pisum sativum* (Pea plant; 38 % identity), *Lycopersicon esculentum* (Tomato plant; 38 % identity), *Solanum tuberosum* (potato plant; 37 % identity) and *Aspergillus fumigatus* (Fungi; 38 % identity).

As LipDH2B had greatest amino acid identity with mitochondrial isoforms of LipDH, I investigated this possible mitochondrial localisation further by analysing the LipDH2B amino acid sequence using sub-cellular localisation prediction programs. The results of these programs suggested that LipDH2 was predicted to be localised to the mitochondrion with the level of confidence depending on the start site used (table 4.1). Analysis of LipDH2A led to poor confidence levels (2 %) and no cleavage site predictions, while the predictions for LipDH2B produced a mitochondrial prediction with 61% confidence and a possible cleavage site after amino acid 27 (highlighted yellow in figure 4.4).

A)

```
Pberg          -------ELKYLSHFIFFHLYSYDVIVIGGGPGGYVCSIRCGQNKLKVLNVNDDNKLGGT 53
Pyoel          --------NIFIIFFFFFYLYSYDVIVIGGGPGGYVCSIRCGQNKLKVLNVNDDKKLGGT 52
Pchab          ---------KYLSHFIFFSFCSYDVIVIGGGPGGYVCSIRCGQNKLKVLNVNEDKKLGGT 51
Pfal           ------------MYIFFPFFCSYDVIVIGGGPGGYVCSIRCAQNKLNVLNVNEDKKLGGT 48
Pvivax         WATLPLYSLPNSPLPLLLPPHSYDVIVIGGGPGGYVCSIRCAQNKLNVLNVNEDKKLGGT 60
                              ::     ********************.****:*****:*:*****
```

B)

```
ATGAACAGCGTTATTTTTAGAGCACATTGTTTTTTTCAGCCATTGCGTAGATGCTTCTCA
  E  Q  R  Y  F  *  S  T  L  F  F  S  A  I  A  *  M  L  L  N +3
  *  T  A  L  F  L  E  H  I  V  F  F  S  H  C  V  D  A  S  Q +2
M  N  S  V  I  F  R  A  H  C  F  F  Q  P  L  R  R  C  F  S  +1

ACTAAGAAAGAGTATACAAATAGAAAATTAAAAAATAAGCATAATTTACTTGTACATAGA
  *  E  R  V  Y  K  *  K  I  K  K  *  A  *  F  T  C  T  *  K +3
  L  R  K  S  I  Q  I  E  N  *  K  I  S  I  I  Y  L  Y  I  E +2
T  K  K  E  Y  T  N  R  K  L  K  N  K  H  N  L  L  V  H  R  +1

AAAAAAAAAATATACATATATATAAATATATATATATCTATTTATATATATATATATT
  K  K  N  I  H  I  Y  K  Y  I  Y  I  S  I  Y  I  Y  I  Y  L +3
  K  K  K  Y  T  Y  I  *  I  Y  I  Y  I  Y  L  Y  I  Y  I  F +2
K  K  K  I  Y  I  Y  I  N  I  Y  I  Y  L  F  I  Y  I  Y  I  +1

TATATATATGTATATTTTTTTTTTTCTATTTTGTCCCACATTTACATTATCACACGCACA
  Y  I  C  I  F  F  F  F  Y  F  V  P  H  L  H  Y  H  T  H  I +3
  I  Y  M  Y  I  F  F  F  L  F  C  P  T  F  T  L  S  H  A  H +2
Y  I  Y  V  Y  F  F  F  S  I  L  S  H  I  Y  I  I  T  R  T  +1

TATATATATATATATATATATAATATATGTATATATTTTTTTCCTTTTTTTTTGTAGCTATG
  Y  I  Y  I  Y  I  *  Y  H  Y  I  F  F  P  F  F  C  S  [  ]  +3
  I  Y  I  Y  I  Y  I  I  Y  V  Y  I  F  S  F  F  L  *  L  *  +2
Y  I  Y  I  Y  I  Y  N  I  C  I  Y  F  F  L  F  F  V  A  M  +1

ATGTTATAGTCATTGGAGGAGGGCCAGGTGGTTACGTGTGCAGTATTCGATGTGCTC
  [                                                       ]  +3
  C  Y  S  H  W  R  R  A  R  W  L  R  V  Q  Y  S  M  C  S  +2
M  L  *  S  L  E  E  G  Q  V  V  T  C  A  V  F  D  V  L  +1
```

C)

```
Pberg          MNGFFSKTNKVFFYPLRRNFSTNKDYDVIVIGGGPGGYVCSIRCGQNKLKVLNVNDDNKL 60
Pyoel          MNSFFSKTNKVFFYPLRRNFSTNKDYDVIVIGGGPGGYVCSIRCGQNKLKVLNVNDDKKL 60
Pchab          MKGIFSKTNKIFFYPVRRNFSTNKDYDVIVIGGGPGGYVCSIRCGQNKLKVLNVNEDKKL 60
Pfal           MNSVIFRA-HCFFQPLRRCFSTKKDYDVIVIGGGPGGYVCSIRCAQNKLNVLNVNEDKKL 59
Pvivax         MSAVVSRT-RVAFFPGRRHFSSKKDYDVIVIGGGPGGYVCSIRCAQNKLNVLNVNEDKKL 59
               *.... :: :  * * ** **::**********************.****:*****:*:**

Pberg          GGTCLNRGCIPSKALLHIAHNYYESKNKFKECGILIDNVKLDIEQVHKHKNKCMGSLADG 120
Pyoel          GGTCLNRGCIPSKSLLHIAHNYYESKNKFKECGILIDNVKLDIEQVHKHKNKCMGNLADG 120
Pchab          GGTCLNRGCIPSKSLLHIAHNYYESKNKFKECGILVDNVKLDIEQLHKHKNKCMGNLADG 120
Pfal           GGTCLNRGCIPSKSLLHISHNYYEAKTRFKECGILVDNVKLDIETMHKHKNKCMGNLSDG 119
Pvivax         GGTCLNRGCIPSKALLHIAHKYYEAKNKFKQSGIIIDSVHLDVEAIHKQKNKCMGNLADG 119
               *************:****:*:***:*.:**:.**::*.*:**:* :**:******.*:**
```

```
Pberg           ISFLYKKNNVKHIIGRGSIIDSNTIIVKTENEGQKKYTAERIVIATGSKPIEIPLKKLND 180
Pyoel           ISFLYKKNNVKHIIGRGSIIDGNTILVETENEGQKKYTAERIVIATGSKPIEIPLKKLDD 180
Pchab           ISFLYKKNNVKHIIGRGSIVDGNTILVETENEGQKKYTAERIVIATGSKPIEIPLKKLDD 180
Pfal            INFLYKKNNVNHIIGHGSLVDEHTVLIKTEKE-EKKVTAERIVIATGSKPIEIPLKKLND 178
Pvivax          ITYLYKTNKVNHLVGHGSIVDGNTVLVNSEGK-EKLVTAERIVIATGSKPIEIPLKKLNE 178
                *.:***.*:*:*::*:**::* :*:*:::* : :*  **********************::

Pberg           DNINDVETVKDILEYDHKLIQTSDDILNFKEIPKTMSIIGGGVIGLEIGSVFSKFGSDVT 240
Pyoel           NNINDVENVKDILEYDHKLIQTSDDILNFKEIPKTMSIIGGGVIGLEIGSVFSKFGSDVT 240
Pchab           NNINDVENVKDILEYDHKLIQTSDDILNFKEIPKTMSIIGGGVIGLEIGSVFSKFGSDVT 240
Pfal            NNFNDADNVNDILEYDHEIIQNSDDILNFKKVPHNISIIGGGVIGLEIGSVFSKLGSDVT 238
Pvivax          DIAKEAETVEELLQYDHQVIQTSDDILNFKKIPKKISIIGGGVIGLEIGSVFARMGSDVT 238
                :. ::.:.*:::*:***:::*.********::*:.:*****************:*:*****

Pberg           VYEYNSRLCGFLDPDVSKVLQKVLEKVKIKFMFNTSIVGGNLNTTNNEAILYARDNKTNK 300
Pyoel           VYEYNSRLCGFLDPDVSKVLQKVLEKIKIKFMFNTSIIGGNLNTTNNEAILYARDNKTNQ 300
Pchab           VYEYNSRLCGFLDPDVSKVLQKVLEKIKMKFMFNTSIIGGNLNTSNNEAILYARDNKTNQ 300
Pfal            VFEYNERLCGFLDADVSKVLQKTLEKIKMKTVFNTSVIGG--NIENNQAALFAKNKKTNE 296
Pvivax          IFEYNSRLCTFLDADVSKVLQKTLEKVKIKFAFNSSVVGG--SVQNGEATLFAKNAKTKE 296
                ::***.*** ***.********.***:*:** **:*::**  .  *.:* *:*:: **::

Pberg           IKKVKSDIVLVCVGRKANLENINLEKLNIELNKNKKIQVDEYFNVKSQPTIKAIGDAIDG 360
Pyoel           IKKVKSDIVLVCVGRKANLENINLEKLSIELNKNKKIQVDEYFNVKSQPTIKAIGDAIDG 360
Pchab           INKIKSDIVLVCVGRKANLENINLENLSIELNKNKKIQVDEYFNVQSQPTIKAIGDAIDG 360
Pfal            IKKTTSEIVLICIGRKANFDNLNLHLLNIELNKNKKIPVDEYFNVIAQPTIKAIGDAIDG 356
Pvivax          IQKVTSDVVLVCVGRRPNFDDLNLEKLNVQLGQNKRLQVDASFGVASHPTIKAIGDAIDG 356
                *:* .*::**:*:**:.*::::**. *.::*.:**:: **  *.* ::************

Pberg           SMLAHKAEEEGYIVADMIFNELKNNKKKKNHINYDLIPSVIYTHPEVASVGYNEQKCKEL 420
Pyoel           SMLAHKAEEEGYIVADMLFNELKNNKKKKNHINYDLIPSVIYTHPEVASVGYNEQKCKEL 420
Pchab           AMLAHKAEEEGYIVADMIFNELKNNNKKKNHINYDLIPSVIYTHPEVASVGYTEQKCKEL 420
Pfal            PMLAHKAEEEGYLLANILFDELKNNKKKKAHINYDLVPSVIYTHPEVATVGYNEAKCKEL 416
Pvivax          PMLAHKAEEEGYILANILLSELKLKKPKKSHINYDLVPSVIYTHPEVASVGLNEERCKQR 416
                .***********::*:::::.*** :: ** *:****:**********:** .* :**:

Pberg           KLNYKTVSFPFAANSRSRTIDDYDGLIKLIVEKDTNVILGSQIIGNNASDLILPLSIYAS 480
Pyoel           KLNYKSVSFPFAANSRSRTIDDYDGLIKLIVEKDTNVILGSQIIGNNASELILPLSIYAS 480
Pchab           KLNYKAVSFPFAANSRSRTIDDYDGLIKLIVEKDTNVVLGSQIIGNNASELILPLSIYAS 480
Pfal            NMNFKSVSFPFAANSRSRTIDDYDGLIKLIVEKDTNRILGSQIIGNNASDLILPLSIYVA 476
Pvivax          NLPYKAVTFPFAANSRSRTIDDFDGVIKLLVESNSNKILGSQIVGNNASDLILPLSIYVA 476
                :: :*:*:***************:**:***:**.::* :*****:*****:********.:

Pberg           HKGTSKNLSKIIYPHPTFSEVIKEVALQSFDKAIHM 516
Pyoel           HKGTSKNLSKIIYPHPTFSEVIKEVALQSFDKAIHM 516
Pchab           HKGTSKNLSKIIYPHPTFSEVIKEVALQSFDKAIHM 516
Pfal            NNGSSKSLSKIIYAHPTFSEVIKEVALQSFDKPIEM 512
Pvivax          NGGSSKSLSKVIYAHPTFSEVIKEVALQSFDRAIEM 512
                : *:**.***:**.*****************:.***
```

Figure 4.3 – ClustalW alignment of LipDH2 homologues from *Plasmodium* species.

The *P. falciparum* LipDH2 deduced amino acid sequence was compared by ClustalW alignment with homologues from *Plasmodium berghei* (Pberg), *Plasmodium yoelii yoelii* (Pyoel), *Plasmodium chabaudi* (Pchab) and *Plasmodium vivax* (Pvivax). (*) Identical residues, (:) conserved residues, (.) homologous residues. **Panel A** displays an N-terminal section of the ClustalW alignment of the open reading frames (ORF) of exon 2 of LipDH2A homologues from *Plasmodium* species genomes. This displays that the in-frame methionine present in the *P. falciparum* exon 2 ORF is not conserved among the different *Plasmodium* LipDH2 sequences which suggests the presence of a possible intron. The most likely site for the position of the intron/exon boundary is highlighted in grey. **Panel B** displays the *LipDH2* genomic DNA sequence with the deduced amino acid sequence for each of the three frames. Exon 1 starts in frame +1 and is highlighted in yellow; exon 2 starts in frame +3 and is highlighted in dark blue. The predicted intron/exon boundaries are highlighted in red and were found to be conserved between *Plasmodium* species. The translation initiation methionine for LipDH2A (green) and LipDH2B (yellow, at the start of exon 1) are displayed. **Panel C** displays a ClustalW alignment of the predicted amino acid sequences of LipDH2B and homologues from other *Plasmodium* species following splicing of the potential intron. The position of the intron/exon boundary is highlighted in grey, and displays that splicing of the intron produces a change in residue from serine to aspartate (compare panel A with panel C).

```
ATGAACAGCGTTATTTTTAGAGCACATTGTTTTTTTCAGCCATTGCGTAGATGCTTCTCA 60
M  N  S  V  I  F  R  A  H  C  F  F  Q  P  L  R  R  C  F  S   20

ACTAAGAAAGACTATGATGTTATAGTCATTGGAGGAGGGCCAGGTGGTTACGTGTGCAGT 120
T  K  K  D  [ ]  D  V  I  V  I  G  G  G  P  G  G  Y  V  C  S   40

ATTCGATGTGCTCAAAATAAACTAAACGTTTTGAACGTGAACGAAGACAAGAAACTAGGA 180
I  R  C  A  Q  N  K  L  N  V  L  N  V  N  E  D  K  K  L  G   60

GGCACATGTTTAAATAGAGGATGCATTCCATCCAAATCTCTATTACACATTTCTCATAAT 240
G  T  C  L  N  R  G  C  I  P  S  K  S  L  L  H  I  S  H  N   80

TATTATGAAGCAAAGACTAGATTCAAAGAATGTGGTATATTAGTTGATAATGTTAAGTTG 300
Y  Y  E  A  K  T  R  F  K  E  C  G  I  L  V  D  N  V  K  L   100

GATATAGAGACTATGCATAAACATAAAAATAAGTGCATGGGTAATTTATCTGATGGAATT 360
D  I  E  T  M  H  K  H  K  N  K  C  M  G  N  L  S  D  [ ]  I   120

AATTTTTTATACAAAAAGAACAATGTGAATCATATTATAGGCCATGGGAGTTTAGTAGAT 420
N  F  L  Y  K  K  N  N  V  N  H  I  I  G  H  G  S  L  V  D   140

GAACATACTGTTTTAATAAAAACCGAAAAAGAAGAAAAGAAGGTAACAGCTGAACGTATT 480
E  H  [ ]  V  L  I  K  T  E  K  E  E  K  K  V  T  A  E  R  I   160

GTTATAGCTACTGGATCCAAACCGATTGAAATACCTTTAAAAAAATTAAATGATAATAAT 540
V  I  A  T  G  S  K  P  I  E  I  P  L  K  K  L  N  D  N  N   180

TTTAATGATGCTGATAATGTGAACGATATATTAGAATATGACCATGAAATAATACAAAAT 600
F  N  D  A  D  N  V  N  D  I  L  E  Y  D  H  E  I  I  Q  N   200

TCAGATGATATTTTAAATTTTAAAAAGGTTCCTCATAATATATCTATTATTGGAGGTGGT 660
S  D  D  I  L  N  F  K  K  V  P  H  N  I  S  I  I  G  G  G   220

GTTATAGGGTTAGAAATTGGATCTGTGTTTTCAAAACTTGGTTCTGATGTTACTGTATTT 720
V  I  G  L  E  I  G  S  V  F  S  K  L  G  S  D  V  T  V  F   240

GAATATAATGAAAGGTTATGTGGTTTTCTTCATGCTGATGTAAGTAAAGTTTTACAAAAA 780
E  Y  N  E  R  L  C  G  F  L  D  A  D  V  S  K  V  L  Q  K   260

ACATTAGAAAAAATAAAAATGAAATTTGTATTTAATACTTCAGTAATAGGTGGAAATATA 840
T  L  E  K  I  K  M  K  F  V  F  N  T  S  V  I  G  G  N  I   280

GAAAATAACCAAGCTGCTTTATTTGCTAAAAATAAAAAAACTAATGAAATAAAAAAAACA 900
E  N  N  Q  A  A  L  F  A  K  N  K  K  T  N  E  I  K  K  T   300

ACATCAGAAATTGTACTTATATGTATTGGTAGGAAAGCTAACTTTGATAATTTAAATTTA 960
T  S  E  I  V  L  I  C  I  G  R  K  A  N  F  D  N  L  N  L   320

CATTTACTTAATATAGAATTAAATAAAAATAAAAAAATACCTGTAGATGAATATTTTAAT 1020
H  L  L  N  I  E  L  N  K  N  K  K  I  P  V  D  E  Y  F  N   340

GTAATAGCACAACCAACTATTAAAGCAATTGGCGATGCTATAGATGGACCTATGTTGGCC 1080
V  I  A  Q  P  T  I  K  A  I  G  D  A  I  D  G  P  M  L  A   360

CATAAAGCAGAAGAAGAAGGATATTTATTAGCTAACATCTTATTTGATGAATTAAAAAAT 1140
H  K  A  E  E  E  G  Y  L  L  A  N  I  L  F  D  E  L  K  N   380

AATAAGAAAAAAAAAGCACATATAAATTATGACCTAGTTCCAAGTGTTATATATACACAT 1200
N  K  K  K  K  A  H  I  N  Y  D  L  V  P  S  V  I  Y  T  H   400

CCTGAGGTAGCTACTGTAGGATATAATGAAGCAAAATGTAAAGAACTCAATATGAATTTC 1260
P  E  V  A  T  V  G  Y  N  E  A  K  C  K  E  L  N  M  N  F   420

AAATCAGTTAGTTTTCCCTTTGCTGCTAACAGTAGATCAAGGACAATTGATGATTATGAT 1320
K  S  V  S  F  P  F  A  A  N  S  R  S  R  T  I  D  D  Y  D   440
```

```
GGACTCATAAAATTAATTGTAGAAAAAGATACTAACAGAATTTTAGGTTCTCAAATTATA 1380
G  L  I  K  L  I  V  E  K  D  T  N  R  I  L  G  S  Q  I  I   460

GGAAATAATGCTAGTGATTTAATATTACCTTTGTCAATATATGTCGCTAATAATGGCTCA 1440
G  N  N  A  S  D  L  I  L  P  L  S  I  Y  V  A  N  N  G  S   480

TCTAAAAGTCTGAGTAAAATTATATATGCGCATCCAACATTTTCAGAAGTAATCAAAGAA 1500
S  K  S  L  S  K  I  I  Y  A  H  P  T  F  S  E  V  I  K  E   500

GTGGCTTTACAATCTTTTGATAAACCTATACACATGTAA                      1539
V  A  L  Q  S  F  D  K  P  I  H  M  *                        513
```

Figure 4.4 – LipDH2B nucleotide and amino acid sequence

The proposed cDNA sequence of *LipDH2B* following splicing of the potential intron is shown, with the deduced amino acid sequence shown underneath. Amino acid numbering includes the initiator methionine; the termination codon is marked with an asterisk. The amino acid residues highlighted are the predicted mitochondrial cleavage site (highlighted yellow), the start site of the recombinantly expressed protein (highlighted red) and the GFP-fusion sites (highlighted green) for LipDH2A (Thr143) and LipDH2B (Gly119).

Gene	Chromosome	Prediction	Score[a]	Cleavage site
LipDH1	8	Apicoplast	+[b]	33[d]
LipDH2A	12	Mitochondrion	1.95 %[c]	n.d.
LipDH2B	12	Mitochondrion	60.7%[c]	27[c]

Table 4.1 - Subcellular localisation predictions

Predictions for the presence of potential mitochondrial targeting sequences or the presence of a potential signal peptide and apicoplast transit peptide were performed using PlasMit, MitoProt, TargetP, SignalP and PlasmoAP (Claros *et al.*, 1996; Nielsen *et al.*, 1997; Emanuelsson O, 2000; Bender *et al.*, 2003; Foth *et al.*, 2003) Bender *et al.*, 2003; Foth *et al.*, 2003)[a]. Prediction scores refer to PlasmoAP[b] and MitoProt[c] predictions. Cleavage site predictions were provided by Mitoprot[c] and SignalP[d].

4.2.2 Identification of conserved motifs

The identified amino acid sequences were further analysed by ClustalW alignment with LipDH proteins from other organisms to analyse the conserved motifs. Alignment of both *P. falciparum* LipDH amino acid sequences identified several regions of conserved motifs and residues, despite the low level of overall identity (20%) between the two full-length proteins. This low level of identity may be partially due to the presence of a long N-terminal extension (the predicted bipartite plastid targeting sequence) in the LipDH1 amino acid sequence, and the amino acid identity was increased to 24 % when the N-terminal sequence was omitted from the alignment.

The conserved motifs displayed in the alignment are characteristic of this class of disulphide oxidoreductases and include motifs for the binding of the FAD-cofactor and NADH, as well as residues important in catalysis such as active site cysteines and acid base catalysts (figure 4.5). The LipDH amino acid sequence can be separated into four domains in analogy with those identified in glutathione reductase (Thieme *et al.*, 1981; Argyrou *et al.*, 2004); the FAD-domain, the NADH-domain, the central domain and the interface domain (Toyoda *et al.*, 1998).

The FAD-domain is located at positions 20 – 164 and 110 – 291 in the LipDH2 and LipDH1 respectively. The FAD-domain contains a conserved motif (GXGXXG) for binding of the pyrophosphate moiety of FAD in part of a Rossman-like fold (Schierbeek *et al.*, 1989; Mattevi *et al.*, 1991), and is located at positions 132-137 in the LipDH1 and 31-36 in the LipDH2 (highlighted blue in figure 4.5). The FAD-domain also contains the redox active cysteine motif (Thorpe *et al.*, 1976) that is present in a CV/LXGC motif that is located at positions 165-170 in the LipDH1 and 63 – 68 in the LipDH2 (highlighted yellow in figure 4.5).

The NADH-domain is located at positions 111 – 435 and 165 – 295 in the LipDH1 and LipDH2, respectively and contains a conserved dinucleotide-binding motif (GXGXIGXEXXXV/IXXXXG). This NADH-binding motif (Mattevi *et al.*, 1992; Bocanegra *et al.*, 1993; Brautigam *et al.*, 2005) is located at positions 328 – 344 in LipDH1 and 219 – 235 in LipDH2 (highlighted green in figure 4.5).

The central domain is located at positions 436 – 517 and 296 – 391 in LipDH1 and LipDH2, respectively. This domain also contains further residues that are important for binding of the flavin co-factor.

The interface domain is located at positions 518 – 667 and 392 – 513 in the LipDH1 and LipDH2, respectively and is the site of dimerisation. The interface also contains the acid/base catalyst motif (HPTXXE) which is an essential component of the active site (Williams, 1976) and is located at positions 646 – 651 in LipDH1 and 491 – 496 in the LipDH2 (highlighted red in figure 4.5).

```
Human          ------------------------------------------------------------
Pea            ------------------------------------------------------------
Yeast          ------------------------------------------------------------
LipDH2         ------------------------------------------------------------
LipDH1         MVIRQNIKHIVKLNVVTLIWLSYLFLLKPHGTLKNMMVCNAVLLPFNEKNKGINNFVYIN 60
Rice           ----------------------------MYSTAISLSAAATAAAAVGGARPGA-----IR 27

Human          ---MQSWSRVYCSLAKRG--------------------HFNRISHG-LQGLSAVPLRTYA 36
Pea            --------MAMANLARRK--------------------GYSLLSSETLRYSFSLRSRAFA 32
Yeast          ------------MLRIR----------------------SLLNNK---RAFSSTVRTLT 22
LipDH2         ----------MNSVIFR----------------------------AHCFFQPLRRCF 19
LipDH1         PKNIILNKIKKDVIKLEKDNIILCQHNRKRDNYIKQQKRKEKNANNFTFMLKGSTQNIMN 120
Rice           PAELRFCGLRREALGLR------------------SLRAPPRAAATPRKAAAATGNGAAG 69
                                 :  .

Human          DQPIDADVTVI[SGPA]YVAAIKAAQLGFKTVCIEKN-ETLGGTCLNVGCIPSKALLNNS 95
Pea            SGSDENDVVII[GGPG]YVAAIKAAQLGFKTTCIEKR-GALGGTCLNVGCIPSKALLHSS 91
Yeast          IN-KSHDVVII[GGPA]YVAAIKAAQLGFNTACVEKR-GKLGGTCLNVGCIPSKALLNNS 80
LipDH2         STKKDYDVIVI[GGPG]YVCSIRCAQNKLNVLNVNED-KKLGGTCLNRGCIPSKSLLHIS 78
LipDH1         INEKEYDLAII[GGVG]HAAAINAMERNLKVIIFAGDENCIGGTCVNVGCIPSKALLYAT 180
Rice           SGGFDYDLVII[GAGG]HGAALHAVEKGLKTAIIEG--DVVGGTCVNRGCVPSKALLAVS 127
                   . *: :** * .*: .::.. :  ::.  .      :****:* **:***:**  :

Human          HYYHMAHGTDFASRG------------------IEMS-EVRLNLDKMMEQKSTAVKALTG 136
Pea            HMYHEAK-HSFANHG------------------VKVS-NVEIDLAAMMGQKDKAVSNLTR 131
Yeast          HLFHQMH-TEAQKRG------------------IDVNGDIKINVANFQKAKDDAVKQLTG 121
LipDH2         HNYYEAK-TRFKECG------------------ILVD-NVKLDIETMHKHKNKCMGNLSD 118
LipDH1         NKYRELKNLDKLYYYGIHSNIFQNNKNTEIENNQLVSNSFQINITKLKEYTQSVIDKLRN 240
Rice           GRMRELQDEHHMKSLGLQVSTAGYDR-------QAVADHANNLASKIRSNLTNSMKALGV 180
                     :                          :    .    :       :  *

Human          GIAHLFKQNKVVHVNGYGKITGKNQVTATKADG------GTQVIDTKNILIATGSEVTPF 190
Pea            GIEGLFKKNKVTYVKGYGKFVSPSEISVDTIEG------ENTVVKGKHIIIATGSDVKSL 185
Yeast          GIELLFKKNKVTYYKGNGSFEDETKIRVTPVDGLEGTVKEDHILDVKNIIVATGSEVTPF 181
LipDH2         GINFLYKKNNVNHIIGHGSLVDEHTVLIKTEKE-------EKKVTAERIVIATGSKPIEI 171
LipDH1         GISHGFKTLKFNKNSEHVQVIYEHGQLLDKNTI--KSKKSGNTYKVKNIIIATGSVPNIP 298
Rice           DILTGFGTIVGKQKVRYGKVGFPD----------------NEITARNIIIATGSVPFVP 223
               .*   :              ..                        ..*::****

Human          PG-----------------ITIDEDTIVSSTGALSLKKVPEKMVVI[GAGVIGVELGSVWQ 233
Pea            PG-----------------VTIDEKKIVSSTGALALSEIPKKLVVI[GAGYIGLEMGSVWG 228
Yeast          PG-----------------IEIDEEKIVSSTGALSLKEIPKRLTII[GGGIIGLEMGSVYS 224
LipDH2         PLKKLNDNNFNDADNVNDILEYDHEIIQNSDDILNFKKVPHNISII[GGGVIGLEIGSVPS 231
LipDH1         NN-----------------VEIDDKSVFTSDMAVKLVGLKNYMSII[GMGIIGLEFADIYT 341
Rice           NG-----------------IEIDGKTVFTSDHALKLESVPDWIAIV[GSGYIGLEFSDVYT 266
                 .. .... ..:.. .. :  * . : .*   : :  : . : ::* * **:*:..::

Human          RLGADVTAVEFLGHVGGVGIDMEISKNFQRILQK-QGFKFKLNTKVTGATKKSDGK-IDV 291
Pea            RIGSEVTVVEFASEIVPT-MDAEIRKQFQRSLEK-QGMKFKLKTKVVGVDTSGDG--VKL 284
Yeast          RLGSKVTVVEFQPQIGAS-MDGEVAKATQKFLKK-QGLDFKLSTKVISAKRNDDKNVVEI 282
LipDH2         KLGSDVTVFEYNERLCGF-LDADVSKVLQKTLEK-IKMKFVFNTSVIGGNIENNQ--AAL 287
LipDH1         ALGSEITFLEYSSELLPI-IDNDVAKYFERVFLKNKPVNYHLNTEVKYIKASKNNNPVII 400
Rice           ALGSEVTFVEALDQLMPG-FDPEIAKLAQRILINPRKIDYHTGVFASKITPAKDGKPVLI 325
                :*:.:* .*    .:    :* :: *  :: : :   ..:   . .       :     :

Human          SIEAASGGK---------AEVITCDVLLVCIGRRPFTKNLGLEELGIELDPRGRIPVNTR 342
Pea            TVEPSAGGE---------QTIIEADVVLVSAGRTPFTSGLNLDKIGVETDKLGRILVNER 335
Yeast          VVEDTKTNK---------QENLEAEVLLVAVGRRPYIAGLGAEKIGLEVDKRGRLVIDDQ 333
LipDH2         FAKNKKTNE---------IKKTTSEIVLICIGRKANFDNLNLHLLNIELNKNKKIPVDEY 338
LipDH1         GYSHRTGNDDNEKKNMTDVKELYVDSCLVATGRNPNTQNLGLEKLKIQMNR-GYVSVNDN 459
Rice           ELIDAKTKEHKET--------LEVDAALIATGRAPFTKGLGLENVNVVTQR-GFVPVDER 376
                        .               :  *:. ** .   .*. . : :  :    : ::

Human          FQTKIP------NIYAIGDVVAGPMLAHKAEDEGIICVEGMAGGAV-------HIDYN-- 387
Pea            FSTNVS------GVYAIGDVIPGPMLAHKAEEDGVACVEYLAGKVG-------HVDYD-- 380
Yeast          FNSKFP------HIKVVGDVTFGPMLAHKAEEEGIAAVEMLKTGHG-------HVNYN-- 378
LipDH2         FNVIAQP-----TIKAIGDAIDGPMLAHKAEEEGYLLANILFDELKNNKKKKAHINYD-- 391
LipDH1         LQVKMENNEIYDNIFCIGDANGKQMLAHTASYQALKVIDFIEKKEKKNVNINVENNLSKP 519
Rice           MQVMDADGNAVPNLYCIGDANGKLMLAHAASAQGISVVERISGKD------NILNHLS-- 428
               :.           :  :**.    **** *. :.    : :     .. . .   . .
```

```
Human           ----CVPSVIYTHPEVAWVGKSEEQLKEEGIEYKVGKFPFAAN------------SRAK 430
Pea             ----KVPGVVYTNPEVASVGKTEEQVKETGVEYRVGKFPFMAN------------SRAK 423
Yeast           ----NIPSVMYSHPEVAWVGKTEEQLKEAGIDYKIGKFPFAAN------------SRAK 421
LipDH2          ----LVPSVIYTHPEVATVGYNEAKCKELNMNFKSVSFPFAAN------------SRSR 434
LipDH1          ILYKNIPSVCYTNPELAFIGLTEKEAKVLYPDNVGVEISYYKSNSKILCENNISLNNNKK 579
Rice            -----IPAACFTHPEISMVGLTEPQAREK-ADKEGFEISVVKT----------SFKANTK 472
                     :*.. :::**:: :* .* : :    :    .:.   .             . :

Human           TNADTDG----------MVKILGQKSTDRVLGAHILGPGAGEMVNEAALALEYGASCEDI 480
Pea             AIDNAEG----------LVKIIAEKETDKILGVHIMAPNAGELIHEAAIALQYDASSEDI 473
Yeast           TNQDTEG----------FVKILIDSKTERILGAHIIGPNAGEMIAEAGLALEYGASAEDV 471
LipDH2          TIDDYDG----------LIKLIVEKDTNRILGSQIIGNNASDLILPLSIYVANNGSSKSL 484
LipDH1          NNSYNKGQYNINDNTNGMVKIIYKEDTKEILGMFIVGNYASVLIHEAVLAINLKLSAFDL 639
Rice            ALAENEG--------DGLAKMIYRPDTGEILGVHILGLHAADLIHEASNAIALGTRVQDI 524
                     .*          : *::   .* .:**  *:.  *. ::      :       .:

Human           ARVCHAHPTLSEAFREANLAASFGKSINF-------- 509
Pea             ARVCHAHPTMSEAIKEAAMAT-YDKPIHI-------- 501
Yeast           ARVCHAHPTLSEAFKEANMAA-YDKAIHC-------- 499
LipDH2          SKIIYAHPTFSEVIKEVALQS-FDKPIHM-------- 512
LipDH1          AYMVHSHPTVSEVLDTAFKSISKIRTH---------- 666
Rice            KFAVHAHPTLSEVLDELFKAAKVNAGVPHSVNEPVAA 561
                    ::***.**.:
```

Figure 4.5 - ClustalW alignment of LipDH proteins

ClustalW alignment of the deduced of amino acid sequences from *P. falciparum* LipDH1 & LipDH2B, *Homo sapiens* (Human; P09622), *Saccharomyces cerevisiae* (Yeast; P09624), *Pisum sativum* mitochondrion (Pea; P41023) and Oryza sativa chloroplast (Rice; NM_183915) are displayed. The conserved motifs for FAD binding (blue), NADH-binding (green), active site cysteines (yellow) and acid base catalyst (red) are highlighted. (*) Identical residues, (:) conserved residues, (.) homologous residues.

4.2.3 Both LipDH are expressed in erythrocytic stages

In order to test if both of the LipDH genes/proteins were expressed in the blood stages, I investigated their transcription and translation by reverse transcriptase PCR (RT-PCR) and western blot analyses.

RT-PCR analyses were carried out on total RNA isolated from erythrocytic stage parasites and confirmed the transcription of both *LipDH* genes in these stages of the life cycle (figure 4.6) Negative controls that tested for the presence of genomic DNA in total RNA preparations were observed to be blank, suggesting that the bands amplified were exclusively from the isolated RNA. The presence of the predicted intron in *LipDH2* was experimentally verified by DNA sequencing of the *LipDH2* RT-PCR product, suggesting that LipDH2B was the protein expressed by the parasites (and not LipDH2A).

Protein expression analysis was carried out on parasite protein extracts by Western blotting using polyclonal antibodies raised against recombinantly expressed LipDH2 (See section 4.2.6). This analysis identified a single band of an estimated size of 55 kDa, which was in good agreement with the predicted mature LipDH2 size of 54 kDa (table 4.1). The protein expression of LipDH1 was not analysed by western blotting, as polyclonal antibodies have not yet been raised against the recombinant LipDH1 protein. The combined data from the RT-PCR and western blot analyses confirmed that both LipDH were expressed during in the erythrocytic stages.

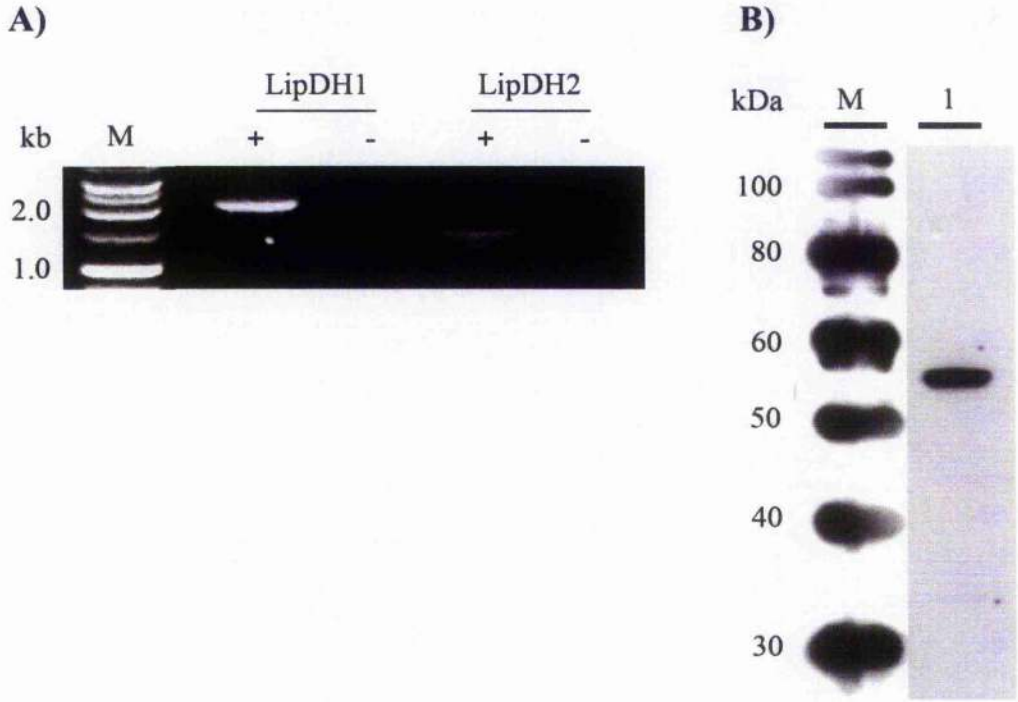


Figure 4.6 – Expression of LipDH in erythrocytic stages

The transcription of both *LipDH* genes in the erythrocytic stages of *P. falciparum* was investigated by RT-PCR (Panel A) and western blotting (Panel B). **Panel A** displays the results of RT-PCR on total RNA isolated from parasites, with PCR performed in the presence (+) or absence (-) of reverse transcriptase. Both LipDH transcripts were amplified from the RNA and resulted in products of 2 kb and 1.6 kb for *LipDH1* and *LipDH2*, respectively. The products were verified by DNA sequencing and confirmed the presence of the predicted intron in *LipDH2*. Samples amplified in the absence of reverse transcriptase (-) were used as negative controls, in order to test for genomic DNA contamination in the RNA preparations. **Panel B** displays the result of a western blot with 10μg of parasite protein extract using polyclonal antibodies raised against recombinantly expressed LipDH2. The antibody reacted with a protein of 56 kDa, which is in good agreement with the predicted size LipDH2 after the cleavage of the potential mitochondrial targeting sequence.

4.2.4 Confirming localisation with GFP-fusion proteins

The homology of the two LipDH proteins to different isoforms (plastid or mitochondrial) combined with the predicted subcellular localisations suggested that the LipDH proteins might be targeted to different organelles in the parasites. LipDH1 displayed highest amino acid identity with plastid isoforms of LipDH (see section 4.2.1.1) and was predicted to be localised to the apicoplast (table 4.1), while LipDH2 displayed highest amino acid sequence identity with mitochondrial isoforms (see section 4.2.1.2) and was predicted to be localised to the mitochondrion (table 4.1).

These predicted localisations were experimentally tested by fusing their respective N-terminal targeting sequences in frame with green fluorescent protein (GFP) and expressing the fusion protein in *P. falciparum* cells. Four constructs were designed to analyse the subcellular localisation of both LipDH, one for LipDH1 and three for LipDH2. The reason that three constructs were necessary to investigate the subcellular localisation of LipDH2 was because of the presence of two possible translation initiation sites. I have already confirmed LipDH2B as being the correctly spliced product of the *LipDH2* gene (see section 4.2.3), but wanted to investigate the localisation of an incorrectly spliced gene product that used the in-frame methionine that is located in the confirmed intron (LipDH2A). The four constructs designed were as follows:

1. The LipDH1 fusion protein consisted of amino acids 1 – 120 in frame with the GFP ORF (highlighted green in figure 4.2).

2. LipDH2A was designed to use the in-frame methionine contained within the confirmed intron and consisted of the first 129 amino acids from the resultant protein (highlighted green in figure 4.4).

3. LipDH2B (C) used the methionine identified in exon 1, but was amplified from total RNA by RT-PCR and therefore lacked the predicted intron. This consisted of the first 119 amino acids from the LipDH2B protein (highlighted green in figure 4.4)

4. LipDH2B (G) used the methionine identified in exon 1, but was amplified from isolated genomic DNA and would contain the intron and both possible translation initiation sites.

The inserts were amplified by PCR, cloned into the TOPO TA cloning vector and finally sub-cloned into *P. falciparum* transfection vector pHH2. Cloning into pHH2 was laborious due to the large size of the destination vector (7.3 kb), therefore positive pHH2 clones were initially identified by colony PCR to allow high throughput screening of clones (figure 4.7). Typically it was necessary to screen over a hundred colonies by colony PCR to identify a positive clone. Positive clones were verified by DNA sequencing and were transfected into *P. falciparum* under drug selection (WR99210) and parasites were observed in Giemsa stained blood smears between 6 and 12 weeks following transfection.

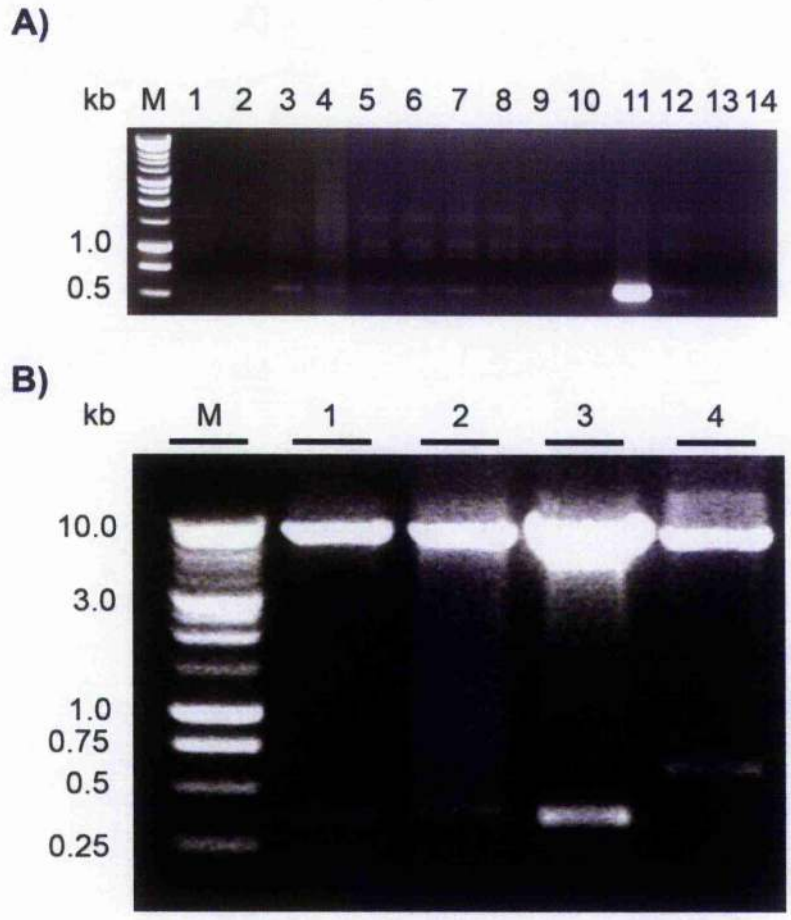


Figure 4.7 - Cloning of pHH2 transfection constructs

Panel A displays an example of colony PCR screening for pHH2 LipDH1-GFP clones. Lane M contains 1kb ladder (Promega); Lanes 1 and 2 are negative controls; Lanes 3 – 14 contain products from colonies screened by colony PCR, with colony 11 identified as a possible positive clone with the expected product size of 567 bp. **Panel B** displays an overview of pHH2-GFP fusion constructs digested with *Avr* II / *Bgl* II. Lane M contains 1 kb ladder (Promega); Lane 1 contains the digest of pHH2-LipDH1 displaying the 360 bp insert; Lane 2 contains the pHH2-LipDH2A digest and displays the 387 bp insert; Lane 3 contains the digest of pHH2-LipDH2B (C) and displays the 357 bp insert; Lane 4 contains the digest of pHH2-LipDH2B (G) and displays the insert of 581 bp. Lanes 1 – 4 all contain a 7.3 kb band that corresponds to digested pHH2.

The subcellular localisations of the GFP-fusion proteins were visualised in live parasites by fluorescent microscopy under 100x magnification. Each image contains an erythrocyte containing a single parasite, whose mitochondrion was selectively stained using Mitotracker CMXRos (Molecular probes). Images of transfected parasites were taken in phase-contrast and using the fluorescent filters for rhodamine (excitation at 570 nm and emission at 590 nm) and GFP (excitation at 488 nm and emission at 505-530 nm) or FITC (excitation at 494 nm and emission at 518 nm).

The localisation of the LipDH1-GFP fusion protein was closely associated to the mitochondrial staining, but showed no clear co-localisation (figure 4.8). Analysis of the LipDH1-GFP fusion protein fluorescence through the ring (figure 4.8A), trophozoite (figure 4.8B) and schizont (figure 4.8C) stages of the erythrocytic life cycle were consistent with those observed in previous studies using GFP fusions and immuno-fluorescence analyses with known apicoplast markers (Foth *et al.*, 2003; Tonkin *et al.*, 2004; van Dooren *et al.*, 2005).

In the ring stages of the parasite's erythrocytic cycle (0 - 15 hours following invasion) the apicoplast appears small, rounded and is closely associated with the mitochondrion (figure 4.8A). During the trophozoite stage the apicoplast elongates (15 - 28 hours post-invasion) and has few points of association with the elongated mitochondrion (figure 4.8B). In the schizont stage (28 - 48 hours post-invasion) the apicoplast begins to branch and has many points of association with the branched mitochondrion, in a process which allows each resultant daughter merozoite to contain both an apicoplast and a mitochondrion (figure 4.8C). These images are duplicated and discussed further in a more in depth investigation into organelle development in chapter 5 (figure 5.16).

The three LipDH2-GFP fusion proteins all displayed mitochondrial localisation, as determined by the observed co-localisation of the GFP-fluorescence with that of the Mitotracker (figure 4.9). The mitochondrial co-localisation of the LipDH2A-GFP fusion construct was extremely interesting, as it was not predicted to contain a mitochondrial targeting peptide (table 4.1). However, the targeting of the LipDH2A-GFP fusion protein was less efficient than that observed for either LipDH2B construct, as revealed by the lower number of parasites that exhibited GFP-expression and mitochondrial co-staining (data not shown). Both the LipDH2B constructs display co-localisation with the elongated mitochondrion of trophozoite stage parasites. Both of these constructs were commonly co-localised with the mitochondrion and the majority of cells expressed

GFP, unlike the situation in the LipDH2A-GFP fusion protein, further confirming that LipDH2B is the correct product of the *LipDH2* gene.

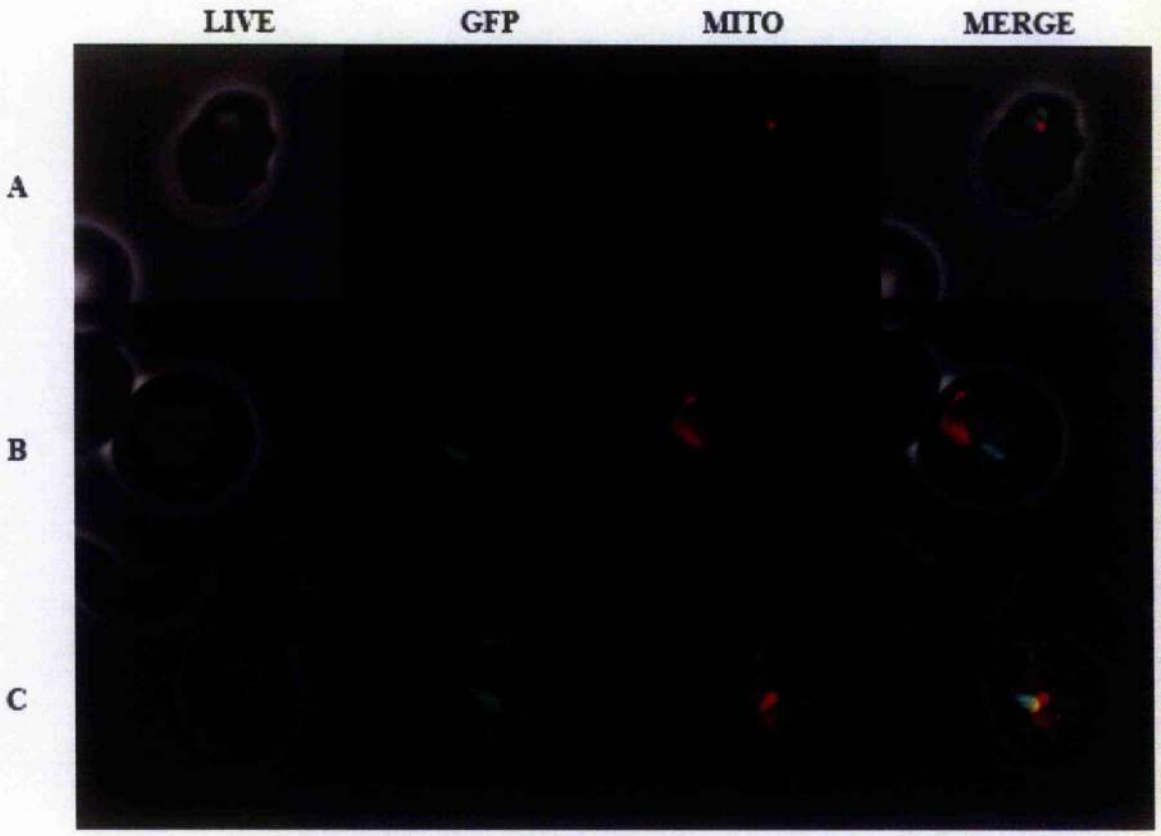


Figure 4.8 - Subcellular localisation of LipDH1

The subcellular localisation of LipDH1 was investigated by fusing the first 120 amino acids in frame with GFP and expressing the resulting protein in erythrocytic stage *P. falciparum*. The parasites mitochondrion was selectively stained with Mitotracker CMXRos (molecular probes) and images of live cells were obtained using a Zeiss Axioplan 2 microscope at 100 x magnification. The parasites were analysed using phase contrast (LIVE), GFP/FITC (GFP) and rhodamine (MITO) fluorescent filters. Overlays of the images obtained from each channel are displayed for determination of co-localisation of the two fluorescent signals (MERGE). The fluorescence pattern is displayed for ring (A), trophozoite (B), and schizont (C) stages in the erythrocytic cycle. The fluorescent signals show a close association but not clear co-localisation throughout the erythrocytic life cycle of the parasites.

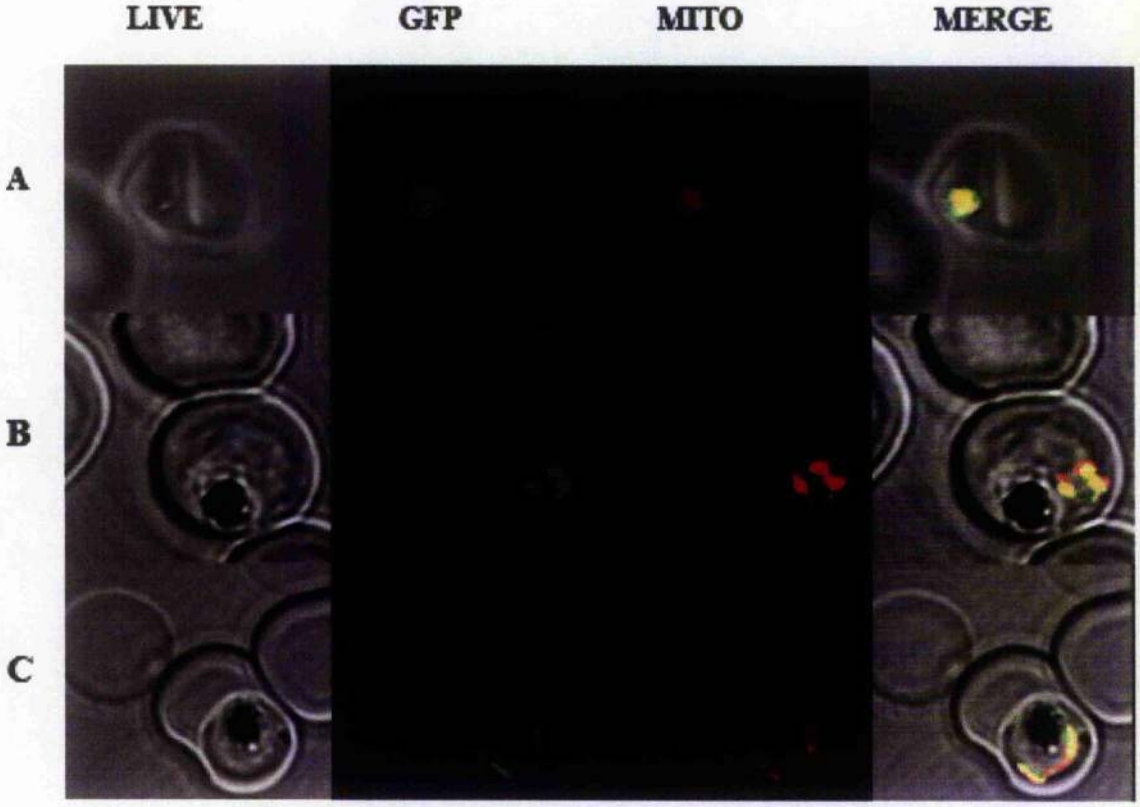


Figure 4.9 - Subcellular localisation of LipDH2

The subcellular localisation of LipDH2 was investigated using three different GFP-fusion proteins. **Panel A** displays the results obtained with the LipDH2A-GFP fusion protein. The LipDH2A-GFP fusion construct utilises the in-frame methionine within the intron (highlighted green in figure 4.3B) and contained the first 129 amino acids from LipDH2A (analogous to Thr143 in LipDH2B; highlighted green in figure 4.4). **Panel B** displays the results obtained with the LipDH2B (C)-GFP fusion protein. The LipDH2B (C)-GFP fusion construct utilises the in-frame methionine identified at the start of exon 1 in the *LipDH2B* gene (highlighted light blue in figure 4.3B) and was amplified from total RNA by RT-PCR. The construct therefore lacked the predicted intron and contained the first 119 amino acids of the LipDH2B protein (highlighted in green in figure 4.4). **Panel C** displays the results obtained with the LipDH2B (G)-GFP fusion protein. The LipDH2B (G)-GFP fusion construct starts at position 1 of LipDH2B and contains the following 581 nucleotides including the predicted intron to allow the parasites to "choose" between the potential start codon in the predicted exon 1 or the potential start codon in the predicted intron of the gene. The parasites mitochondrion was selectively stained with Mitotracker CMXRos (molecular probes) and images of live cells were obtained using a Zeiss Axioplan 2 microscope at 100 x magnification. The parasites were analysed using phase contrast (LIVE), GFP/FITC (GFP) and rhodamine (MITO) fluorescent filters. Overlays of the images obtained from each channel are displayed for determination of co-localisation of the two fluorescent signals (MERGE). The fluorescence patterns are displayed for ring (A), trophozoite (B), and schizont (C) stages in the erythrocytic cycle. The fluorescent signals for all three constructs display a clear co-localisation.

4.2.5 Cloning of both LipDH for recombinant expression

Previous studies in our laboratory had identified that in order for soluble recombinant expression to occur in *E. coli* it was necessary to prepare recombinant expression constructs without the hydrophobic N-terminal targeting sequences.

The LipDH1 expression construct was designed by utilising the LipDH1-GFP fusion protein to allow me to estimate the size of the mature protein. By immuno-precipitating the fusion protein with anti GFP-antibodies (AbCam), I could determine the approximate size of the mature protein and therefore the potential cleavage site of the bipartite signal/transit peptide. The full-length LipDH1-GFP fusion protein had a predicted molecular weight of 42 kDa and the GFP protein of 27 kDa (figure 4.10A); the mature LipDH1-GFP fusion protein isolated by immunoprecipitation was estimated to be 29 kDa (Lane 4; figure 4.10B). This mature protein size corresponded to a transit peptide cleavage site around amino acid 110 (highlighted red in figure 4.2). Therefore the construct for recombinant expression in *E. coli* was designed to be 1674 bp in length and ran from amino acid 111 – 667, which when cloned into pQE30 (a vector that attaches a His_6-tag to the N-terminus of the protein) would result in a recombinant protein of 64 kDa.

The LipDH2 expression construct was designed taking into account the predicted cleavage site obtained from the sub-cellular predictions programs for the LipDH2B protein. These predictions suggested a mitochondrial targeting sequence cleavage site near the intron/exon 2 boundary (table 4.1). I therefore designed an expression construct from the start of exon 2 (figure 4.3B; highlighted red in figure 4.4) that was 1470 bp in length, encompassing amino acids 25 – 513 (from the LipDH2B sequence) that, when cloned into pJC40 (a vector that attaches a His_{10}-tag to the N-terminus of the expressed protein) would produce an expressed protein of 57 kDa.

Both expression constructs were amplified by PCR, cloned into TOPO and sub-cloned into their relevant expression plasmid (figure 4.11).

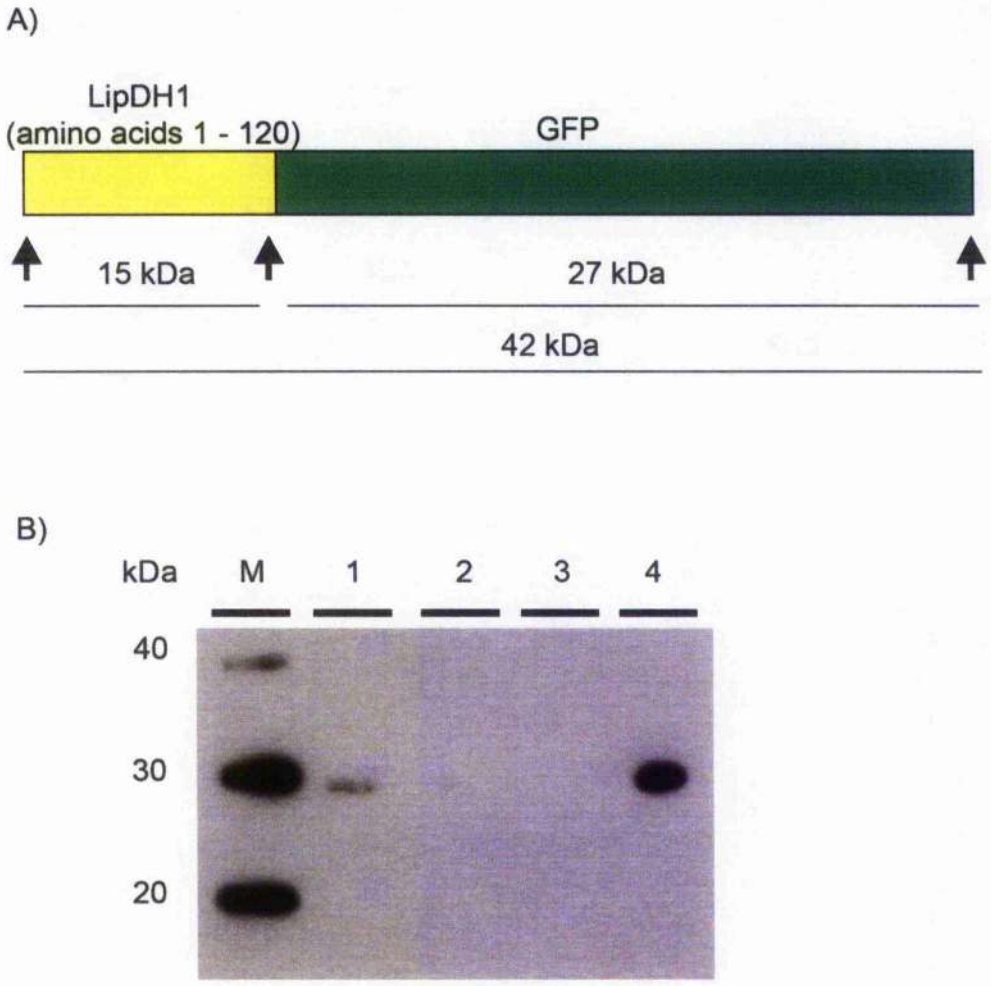


Figure 4.10 - Immunoprecipitation of LipDH1 GFP-fusion protein

The possible size of the mature LipDH1, following cleaving of the signal peptide and transit peptide was determined by immunoprecipitation of the LipDH1 GFP-fusion protein. **Panel A** displays a schematic overview of the 42 kDa fusion protein that was generated by the pHH2 LipDH1-GFP transfected parasites. This consisted of the 27 kDa GFP protein, fused with the 15 kDa LipDH1 N-terminal section (amino acids 1 – 120). **Panel B** displays the immunoprecipitation experiment where proteins extracted from parasites (lane 1) were initially pre-cleared in protein A sepharose CL-4B beads (without antibodies bound) to remove any unspecific binding (lane 2). The lysate was then incubated in beads (with anti GFP antibodies bound) and samples being separated into flow through fractions (lane 3) and the elution fraction (lane 4). All fractions were run on SDS-PAGE for western blotting with anti GFP antibodies; Lane M contains Magicmark protein standards (Invitrogen). Mature length GFP fusion proteins (with N-terminal signal/transit peptides cleaved) can be observed in both lysate (Lane 1) and elution (Lane 4) fractions, and was estimated to be 29 kDa in size; no full-length (42 kDa) proteins were observed in the western blots.

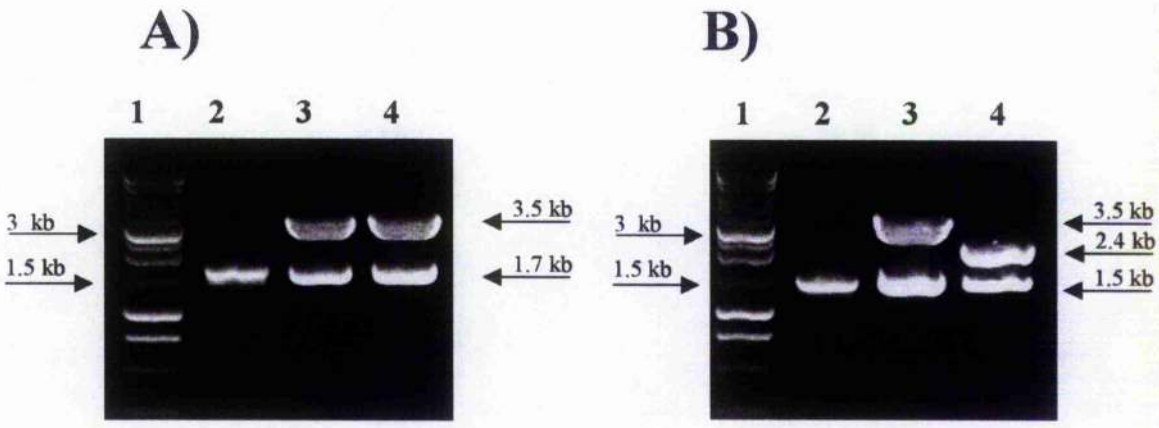


Figure 4.11 - Cloning of PfLipDH for recombinant expression

Both LipDH genes were amplified from gDNA using Taq supermix (Invitrogen) and specific oligonucleotides (table 2.1). **Panel A** displays an overview of the cloning of *LipDH1* for recombinant expression. The amplified PCR product was initially cloned into the TOPO TA PCR cloning vector for sequence verification before being sub-cloned into the recombinant expression vector pQE30. Lane 1 contains 10 μl 1kb DNA ladder (Promega); lane 2 contains the 1.7 kb PCR product of *LipDH1*; lane 3 contains the *Bam* HI / *Hind* III digest of the TOPO TA clone used for sub-cloning which resulted in a 3.5 kb band corresponding to the pCR2.1-TOPO vector and a 1.7 kb band corresponding to *LipDH1*; lane 4 contains a *Bam* HI / *Hind* III digest of the pQE30 recombinant expression clone and resulted in a 3.5 kb band corresponding to the pQE30 vector and a 1.7 kb band corresponding to *LipDH1*. **Panel B** displays an overview of the cloning of *LipDH2* for recombinant expression. The amplified PCR product was initially cloned into the TOPO TA PCR cloning vector for sequence verification before being for sub-cloned into the recombinant expression vector pJC40. Lane 1 contains 10 μl 1kb DNA ladder (Promega); lane 2 contains the 1.5 kb PCR product of *LipDH2*; lane 3 contains the *Nde* I / *Xho* I digest of the TOPO TA clone used for sub-cloning which resulted in a 3.5 kb band corresponding to the pCR2.1-TOPO vector and a 1.5 kb band corresponding to *LipDH2*; lane 4 contains a *Nde* I / *Xho* I digest of the pJC40 recombinant expression clone and resulted in a 2.4 kb band corresponding to the pJC40 vector and a 1.5 kb band corresponding to *LipDH2*.

4.2.6 Recombinant expression and purification of both PfLipDH

4.2.6.1 Recombinant expression of PfLipDH

In order to maximise parameters for the expression of both LipDH proteins, test expressions were set up to analyse the *E. coli* strains and temperatures following induction that led to the best expression levels. The use of terrific broth (TB) allowed the bacterial expression clones to grow to higher densities, therefore producing more recombinant protein.

The highest levels of expressions in LipDH1 test expressions were observed in *E. coli* strain Novablue (Novagen) in TB/Amp with expression induced with 1mM IPTG during mid-log phase and an overnight incubation at 30°C, 200 rpm (data not shown).

Although test expressions of LipDH2 initially suggested that no expression had occurred, it was later discovered that expression of LipDH2 was below the levels of detection for this method (see section 2.2.5). Larger scale expressions were required and LipDH2 expression analysed by purifying protein from these test expressions. The parameters determined to be optimal for LipDH2 were expression in terrific broth with *E. coli* strain BL21 (DE3) RIL with 1mM IPTG induction during mid-log phase and overnight expression at 37°C, 200 rpm (data not shown).

Large-scale expressions were then performed using the optimal conditions identified during the test expressions, to express sufficient recombinant proteins for characterisation of both LipDH proteins. The next section presents the purification of recombinant proteins from these large-scale expressions.

4.2.6.2 Purification of PfLipDH

The vectors used for the expression of both LipDH attach N-terminal His-tags to the recombinant proteins allowing purification using Ni-NTA affinity chromatography. Initial purification of both LipDH was carried out by batch purification using Ni-NTA agarose.

LipDH1 was purified to 95% homogeneity by the Ni-NTA batch purification process, with the purification process and purity of recombinant LipDH1 being analysed by SDS-PAGE (figure 4.12A). Expression levels of LipDH1 was in the range of 1-2 mg of

purified protein per 1L *E.coli* expression, although purified LipDH1 was found unstable with storage (greater than two weeks) at 4°C and may be susceptible to proteolysis.

LipDH2 was further purified after batch purification (following dialysis) by Ni-NTA affinity chromatography utilising an ÅKTA FPLC (Amersham biosciences) to elute proteins across a linear gradient of imidazole (10 – 500 mM), with the LipDH2 protein eluting between 400 - 500 mM imidazole (figure 4.12C). The purification process and purity of LipDH2 was analysed by SDS-PAGE (figure 4.12B). Expression levels of LipDH2 were extremely low, with 100 µg LipDH2 purified from a 1L *E. coli* expression.

Purification and activity levels of both LipDH proteins were found to be increased by the addition of 100 µM FAD (Sigma) to the harvested *E. coli* pellets. This addition increased the yield of enzymatically active LipDH2 from 57% to 87%, based on a comparison of the total protein (observed at 280 nm) to that of FAD containing protein (observed at 460 nm).

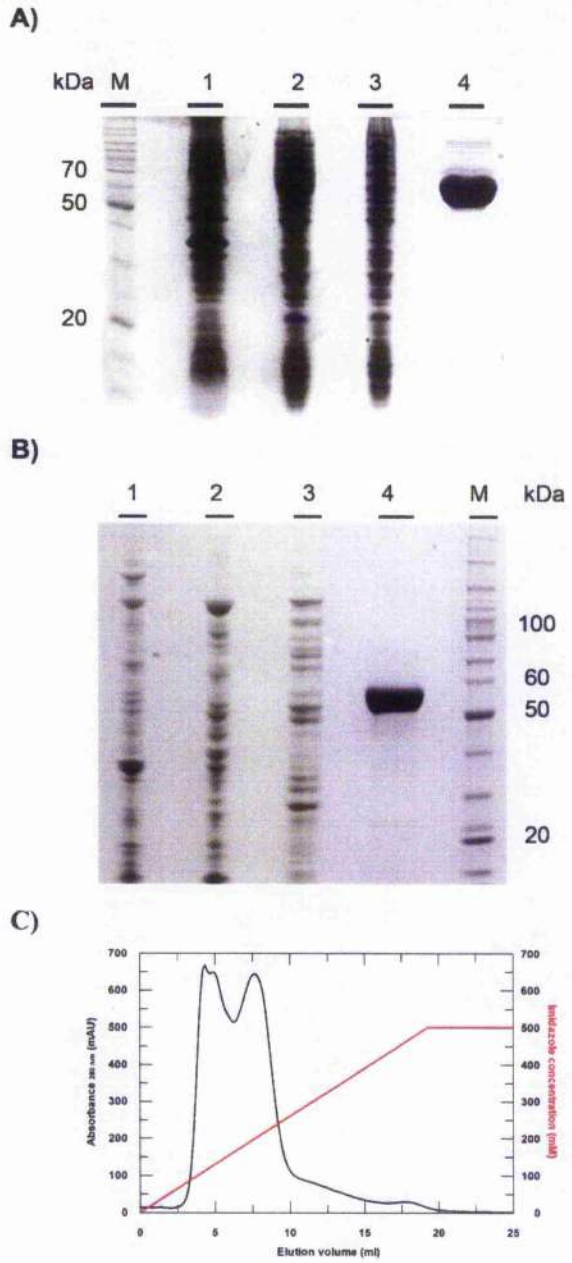


Figure 4.12 - Purification of recombinantly expressed PfLipDH

Both LipDH proteins were purified from soluble *E. coli* extract by Ni-NTA affinity chromatography. **Panel A** displays an overview of the batch purification of LipDH1, run on 15 % SDS-PAGE and stained with Coomassie stain. The gel displays the Benchmark protein ladder (Lane M; Invitrogen) and 10μg of the soluble *E. coli* fraction (Lane 1), insoluble *E. coli* fraction (Lane 2), Ni-NTA beads flow through fraction (lane 3) and elution fraction (Lane 4). The elution fraction contains a main band of 64 kDa that is in good agreement with the predicted size of the recombinant LipDH1 protein. **Panel B** displays an overview of the two-step Ni-NTA batch and FPLC purification of LipDH2, run on 4 - 12 % SDS-PAGE and stained with Coomassie stain. The gel was loaded with 10μg of the insoluble *E. coli* fraction (Lane 1), soluble *E. coli* fraction (Lane 2), batch purification elution fraction (Lane 3) and elution fraction from FPLC purification (Lane 4); Lane M contains Benchmark protein standards (Invitrogen). **Panel C** displays the FPLC purification of LipDH2 following dialysis of the batch binding elution. The sample (lane 3 in panel B) was loaded onto a 1 ml HiTRAP Ni-chelating column (Amersham pharmacia) and was run on an ÄKTA purifier FPLC (Amersham pharmacia) to elute the recombinant protein over a linear range between 10 mM and 500 mM imidazole. Recombinant LipDH2 was eluted between 450 – 500 mM Imidazole (small peak eluted after 16 ml), and the purified protein was analysed on SDS-PAGE (lane 4 in panel B).

4.2.7 Characterisation of LipDH

In order to confirm that the recombinantly expressed proteins were functional as LipDH, I analysed their oligomeric state, spectral properties, kinetic parameters and catalytic mechanism.

4.2.7.1 Analyses of oligomeric state

The oligomeric state of both LipDH was studied in order to confirm whether or not they formed homodimers. Purified samples of LipDH1 and LipDH2 were separated according to molecular weight by applying to gel filtration chromatography columns (either Sephadex S200 16/60 or Superose 6 10/30). The columns were also calibrated with molecular weight markers (Biorad), and the resulting calibration curve allowed the calculation of the molecular size of the recombinant proteins (insets in figure 4.13).

The gel filtration of LipDH1 produced two observed peaks; the first peak eluted after 46 ml and the second after 74 ml (figure 4.13A). The first peak corresponded to the void volume of the column and presumably consisted of large multimers or precipitated protein, while the second peak (which contained the majority of the protein) had a calculated molecular weight of 126 kDa. This second peak was close to the expected size for a homodimeric LipDH1 protein (130 kDa including the His_6 tag), confirming that LipDH1 was likely to form homodimers.

The gel filtration of LipDH2 produced one major peak that eluted after 14.5 ml. This elution volume corresponded to a calculated molecular weight of LipDH2 of 77.6 kDa, unfortunately this lay between the sizes expected for monomeric (57 kDa) and homodimeric (114 kDa) proteins.

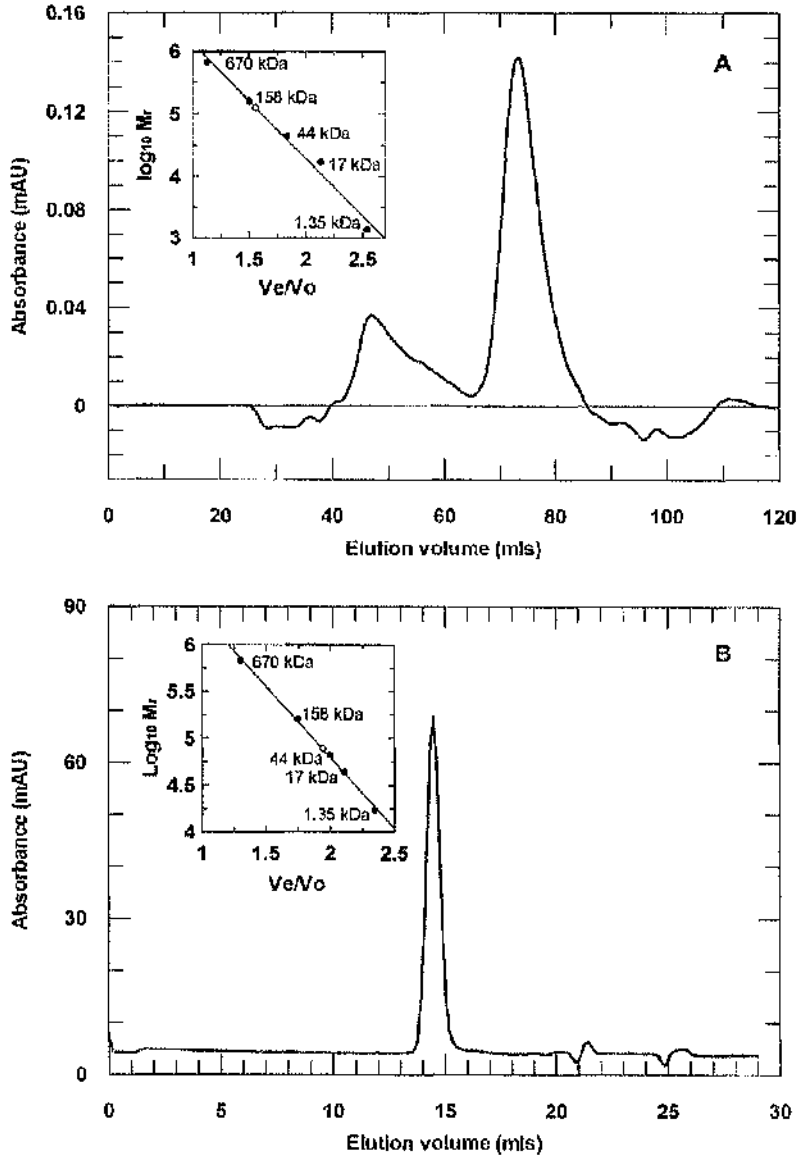


Figure 4.13 - Oligomeric state analyses of recombinant LipDH

The oligomeric state of both purified LipDH proteins was analysed by either gel filtration. **Panel A** displays the gel filtration elution of 1.5 mg of purified LipDH1 on a Sephadex S200 16/60 size exclusion column. The inset displays the calibration curve for the column based on the elution volumes of protein standards (Filled circles; Biorad) with the position of the eluted LipDH1 protein (Open circle). **Panel B** displays the gel filtration elution of 300 µg of purified LipDH2 on a Superose 6 10/30 size exclusion column. The inset displays the calibration curve based on the elution volumes of protein standards (Filled circles; Biorad) with the position of the eluted LipDH2 protein (Open circle).

4.2.7.2 Spectral properties of PfLipDH

The spectral properties of the purified proteins were analysed in order to confirm their role as active LipDH (figure 4.14). Both proteins were purified in the oxidised form and displayed characteristic spectra of flavoprotein disulphide oxidoreductases (Ghisla *et al.*, 1974). The peak observed at 460 nm along with the shoulders at 430 nm and 480 nm are consistent with the presence of a bound flavin adenine dinucleotide (FAD) cofactor.

The absorption spectra of the purified proteins change following reduction with 20 molar equivalents of NADH (figure 4.14; line 2). The maximum at 460 nm decreases in absorbance and shifts to a shorter wavelength (approximately 450 nm), which is consistent with reduction of the FAD prosthetic group. Another peak in the reduced spectra is centred around 530 nm and is due to the formation of a charge transfer complex between the FAD and a thiolate formed on one of the active site cysteine residues. Both of these spectral features are consistent with those observed for other members of the disulphide oxidoreductases (Ghisla *et al.*, 1974; Thorpe *et al.*, 1976). The large absorption increase at shorter wavelengths, which has a maximum centred at 340 nm (off scale, figure 4.14) was due to the presence of excess NADH used as the reductant. The lower level of flavin reduction and thiolate-flavin charge transfer observed in reduced LipDH1 (figure 4.14A; line 2) than that observed in reduced LipDH2 (figure 4.14B; line 2), was due to rapid re-oxidation of the flavin by oxygen. Ideally this experiment should have been performed under anaerobic conditions.

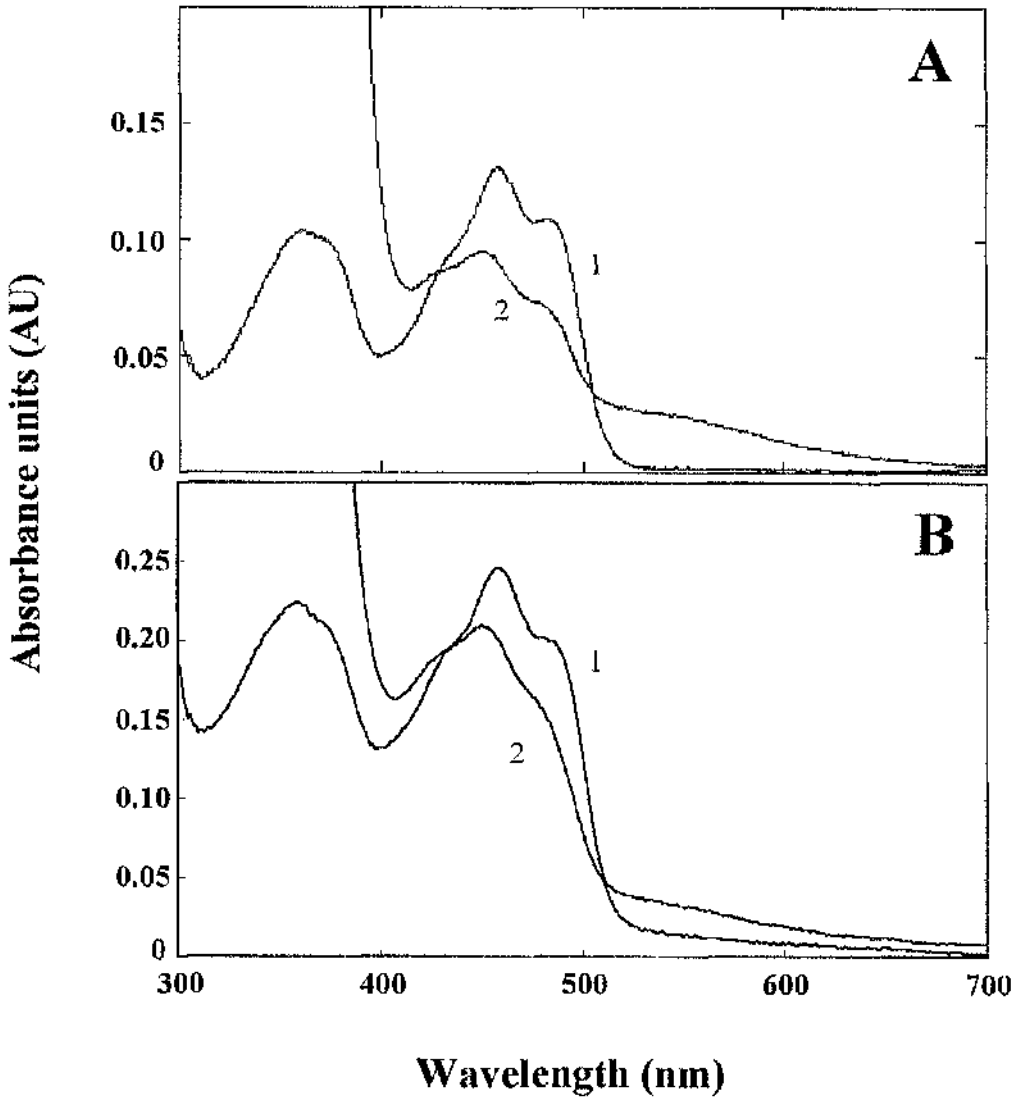


Figure 4.14 - Oxidised and reduced spectra of PfLipDH

Purified *Pf*LipDH was analysed spectrally between 300 nm and 700 nm using a UV-2501 PC spectrophotometer (Shimadzu). Purified recombinant protein were analysed in the oxidised state (line 1) and reduced state (line 2); following reduction with a 20 molar excess (NADH/FAD) of NADH under aerobic conditions. **Panel A** displays spectra for 11 μM LipDH2.; **Panel B**, displays spectra for 21.8 μM LipDH1

4.2.7.3 Kinetic analyses of LipDH1

In order to confirm that LipDH1 was enzymatically active, I tested the recombinant protein with LipDH-specific substrates. Before kinetic analyses could be carried out I verified the pH-optimum for the LipDH1 reverse reaction. Analyses of the activity of LipDH1 between pH 7.0 and pH 8.0 revealed a broad shoulder with a peak at approximately pH 7.6 (data not shown). LipDH1 displayed Michaelis-Menten kinetics for both substrates that lead to $K_{m\ app}{}^{NADH}$ of 15.5 ± 3.0 μM and $K_{m\ app}{}^{lipoamide}$ 1211 ± 260 μM (figure 4.15). The calculated apparent k_{cat} was determined to be between 27.3 and 37.1 s^{-1} (table 4.2).

4.2.7.4 Kinetic analyses of LipDH2

In order to confirm that LipDH2 was enzymatically active I tested the recombinant protein with LipDH-specific substrates. Before in depth kinetic analyses could be carried out on LipDH2 it was necessary to determine the pH-profiles for both the forward (physiological) and reverse reactions (non-physiological). The forward reaction enzymatic activity was maximal at pH 9.0 and the reverse reaction enzymatic activity was maximal at pH 7.0.

The kinetic parameters of the LipDH2 forward reaction were determined using stopped flow fast kinetics to measure initial rates as the reaction was inhibited by NADH (Sahlman *et al.*, 1989). The apparent K_m values for both substrates (dihydrolipoamide and NAD^+) and the apparent k_{cat} were calculated for the forward reaction at pH 8.0 and 25°C; pH 8.0 was used to minimise the spontaneous oxidation of dihydrolipoamide observed at pH 9.0 (Reed *et al.*, 1958). The apparent k_{cat} was calculated to be between 135.5 and 337 s^{-1} (depending on enzyme preparation), the apparent $K_m{}^{dihydrolipoamide}$ was 146 ± 15 μM and the apparent $K_m{}^{NAD+}$ was 450 ± 30 μM (figure 4.16; table 4.2).

The pH-optimum determined (pH 7.0) for the reverse reaction was used in all activity and kinetic studies of this reaction. The assay measured the oxidation of NADH at 340 nm using stopped flow fast kinetics. The apparent K_m for both substrates (NADH and lipoamide) and the k_{cat} were calculated for the reverse reaction at 25°C. The apparent k_{cat} was calculated to be 448 s^{-1} and the $K_{m\ app}{}^{NADH}$ and $K_{m\ app}{}^{lipoamide}$ were 20.8 ± 5.6 μM and 870 ± 270 μM, respectively (figure 4.17 and table 4.2).

The catalytic mechanism of LipDH2 was investigated using the reverse reaction. By measuring enzyme activity over a range of lipoamide concentrations (between 0.5 and 2 mM) at three NADH concentrations (10, 20 and 40 μM) it is possible to determine the enzyme mechanism of LipDH2. A Lineweaver-Burk (double reciprocal) plot of the resultant activities produces data that can be fitted with parallel lines, suggesting that LipDH2 acts through a ping-pong mechanism (figure 4.18; in a sequential reaction these lines will converge on a single point). In a ping-pong mechanism a reduced substrate (in this case NADH) must first bind and reduce the enzyme before the second substrate (in this case lipoamide) can bind and oxidise the enzyme. Further analysis of the data from this experiment allowed me to calculate the true K_m for both substrates. The calculated K_m^{NADH} is 28 ± 4 μM, $K_m^{lipoamide}$ is 1,720 ± 230 μM (table 4.2).

	LipDH1	**LipDH2**	***H. sapiens*[a]**
pH optimum (forward)	n.d.	pH 9.0	n.d.
pH optimum (reverse)	pH 7.6	pH 7.0	n.d.
$K_{m\,app}$ NAD^+	n.d.	0.45 ± 0.03 mM	0.29 mM
$K_{m\,app}$ NADH	15.5 ± 3.0 μM	20.8 ± 5.6 μM	n.d.
K_m NADH	n.d.	28 ± 4 μM	51 μM
$K_{m\,app}$ Lipoamide	1.21 ± 0.26 mM	0.87 ± 0.27 mM	n.d.
K_m Lipoamide	n.d.	1.72 ± 0.23 mM	1.01 mM
$K_{m\,app}$ Dihydrolipoamide	n.d.	146 ± 15 μM	570 μM
k_{cat} forward	n.d.	135.5 – 337.0 s^{-1} [b]	382 s^{-1}
k_{cat} reverse	27.3 – 37.1 s^{-1} [c]	448 s^{-1}	167 s^{-1}

Table 4.2 - Kinetic parameters of LipDH

Overview of the determined kinetic parameters for both LipDH proteins, compared to those of the human mitochondrial isoform. [a] Kinetic data for the Human LipDH was taken from Kim *et al* 1991 and are true Km. [b] The k_{cat} determined for the LipDH2 forward reaction was dependent on the enzyme preparation and the assay system used. [c] The k_{cat} determined for the LipDH1 forward reaction was dependent on the parameters of the reaction used to determine it. n.d. parameter was not determined.

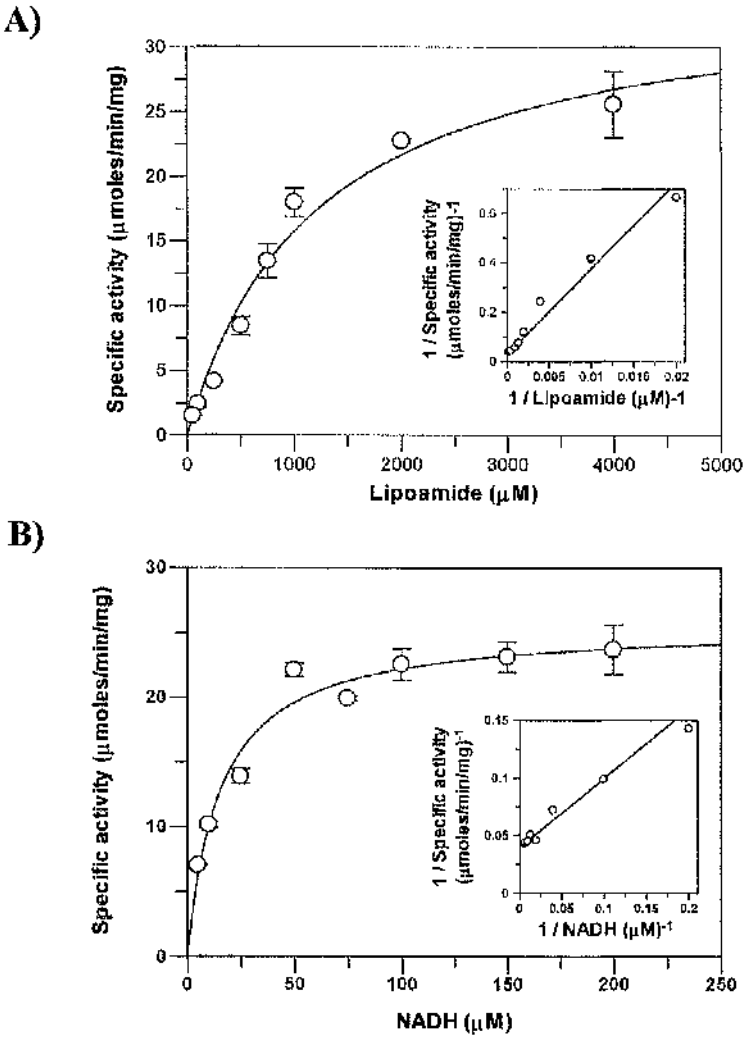


Figure 4.15 – Determining the kinetic parameters of the LipDH1 reverse reaction

The kinetic parameters of the LipDH1 reverse reaction were determined at 25 °C in 50 mM potassium phosphate, 1 mM EDTA pH 7.6. The assay was followed at 340 nm for the oxidation of NADH using the extinction coefficient of 6220 M^{-1} cm^{-1}. Each point represents the mean ± standard error for 2 to 3 independent measurements, with apparent Km and Vmax determined using the Grafit 5.0 software (Erithacus software). **Panel A** displays the Michaelis-Menten plot of LipDH1 reverse reaction for lipoamide. The reaction mixture contained saturating levels of NADH (200µM) and variable concentrations of lipoamide (50 µM to 4 mM). **Panel A inset** displays the Lineweaver-Burk (Double reciprocal) plot of the data. The apparent $K_m^{\text{lipoamide}}$ and the V_{max} of LipDH1 were determined to be 1210 ± 260 µM and 37.1 units mg^{-1}, respectively. **Panel B** displays the Michaelis-Menten plot of the LipDH1 reverse reaction for NADH. The reaction mixture contained 2 mM lipoamide and variable concentrations of NADH (5 µM to 200µM) lipoamide. **Panel B inset** displays the Lineweaver-Burk (Double reciprocal) plot of the data obtained. The apparent K_m^{NADH} and the V_{max} of LipDH1 were determined to be 15.5 ± 3.0 µM and 27.3 units mg^{-1}, respectively.

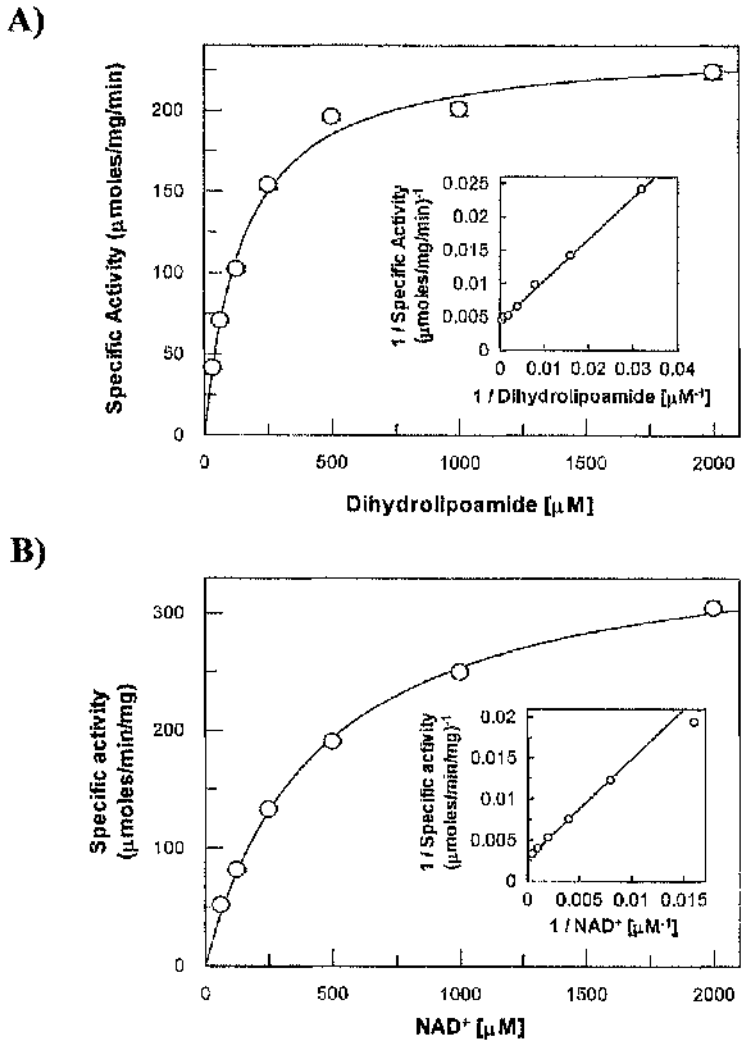


Figure 4.16 – Determining the kinetic parameters of the LipDH2 forward reaction

The kinetic parameters of the LipDH2 forward reaction were determined at 25°C in 50 mM potassium phosphate, 1 mM EDTA pH 8.0. The activity of 1 µg of recombinantly expressed LipDH2 was followed at 340 nm for the reduction of NAD^+ using the extinction coefficient of 6220 M^{-1} cm^{-1}. Each point represents the mean ± standard error for 3 to 5 independent measurements, with apparent Km and Vmax determined using the Grafit 5.0 software (Erithacus software). **Panel A** displays the Michaelis-Menten plot of LipDH2 forward reaction for dihydrolipoamide. The reaction mixture contained 2 mM NAD^+ and variable concentrations of dihydrolipoamide (50 µM to 2mM). **Panel A inset** displays the Lineweaver-Burk (Double reciprocal) plot of the data obtained. The apparent K_m ^{dihydrolipoamide} and the V_{max} of LipDH2 were determined to be 146 ± 15µM and 240 ± 6 units mg^{-1}, respectively. **Panel B** displays the Michaelis-Menten plot of the LipDH2 forward reaction for NAD^+. The reaction mixture contained saturating levels of dihydrolipoamide (2mM) and variable concentrations of NAD^+ (62.5µM to 2mM). **Panel B inset** displays the Lineweaver-Burk (Double reciprocal) plot of the data obtained. The apparent K_m $^{NAD+}$ and the V_{max} of LipDH2 were determined to be 450 ± 30µM and 368 ± 8 units mg^{-1}, respectively.

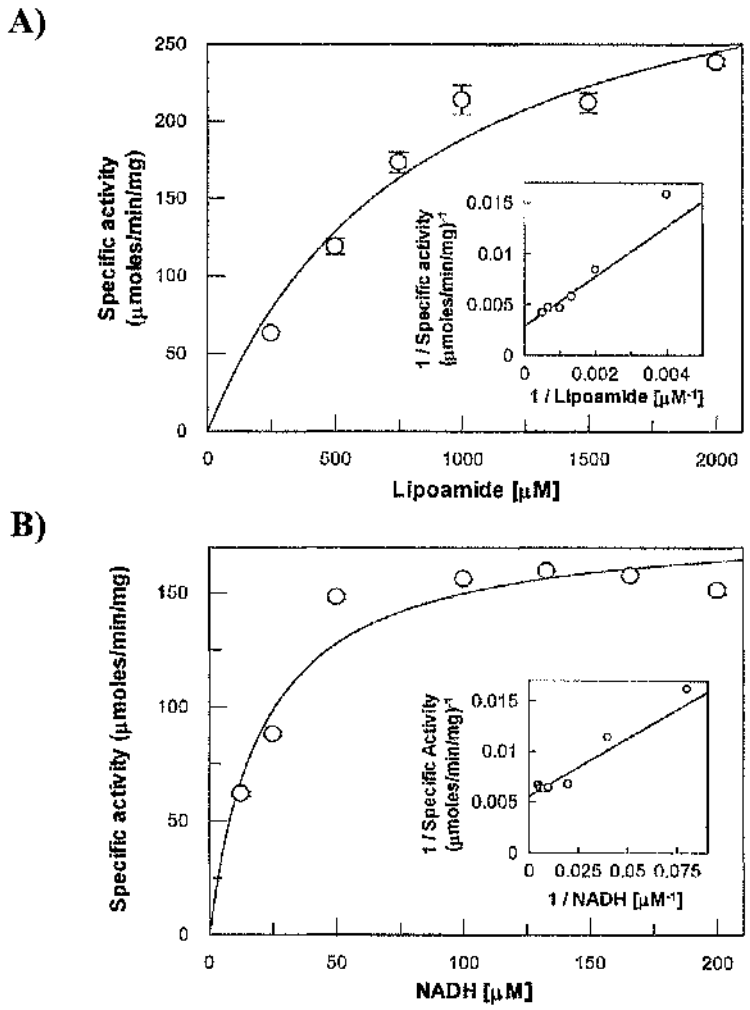


Figure 4.17 – Determining the kinetic parameters of the LipDH2 reverse reaction

The kinetic parameters of the LipDH2 reverse reaction were determined at 25 °C in 50 mM potassium phosphate, 1 mM EDTA pH 7.0. The activity of 1 µg of recombinantly expressed LipDH2 was followed at 340 nm for the oxidation of NADH using the extinction coefficient of 6220 M^{-1} cm^{-1}. Each point represents the mean ± standard error for 3 to 5 independent measurements, with apparent Km and Vmax determined using the Grafit 5.0 software (Erithacus software). **Panel A** displays the Michaelis-Menten plot of LipDH2 reverse reaction for lipoamide. The reaction mixture contained saturating levels of NADH (200µM) and variable concentrations of lipoamide (50 µM to 2 mM) lipoamide. **Panel A inset** displays the Lineweaver-Burk (Double reciprocal) plot of the data obtained. The apparent K_m ^{lipoamide} and the V_{max} of LipDH2 were determined to be 870 ± 270µM and 351 ± 50 units mg^{-1}, repectively. **Panel B** displays the Michaelis-Menten plot of the LipDH2 reverse reaction for NADH. The reaction mixture contained 2 mM lipoamide and variable concentrations of NADH (10 µM to 200µM). **Panel B inset** displays the Lineweaver-Burk (Double reciprocal) plot of the data obtained. The apparent K_m ^{NADH} and the V_{max} of LipDH2 were determined to be 20.8 ± 5.6µM and 181 ± 11 units mg^{-1} units mg^{-1}, respectively.

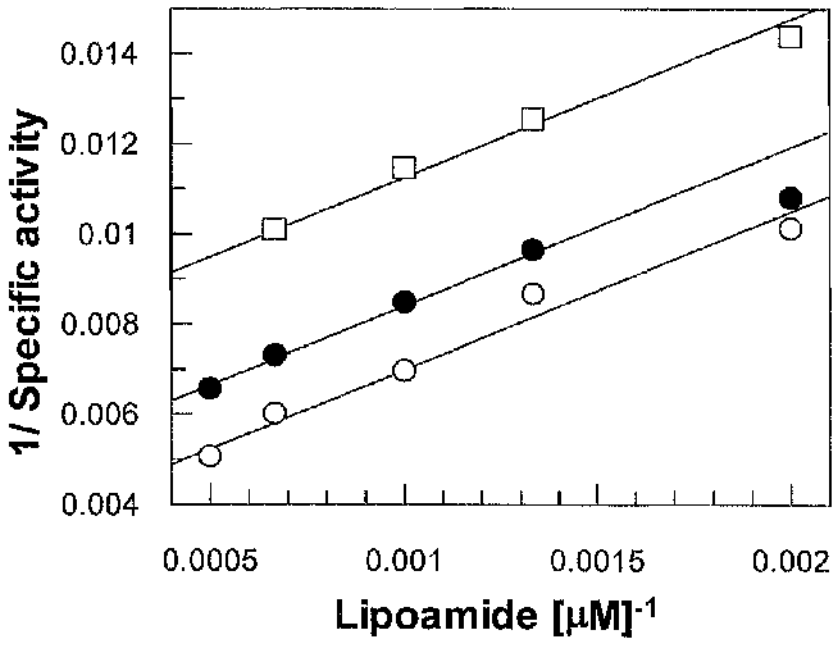


Figure 4.18 - LipDH2 displays ping-pong kinetics

The catalytic mechanism of LipDH2 was analysed using the reverse reaction. Specific activities were measured for a range of lipoamide concentrations (between 0.5 and 2 mM) at three NADH concentrations (○ 10 μM; ● 20 μM; □ 40 μM), in all combinations. The kinetic data determined was in good agreement with those observed in previous experiments (table 4.2, figures 4.16 & 4.17), with $K_m^{Lipoamide}$ determined as 1.72 ± 0.23 mM and K_m^{NADH} was estimated to be 28 ± 4 μM. The calculated V_{max} of LipDH2 was 489 ± 45 units mg^{-1}.

4.3 Discussion

4.3.1 *P. falciparum* possess two distinct *LipDH* genes

The completion of the *Plasmodium falciparum* genome sequencing project has provided researchers with a large number of resources to aid their research of this human parasite (www.plasmodb.org). Using these resources I have identified and characterised two genes that encode LipDH proteins in the *P. falciparum* genome. The presence of two *LipDH* genes in the *P. falciparum* genome is in contrast to the single *LipDH* identified in the genome of the human host.

One of these genes (*LipDH2*) was found to contain an intron at its 5' end separating two exons; the first of which contained the predicted mitochondrial targeting sequence. Analyses of the *P. falciparum* genome displayed that 53.9 % of predicted genes contained introns and that the predicted intron sequences displayed a higher A/T content (86.5 %) than exon (76.3 %) sequences (Gardner *et al.*, 2002). This higher A/T content of introns was also observed in the LipDH2B gene and was further confirmed in a recent study into the splice sites from *P. falciparum*. This study identified that the parasites use the canonical GT-AG boundaries with a high preference for A/T rich sequences in the intervening intronic sequences (Szafranski *et al.*, 2005).

4.3.2 *P. falciparum* LipDH proteins are organelle specific

The predicted LipDH proteins displayed greatest homology to different LipDH isoforms (mitochondrial or plastid), suggesting that the proteins were organelle specific. This theory was further strengthened by detailed phylogenetic analysis of both LipDH proteins with those from other species (Foth *et al.*, 2005). This study confirmed that the LipDH1 was grouped with other plastid isoforms and those of Cyanobacteria, while the LipDH2 was grouped with mitochondrial isoforms and those of α-protobacteria (Foth *et al.*, 2005). This suggested that these nuclear encoded genes arose by lateral gene transfer from the endosymbionts that were the origins of these organelles (Martin *et al.*, 1998; Zhu *et al.*, 2002; Adams *et al.*, 2003).

Both *P. falciparum* proteins exhibited N-terminal extensions that contained targeting peptides required for targeting to their respective organelles. The *P. falciparum* LipDH1 amino acid sequence contained a bipartite signal/transit peptide, in which the hydrophobic signal peptide was responsible for placing the protein into the secretory

system (via the endoplasmic reticulum) and the transit peptide was responsible for the transport of the protein into the apicoplast. The features of the bipartite signal/transit peptide have been investigated in depth and have led to the production of the subcellular localisation prediction programs used in this study (Nielsen *et al.*, 1997; Emanuelsson O, 2000; Foth *et al.*, 2003). The *P. falciparum* LipDH2 protein contained a mitochondrial targeting peptide, which was consistent with studies into mitochondrial targeting sequences for *Plasmodium* (Bender *et al.*, 2003) and other species (Claros *et al.*, 1996).

I have experimentally determined that LipDH1 is localised in the apicoplast and that LipDH2 is localised in the mitochondrion and these subcellular localisations have also been confirmed independently in another lab (using both GFP-fusion protein and immuno-fluorescence analyses; Foth & McFadden personal communication). However, the localisation of the LipDH2A-GFP fusion protein (which was predicted not to contain a targeting peptide) to the mitochondrion identified the possibility that further signals may be responsible for transfer of proteins to this organelle. This was consistent with previous studies that have identified mitochondrial-localised proteins that lack the classical N-terminal targeting peptide (Stan *et al.*, 2003). In these cases internal targeting sequences have been proposed to be sufficient for mitochondrial localisation (Pfanner *et al.*, 1987; Truscott *et al.*, 2001).

The observed organelle specific subcellular localisations of the *P. falciparum* LipDH proteins was reminiscent of the situation in plants, where cells contain both mitochondria and plastids. In *Pisum sativum* (Pea plant) two LipDH proteins have been identified, one of which is located in the mitochondrion and the other in the plastid. The *P. sativum* mitochondrial LipDH is shared between the mitochondrial KADH (Bourguignon *et al.*, 1996), while the *P. sativum* plastid LipDH is a component of the plastid PDH (Conner *et al.*, 1996). However, the situation is complicated in *Arabidopsis thaliana*, by the presence of two mitochondrial and two plastid LipDH isoforms (Lutziger *et al.*, 2000; Lutziger *et al.*, 2001). Gene knockout studies demonstrated that one mitochondrial LipDH was able to substitute for the loss of the other, suggesting that the two mitochondrial LipDH can be shared between the mitochondrial KADH (Lutziger *et al.*, 2001). The two plastid LipDH isoforms identified in the *A. thaliana* genome have 81% amino acid sequence identity and their subcellular localisation was confirmed by chloroplast import assays (Lutziger *et al.*, 2000). The predicted localisation of organelle-specific LipDH proteins would suggest that there are likely to be KADH located in both organelles (Johnston *et al.*, 1997; Lutziger *et al.*, 2000). The

localisation of KADH and the consequences of their distribution will be investigated and discussed in detail in chapter 5.

4.3.3 Recombinantly expressed PfLipDH are active

Both LipDH from *P. falciparum* were recombinantly expressed and displayed enzymatic activity using LipDH-specific substrates. The k_{cat} of the LipDH1 reverse reaction was determined to be more than ten fold lower than that of the LipDH2. The lower specific activity of the plastid LipDH when compared to its mitochondrial counterpart was analogous to the situation observed in *Pisum sativum* (pea plant), where the plastid LipDH has a 5 fold lower specific activity to that of the mitochondrial LipDH (Conner *et al.*, 1996). The pea plastid LipDH also displayed a 7 fold higher affinity for NAD^+ than its mitochondrial isoform (G. Lindsay, personal communication). The steady state kinetic parameters determined for both LipDH were in good agreement with those determined previously for LipDH from other organisms (Kim *et al.*, 1991; Else *et al.*, 1994; Schoneck *et al.*, 1997; Argyrou *et al.*, 2001) and the inhibition of LipDH2 by NADH that was observed in the forward reaction, has also been reported previously (Sahlman *et al.*, 1989).

The apparent instability displayed in the recombinantly expressed LipDH1 suggests that this recombinant protein may be susceptible to proteolysis. The correct folding of these proteins is essential for catalysis and this was confirmed by the findings of Lindsay *et al*, where they displayed that FAD-binding was essential for the correct folding of denatured LipDH and enzymatic activity (Lindsay *et al.*, 2000). Further analyses of the mature forms of both *P. falciparum* LipDH would perhaps allow for more stable protein expression; this may be achieved by Edman sequencing (N-terminal sequencing) or tryptic digestion of immuno-precipitated GFP-fusion proteins, and could be tested by circular dichroism of the recombinant proteins. Re-cloning of the both LipDH constructs into a different expression system, perhaps using a C-terminal His-tag or a Streptavidin-tag, may also result in higher levels of protein expression.

Other factors important for LipDH-activity could be analysed from the mutations observed in patients displaying LipDH-deficiency, which leads to numerous clinical symptoms including lactic acidosis, neurological dysfunctions and maple syrup urine disease (Liu *et al.*, 1993; Hong *et al.*, 1996; Hong *et al.*, 1997; Shany *et al.*, 1999; Cerna *et al.*, 2001; Grafakou *et al.*, 2003). Ten of the mutations that caused LipDH deficiency were investigated in the crystal structure of the human LipDH, with nine of them

observed to cluster in three areas; the homodimer interface, the FAD and NAD^+ binding sites and the disulphide interface (Brautigam *et al.*, 2005); the other produced a nonsense mutation that truncated the human LipDH protein at tyrosine 35 (Hong *et al.*, 1996).

The clinical manifestations of LipDH-deficiency result from a decrease in activity of the KADH, with maple syrup urine disease the product of an accumulation of branched chain α-keto acids brought about by the loss of BCKADH activity (Kuhara *et al.*, 1983). As an essential component of the α-keto acid dehydrogenase complexes, LipDH plays an important role in the biochemistry of the parasites during their erythrocytic stages. Although the main roles of KADH in these stages will be further discussed in chapter 5, it has also been suggested that LipDH/KADH may also interact with other pathways.

Indeed, LipDH has been shown to act with one of the major antioxidant systems in the human pathogen *Mycobacterium tuberculosis*. This interaction has been determined to involve a LipDH (Lpd), an acetyl transferase (DlaT), a thioredoxin-like protein (AhpD) and a peroxiredoxin-like protein (AhpC) and results in the NADH-specific reduction of hydroperoxides and peroxynitrite (Bryk *et al.*, 2002; Koshkin *et al.*, 2004; Tian *et al.*, 2005). The LipDH transfers reducing equivalents, via the protein bound lipoamide, to the AhpD and onto the AhpC, in a system analogous with the thioredoxin cascade (figure 4.19).

The presence of an antioxidant function in the *Plasmodium* mitochondrion would be required, because of the presence of an active electron transport chain that may produce reactive oxygen species (Krungkrai, 2004; Uyemura *et al.*, 2004). A mitochondrially localised superoxide dismutase has already been identified (Sienkiewicz *et al.*, 2004), that could produce hydrogen peroxide (H_2O_2) from superoxide (O_2^-). Immunofluorescence studies have recently identified a mitochondrially localised thioredoxin peroxidase (Yano *et al.*, 2005) and a possible mitochondrial thioredoxin has been proposed (Muller, 2004), which may be involved in a peroxiredoxin linked antioxidant system similar to those identified in humans and yeast (Miranda-Vizuete *et al.*, 2000). No mitochondrially localised thioredoxin reductase (homologous to those identified in mammals) has yet been identified, suggesting that the LipDH/KADH antioxidant function may play an important role in the mitochondrion. Work is ongoing in our laboratory to try and identify possible proteins involved in antioxidant functions in the parasites organelles.

In addition to acting on protein bound lipoamide, LipDH may also use free lipoic acid as a substrate. Although the level of free lipoic acid is usually extremely low, exogenous stressing of the parasites with hydrogen peroxide resulted in a 10-fold increase in protein bound lipoic acid and equivalent levels of free lipoic acid (T. Smith and S. Müller, unpublished data). Many studies have identified the antioxidant functions of both the oxidised and reduced forms of lipoic acid. The oxidised form (lipoic acid) can be reduced by thioredoxin reductase (Arner *et al.*, 1996), glutathione reductase (Pick *et al.*, 1995) and can also scavenge hydroxyl radicals, hypochlorous acid, singlet oxygen and chelates iron, copper and other transition metals (Moini *et al.*, 2002a). The reduced form (dihydrolipoic acid) has been shown to reduce other antioxidants such as glutathione disulphide, thioredoxin, vitamin C and vitamin E (Jocelyn, 1967; Packer *et al.*, 1995). Dihydrolipoic acid may also act as a pro-oxidant via the reduction of iron (Fe^{3+} to Fe^{2+}) producing hydroxyl radicals (Moini *et al.*, 2002b).

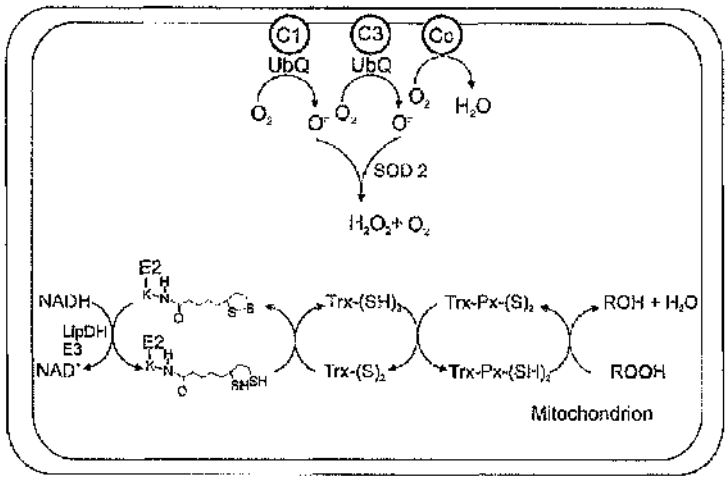


Figure 4.19 – Mitochondrial antioxidant function

The presence of an active electron transport chain can produce reactive oxygen species that need to be detoxified. The electron transport chain is represented in this figure by complexes I and III (C1 & C3), that both utilise ubiquinone (UbQ) as a substrate. Both of these complexes leak superoxide anions (O^-), which can be dismutated to hydrogen peroxide (H_2O_2) by a mitochondrially localised superoxide dismutase (SOD2). The *Plasmodium* genome contains genes that encode for a mitochondrial peroxiredoxin (Trx-Px) and a potential mitochondrially localised thioredoxin (Trx) that can reduce the H_2O_2 produced by SOD2. However, the *Plasmodium* genome lacks a mitochondrially localised thioredoxin reductase and therefore a reducing partner for the thioredoxin. We postulate that the mitochondrially localised LipDH2 (LipDH E3) and acyl transferases of the KADH (E2) may be able to provide the reducing equivalents for the pathway to be functional.

4.3.4 Inhibition of LipDH

LipDH plays a major role in integral biochemical processes of the *P. falciparum* parasite via its involvement in the KADH (see chapter 5) and the possible anti-oxidant role proposed earlier. Any inhibition of LipDH may have extreme consequences for the parasite. Inhibition of LipDH activity has been observed by NADH (Sahlman *et al.*, 1989) and has been proposed to be involved in regulation of KADH activity (Tovar-Mendez *et al.*, 2003). This regulation has also been observed in the inhibition of α – ketoglutarate dependent mitochondrial respiration by competitive inhibition by sub-micromolar levels of zinc (Brown *et al.*, 2000). The zinc (in the form of Zn^{2+}) interacts with the reduced active site thiols (where it competes with lipoamide) and increases the NADH oxidase activity of the LipDH protein, thereby producing reactive oxygen species (Gazaryan *et al.*, 2002).

The inhibition of LipDH proteins (along with other disulphide oxidoreductases) has also been investigated to provide chemotherapeutic drugs against human pathogens. Possible drug inhibitors of the disulphide oxidoreductases have been extensively investigated in the *Trypanosome* species (due to the presence of the parasite specific trypanothione reductase) as well as in *P. falciparum* (Krauth-Siegel *et al.*, 2005). In addition to the numerous inhibition studies that have investigated trypanothione reductase as a drug target (Krauth-Siegel *et al.*, 2005), many studies have also investigated the inhibition of LipDH in *Trypanosomes*. Trivalent arsenicals such as melarsenoxide have been demonstrated to interact with the active site cysteines in the reduced form of the protein (Fairlamb *et al.*, 1992; Cunningham *et al.*, 1994). Nitrosoureas such as carmustine also act via covalent inhibition of the protein by carbomylation of the interchange thiol (Lohrer *et al.*, 1990). Several other compounds display what has been termed turncoat inhibition, where they are substrates for the protein that induce one electron transfer to oxygen, which leads to the production of reactive oxygen species; examples of turncoat inhibitors active against LipDH are napthoquinones/benzoquinones derivatives (Cenas *et al.*, 1994; Cenas *et al.*, 1994) and nitrofurans such as nifurtimox (Cenas *et al.*, 1994; Blumenstiel *et al.*, 1999). Phenothiazine radicals were also shown to produce inhibition by irreversible modification of the LipDH protein structure (Gutierrez-Correa *et al.*, 2003).

In *Plasmodium*, these disulphide oxidoreductase inhibition studies have been based on thioredoxin reductase and glutathione reductase (see chapters 1 & 3), as the dihydrolipoamide dehydrogenase had not been recombinantly expressed until now.

Therefore the recombinant LipDH proteins that I have generated may also now be utilised for anti-malarial drug screening. I am currently working to validate whether both LipDH proteins are essential for parasite survival by targeted gene knockout, although the analyses are at too early a stage to be included in this thesis.

4.4 Conclusions

I have identified two LipDH genes in the *Plasmodium* genome, one of which (LipDH2) contains an intron. The two LipDH proteins were determined to be organelle-specific using GFP-fusion proteins, with one LipDH located in the apicoplast (LipDH1) and the other in the mitochondria (LipDH). This organelle-specific localisation suggests that there should be KADH located in both organelles and this theory will be further investigated in Chapter 5.

Analyses of the mature forms of both LipDH allowed me to generate constructs that display soluble recombinant expression of both LipDH proteins using the *E. coli* system, although the observed instability of LipDH1 and low expression levels of LipDH2 requires further optimisation.

Analysis of both purified LipDH proteins demonstrated that they were enzymatically active with typical LipDH substrates and displayed the relevant biophysical parameters of LipDH proteins. Although the specific activity of the plastid isoform was observed to be lower than that observed for the mitochondrial isoform, this was consistent with previous studies on the two different isoforms. Further studies are necessary to investigate the basis of the reduced activity in the plastid isoforms.

Chapter 5 The α-keto acid dehydrogenase complexes (KADH) of *Plasmodium falciparum*

5.1 Introduction

5.1.1 α-Keto acid dehydrogenase complexes

There are three α-keto acid dehydrogenase complexes (KADH) in humans that are all located in the mitochondrion. The pyruvate dehydrogenase complex (PDH) acts as a link between glycolysis and the citric acid cycle (TCA cycle), the α-ketoglutarate dehydrogenase complex (KGDH) acts within the TCA cycle and the branched chain α - keto acid dehydrogenase complex (BCKDH) acts in the catabolism of branched chain amino acids such as valine, leucine and isoleucine. The KADH are large multi-enzyme complexes that form from multimers of three independent enzymes.

The first enzyme of the KADH (referred to as E1) is a decarboxylase that catalyses the oxidative decarboxylation of α-keto acids using a thiamine diphosphate cofactor forming a hydroxyl-acyl thiamine diphosphate intermediate. It is this protein that confers the substrate specificity of the complexes and its activity can be regulated by phosphorylation and de-phosphorylation (Randall *et al.*, 1977; Odessey, 1982). The E1 protein can occur as a homodimeric protein of about 100 - 120 kDa or as a hetero-tetrameric protein that consist of two α and two β subunits ($\alpha_2\beta_2$) of between 50 kDa and 70 kDa.

The second enzyme (referred to as E2) is a dihydrolipoamide acyl transferase and forms the catalytic core of the KADH. The protein is organised into three domains, the lipoyl domain, subunit binding domain and catalytic domain that are separated by flexible linker regions. The lipoyl domain contains a lipoic acid moiety covalently attached to an ε-amino group of a conserved lysine residue, which accepts the acyl group transferred from the hydroxyl-acyl thiamine diphosphate of E1. The subunit-binding domain of E2 is responsible for the attachment of E1 and the third component of KADH, a LipDH protein. The catalytic domain is the region of self-association of E2 and is responsible for the generation of the E2 core – the E2 protein forms homo-trimers that associate to form 24 mers (KGDH) or 60 mers (PDH). The catalytic domain also catalyzes the transfer of the acyl group bound to the lipoamide cofactor of the lipoyl domain to coenzyme A, producing acyl CoA and dihydrolipoamide (see figure 5.1 & section 1.4).

Mammalian PDH also contains a fourth component, the so-called protein X that has a high degree of homology to the E2 protein but lacks the catalytic activity (Harris *et al.*, 1997). Protein X is involved in forming the core of the PDH together with the E2 protein and is the site of LipDH association (McCartney *et al.*, 1997).

The third enzyme (referred to as LipDH, dihydrolipoamide dehydrogenase) is responsible for the re-oxidation of dihydrolipoamide producing lipoamide and NADH (figure 5.1). The *P. falciparum* LipDH proteins have been investigated/discussed in depth in chapter 4.

5.1.2 Aims

The aims of this part of the study were to identify the genes encoding the α-keto acid dehydrogenase complexes (KADH) in the *Plasmodium falciparum* genome and determine whether they have a role during the erythrocytic development of the parasite. These studies included investigation of their expression levels, subcellular localisation and preliminary biochemical analyses utilising recombinantly expressed protein.

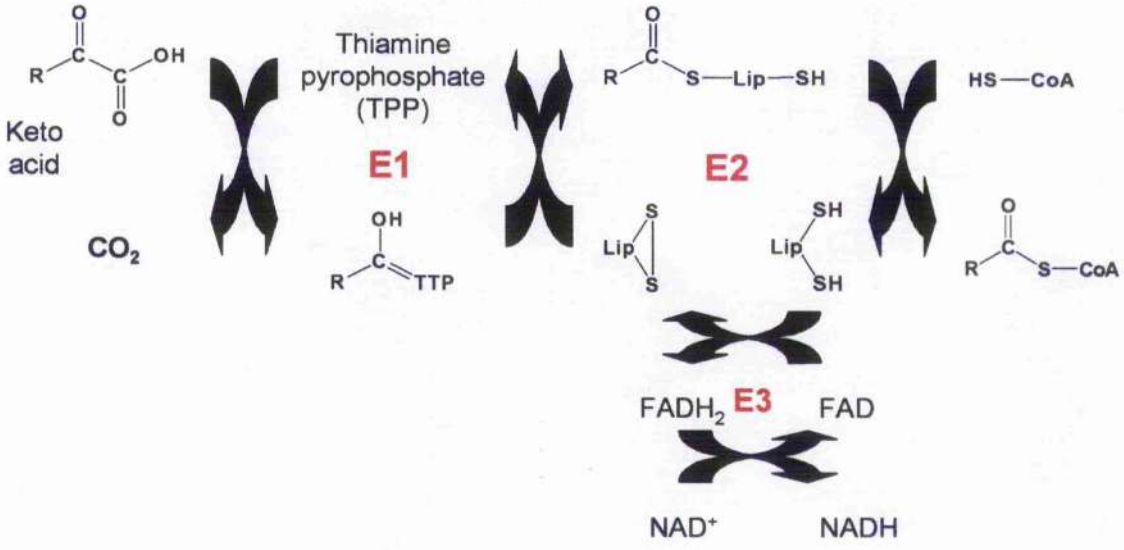


Figure 5.1 – Catalytic mechanism of KADH

The catalytic mechanism of KADH involves the three proteins that form this complex; the α – keto acid dehydrogenase (E1), acyl transferase (E2) and dihydrolipoamide dehydrogenase (E3 or LipDH). The E1 protein is responsible for the oxidative decarboxylation of an α – keto acid using thiamine pyrophosphate as a cofactor. This reaction produces a hydroxyl-acyl thiamine pyrophosphate and CO_2. The acyl group from hydroxyl-acyl thiamine pyrophosphate is transferred onto the lipoamide moiety that is covalently attached to the E2 protein, before being transferred onto coenzyme A producing acyl CoA and dihydrolipoamide. The dihydrolipoamide is oxidised to lipoamide by E3, allowing it to accept another acyl group from the E1 protein, with the reaction of E3 requiring a protein bound flavin cofactor (FAD) and producing NADH.

5.2 Results

5.2.1 Identification of components of KADH in the *P. falciparum* genome

The *P. falciparum* genome database (www.plasmodb.org) was searched by TBLASTP using the human, yeast or *E. coli* amino acid sequences as queries to identify possible homologues in the *P. falciparum* genome. These searches identified that the *P. falciparum* genome contained components of all three KADH:

1. Three genes were identified that encoded components of the α–ketoglutarate dehydrogenase complex (KGDH); an α-ketoglutarate dehydrogenase (KGDH-E1), a succinyl transferase (KGDH-E2) and one of the LipDH described in chapter 4.

2. Four genes were identified that encoded components of the branched chain α – keto acid dehydrogenase complex (BCKDH); the α and β subunits of the BCKDH dehydrogenase (BCKDH-E1α & BCKDH-E1β), a branched chain acyl transferase (BCKDH-E2) and one of the LipDH described in chapter 4.

3. The pyruvate dehydrogenase complex (PDH) was found to be encoded by four genes; the α and β subunits of the PDH dehydrogenase (PDH-E1α and PDH-E1β), the acetyl transferase (PDH-E2) and one of the LipDH described in Chapter 4. The searches failed to identify a homologue for the protein X (O00330) component, suggesting this does not form a part of the parasite PDH.

The following sections detail the sequence analyses of the *Plasmodium* genes and proteins identified in these searches. This does not include the LipDH components of the complexes, as these have already been analysed/discussed in Chapter 4.

5.2.1.1 α–Ketoglutarate dehydrogenase

Searching the *Plasmodium* genome with the α - ketoglutarate dehydrogenase (KGDH-E1) amino acid sequences of *Homo sapiens* (Q02218) and *E. coli* (P07015) identified a *P. falciparum* homologue with 39 % and 38 % identity respectively. The *P. falciparum* gene I identified is located on chromosome 8 and was later annotated PF08_0045 in PlasmoDB (www.plasmodb.org).

Alignment of the *P. falciparum* KGDH- E1 with open reading frames of homologues identified in other *Plasmodium* species (*P. yoelii yoelii*, *P. berghei* and *P. vivax*) demonstrated that the predicted start site for the *P. falciparum* protein was not conserved in the proteins identified in the other species (figure 5.2A). Comparative analyses of the genomic sequences of all the *Plasmodium* genes identified a possible intron/exon boundary that was conserved at the same amino acid position in all *Plasmodium* proteins analysed (highlighted grey in figure 5.2A and red in figure 5.2C). Alignment of the resulting proteins (with the intron removed) produced proteins with conserved sequences and in-frame translation initiation sites (figure 5.2B). The predicted intron (boundaries are highlighted red in figure 5.2C) resulted in a *P. falciparum* KGDH-E1 gene that consisted of a 157 bp exon 1 (highlighted yellow in figure 5.2C; 74.7 % A/T), a 190 bp intron (85.2 % A/T) and a 2958 bp exon 2 (highlighted grey in figure 5.2C; 71.6 % A/T). The 3306 bp α-*ketoglutarate dehydrogenase* gene (which contains the single intron) encoded a protein consisting of 1038 amino acids with a predicted molecular size of 121.2 kDa.

ClustalW alignment with homologues identified in the other *Plasmodium* species *P. berghei*, *P. yoelii yoelii*, *P. chabaudi* confirmed the predicted start site and displayed amino acid identities of 85 %, 84 % and 86 %, respectively (data not shown). Reciprocal BLASTP analyses of the KGDH-E1 amino acid showed that the *P. falciparum* protein had the highest amino acid identity with slime mould (*Dictyostelium discoideum*; 45 %), plant (*Arabidopsis thaliana*; 44 % and *Oryza sativa*; 44 %), fungi (*Aspergillus nidulans*; 43 % and *Neurosporra crassa*; 42 %) and yeast (*Saccharomyces cerivisiae*; 43 %). Comparison of the *P. falciparum* KGDH-E1 with amino acid sequences of KGDH-E1 from other organisms identified the conserved thiamine pyrophosphate binding motif (Koike, 1998) is present in the parasite protein sequence. Further identification of residues important in the catalysis of the KGDH-E1 protein has been limited due to the inherent difficulties in recombinantly expressing this large protein from other sources.

Analysis of the KGDH-E1 amino acid sequence in subcellular prediction programs (see section 2.2.2.3) suggests that the KGDH-E1 is localised to the mitochondrion, with a relatively high level of confidence (74 %), although the programmes did not predict a mitochondrial targeting peptide cleavage site (table 5.2).

A)

```
Pyoel          ----------------------HFLTFFMICTYIFFYWHFLSDLFRNTLDKSWDRYFT 36
Pberg          FKKKNYVRVNEEYTFFCIINLNCSYFLIFFMICTYIFFYYHFLSDLLRNTLDKSWDRYFT 60
Pfal           -----------YLYICTL----LHIYIYIYIYVYIFFFFFFFFLPFRNNLHKSWDSLFS 44
Pviv           --------------------GAPPRDTPDVLLPLLNAALPLPYPLRRNSLHKSWDVYFA 39
                                .             :   :     :     **.*.****  *:

Pyoel          NICNGMPN---NSLLDENSRIIYKNCDREYG--ERHNNLRITYVNEEMIEKGKTGNIYDI 91
Pberg          NICNGMPN---NSLIDENSRVIYKNRDREYG--EKHNNLRITYVNEEMIEKGKTGNIYDI 115
Pfal           MYPHAEMDN-YNNPIKINRKTDNYNNSNCFQDVLKNNNLRITYVNNEMLEKGKTENIYDL 103
Pviv           EMAEEAGPLGSAPPRVLSTREVRSRMGASVPQNRSSSNLRITYVNKEMLDKGRMGNIYDI 99
                  .              . :    . .    .   .********:**::**:  ****:
```

B)

```
Pyoel          ----MQTS---RKNGNNIK-KYIHTSSVLYNEEFNPSMASYLENVYRLWKSDRNTLDKSW 52
Pberg          ----MQTS---RKNGNNIK-KYIHTSSVLYNEEFNPSMASYLESVYRLWKSDRNTLDKSW 52
Pfal           MRKFFYTNNLVRNNGKDIRTRSFHLSSCLYDYNFNPSMTSYIENTYKIWKEDRNNLHKSW 60
Pviv           MRKILHSNKLLKENGDMLR-RRFHVGGCAGSDNFNPSMAAYIEGAYRMWRQDRNSLHKSW 59
                 . : :..  ::**. :::: :* ..   . :*****::*:*..*::*:.***.*.***

Pyoel          DRYFTNICNGMPN---NSLLDENSRIIYKNCDREYG--ERHNNLRITYVNEEMIEKGKTG 107
Pberg          DRYFTNICNGMPN---NSLIDENSRVIYKNRDREYG--EKHNNLRITYVNEEMIEKGKTG 107
Pfal           DSLFSMYPHAEMDN-YNNPIKINRKTDNYNNSNCFQDVLKNNNLRITYVNNEMLEKGKTE 119
Pviv           DVYFAEMAEEAGPLGSAPPRVLSTREVRSRMGASVPQNRSSSNLRITYVNKEMLDKGRMG 119
               *  *:   .              . :    . .    .   .********:**::**:
```

C)

```
DNA: ATGAGAAAATTTTTTTACACAAATAATTTGGTAAGAAATAATGGAAAAGATATAAGAACACGATCATTT
 +3:   E  K  I  F  L  H  K  *  F  G  K  K  *  W  K  R  Y  K  N  T  I  I  S
 +2:  *  E  N  F  F  T  Q  I  I  W  *  E  I  M  E  K  I  *  E  H  D  H  F
 +1: M  R  K  F  F  Y  T  N  N  L  V  R  N  N  G  K  D  I  R  T  R  S  F

DNA: CATTTAAGTTCTTGCTTATATGATTATAATTTTAACCCAAGTATGACCTCGTATATTGAGAATACATAT
 +3:   F  K  F  L  L  I  *  L  *  F  *  P  K  Y  D  L  V  Y  *  E  Y  I  *
 +2:  I  *  V  L  A  Y  M  I  I  I  L  T  Q  V  *  P  R  I  L  R  I  H  I
 +1: H  L  S  S  C  L  Y  D  Y  N  F  N  P  S  M  T  S  Y  I  E  N  T  Y

DNA: AAGATATGGAAAGAAGACAGGTATAATGTGCAACAGGCACAACATAGAAATAAACTTTCATGTATACAT
 +3:   D  M  E  R  R  Q  V  *  C  A  T  G  T  T  *  K  *  T  F  M  Y  T  Y
 +2:  R  Y  G  K  K  T  G  I  M  C  N  R  H  N  I  E  I  N  F  H  V  Y  I
 +1: K  I  W  K  E  D  R  Y  N  V  Q  Q  A  Q  H  R  N  K  L  S  C  I  H

DNA: ATATATATATATATTTATAACTATACATACCTATATTTATATTAATATTTATATATATGTACATTGTTG
 +3:   I  Y  I  Y  L  *  L  Y  I  P  I  F  I  L  I  F  I  Y  M  Y  I  V  A
 +2:  Y  I  Y  I  F  I  T  I  H  T  Y  I  Y  I  N  I  Y  I  Y  V  H  C  C
 +1: I  Y  I  Y  I  Y  N  Y  T  Y  L  Y  L  Y  *  Y  L  Y  I  C  T  L  L

DNA: CATATATATATATATATATATATATATGTATATATTTTTTTTTTTTTTTTTTTTTTTTTTTACCTTTT
 +3:   Y  I  Y  I  Y  I  Y  I  C  I  Y  F  F  F  F  F  F  F  F  F  T  F  *
 +2:  I  Y  I  Y  I  Y  I  Y  M  Y  I  F  F  F  F  F  F  F  F  F  Y  L  L
 +1: H  I  Y  I  Y  I  Y  I  Y  V  Y  I  F  F  F  F  F  F  F  F  L  P  F

DNA: AGAAACAATTTACACAAATCGTGGGACTCTTTATTTAGCATGTATCCACATGCAGAAATGGATAATTAT
 +3:   K  Q  F  T  Q  I  V  G  L  F  I  *  H  V  S  T  C  R  N  G  *  L  *
 +2:  E  T  I  Y  T  N  R  G  T  L  Y  L  A  C  I  H  M  Q  K  W  I  I  I
 +1: R  N  N  L  H  K  S  W  D  S  L  F  S  M  Y  P  H  A  E  M  D  N  Y
```

Figure 5.2 – Determining intron/exon boundaries for KGDH-E1

The intron/exons boundaries of *P. falciparum* KGDH-E1 were determined by ClustalW alignment with homologues identified in the other *Plasmodium* species, *P. yoelii yoelii*, *P. berghei* and *P. vivax*. (*) Identical residues, (:) conserved residues, (.) homologous residues. All alignments display only the N-terminus of the predicted proteins. **Panel A)** ClustalW alignment open reading frames of KGDH-E1 homologues from *Plasmodium* species. The predicted intron/exon boundary is highlighted in grey. **Panel B)** ClustalW alignment of the predicted proteins of the *Plasmodium* KGDH-E1 with the intron removed, with the intron/exon boundary highlighted in grey. **Panel C)** The intron/exon boundaries of the *P. falciparum KGDH-E1* are marked in red with exon 1 highlighted in yellow and exon 2 highlighted in grey.

```
DNA: ATGAGAAAATTTTTTTTACACAAATAATTTGGTAAGAAATAATGGAAAAGATATAAGAACA  60
 +1: M  R  K  F  F  F  Y  T  N  N  L  V  R  N  N  G  K  D  I  R  T   20

DNA: CGATCATTTCATTTAAGTTCTTGCTTATATGATTATAATTTTAACCCAAGTATGACCTCG  120
 +1: R  S  F  H  L  S  S  C  L  Y  D  Y  N  F  N  P  S  M  T  S   40

DNA: TATATTGAGAATACATATAAGATATGGAAAGAAGACAGAAACAATTTACACAAATCGTGG  180
 +1: Y  I  E  N  T  Y  K  I  W  K  E  D  R  N  N  L  H  K  S  W   60

DNA: GACTCTTTATTTAGCATGTATCCACATGCAGAAATGGATAATTATAATAACCCAATAAAA  240
 +1: D  S  L  F  S  M  Y  P  H  A  E  M  D  N  Y  N  N  P  I  K   80

DNA: ATAAATCGAAAGACTGATAATTATAATAATAGTAATTGTTTTCAAGATGTGCTTAAGAAT  300
 +1: I  N  R  K  T  D  N  Y  N  N  S  N  C  F  Q  D  V  L  K  N  100

DNA: AATAATCTTCGAATTACATATGTGAATAATGAAATGCTTGAAAAAGGGAAAACAGAAAAT  360
 +1: N  N  L  R  I  T  Y  V  N  N  E  M  L  E  K  G  K  T  E  N  120

DNA: ATTTATGATCTAGCTAGGATTGTTCAATTAATTAGATGGTATCAAAAGAAAGGACATTTA  420
 +1: I  Y  D  L  A  R  I  V  Q  L  I  R  W  Y  Q  K  K  G  H  L  140

DNA: TATGCTAATATAAATCCTTTGCCATTACCAAAAGAACCTCCCTATAGTAGTGTTTGTTAT  480
 +1: Y  A  N  I  N  P  L  P  L  P  K  E  P  P  Y  S  S  V  C  Y  160

DNA: GAACCTTGTAAAAGGAAAATGTCTTACGTGGATTTTGGTTTTAATGAGGATGATTTAGAT  540
 +1: E  P  C  K  R  K  M  S  Y  V  D  F  G  F  N  E  D  D  L  D  180

DNA: AAAGAGTTTTTTTTTGATTTGCCGTCGATTAGTGGTTTTTCAAGTAATGGTATGAAAAAA  600
 +1: K  E  F  F  F  D  L  P  S  I  S  G  F  S  S  N  G  M  K  K  200

DNA: TGTAATTTAAGAAATTTATTAAAAAGATTAGAAGAAACATATTGTGGTACTATTGGTTTT  660
 +1: C  N  L  R  N  L  L  K  R  L  E  E  T  Y  C  G  T  I  G  F  220

DNA: GAATATATGCATATAACAAATGAAAATATAGTAAATTATATAATACAAAGAATAGAGAAG  720
 +1: E  Y  M  H  I  T  N  E  N  I  V  N  Y  I  I  Q  R  I  E  K  240

DNA: GACAAAAAATATGAATATGATACAAAGATGAAAAAAAGAATTTTAGAATATACTGCTCGA  780
 +1: D  K  K  Y  E  Y  D  T  K  M  K  K  R  I  L  E  Y  T  A  R  260

DNA: GCATTTATATTTGAAAATTATATGGCAGCTAAATTTGCTACAACAAAAAGATTTGGAGTT  840
 +1: A  F  I  F  E  N  Y  M  A  A  K  F  A  T  T  K  R  F  G  V  280

DNA: GATGGATGTGAAACATTAATTACTGGTATGAAAGCATTAATAAAAAGGGCTGCTCAGTTA  900
 +1: D  G  C  E  T  L  I  T  G  M  K  A  L  I  K  R  A  A  Q  L  300

DNA: GATGTAGATAGTGTATTGATGAGTATGTCTCATCGAGGTCGATTGAATGTTTTATTTAAT  960
 +1: D  V  D  S  V  L  M  S  M  S  H  R  G  R  L  N  V  L  F  N  320

DNA: GTATTACATAAACCATTAGAACAAATGATGTCAGAATTTCGAGGGAAAACTGGTTTTAGT 1020
 +1: V  L  H  K  P  L  E  Q  M  M  S  E  F  R  G  K  T  G  F  S  340

DNA: GATAATATATGGGGAAATACCGGAGATGTAAAATACCATTTAGGTGTTGAAATTGATTAT 1080
 +1: D  N  I  W  G  N  T  G  D  V  K  Y  H  L  G  V  E  I  D  Y  360

DNA: TATGATGAAGATTCACAAAGATATATACATATGGGTATAGTAGACAATTCGTCTCATTTA 1140
 +1: Y  D  E  D  S  Q  R  Y  I  H  M  G  I  V  D  N  S  S  H  L  380

DNA: GAATCTGTTGATCCTATTTTAATGGGTCAAGCTAGAGCACAACAATATTATTGTAATGAT 1200
 +1: E  S  V  D  P  I  L  M  G  Q  A  R  A  Q  Q  Y  Y  C  N  D  400

DNA: AAAGAAAAAAAGAAAGTGTTACCAATAACTATACATGGAGATGCTTCAATCGCTGGACAA 1260
 +1: K  E  K  K  K  V  L  P  I  T  I  H  G  D  A  S  I  A  G  Q  420

DNA: GGAATTGCATATGAAACATTTCAAATGTCTAAATTACCAAGTTATAATGTGGGTGGTACT 1320
 +1: G  I  A  Y  E  T  F  Q  M  S  K  L  P  S  Y  N  V  G  G  T  440

DNA: ATACATATTGTGGTTAATAATCAAATAGGATTTACCACATATCCAATAGATGCAAGGTCA 1380
 +1: I  H  I  V  V  N  N  Q  I  G  F  T  T  Y  P  I  D  A  R  S  460

DNA: GGAAAATATTGTACAGACATAGCTAAATGTATAGATATACCAATTATACATGTAAATGCA 1440
 +1: G  K  Y  C  T  D  I  A  K  C  I  D  I  P  I  I  H  V  N  A  480

DNA: GATGATCCAGAAGCTGTAACATATGTATTTGAATTAGCATTAGATATTCGAAATAAATTT 1500
 +1: D  D  P  E  A  V  T  Y  V  F  E  L  A  L  D  I  R  N  K  F  500

DNA: AATATTGATACAATTATTGATATTGTAGGTTATAGGAGATTTGGGCATAACGAATTAGAT 1560
 +1: N  I  D  T  I  I  D  I  V  G  Y  R  R  F  G  H  N  E  L  D  520

DNA: ATGCCAAAATTTACGAATCCATTATTATATGATGTAATAGCTAGACATAAATCTGTTTTA 1620
 +1: M  P  K  F  T  N  P  L  L  Y  D  V  I  A  R  H  K  S  V  L  540
```

```
DNA: GATATATATAGTAAAAAATTAATTGATGAAAATATTATTAGTCTTAAAGAATTTGAAGAT 1680
+1: D  I  Y  S  K  K  L  I  D  E  N  I  I  S  L  K  E  F  E  D   560

DNA: AATAAAACTGAAATATTTAATTTTTATGAACAAGTGTATGAGAAATCCAAATCATTTGTA 1740
+1: N  K  T  E  I  F  N  F  Y  E  Q  V  Y  E  K  S  K  S  F  V   580

DNA: CCAACACCTAAGGAAAAATATTTACCTCAATGGGAACATATGGTAACACCCCAAAAATTT 1800
+1: P  T  P  K  E  K  Y  L  P  Q  W  E  E  M  V  T  P  Q  K  F   600

DNA: TCTCCCTCACGAAAAACAGGAGTAGAAAAAGATGTATTAATAAATCTTGGGAAAAAGATA 1860
+1: S  P  S  R  K  T  G  V  E  K  D  V  L  I  N  L  G  K  K  I   620

DNA: TTTACATTACGTGAAAATTTCACTGCTCATCCAATAATAACAAAATTATTTAAAAGTAGA 1920
+1: F  T  L  R  E  N  F  T  A  H  P  I  I  T  K  L  F  K  S  R   640

DNA: ATAGATAGTTTAGAAACAGGCAAAAATATAGATTTTGGTACCGCTGAATTATTAGCTTAT 1980
+1: I  D  S  L  E  T  G  K  N  I  D  F  G  T  A  E  L  L  A  Y   660

DNA: GCAACTTTATTATCTGATGGATTCCATGCACGTTTATCTGGTCAGGATTCACAAAGAGGA 2040
+1: A  T  L  L  S  D  G  F  H  A  R  L  S  G  Q  D  S  Q  R  G   680

DNA: ACTTTTTCTCATCGACACGCAGTATTACATGATCAAATAACATATGAATCATATAATATA 2100
+1: T  F  S  H  R  H  A  V  L  H  D  Q  I  T  Y  E  S  Y  N  I   700

DNA: TTTGATTCATTAAAAACACCACATACCATAGAAGTAAATAATTCTCTTTTATCTGAATAC 2160
+1: F  D  S  L  K  T  P  H  T  I  E  V  N  N  S  L  L  S  E  Y   720

DNA: GCATGTTTAGGTTATGAGATAGGATATAGTTATGAACATCCAGATGCTCTTGTTATTTGG 2220
+1: A  C  L  G  Y  E  I  G  Y  S  Y  E  H  P  D  A  L  V  I  W   740

DNA: GAAGCTCAATTTGGTGATTTTGCTAATGGAGCTCAGGTTATGATTGACAATTATATTGCA 2280
+1: E  A  Q  F  G  D  F  A  N  G  A  Q  V  M  I  D  N  Y  I  A   760

DNA: TCTGGTGAAACAAAATGGAATAAACAATCGGCTATTGTTATGTTATTACCTCATGGATAT 2340
+1: S  G  E  T  K  W  N  K  Q  S  G  I  V  M  L  L  P  H  G  Y   780

DNA: GATGGACAAGGTCCTGAGCATTCATCTGCTCGGATAGAACGTTTTTTACAATTGTGTGAT 2400
+1: D  G  Q  G  P  E  H  S  S  A  R  I  E  R  F  L  Q  L  C  D   800

DNA: GATAGAGAAGATATAGCTACCTATTCTGTTGAAAAAGATAATAAGATTATTCAGCAGCAT 2460
+1: D  R  E  D  I  A  T  Y  S  V  E  K  D  N  K  I  I  Q  Q  H   820

DNA: AATATGCAGGTTATAAATTGTAGTAAACCATCTAATTTTTTCCATGCATTAAGAAGACAA 2520
+1: N  M  Q  V  I  N  C  S  K  P  S  N  F  F  H  A  L  R  R  Q   840

DNA: ATGCATAGATCTTTTAGGAAACCATTAATTGTAATAACACCTAAGAAAATGTTGAAAATG 2580
+1: M  H  R  S  F  R  K  P  L  I  V  I  T  P  K  K  M  L  K  M   860

DNA: AGAATGGCATTTGATAAGATTGAAAATTTTTTAACGAACACTGAATTTTTACCTTATTTA 2640
+1: R  M  A  F  D  K  I  E  N  F  L  T  N  T  E  F  L  P  Y  L   880

DNA: CCAGAAGAAGTTGGTCATAAATTGAAAGATAAAAAAGAAATCAAAAGAATTATATTATGT 2700
+1: P  E  E  V  G  H  K  L  K  D  K  K  E  I  K  R  I  I  L  C   900

DNA: TCAGGACAAGTACATTATGATTTATTAAATTATAGATATACAAATAAAATTGATGATGTA 2760
+1: S  G  Q  V  Y  Y  D  L  L  N  Y  R  Y  T  N  K  I  D  D  V   920

DNA: GCTATTGCCAGAATAGAACAATTATCTCCTTTCCCCTTTAAACAAATAATGAATGATCTG 2820
+1: A  I  A  R  I  E  Q  L  S  P  F  P  F  K  Q  I  M  N  D  L   940

DNA: CAAACTTATCCAAATTTGAGAGATATTATATGGGCACAAGAAGAACATATGAATATGGGT 2880
+1: Q  T  Y  P  N  L  R  D  I  I  W  A  Q  E  E  H  M  N  M  G   960

DNA: CCATGGTTTTATGTCTCTCGTCGTATTGAAGCATCTATAAAACAATTAAAAAAAGATAAA 2940
+1: P  W  F  Y  V  S  R  R  I  E  A  S  I  K  Q  L  K  K  D  K   980

DNA: CCAAAATGGAATATTCAAATTCCTCAGGTACGTTATTCAGGAAGGGATGTATATGCAGCG 3000
+1: P  K  W  N  I  Q  I  P  Q  V  R  Y  S  G  R  D  V  Y  A  A  1000

DNA: CAATCAGCAGGTGATCTTAATTTACATCTTTATCAATTAGATGAATTTTTAGTTGATGCA 3160
+1: Q  S  A  G  D  L  N  L  H  L  Y  Q  L  D  E  F  L  V  D  A  1020

DNA: TTTAATTTGGATAAAAAATATAACATGCATGTTCAAAAATATACCGATGCGTTATAA     3217
+1: F  N  L  D  K  K  Y  N  M  H  V  Q  K  Y  T  D  A  L  *      1039
```

Figure 5.3 – *Plasmodium falciparum* KGDH-E1

The complementary DNA sequence (with the intron removed) is displayed with the resulting amino acid sequence shown below it. The conserved domain involved in thiamine pyrophosphate is highlighted in grey.

5.2.1.2 Succinyl transferase

The deduced amino acid sequence of the open reading frame of *P. falciparum succinyl transferase (KGDH-E2)* gene displayed 42 % and 39 % identity to the human (P36957) and *E. coli* (P07016) proteins, respectively; I named this gene *KGDH-E2A*. However, ClustalW alignment of the predicted protein with open reading frames identified in other *Plasmodium* species (*P. berghei*, *P. knowlesi*, *P. gallinecium* and *P. vivax*) identified the presence of two possible introns in the 5'region of the gene (see below). The *P. falciparum KGDH-E2* gene is located on chromosome 13 and was later annotated PF13_0121 in PlasmoDB (www.plasmodb.org).

Intron 2 was predicted because the initiator methionine identified in the *P. falciparum* KGDH-E2A protein (highlighted green in figure 5.4A) was not conserved in the predicted proteins from the other *Plasmodium* species. Analyses of the genomic sequence of *KGDH-E2* from *P. falciparum* and other *Plasmodium* species identified the possible intron/exon boundaries (highlighted grey in figure 5.4A and B, and red in figure 5.4D), with their position being conserved in the sequences of all the *Plasmodium* species investigated.

Intron 1 was identified in the genomic sequences of the different *Plasmodium KGDH-E2* genes because there was no consensus in-frame methionine initiation codon within the predicted intron 2 (highlighted red in figure 5.4B & C). Further analysis of the genomic sequences of *Plasmodium KGDH-E2* identified the possible intron/exon boundaries (highlighted red in figure 5.4D), with their position being conserved in the sequences of all the *Plasmodium* species analysed (highlighted red in figure 5.4B and C).

As demonstrated previously in *LipDH2B* (see section 4.2.1.2) the A/T content of intron sequences was higher than that observed in coding regions. This a common feature of the *Plasmodium* genome sequences as has been shown by whole genome analyses by Gardner *et al.* (Gardner *et al.*, 1999; Szafranski *et al.*, 2005). In the *P. falciparum KGDH-E2B* sequence exon 1 was 30 bp long and had an A/T content of 76.7 % (highlighted yellow in figure 5.4D), intron 1 was 142 bp in length with an A/T content of 88 %, exon 2 was 106 bp long and had an A/T content of 76.4 % (highlighted blue in figure 5.4D), intron 2 was 124 bp with an A/T content of 82.7 %, and exon 3 was 1130 bp in length and had an A/T content of 72.4 % (highlighted grey in figure 5.4D).

Removal of both introns resulted in *Plasmodium* deduced proteins that contained a conserved in-frame methionine (data not shown); I named this gene *KGDH-E2B* (figure 5.4C and 5.5). ClustalW alignment with the proteins (with the introns spliced out) identified in the other *Plasmodium* species *P. berghei*, *P. yoelii yoelii*, *P. chabaudi* displayed amino acid identities of 73 %, 73 % and 72 %, respectively. BLASTP analyses of the deduced *P. falciparum* KGDH-E2 amino acid sequence displayed greatest identity to homologous proteins from other apicomplexa (such as *Theileria parva*; 50 % identity), slime mould (*Dictyostelium discoideum*; 46 % identity), yeast (*S. cerevisiae*; 45 % identity) and fungi (*Aspergillus nidulans*; 45 % identity).

The *P. falciparum succinyl transferase* gene (*KGDH-E2B*) was encoded by 1532 bp (including the two predicted introns). The predicted full-length protein was 421 amino acids in length with a predicted molecular mass of 47.6 kDa. The KGDH-E2B protein can be divided into three domains that are separated by flexible linker regions; the lipoyl domain, subunit- binding domain and the catalytic/inner core domain. The lipoyl domain runs from residues 46 to 125 and contains the conserved lysine residue (residue 86 in the *P. falciparum* protein; highlighted red in figure 5.5) that becomes lipoylated via an amide linkage. Another residue that is essential for the lipoylation reaction is glycine 97 (in *P. falciparum* protein), which is conserved between the organisms (highlighted yellow in figure 5.5). The subunit-binding domain is located between residues 139 and 172 in the *P. falciparum* protein and is responsible for the binding of E1and LipDH. The catalytic domain of the *P. falciparum* protein runs between residues 197 and 421 and is the site of self-association and acylation of coenzyme A. Residues His392 and Ser340 (highlighted green in figure 5.5) have been implicated in catalysis of the acylation reaction. His392 (highlighted blue in figure 5.5) is thought to act as a base catalyst and is located in a conserved motif (DHRXXDG).

Analysis of the two possible KGDH-E2 amino acid sequences by subcellular localisation predictions programs resulted in very different results (table 5.2). The KGDH-E2A protein (which uses the in-frame methionine within the identified intron 2) was predicted to be mitochondrial, although with a very low level of confidence of 3.7 %. The KGDH-E2B protein, which uses the methionine identified in exon 1 with both predicted introns removed, was suggested to be mitochondrial with an 85.2% confidence level, although the mitochondrial targeting peptide cleavage site was not determined by the algorithms used. This difference in the subcellular localisation prediction data suggests that KGDH-E2B is the more likely gene product, although this will be further investigated later in this chapter (see sections 5.2.2.1 & 5.2.4).

Alignment of the *P. falciparum* KGDH-E2 with its *E. coli* homologue displays amino acid identity from amino acid 44 (highlighted grey in figure 5.5), which suggests that the mature protein size (minus the mitochondrial targeting peptide) will be approximately 42.5 kDa.

A)

```
Pberg          -------------------------YVLLCVY---RYFAIYYYIFIYADTLKVPRLGDSI 32
Pgall          ------------------IFIIFRRCLILLIYNILFYIILFYYFFFFLDTIKVPRLGDSI 42
Pvivax         ---------------------------------PLHMHLIAYSHPSNADTIKVPRLGDSI 27
Pknow          CHYRNKKRDSPAPCLHLIWTRACMFVGHQCSFLPLFMYLIALSPSLNAETIKVPRLGDSI 60
PfalA          ----------------------------MCVCIFFFFFLLPF-----VETIKVPRLGDSI 27
                                                  :        :*:*********

Pberg          TEGVINEWKKKVGDYVYSDETLAVIDTDKVSVDINSKSSGALYKIFAEAGDTVLVDSPLC 92
Pgall          TEGTINEWKKKVGDYVKNDETLAIIDTDKVSVDINSKSSGTLHKIFAEAGDIILVDSPLC 102
Pvivax         TEGTISEWKKKVGDYVKVDETITIIDTDKVSVDINSKSSGALSKIFAEAGDIVLVDAPLC 87
Pknow          TEGTISEWKKKVGDYVKMDETITIIDTDKVSVDINSQFSGELSNIFAEAGEIVLVDAPLC 120
PfalA          TEGTINEWKKKVGDYVKADETITIIDTDKVSVDINSKVSGGLSKIFADVGDVVLVDAPLC 87
               ***.*.**********  ***:::************: ** * :***:.*: :***:***
```

B)

```
Pfal           KSLLHHVKSFNRTFLNHKYNSNANIEGSLKRYFSIETIKVPRLGDSITEGTINEWKKKVG 60
Pberg          KQLIYNGNNVKRAFFNVEFRQLVNNYITCKRHFSIDTLKVPRLGDSITEGVINEWKKKVG 60
Pgall          KSSLYNLNSVKRTFFSLKSRQNEKYRFVFRRYFSIDTIKVPRLGDSITEGTINEWKKKVG 60
pvivax         KPLFYNAKGISR-ILNLECKQCVNCHSTFKRCFSIDTIKVPRLGDSITEGTISEWKKKVG 59
Pknow          KPLFYNAKGISR-ILNLECKQSVKCHSKFKRCFSIETIKVPRLGDSITEGTISEWKKKVG 59
               *  ::: :...*:::. : ..  :     :* ***:*:************.*.*******
```

C)

```
Pvivax         MTQNLVLRLNKPLFYNAKGISR-ILNLECKQCVNCHSTFKRCFSIDTIKVPRLGDSITEG 59
Pknow          MTQNLVLRLNKPLFYNAKGISR-ILNLECKQSVKCHSKFKRCFSIETIKVPRLGDSITEG 59
Pfal           MTKNLVFRLNKSLLHHVKSFNRTFLNHKYNSNANIEGSLKRYFSIETIKVPRLGDSITEG 60
Pgall          MTQNLILRLNKSSLYNLNSVKRTFFSLKSRQNEKYRFVFRRYFSIDTIKVPRLGDSITEG 60
Pberg          MNQTLINRLNKQLIYNGNNVKRAFFNVEFRQLVNNYITCKRHFSIDTLKVPRLGDSITEG 60
               *.:.*: ****  ::: :...*:::. : ..  :     :* ***:*:************

Pvivax         TISEWKKKVGDYVKVDETITIIDTDKVSVDINSKSSGALSKIFAEAGDIVLVDAPLCEID 119
Pknow          TISEWKKKVGDYVKMDETITIIDTDKVSVDINSQFSGELSNIFAEAGEIVLVDAPLCEID 119
Pfal           TINEWKKKVGDYVKADETITIIDTDKVSVDINSKVSGGLSKIFADVGDVVLVDAPLCEID 120
Pgall          TINEWKKKVGDYVKNDETLAIIDTDKVSVDINSKSSGTLHKIFAEAGDIILVDSPLCEID 120
Pberg          VINEWKKKVGDYVYSDETLAVIDTDKVSVDINSKSSGALYKIFAEAGDTVLVDSPLCEID 120
               .*.**********  ***:::************: ** * :***:.*: :***:******
```

D)

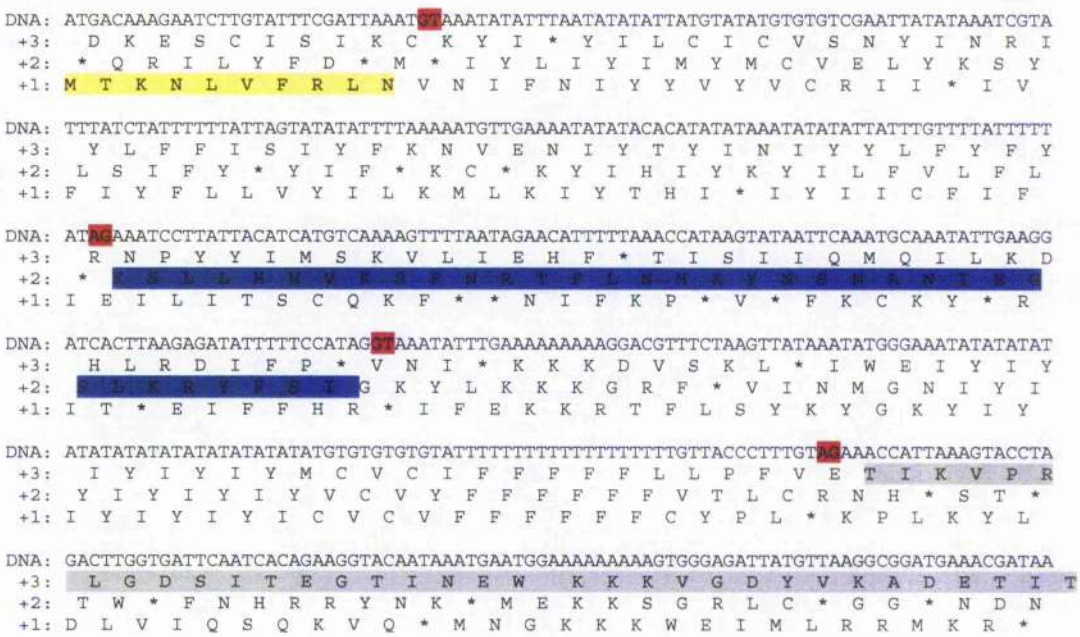


Figure 5.4 – Determining the intron/exon boundaries of KGDH-E2

N-terminal sections of ClustalW alignments of the *Plasmodium falciparum* KGDH-E2 with homologues identified in other *Plasmodium* species *P. vivax* (Pvivax), *P. knowlesi* (Pknow), *P. falciparum* (Pfal), *P. gallinecium* (Pgall) and *P. berghei* (Pberg). The alignment displays only the N-terminal sections and shows identical residues (*), conserved residues (:) and homologous residues (.). **Panel A**) The presence of an intron was predicted from alignment of the KGDH-E2A amino acid sequence open reading frames identified in other *Plasmodium* species. This identified that the in-frame methionine (marked in green) was not conserved in all species. The predicted intron/exon boundary is highlighted in grey. **Panel B**) ClustalW alignment of the open reading frames that resulted from the removal of the intron (intron 2) identified in panel A displays that there must be another intron within this gene. This was because the exon did not contain an in-frame methionine in any of the *Plasmodium* genes. The predicted intron/exon boundary (for intron 1) is marked in red. **Panel C**) ClustalW alignment of *Plasmodium* KGDH-E2 sequences with both introns removed; the intron 1 boundary is marked in red and the intron 2 boundary is marked in grey. **Panel D**) The genomic DNA sequence from the KGDH-E2 gene highlights the intron/exon boundaries (red), exon 1 (yellow), exon 2 (blue) and exon 3 (continues for the rest of the amino acid sequence; highlighted in grey).

```
DNA: ATGACAAAGAATCTTGTATTTCGATTAAATAAATCCTTATTACATCATGTCAAAAGTTTT      60
 +1: M  T  K  N  L  V  F  R  L  N  K  S  L  L  H  H  V  K  S  F       20

DNA: AATAGAACATTTTTAAACCATAAGTATAATTCAAATGCAAATATTGAAGGATCACTTAAG     120
 +1: N  R  T  F  L  N  H  K  Y  N  S  N  A  N  I  E  G  S  L  K       40

DNA: AGATATTTTTCCATAGAAACCATTAAAGTACCTAGACTTGGTGATTCAATCACAGAAGGT     180
 +1: R  Y  F  S  I  E  T  I  K  V  P  R  L  G  D  S  I  T  E  G       60

DNA: ACAATAAATGAATGGAAAAAAAAAAGTGGGAGATTATGTTAAGGCGGATGAAACGATAACA     240
 +1: T  I  N  E  W  K  K  K  V  G  D  Y  V  K  A  D  E  T  I  T       80

DNA: ATTATCGATACAGATAAAGTTAGTGTTGATATAAATTCTAAAGTAAGTGGAGGGTTATCC     300
 +1: I  I  D  T  D  K  V  S  V  D  I  N  S  K  V  S  G  G  L  S      100

DNA: AAAATTTTTGCGGATGTAGGAGATGTTGTTTTGGTAGATGCACCTTTATGTGAAATTGAT     360
 +1: K  I  F  A  D  V  G  D  V  V  L  V  D  A  P  L  C  E  I  D      120

DNA: ACATCTGTAGAACCACCTGAAGATATTTGTAAAACTAAAGAGGAAGTAGGAGAAAGTAAA     420
 +1: T  S  V  E  P  P  E  D  I  C  K  T  K  E  E  V  G  E  S  K      140

DNA: AATAATGAAAATAATTATACATTTAATCAATTAAATAGAGATATTAAAGATGAAGCACAT     480
 +1: N  N  E  N  N  Y  T  F  N  Q  L  N  R  D  I  K  D  E  A  H      160

DNA: ATAAAAGATGAAGTTAGTAAAAATGAAAAAGATATTTTTGTAAAAGATCCTATATGTTTT     540
 +1: I  K  D  E  V  S  K  N  E  K  D  I  F  V  K  D  P  I  C  F      180

DNA: GGAAATGATTATGAATCAATAAATGAAAGAACAGAAAGAAGAGTCCGTATGTTACCTATA     600
 +1: G  N  D  Y  E  S  I  N  E  R  T  E  R  R  V  R  M  L  P  I      200

DNA: AGAAAAAGAATTGCAGAACGATTAAAGGAATCACAAAATACATGTGCTTTATTAACAACA     660
 +1: R  K  R  I  A  E  R  L  K  E  S  Q  N  T  C  A  L  L  T  T      220

DNA: TTTAATGAATGTGATATGTCAAAGGCAATGTTATTAAGAAGTGAATTAAATGATATATTT     720
 +1: F  N  E  C  D  M  S  K  A  M  L  L  R  S  E  L  N  D  I  F      240

DNA: CAAAAAAAATATTCATGTAAATTAGGATTTGTATCATTATTTATGTATGCATCAACATTG     780
 +1: Q  K  K  Y  S  C  K  L  G  F  V  S  L  F  M  Y  A  S  T  L      260

DNA: GCTTTAAAAAAAATGCCAAATGTAAATGCATATATAGAAAATGATGAAATTGTATATAAA     840
 +1: A  L  K  K  M  P  N  V  N  A  Y  I  E  N  D  E  I  V  Y  K      280

DNA: AATTATATAGATATATCTGTAGCAGTAGCTACACCAAATGGCTTAACCGTTCCAGTTATA     900
 +1: N  Y  I  D  I  S  V  A  V  A  T  P  N  G  L  T  V  P  V  I      300

DNA: AGAAATTGTCAAAATAAAAATTTACCACAATTAGAATTAGCTTTATCTGATTTAGCTACA     960
 +1: R  N  C  Q  N  K  N  L  P  Q  L  E  L  A  L  S  D  L  A  T      320

DNA: AAAGCACGATCAAATAAATTATCGATTGATGATTTTAGTGGTGGTACATTTACAATATCT    1020
 +1: K  A  R  S  N  K  L  S  I  D  D  F  S  G  G  T  F  T  I  S      340

DNA: AATGGAGGTGTTTTTGGTAGTATGTTAAGTACTCCTATTATTAATATGCCACAGTCAGCT    1080
 +1: N  G  G  V  F  G  S  M  L  S  T  P  I  I  N  M  P  Q  S  A      360

DNA: ATTCTAGGAATGCATACAATAAAAAATAGACCAGTAGTTGTCAATAATGAAATTGTTATT    1140
 +1: I  L  G  M  H  T  I  K  N  R  P  V  V  V  N  N  E  I  V  I      380

DNA: AGACCAATTATGTACCTTGCCCTTACATATGATCATAGATTATTAGACGGAAGAGAAGCC    1200
 +1: R  P  I  M  Y  L  A  L  T  Y  D  H  R  L  L  D  G  R  E  A      400

DNA: GTACAATTTTTATGTGCAATAAGAGATTATATAGAAAACCCTAATCTAATGTTAATTGAT    1260
 +1: V  Q  F  L  C  A  I  R  D  Y  I  E  N  P  N  L  M  L  I  D      420

DNA: TGTTAA                                                          1266
 +1: C  *                                                            422
```

Figure 5.5 – *Plasmodium falciparum* KGDH-E2B

The complementary DNA sequence (with both introns removed) is displayed with the resulting amino acid sequence shown below. The initiator methionine is included in the numbering and the termination codon is marked with an asterisk (*). The predicted site of mitochondrial targeting peptide cleavage (S44) is highlighted in grey. The last amino acid of the N-terminal sections used to produce the GFP-fusion proteins is marked in pink, with the KGDH-E2A construct using the residue equivalent to valine 195 and both KGDH-E2B constructs (C & G) using alanine 114. The conserved lysine residue (K86) that is the site of lipoylation is highlighted in red, with another residue (G97) that is essential for lipoylation to occur highlighted in yellow. Residues involved in the acylation reaction in the catalytic domain are highlighted blue (H392) and green (S340).

5.2.1.3 Branched chain α - keto acid dehydrogenase alpha subunit

The branched chain α – keto acid dehydrogenase alpha subunit (BCKDH E1α) displayed 49 % amino acid sequence identity with the human protein (012694). The *P. falciparum* gene I identified was located on chromosome 13 and was later annotated PF13_0070 in PlasmoDB (www.plasmodb.org). ClustalW alignment with homologues identified in the other *Plasmodium* species *P. berghei*, *P. yoelii yoelii*, *P. chabaudi* confirmed the predicted start site of the open reading frame and the deduced amino acid sequences displayed identities of 73 %, 65 % and 66 %, respectively (data not shown). Therefore it was assumed that the *P. falciparum branched chain α - keto acid dehydrogenase alpha* gene contained a single open reading frame of 1290 bp, which encoded a protein consisting of 430 amino acids with a predicted molecular mass of 49.9 kDa. BLASTP analyses of the predicted amino acid sequence displayed greatest amino acid identity to homologous proteins from other apicomplexa such as *Theileria parva* (52 %) and fungi such as *Yarrowia lipolytica* (44 %) and *Neurospora crassa* (45 %).

Comparison of the BCKDH-E1α amino acid sequence with those previously studied from other organisms has identified a number of important residues involved in protein function (Fisher *et al.*, 1991; Denoya *et al.*, 1995; Aevarsson *et al.*, 1999; Ono *et al.*, 2001). Residues responsible for the binding of the thiamine pyrophosphate co-factor are present in the *Plasmodium* protein in a conserved motif ($GDGX_{28}NN$; highlighted grey in figure 5.6). A conserved leucine residue (Leu 191; highlighted red in figure 5.6) is involved in orientation of the TPP co-factor into a V-conformation, while a conserved histidine residue (His 318; highlighted green in figure 5.6) acts as a proton donor to one of the sulphur atoms of the lipoyl-arm of BCKDH-E2 (while a conserved histidine in BCKDH-E1β acts as a proton donor to the other; see section 5.2.4). Of the two phosphorylation sites present in mammalian BCKDH-E1α, only the first is conserved in the *P. falciparum* amino acid sequence (Ser 319; highlighted yellow in figure 5.6) with the other being replaced by a proline (Pro 329). A conserved tyrosine residue (Y422; highlighted blue in figure 5.6) has been shown to be essential for formation of the $\alpha_2\beta_2$ hetero-tetramer in human BCKDH (Fisher *et al.*, 1991).

Analysis of the BCKDH-E1α with subcellular prediction programs identified that the protein was likely to be mitochondrially localised with high levels of confidence (99.4 %) and a predicted mitochondrial targeting peptide cleavage site at amino acid 21

(highlighted pink in figure 5.6; table 5.2); this would result in a predicted mature protein size of 47.5 kDa.

```
DNA: ATGAGAAATATTGTTCAGAAATACTTACAAAGGAATAGCACAAAGTTATTTAACAGAACA      60
 +1: M  R  N  I  V  Q  K  Y  L  Q  R  N  S  T  K  L  F  N  R  T      20

DNA: AGTGTATTCAATTTATATAAAAAATGTAATTTTTCTGGATACAAAATTTACAGTGACGGA     120
 +1: S  V  F  N  L  Y  K  K  C  N  F  S  G  Y  K  I  Y  S  D  G      40

DNA: CTAGTTCACTCTGAATTTTCTACTGAGTTAAAAACTGTAAATGAAGTTATAAAAATGCCT     180
 +1: L  V  H  S  E  F  S  T  E  L  K  T  V  N  E  V  I  K  M  P      60

DNA: ATTTATAGAATATTGGATACGAATGGACATTTATTAGACGGTCATGAAGCGCCATTTAAG     240
 +1: I  Y  R  I  L  D  T  N  G  H  L  L  D  G  H  E  A  P  F  K      80

DNA: GATGAAGAAGTATTAAAAATTTACAAGGATATGGTTGAATTTTCTATATGGGATGAAATA     300
 +1: D  E  E  V  L  K  I  Y  K  D  M  V  E  F  S  I  W  D  E  I     100

DNA: TTTTATGGAATACAAAGACAAGGTAGAATATCATTTTATATAGTTAATGAAGGAGAAGAA     360
 +1: F  Y  G  I  Q  R  Q  G  R  I  S  F  Y  I  V  N  E  G  E  E     120

DNA: GGTTTACAATTTGGTATGGGTAAAGCATTAAGTGTTGATGATCATTTATATTGTCAATAT     420
 +1: G  L  Q  F  G  M  G  K  A  L  S  V  D  D  H  L  Y  C  Q  Y     140

DNA: AGAGAAACAGGTGTTTTATTATCCAGAGGTTTTACATACACTGATATATTAAATCAATTA     480
 +1: R  E  T  G  V  L  L  S  R  G  F  T  Y  T  D  I  L  N  Q  L     160

DNA: TTCGGTACCAAATATGATGAAGGTAAAGGTAGACAAATGTGTATATGCTATACTAAAAAA     540
 +1: F  G  T  K  Y  D  E  G  K  G  R  Q  M  C  I  C  Y  T  K  K     180

DNA: GATTTAAATATTCATACTATTACTACACCATTAGGATCTCAATTATCACATGCTGCCGGT     600
 +1: D  L  N  I  H  T  I  T  T  P  R  G  S  Q  L  S  H  A  A  G     200

DNA: TGTGGTTATGCATTAAAACTAAAAAATCAAAAAGCTGTTGCTGTTACTTATTGTGGTGAT     660
 +1: C  G  Y  A  L  K  L  K  N  Q  K  A  V  A  V  T  Y  C  G  D     220

DNA: GGTTCTTCATCAGAAGGAGATTTTTATGCTGCTCTAAATTTTGCTTCTGTAAGACAATCA     720
 +1: G  S  S  S  E  G  D  F  Y  A  A  L  N  F  A  S  V  R  Q  S     240

DNA: CAAACCATGTTTGTATGCAAAAATAATTTGTATGCTATATCCACATCTATTAAAGATCAA     780
 +1: Q  T  M  F  V  C  K  N  N  L  Y  A  I  S  T  S  I  K  D  Q     260

DNA: TATAGAGGTGATGGTATTGCACCAAGAGCATTAGCCCTAGGAATAGAATCCATAAGAGTA     840
 +1: Y  R  G  D  G  I  A  P  R  A  L  A  L  G  I  E  S  I  R  V     280

DNA: GATGGAAATGATTTATTCGCAAGTTATCTAGCTACAAAAAAATTAAGAGATATTTGTATT     900
 +1: D  G  N  D  L  F  A  S  Y  L  A  T  K  K  L  R  D  I  C  I     300

DNA: CAAGAATCAAAACCAGTTTTTATTGAATTCATGTCCTATAGATATGGTCACCATAGTACT     960
 +1: Q  E  S  K  P  V  F  I  E  F  M  S  Y  R  Y  G  H  E  S  T     320

DNA: TCTGATGATTCTAGTTTGTACAGACCAAAAGAAGAAAATGAAGCTTGGAGACAAGAAGGA    1020
 +1: S  D  D  S  S  L  Y  R  P  K  E  E  N  E  A  W  R  Q  E  G     340

DNA: GTACACCCAATTAGTAGAATCTTTTTATATCTCAAGAATAAAAACTTATATAGTGAAAAA    1080
 +1: V  H  P  I  S  R  I  F  L  Y  L  K  N  K  N  L  Y  S  E  K     360

DNA: GAAGATCAAGAACACAGAAAAAGTGTCAAAGAAAATGTATTAAAAGAATTAAAAAAATAT    1140
 +1: E  D  Q  E  H  R  K  S  V  K  E  N  V  L  K  E  L  K  K  Y     380

DNA: GAAAGTGTTAAAAGATATAATATTGTTGGTGGATTATTTGAAGATGTATATCATGAAGAA    1200
 +1: E  S  V  K  R  Y  N  I  V  G  G  L  F  E  D  V  Y  H  E  E     400

DNA: GACTGGAACATTAAAGAACAAAGAGAAAACTTTGAGCAGTTTTTCAAAGAAAATAAACAT    1260
 +1: D  W  N  I  K  E  Q  R  E  N  F  E  Q  F  F  K  E  N  K  H     420

DNA: AATTATGATACATCAAAATTTGAGCGATGA                                   1310
 +1: N  Y  D  T  S  K  F  E  R  *                                    430
```

Figure 5.6 – *Plasmodium falciparum* BCKDH-E1α

The DNA sequence of the *BCKDH-E1α* gene is displayed with the resulting amino acid sequence shown below it. The predicted mitochondrial targeting peptide cleavage site is highlighted in pink, while the thiamine pyrophosphate binding motif is highlighted in grey. A conserved histidine residue (His318) residue that acts as a proton donor to the lipoic acid attached to the BCKDH-E2 is highlighted green. Leucine 191 is responsible for the V-conformation of the bound thiamine pyrophosphate and is highlighted red. A single phosphorylation site (Ser319) is highlighted in yellow, and a residue (Tyr422) that is essential for formation of the hetero-tetramer is highlighted in blue.

5.2.1.4 Branched chain α - keto acid dehydrogenase beta subunit

The *P. falciparum* branched chain α – keto acid dehydrogenase beta subunit (BCKDH-E1β) displayed 61 % amino acid sequence identity with the human protein (P21953). The *P. falciparum* gene I identified was located on chromosome 5 and was later annotated PFE0225w in PlasmoDB (www.plasmodb.org). The 1146 bp *BCKDH-E1β* gene coded for a protein of 382 amino acids with a predicted molecular size of 43.0 kDa. Alignment with homologues from other *Plasmodium* species *P. yoelii yoelii* and *P. berghei* confirmed the predicted start site of the open reading frame and displayed amino acid identities of 82 % and 77 %, respectively (data not shown). BLASTP analyses of the predicted amino acid sequence of *P. falciparum* BCKDH-E1β displayed greatest amino acid identities to homologous proteins of mammals such as *Bos taurus* (Cow; 61 % identity), *Gallus gallus* (chicken; 62 % identity), *Rattus norvegicus* (rat; 61 % identity) and *Mus musculus* (house mouse; 61 % identity).

Comparison of the *P. falciparum* BCKDH-E1β amino acid sequence with those from organisms previously investigated identified some of the conserved and important residues (Aevarsson *et al.*, 1999). The active site histidine that acts as a proton donor to one of the sulphurs from the lipoyl-arm of the BCKDH-E2 was conserved identified in the *P. falciparum* sequence (His 186; highlighted red in figure 5.7). Two residues involved in binding of the TPP co-factor were also conserved; the conserved tyrosine (Tyr 142) and glycine (Gly 118) are involved in the binding of the aminopyrimidine ring of the TPP co-factor (highlighted blue in figure 5.7).

Analysis of the BCKDH-E1β amino acid sequence with subcellular localisation prediction programs suggested that the protein was mitochondrial with a high level of confidence (99.5 %), although a mitochondrial targeting peptide cleavage site was not predictable by the programs used (table 5.2). Alignment of the *P. falciparum* BCKDH-E1β with its *E. coli* homologue displayed amino acid identity starting from amino acid 57 (highlighted yellow in figure 5.7), which suggests that the possible mature protein (minus the mitochondrial targeting peptide) was about 36.2 kDa in size.

```
DNA: ATGATGAGACTATTAAGAAATAACGTAAAAAATATGTTCTTAAAAAATTGCTTTCTTGTG  60
 +1: M  M  R  L  L  R  N  N  V  K  N  M  F  L  K  N  C  F  L  V   20

DNA: AGGAATACAAATGTAAAATGTTGGAGAACCGATAGGTTTCATAATTACAATATGTGGAAT 120
 +1: R  N  T  N  V  K  C  W  R  T  D  R  F  H  N  Y  N  M  W  N   40

DNA: ACTGTTGAAAATAATGGCATTCGTGGTTTTTCTTCCTCTAGTTTTGAAGAGATAAAAAAA 180
 +1: T  V  E  N  N  G  I  R  G  F  S  S  S  S  F  E  E  I  K  K   60

DNA: ATGAATATGTTTACAGCAATAAATTCAGCTATGCATAATGTATTTGAAAGTAACCCTAAT 240
 +1: M  N  M  F  T  A  I  N  S  A  M  H  N  V  F  E  S  N  P  N   80

DNA: TCTGTATTATTAGGAGAAGATGTTGCTTTCGGAGGTGTGTTTAGATGTTCTTTAGATTTA 300
 +1: S  V  L  L  G  E  D  V  A  F  G  G  V  F  R  C  S  L  D  L  100

DNA: TTAAAAAAATATGGAAACATGAGAGTTTTCAATACTCCCTTATGTGAACAAGGAATTATA 360
 +1: L  K  K  Y  G  N  M  R  V  F  N  T  P  L  C  E  Q  E  I  I  120

DNA: GGATTTGCTATAGGTCTAGCTGAAAATGGATTTACAACAATAGCAGAAATACAATTTGGT 420
 +1: G  F  A  I  G  L  A  E  N  G  F  T  T  I  A  E  I  Q  F  G  140

DNA: GATTACATATTTCCAGCTTTTGATCAAATAGTAAATGATGTAGCTAAATATAGATATAGA 480
 +1: D  H  I  F  P  A  F  D  Q  I  V  N  D  V  A  K  Y  R  Y  R  160

DNA: TCAGGTAGTAGTTTCGATGTAGGTAAATTAACTATAAGATCGACGTGGGGTGCTGTTGGT 540
 +1: S  G  S  S  F  D  V  G  K  L  T  I  R  S  T  W  G  A  V  G  180

DNA: CATGGTGGTTTATATCATTCACAAAGTCCTGAAGCCTTTTTTGCACATGCAGCTGGTATA 600
 +1: H  G  G  L  Y  H  S  Q  S  P  E  A  F  F  A  H  A  A  G  I  200

DNA: AAAATTATTGTACCAAGTGATGCATATAAAGCTAAAGGATTGTTATTGTCAGCTATTAAT 660
 +1: K  I  I  V  P  S  D  A  Y  K  A  K  G  L  L  L  S  A  I  N  220

DNA: GATCCGAACCCATGCTTATTTTTTGAACCCAAAATTTTATATAGATCATCAGTGTGCGAT 720
 +1: D  P  N  P  C  L  F  F  E  P  K  I  L  Y  R  S  S  V  C  D  240

DNA: GTACCTACGGGACCATATCAATTAGAATTGGGTAAGGCAGATGTTGTAAGACAAGGTTCA 780
 +1: V  P  T  G  P  Y  Q  L  E  L  G  K  A  D  V  V  R  Q  G  S  260

DNA: GATGTGACTATAGTAACATGGGGATCTTTAGTACATAAAATGAAAAATGCTGCAGAGATT 840
 +1: D  V  T  I  V  T  W  G  S  L  V  H  K  M  K  N  A  A  E  I  280

DNA: TTATCCAAAAAACATAATATCGAATGTGAAGTTATTGATCTACAATCTATCATACCATGG 900
 +1: L  S  K  K  H  N  I  E  C  E  V  I  D  L  Q  S  I  I  P  W  300

DNA: GATATTGAAACTGTACAAAAATCTGTCGAAAAAACGGGAAGACTTTTAATTACACATGAA 960
 +1: D  I  E  T  V  Q  K  S  V  E  K  T  G  R  L  L  I  T  H  E  320

DNA: GCTCAGCTAACCAATGGTTTTGGAGCTGAAATTGCAGCAAAAATTCAAGAAAGATGTTTC 1020
 +1: A  Q  L  T  N  G  F  G  A  E  I  A  A  K  I  Q  E  R  C  F  340

DNA: TATAATTTGCATACACCTATAAAAAGAGTCTGTGGTTATGATACTCCTTTCCCCCATGTC 1080
 +1: Y  N  L  H  T  P  I  K  R  V  C  G  Y  D  T  P  F  P  H  V  360

DNA: TACGAGCCATTTTACATGCCCGATGCACATAAAGTTATATACGAAGTCAAAAAAATGATG 1140
 +1: Y  E  P  F  Y  M  P  D  A  H  K  V  I  Y  E  V  K  K  M  M  380

DNA: AAGTAA                                                       1146
 +1: K  *                                                         382
```

Figure 5.7 – *Plasmodium falciparum* BCKDH-E1β

The DNA sequence of the *BCKDH-E1β* gene is displayed with the resulting amino acid sequence shown below it. The first amino acid of the predicted mature protein (Glu57) is highlighted in yellow. An active site Histidine (His186) that acts as a proton donor to the E2 bound lipoic acid is highlighted in red. Two residues (Tyr142 and Gly118) that are involved in the binding of thiamine pyrophosphate are highlighted in blue. The last amino acid (Gly130) of the N-terminal section used to construct the BCKDH-E1β GFP fusion protein is highlighted in green.

5.2.1.5 Branched chain acyl transferase

The *P. falciparum* BCKDH-E2 protein displayed 34 % sequence identity with the human protein (P11182). The *P. falciparum* gene I identified was located on chromosome 3 and was later annotated PFC0170c in PlasmoDB (www.plasmodb.org). The 1347 bp gene coded for a protein of 449 amino acids with a theoretical molecular size of 51.0 kDa. Alignment with homologues from other *Plasmodium* species confirmed the predicted start site and displayed 59 %, 57 % and 57 % amino acid identity to the *P. chabaudi*, *P. berghei* and *P. vivax* homologues, respectively. BLASTP analyses of the predicted *P. falciparum* BCKDH-E2 amino acid sequence displayed greatest amino acid identity to homologous proteins from the nematode *Caenorhabditis elegans* (32 % identity), gammaproteobacteria (*Idiomarina loihiensis*; 30 % identity) and plant *Oryza sativa* (30 % identity).

ClustalW alignment of the *P. falciparum* BCKDH-E2 with mitochondrial isoforms from other organisms displayed that the protein can be divided into three domains that are separated by flexible linker regions; the lipoyl-domain, the subunit-binding domain and the catalytic/inner core domain. The position of these domains in the *P. falciparum* BCKDH-E2 amino acid sequence was identified by alignment with proteins from organisms that have been previously investigated (Griffin *et al.*, 1990; Wexler *et al.*, 1991; Meng *et al.*, 1994; Ono *et al.*, 2001). The lipoyl-domain runs between residues 15 and 102 and contains the lysine residue (residue 74 in the *P. falciparum* protein; highlighted in red in figure 5.8) that is the site of lipoylation. The subunit-binding domain runs between residues 144 and 189 and is the site of E1 and LipDH association. The catalytic domain is the site of self-association and acylation of coenzyme A and runs between residues 212 and 449. Residues His 418 and Ser 358 (highlighted in yellow in figure 5.8) have been implicated in acylation reaction, with His 418 (highlighted in blue in figure 5.8) acting as an acid/base catalyst. Analysis of the flexible linker domains displays that the BCKDH-E2 linkers are not rich in proline and alanine residues as observed in mammalian proteins, but are rich in asparagines, glutamates and aspartate residues (figure 5.8).

Analysis of BCKDH E2 amino acid sequence with subcellular localisation prediction programs identified that the protein was predicted to be mitochondrially localised with high confidence levels (92.2 %) and a mitochondrial targeting peptide cleavage site was suggested after amino acid 14 (highlighted green in figure 5.8; table 5.2); this would correspond to a mature protein size of 49.5 kDa.

```
DNA: ATGTTTGTGAAGAATGTACTAAACGTGCTTAGGAGAATCGAGGGAAAATCATTTAAAGGG  60
 +1: M  F  V  K  N  V  L  N  V  L  R  R  I  E  G  K  S  F  K  G   20

DNA: CGCCATTACCTGAACACGAGTGCAATACATTTTAAGATTGTGAAATGCAAATTATTTGAT 120
 +1: R  H  Y  L  N  T  S  A  I  H  F  K  I  V  K  C  K  L  F  D   40

DNA: ATAGGGGAAGGGATATCAGAAGTGGAAATTACCAAGTGGCATAAAAATGAAGGTGATCAA 180
 +1: I  G  E  G  I  S  E  V  E  I  T  K  W  H  K  N  E  G  D  Q   60

DNA: GTAAGTGAAATGGAGAGTTTGTTGACTGTACAAAGTGACAAAGCTGCTGTAGATATAACT 240
 +1: V  S  E  M  E  S  L  L  T  V  Q  S  D  K  A  A  V  D  I  T   80

DNA: AGTAAATATAATGGTGTATTGGTAAAAAAATATTTAAATGAAAATGATATGTTAAAAGTT 300
 +1: S  K  Y  N  G  V  L  V  K  K  Y  L  N  E  N  D  M  L  K  V  100

DNA: GGATCATATTTCTGTGAGATAGATACAGATGATGATATTATTGAAAGAGATGAAGAAGAA 360
 +1: G  S  Y  F  C  E  I  D  T  D  D  D  I  I  E  R  D  E  E  E  120

DNA: GTAGAAAAGGAGGAAAATAATAAAAAAGAAGAAGATGGTGAAAGTGATTTAAGTTTAAAT 420
 +1: V  E  K  E  E  N  N  K  K  E  E  D  G  E  S  D  L  S  L  N  140

DNA: GATGATATTAGCAATAATGATTATATTAAAGCATCTCCAGGTGTTAAAAGAAAAGCCAAA 480
 +1: D  D  I  S  N  N  D  Y  I  K  A  S  P  G  V  K  R  K  A  K  160

DNA: GAATATAAAGTAAATTTAAATAAAGTTGGTGATTATTTTAATAAAGTTAATATAAGCTTA 540
 +1: E  Y  K  V  N  L  N  K  V  G  D  Y  F  N  K  V  N  I  S  L  180

DNA: GAAGATTTGGAATTATATTATAATAATGTTGTTAAAAATGAATATAGTAATAATATAAAT 600
 +1: E  D  L  E  L  Y  Y  N  N  V  V  K  N  E  Y  S  N  N  I  N  200

DNA: AATAATGATATGGATATTATTGAAGAAGTATCATTAAAAGGTATAAAATTAGCTATGTGT 660
 +1: N  N  D  M  D  I  I  E  E  V  S  L  K  G  I  K  L  A  M  C  220

DNA: AAAAGTATGAATGAATCTTTACAAGTACCATTATTTCATTTAAATGAAATGTGTATTATA 720
 +1: K  S  M  N  E  S  L  Q  V  P  L  F  H  L  N  E  M  C  I  I  240

DNA: AATAATTTAATAAAAATGAGAAAAGAATATAAAGAACAACAAAAAAATTTACAAACAAAA 780
 +1: N  N  L  I  K  M  R  K  E  Y  K  E  Q  Q  K  N  L  Q  T  K  260

DNA: GAAACTAATATAACTATTACATGTATACTTATAAAATTAATATCTAATGTATTAAAAGAA 840
 +1: E  T  N  I  T  I  T  C  I  L  I  K  L  I  S  N  V  L  K  E  280

DNA: TTCCCAATTTTAAATTCCAAATTTAATTTTAAAACTAATACTTATACTATGTATAAGAAT 900
 +1: F  P  I  L  N  S  K  F  N  F  K  T  N  T  Y  T  M  Y  K  N  300

DNA: CATAATATATCTATAGCTGTTGATACTCCTCATGGATTATTAGTACCAAATATAAAAAAT 960
 +1: H  N  I  S  I  A  V  D  T  P  H  G  L  L  V  P  N  I  K  N  320

DNA: GTTCAGAATAAAAATATTTTAGATATTCAAAAAGATTTATTATCATTACGTGATAAAGCA 1020
 +1: V  Q  N  K  N  I  L  D  I  Q  K  D  L  L  S  L  R  D  K  A  340

DNA: AATAATATGCAACTTGATAAAAGTGATATTACTAATGGTACAATAACTGTTAGTAATTTT 1080
 +1: N  N  M  Q  L  D  K  S  D  I  T  N  G  T  I  T  V  S  N  F  360

DNA: GGAGCCATATCAGGAACCTTTGCTACTCCTATAGTATTTGATAATCAAGCATGTATAATA 1140
 +1: G  A  I  S  G  T  F  A  T  P  I  V  F  D  N  Q  A  C  I  I  380

DNA: GGTATAGGCAAAATGGAAAAAAAAATTGCTTTTAAAAGATGAATCTAATAACTTAAATTCT 1200
 +1: G  I  G  K  M  E  K  K  L  L  L  K  D  E  S  N  N  L  N  S  400

DNA: TTAAATGATATATTAGTTGCTGATACTATAAACTTCACCTTTGGAGCTGACCATAGGTAT 1260
 +1: L  N  D  I  L  V  A  D  T  I  N  F  T  F  G  A  D  H  R  Y  420

DNA: ATAGATGGAGCAACCTTAGCTCAGTTTTCTAAAATGTTAAAAATGAATATTGAAAATTGT 1320
 +1: I  D  G  A  T  L  A  Q  F  S  K  M  L  K  M  N  I  E  N  C  440

DNA: GCATCCCTAGGACCACTATTGGAATAA                                   1347
 +1: A  S  L  G  P  L  L  E  *                                    449
```

Figure 5.8 – *Plasmodium falciparum* BCKDH E2

The DNA sequence of the *BCKDH E2* gene is displayed with the resulting amino acid sequence shown below it. The predicted mitochondrial targeting peptide cleavage site (Glu14) is highlighted in green. The conserved lysine residue (Lys74) that is the site of lipoic acid attachment to the protein is highlighted in red. Two residues, Ser358 and His418, which are involved in the acylation reaction are highlighted in yellow and blue, respectively.

5.2.1.6 Pyruvate dehydrogenase alpha subunit

The *pyruvate dehydrogenase alpha* (*PDH-E1α*) gene displayed 32 % identity with the human amino acid sequence (P08559). The *P. falciparum* gene I identified was located on chromosome 11 and was later annotated PF11_0256 in PlasmoDB (www.plasmodb.org). The 1824 bp gene is intron-less and encodes a protein of 608 amino acids with a predicted molecular size of 70.9 kDa. ClustalW alignment with homologues identified in other *Plasmodium* species confirmed the predicted start site of the protein and displayed 73 %, 66 % and 65 % amino acid identity to the *P. berghei*, *P. chabaudi* and *P. yoelii yoelii* proteins, respectively. BLASTP analyses of the predicted amino acid sequence of PDH-E1α displayed greatest amino acid identity to homologous proteins from the cyanobacteria *Thermosynechococcus elongatus* (46 % identity) and *Crocoshpaera watsonii* (44 % identity) and the red algae *Cyanidium caldarium* (48 % identity) and *Porphyra purpurea* (44 % identity).

Comparison of the *P. falciparum* PDH-E1α amino acid sequence with those previously studied in other organism identified a number of key residues (Diefenbach *et al.*, 1992; Russell *et al.*, 1992; Ciszak *et al.*, 2003). The TPP co-factor binding motif ($GDGX_{26}NN$) was observed in the sequence between residues 364 and 395 (highlighted grey in figure 5.9). One of the phosphorylation sites observed in the mammalian proteins was conserved in the *P. falciparum* sequence (Ser 467; highlighted yellow in figure 5.9). The active site histidine residue that acts as a proton donor to one the sulphurs of a lipoyl-arm from the PDH-E2 was conserved (His 466; highlighted blue in figure 5.9).

The *Plasmodium* PDH-E1α deduced amino acid sequence possessed a long N-terminal extension compared to those PDH-E1α from other organisms. Analysis of the PDH-E1α subunit amino acid sequence with subcellular prediction programs identified a typical signal peptide of 17 or 23 amino acids at its *N*-terminus (depending on the prediction program used; highlighted red in figure 5.9), which was followed by a stretch of hydrophilic amino acids. Using the prediction program PlasmoAP it was found that this hydrophilic stretch of amino acids possibly could confer transit into the parasite's apicoplast (++; table 5.2). These results therefore suggest that PDH possesses a typical bipartite targeting sequence responsible for the transfer of *Plasmodium* proteins to the apicoplast (Foth *et al.*, 2003). Given that the parasite genome only contains genes encoding for the components of a single PDH this predicted localisation appears rather

unusual and will be investigated in more detail later. Prediction of the possible mature PDH-E1α was not possible using any of the prediction programs. Also it is impossible to draw any conclusions from comparison with the *E. coli* counterpart because the bacterial PDH contains a homodimeric E1 rather than a hetero-dimeric $\alpha_2\beta_2$ E1 found in eukaryotes (de Kok *et al.*, 1998).

```
DNA: ATGCTGTTAGTAATATGGTTAAACATTTTTTGGTTCATCCCTTATTATTGTTTTAATTAC  60
 +1: M  L  L  V  I  W  L  N  I  F  W  F  I  P  Y  Y  C  F  N  Y   20

DNA: GTTTTATGTGTAGGGAATAAAAATGACATGCTTTTTGTAAAAAATAAAAATCTAATATAT 120
 +1: V  L  C  V  G  N  K  N  D  M  L  F  V  K  N  K  N  L  I  Y   40

DNA: CCTTATAATGATATGAATAGAATAAGGAGACATAAAGGGGTGGAACTAAGCAGAAAGATG 180
 +1: P  Y  N  D  M  N  R  I  R  R  H  K  G  V  E  L  S  R  K  M   60

DNA: GTTGCTGAGAATATTCAAATGAGAAATATGAAGAGTAACAATAATAGTAATAATGTTCAT 240
 +1: V  A  E  N  I  Q  M  R  N  M  K  S  N  N  N  S  N  N  V  H   80

DNA: GAGAATAATAATATTGATCATAGTAACAATTATAATAATAATTTATATTTATCATGTAGT 300
 +1: E  N  N  N  I  D  H  S  N  N  Y  N  N  N  L  Y  L  S  C  S  100

DNA: AATAAAAAAAAAGGAAAAGATGTACAATCATTGTGTGCCATAAAACACGGCAACGTACTG 360
 +1: N  K  K  K  G  K  D  V  Q  S  L  C  A  I  K  H  C  N  V  L  120

DNA: ATCAAAAATATAAATGAAGAAGAAAAAGAAAAGGGAATTAAGAAAAATGATACAGATAAA 420
 +1: I  K  N  I  N  E  E  E  K  E  K  G  I  K  K  N  D  T  D  K  140

DNA: TTAGAATCAAAAAATGATAATGAACATAAAAATGAAGAGAATAATTTCGTTTTTTCATAT 480
 +1: L  E  S  K  N  D  N  E  H  K  N  E  E  N  N  F  V  F  S  Y  160

DNA: GATAAAAAATTAAATAATTATTCTGAATTTAATATATATATGGAAAATAATAATATAGAA 540
 +1: D  K  K  L  N  N  Y  S  E  F  N  I  Y  M  E  N  N  N  I  E  180

DNA: GAATATATATCCGATGTAAATATTAGTAGAGAAGAAATATGTACCTTATATGAAGATATG 600
 +1: E  Y  I  S  D  V  N  I  S  R  E  E  I  C  T  L  Y  E  D  M  200

DNA: TATTTAGGACGACTTTTTGAGAACTTAGTAGCCAAATTATATTATAATAAACGAGTTAAT 660
 +1: Y  L  G  R  L  F  E  N  L  V  A  K  L  Y  Y  N  K  R  V  N  220

DNA: GGTTTTGTTCATTTATATAATGGTCAAGAAGCTGTAAGTACAGGTATTATTAAAAATTTA 720
 +1: G  F  V  H  L  Y  N  G  Q  E  A  V  S  T  G  I  I  K  N  L  240

DNA: AAAAATTCAGATTTTGTTACAAGTACTTATAGAGATCATGTTCATGCTCTAAGTAAAGGT 780
 +1: K  N  S  D  F  V  T  S  T  Y  R  D  H  V  H  A  L  S  K  G  260

DNA: GTGCCTGCACATAAAATATTAAATGAATTATATGGAAATTATTATGGTAGTACAAATAAA 840
 +1: V  P  A  H  K  I  L  N  E  L  Y  G  N  Y  Y  G  S  T  N  K  280

DNA: GGGAAAGGTGGTTCTATGCATATTTATTCAAAAGAAAATAATTTTATAGGAGGTTTTGGT 900
 +1: G  K  G  G  S  M  H  I  Y  S  K  E  N  N  F  I  G  G  F  G  300

DNA: TTTATTGGTGAACAAATACCTATTGCTGTTGGATTAGCATATAGTATCTTATATAAAAAT 960
 +1: F  I  G  E  Q  I  P  I  A  V  G  L  A  Y  S  I  L  Y  K  N  320

DNA: GAATTTCACTATAATCCAAAAAATACCTCTTTTACATCTACTAAAAATAAAAATAATTAT 1020
 +1: E  F  H  Y  N  P  K  N  T  S  F  T  S  T  K  N  K  N  N  Y  340

DNA: ATACAAGAAAATGAAAATATGATACATATGAATAATTCTCAAAACGTAGATGTTGTAGTA 1080
 +1: I  Q  E  N  E  N  M  I  H  M  N  N  S  Q  N  V  D  V  V  V  360

DNA: TGTTTCTTAGGTGATGGTACTACTAATATTGGACAATTCTTTGAATCATTAAATCTTGCT 1140
 +1: C  F  L  G  D  G  T  T  N  I  G  Q  F  F  E  S  L  N  L  A  380

DNA: TCTTCTTACAACTTACCAATAATATTTGTTATAGAAAATAATAATTGGGCTATAGGTATG 1200
 +1: S  S  Y  N  L  P  I  I  F  V  I  E  N  N  N  W  A  I  G  M  400

DNA: GAAAGTTCTAGAAGCTCATCTGATGATCTTATGAATAATTATTCCAAAGGAAAAGCTTTT 1260
 +1: E  S  S  R  S  S  S  D  D  L  M  N  N  Y  S  K  G  K  A  F  420

DNA: AATATAGATACTTTTAAAGTAGATGGAAATGATGTTCTTACTATATATAAACTAGCCAAA 1320
 +1: N  I  D  T  F  K  V  D  G  N  D  V  L  T  I  Y  K  L  A  K  440

DNA: AAGAAAATTCAACAAATTAGAAATAGAACATCCGGACCAATAATTATAGAAGCTATTACA 1380
 +1: K  K  I  Q  Q  I  R  N  R  T  S  G  P  I  I  I  E  A  I  T  460
```

```
DNA: TATCGAGCTAAAGGACATTCTTTAGCAGATCCAGATGAACTTCGTATCAAAGAAGAAAAA 1440
 +1: Y  R  A  K  G  ■  S  L  A  D  P  D  E  L  R  I  K  E  E  K   480

DNA: ACTTCATGGAAAAAAAGAGACCCTATTTTATTCTTATCTAGTTATATGAAGAAATATAAT 1500
 +1: T  S  W  K  K  R  D  P  I  L  F  L  S  S  Y  M  K  K  Y  N   500

DNA: TTAGTACAAGAATCATACTTCGAACAAGTTAAAAAAAATACACAAACATTATTACAACAA 1560
 +1: L  V  Q  E  S  Y  F  E  Q  V  K  K  N  T  Q  T  L  L  Q  Q   520

DNA: GCAGAATTAGATGCTGAACAAAATACAAAGAAAGGGGAACATATAGATATATGTAATGTA 1620
 +1: A  E  L  D  A  E  Q  N  T  K  K  G  E  H  I  D  I  C  N  V   540

DNA: ATTCAACAAAATATATTCGCCCCATCAAAATCTACTCCATATCAAAGTGAATATGAAAAT 1680
 +1: I  Q  Q  N  I  F  A  P  S  K  S  T  P  Y  Q  S  E  Y  E  N   560

DNA: TATAAACAATTTGACGACATTTCTAATGACGAATTAAAAGAATATTATCAAGAAGTACTT 1740
 +1: Y  K  Q  F  D  D  I  S  N  D  E  L  K  E  Y  Y  Q  E  V  L   580

DNA: AAAGAAATTCAAAGAAAAAAAGAAAAGAAAAAACTTGATCCGAACGACAAATTTGATCAA 1800
 +1: K  E  I  Q  R  K  K  E  K  K  K  L  D  P  N  D  K  F  D  Q   600

DNA: AAAAAGTTACCCTTAATAATAGATTAA                                   1827
 +1: K  K  L  P  L  I  I  D  *                                     609
```

Figure 5.9 – *Plasmodium falciparum* PDH-E1α

The DNA sequence of the *PDH-E1α* gene is displayed with the resulting amino acid sequence shown below it. Two signal peptide cleavage sites were predicted (Cys17 or Cys23) and both are highlighted in red. The last amino acid (Gly117) in the N-terminal section (amino acids 1 – 117) used to construct the GFP-fusion protein is highlighted in green. The thiamine pyrophosphate-binding motif ($GDGX_{26}NN$) is highlighted in grey. An active site histidine (His466) that acts as a proton donor to the PDH-E2 bound lipoamide is highlighted in blue. The adjacent residue (Ser467) is a site of PDH control by phosphorylation with pyruvate dehydrogenase kinase and is highlighted in yellow.

5.2.1.7 Pyruvate dehydrogenase beta subunit

The *pyruvate dehydrogenase beta* (*PDH-E1β*) gene displayed 35 % identity with the human amino acid sequence (P11177). The *P. falciparum* gene I identified was located on chromosome 14 and was later annotated PF14_0441 in PlasmoDB (www.plasmodb.org). ClustalW alignment with homologues identified in other *Plasmodium* species confirmed the predicted start site of the *P. falciparum PDH-E1β* open reading frame and displayed 82 % and 77 % amino acid identity to the *P. yoelii yoelii* and *P. berghei* proteins, respectively. BLASTP analyses of the predicted *P. falciparum* PHD-E1β amino acid sequence displayed greatest amino acid identity to homologous proteins from the cyanobacteria *Gleobacter violaceus* (58 % identity), *Anabena variabilis* (57 % identity) and *Prochlorococcus marinus* (57 % identity) and the red algae *Gracilaria tenuistipitata* (57 % identity) and *Porphyra purpurea* (57 % identity). The 1248 bp gene coded for a protein of 416 amino acids with a predicted molecular size of 47.0 kDa.

Comparison of the predicted PDH-E1β amino acid sequence with those previously studied in other organisms (Ciszak *et al.*, 2003) identified a number of important residues conserved in the *P. falciparum* protein. The active site histidine residue that acts as a proton donor to a sulphur on a lipoyl-arm of the PDH-E2 was identified (His 217; highlighted red in figure 5.10). Two conserved residues, a phenylalanine (Phe 151) and a valine (Val 202) both involved in producing the V-conformation of the bound TPP co-factor (highlighted green in figure 5.10) were also conserved in the parasite protein.

Similar to PDH-E1α it was found that the PDH-E1β deduced amino acid sequence contained a signal peptide at its *N*-terminus followed by a hydrophilic stretch of amino acids both possibly comprising a bipartite apicoplast targeting peptide. PlasmoAP analyses supported this suggestion with high confidence (++; table 5.2). However, as opposed to PDH-E1 α, the prediction programs did not suggest a potential cleavage site of the signal peptide . Similar to PDH-E1 α prediction of the possible mature PDH-E1β was not possible as the bacterial PDH contains a homodimeric E1 rather than a heterodimeric $\alpha_2\beta_2$ E1 and a precise cleavage of a potential apicoplast transit peptide cannot be made.

```
DNA: ATGGGGAGAAAAAGAAACAATATTGTAAACATTAGATATAATTTTTTCTTGATTTATTTT  60
 +1: M  G  R  K  R  N  N  I  V  N  I  R  Y  N  F  F  L  I  Y  F   20

DNA: TGGGTGATATTTATGTATCCATGCGAGACAAATAAAGGAAATGTGAAGCCACTAAATTTT 120
 +1: W  V  I  F  M  Y  P  C  E  T  N  K  G  N  V  K  P  L  N  F   40

DNA: ATAAAAGGAAAGAATATTTTTTTTGAATATAACTAAAAACAAAGTGAAGCATTTAAATACA 180
 +1: I  K  G  K  N  I  F  L  N  I  T  K  N  K  V  K  H  L  N  T   60

DNA: ATAAATGGAATTGAAACTATATCGAATGTTGAGAACAAAAATATTTTGAATGATACAAAT 240
 +1: I  N  G  I  E  T  I  S  N  V  E  N  K  N  I  L  N  D  T  N   80

DNA: TATATAAATGAAATGAAGAATATAAAGGTTAGAAGAAATATAAGTGAAGCTTTACATATG 300
 +1: Y  I  N  E  M  K  N  I  K  V  R  R  N  I  S  E  A  L  H  M  100

DNA: GCCATATATGAAGAAATGAAAAAAGATAAAGGTGTATATGTACTTGGAGAAGATGTAGGA 360
 +1: A  I  Y  E  E  M  K  K  D  K  G  V  Y  V  L  G  E  D  V  G  120

DNA: TTGTATGGAGGTTCGTATAAAGTTACTAAAAATTTAGCTCACTTTTTTGGTTTTTCTAGA 420
 +1: L  Y  G  G  S  Y  K  V  T  K  N  L  A  H  F  F  G  F  S  R  140

DNA: GTTTTAGATACACCCATATGTGAAAATGCCTTCATGGGATTAGGTATAGGTTCTGCAATT 480
 +1: V  L  D  T  P  I  C  E  N  A  F  M  G  L  G  I  G  S  A  I  160

DNA: AACGACTTAAGACCTATTATTGAAGGTATGAATTTGTCTTTTTTAATTTTAGCATTTAAT 540
 +1: N  D  L  R  P  I  I  E  G  M  N  L  S  F  L  I  L  A  F  N  180

DNA: CAAATATCAAATAATGCTTGTATGATGAGATATATGTGTGATGGTCAATTTAATATTCCT 600
 +1: Q  I  S  N  N  A  C  M  M  R  Y  M  C  D  G  Q  F  N  I  P  200

DNA: ATTGTTATTAGAGGTCCAGGAGGTATAGGGAAACAATTAGGTCCTGAACATTCTCAAAGA 660
 +1: I  F  I  R  G  P  G  G  I  G  K  Q  L  G  P  E  H  S  Q  R  220

DNA: ATTGAATCGTATTTAATGAGTATACCAGGTATTAAAATCGTTTCATGCTCAACACCTTTT 720
 +1: I  E  S  Y  L  M  S  I  P  G  I  K  I  V  S  C  S  T  P  F  240

DNA: AATGCTAGAGGATTATTAAAATCAGCAATTAGAGATAATAACCCTATATTATTTATAGAA 780
 +1: N  A  R  G  L  L  K  S  A  I  R  D  N  N  P  I  L  F  I  E  260

DNA: CATGTTTTATTATATAATTATGAACAGGAAATTCCTCTTTTACCTTACACCTTACCTATT 840
 +1: H  V  L  L  Y  N  Y  E  Q  E  I  P  L  L  P  Y  T  L  P  I  280

DNA: GATAAAGCAGAAGTTGTTAAAAATGGAAAAGATCTAACCGTTTTATCCTATGGAATAACA 900
 +1: D  K  A  E  V  V  K  N  G  K  D  L  T  V  L  S  Y  G  I  T  300

DNA: AGACATTTAGCTTCCGAAGCAGCAAAAGAATTAACGAAATTTAATATAGATATAGAAGTA 960
 +1: R  H  L  A  S  E  A  A  K  E  L  T  K  F  N  I  D  I  E  V  320

DNA: ATTGATTTAATTTCTTTAAAACCATTTGATATGGAAACTATAGAAAAATCTCTTAAAAAA 1020
 +1: I  D  L  I  S  L  K  P  F  D  M  E  T  I  E  K  S  L  K  K  340

DNA: ACCAAGAAATGTTTAATTTTGGATGAGTCAGCTGGTTTTGGGGGTATAGGAGCTGAATTA 1080
 +1: T  K  K  C  L  I  L  D  E  S  A  G  F  G  G  I  G  A  E  L  360

DNA: TATACACAAGTTATAGAAATGTTTTCTTCATACTTAATAACCAAACCTATTAGATTATGT 1140
 +1: Y  T  Q  V  I  E  M  F  S  S  Y  L  I  T  K  P  I  R  L  C  380

DNA: ACCAAGGATATACCTATAGCTTATTCAAATAAATATGAAGACGCATGTATTATCAAAAAG 1200
 +1: T  K  D  I  P  I  A  Y  S  N  K  Y  E  D  A  C  I  I  K  K  400

DNA: GAGGACATTGTATATATGTCTACTTATCTTCATTCGCTATCATCTTGA              1248
 +1: E  D  I  V  Y  M  S  T  Y  L  H  S  L  S  S  *               416
```

Figure 5.10 – *Plasmodium falciparum* PDH-E1β

The DNA sequence of the *PDH-E1α* gene is displayed with the resulting amino acid sequence shown below it. An active site histidine residue that acts as a proton donor to the lipoamide bound to PDH-E2 is highlighted in red. Two residues (Phe151 and Val202) that are involved in the formation of the bound thiamine pyrophosphate into a V-conformation are highlighted in green.

5.2.1.8 Acetyl transferase

The *P. falciparum acetyl transferase* gene (*PDH-E2*) identified had been previously annotated PF10_0407 in PlasmoDB (www.plasmodb.org). The predicted protein had 26 % and 23 % identity with the human (P10515) and yeast (P12695) amino acid sequences, respectively. The 3192 bp *P. falciparum PDH-E2* gene was found to contain 10 predicted introns (table 5.1); the predicted intron/exon boundaries are displayed in figure 5.11 (highlighted red) with the resulting exons highlighted in grey (figure 5.11). As observed for *LipDH2*, *KGDH-E1* and *KGDH-E2* genes, the *PDH-E2* intron sequences had a higher A/T content (average of 87.5 % in the *PDH E2* introns; table 5.1) than the exon sequences (average of 72.9 % in the *PDH E2* exons; table 5.1).

The predicted full-length protein consisted of 641 amino acids with a theoretical molecular size of 73.9 kDa. BLASTP analyses of the PDH-E2 deduced amino acid sequence displayed greatest amino acid identity to homologous proteins from cyanobacteria *Crocosphaera watsonii* (30 % identity), *Anabena variabilis* (28 % identity) and *Thermosynechococcus elongates* (28 % identity).

The *P. falciparum* PDH-E2 predicted amino acid sequence can be divided into three domains; the lipoyl-domain, the subunit-binding domain and the catalytic domain. The position of these domains in the *P. falciparum* amino acid sequence was identified by alignment with proteins from organisms that have been previously investigated (Thekkumkara *et al.*, 1988; Russell *et al.*, 1992; Harris *et al.*, 1997). The *P. falciparum* protein contains two lipoyl-domains that run between residues 52 and 336. Both domains contain the essential lysine residue (residues 93 and 223) that is the site of lipoylation (highlighted red in figure 5.12). The subunit-binding domain was identified to be between residues 354 and 385 and is responsible for the binding of E1 and LipDH proteins. The catalytic domain was identified to be between residues 414 and 641 and is the site of self-association and acetylation of CoA. The catalytic domain contains conserved histidine (His 612) and tyrosine (Tyr 559) residues that are involved in the acetylation reaction (highlighted grey in figure 5.12). Similar to the other dihydrolipoamide acyl transferase from *Plasmodium*, the flexible linker regions connecting the domains are not proline (P) and alanine (A) rich as found in the human acetyl transferase but are rich in lysine (K) and asparagines (N).

As found for PDH-E1α and PDH-E1β analysis of the deduced PDH-E2 amino acid sequence with subcellular localisation predictions programs suggested that the protein

might be targeted to the parasite's apicoplast with confidence (++). The deduced amino acid sequence was predicted to contain a hydrophobic signal peptide with a potential cleavage site after amino acid 17 (table 5.2; highlighted blue in figure 5.12). The probable length of the apicoplast transit peptide was predicted by aligning the *P. falciparum* PDH-E2 protein sequence with its *E. coli* homologue; this displayed that amino acid identity began around amino acid 51 (highlighted yellow in figure 5.12), which would correspond to a mature protein (minus apicoplast bipartite targeting peptide) size of 67.2 kDa.

```
DNA: ATGTTATACAACTTAATTATATTAATTTTTTATTTAAGATTTTCTAAATGT
 +3:   V  I  Q  L  N  Y  I  N  F  L  F  K  I  F  *  M  Y
 +2:  C  Y  T  T  *  L  Y  *  F  F  I  *  D  F  L  N  V
 +1: M  L  Y  N  L  I  I  L  I  F  Y  L  R  F  S  K  C

DNA: ATTTCTAAGAACAATAATTACGGTTACATCAACTTCGGTACCTTTTCAAAT
 +3:   F  *  E  Q  *  L  R  L  H  Q  L  R  Y  L  F  K  C
 +2:  F  L  R  T  I  I  T  V  T  S  T  S  V  P  F  Q  M
 +1: I  S  K  N  N  N  Y  G  Y  I  N  F  G  T  F  S  N

DNA: GTTGTAAACAACAGTAATAATTTAAGAAATAGAAAAAATGTTGTTTTTTCA          Exon 1
 +3:   C  K  Q  Q  *  *  F  K  K  *  K  K  C  C  F  F  K
 +2:  L  *  T  T  V  I  I  *  E  I  E  K  M  L  F  F  Q
 +1: V  V  N  N  S  N  N  L  R  N  R  K  N  V  V  F  S

DNA: AAAATAGAAATAAAAATGCCAGCTCTATCTAGTACCATGACGACAGGCAAA
 +3:   N  R  N  K  N  A  S  S  I  *  Y  H  D  D  R  Q  N
 +2:  K  *  K  *  K  C  Q  L  Y  L  V  P  *  R  Q  A  K
 +1: K  I  E  I  K  M  P  A  L  S  S  T  M  T  T  G  K

DNA: ATTGTTAAATGGAATAAAAATATAGGAGATTATGTAAATGTAGGATAATAA
 +3:   C  *  M  E  *  K  Y  R  R  L  C  K  C  R  I  I  K
 +2:  L  L  N  G  I  K  I  *  E  I  M  *  M  *  D  N  K
 +1: I  V  K  W  N  K  N  I  G  D  Y  V  N  V  G  *  *

DNA: AAGGGATTCCTTAAAAAAAAAAAAAAAAAAAATATATATATATATATATAA
 +3:   G  I  P  *  K  K  K  K  K  I  Y  I  Y  I  Y  I  I
 +2:  R  D  S  L  K  K  K  K  K  N  I  Y  I  Y  I  Y  N
 +1: K  G  F  L  K  K  K  K  K  K  Y  I  Y  I  Y  I  *

DNA: TAAATATATGTATAATAAATATATGTATAATAAATATATGTATAATAAATA          Intron 1
 +3:   N  I  C  I  I  N  I  C  I  I  N  I  C  I  I  N  I
 +2:  K  Y  M  Y  N  K  Y  M  Y  N  K  Y  M  Y  N  K  Y
 +1: *  I  Y  V  *  *  I  Y  V  *  *  I  Y  V  *  *  I

DNA: TATGTATAATAAATATATGTATAATAAATATATGTATAATAAATATATGTA
 +3:   C  I  I  N  I  C  I  I  N  I  C  I  I  N  I  C  I
 +2:  M  Y  N  K  Y  M  Y  N  K  Y  M  Y  N  K  Y  M  Y
 +1: Y  V  *  *  I  Y  V  *  *  I  Y  V  *  *  I  Y  V

DNA: TATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAGCTAGGTG
 +3:   F  F  F  F  F  F  F  F  F  F  F  F  F  *  L  G  D
 +2:  I  F  F  F  F  F  F  F  F  F  F  F  F  L  A  R  *          Exon 2
 +1: Y  F  F  F  F  F  F  F  F  F  F  F  F  F  S  *  V

DNA: ATATTATAATGACCGTTGAAAGTGATAAAGCAGACATGGACGTGGAAGCAT
 +3:   I  I  M  T  V  E  S  D  K  A  D  M  D  V  E  A  F
 +2:  Y  Y  N  D  R  *  K  *  *  S  R  H  G  R  G  S  I
 +1: I  L  *  *  P  L  K  V  I  K  Q  T  W  T  W  K  H

DNA: TTGATGAAGGTATAACAATTTTAATAAAGATAAAATTATTCCATCTTATTA
 +3:   D  E  G  I  T  I  L  I  K  I  K  L  F  H  L  I  S
 +2:  *  *  R  Y  N  N  F  N  K  D  K  I  I  P  S  Y  *
 +1: L  M  K  V  *  Q  F  *  *  R  *  N  Y  S  I  L  L

DNA: GTCCGATAAAAAAAAATATATTTATTTATATACATATGTGTTGTTACCTAT          Intron 2
 +3:   P  I  K  K  N  I  F  I  Y  I  H  M  C  C  Y  L  L
 +2:  S  D  K  K  K  Y  I  Y  L  Y  T  Y  V  L  L  P  I
 +1: V  R  *  K  K  I  Y  L  F  I  Y  I  C  V  V  T  Y
```

```
DNA: TATTTTGTTAGGCTTTTTAAGGGTTAAGCGTTTAGAAGATGGATGTGAAGC
 +3:   F  C  *  A  F  *  G  L  S  V  *  K  M  D  V  K  Q
 +2:  I  L  L  G  F  L  R  V  K  R  L  E  D  G  C  E  A
 +1: Y  F  V  R  L  F  K  G  *  A  F  R  R  W  M  *  S

DNA: AAATGTTGGAGATGTCCTAGGAGTTTTAACTACAGAGGAAAACGAAAATAT
 +3:   M  L  E  M  S  *  E  F  *  L  Q  R  K  T  K  I  W
 +2:  N  V  G  D  V  L  G  V  L  T  T  E  E  N  E  N  M
 +1: K  C  W  R  C  P  R  S  F  N  Y  R  G  K  R  K  Y

DNA: GGATGAAAAGAAATATAATGATGGGGACATTAACAAGACAGAGAACGAAAT
 +3:   M  K  R  N  I  M  M  G  T  L  T  R  Q  R  T  K  *
 +2:  D  E  K  K  Y  N  D  G  D  I  N  K  T  E  N  E  I
 +1: G  *  K  E  I  *  *  W  G  H  *  Q  D  R  E  R  N

DNA: AAAAGTATTGAATCCTGATAAAGACAAAAGTGAACAAATTATAAAAGAAGA
 +3:   K  Y  *  I  L  I  K  T  K  V  N  K  L  *  K  K  I
 +2:  K  V  L  N  P  D  K  D  K  S  E  Q  I  I  K  E  D
 +1: K  S  I  E  S  *  *  R  Q  K  *  T  N  Y  K  R  R

DNA: TATTCATTTTGTAAAGAAGCATATAAATGATGATGTAAATGAAGAGAAAAT
 +3:   F  I  L  *  R  S  I  *  M  M  M  *  M  K  R  K  Y
 +2:  I  H  F  V  K  K  H  I  N  D  D  V  N  E  E  K  I
 +1: Y  S  F  C  K  E  A  Y  K  *  *  C  K  *  R  E  N

DNA: ATTTATTCCTTTTATAAAGTGTAAAAAAAAGAAAGCAAAAATAAATAAATG
 +3:   L  F  L  L  *  S  V  K  K  R  K  Q  K  *  I  N  G
 +2:  F  I  P  F  I  K  C  K  K  K  K  A  K  I  N  K  W
 +1: I  Y  S  F  Y  K  V  *  K  K  E  S  K  N  K  *  M

DNA: GTTAAAAAATGAGAATGATTTTGTAAAGAAGAATGATTTGTTACTTTATGT
 +3:   *  K  M  R  M  I  L  *  R  R  M  I  C  Y  F  M  *
 +2:  L  K  N  E  N  D  F  V  K  K  N  D  L  L  L  Y  V
 +1: V  K  K  *  E  *  F  C  K  E  E  *  F  V  T  L  C

DNA: AGAGGATGATAAAAGCACCATAGAAGTGGAAAGTCCATATTCTGGTAATAA
 +3:   R  M  I  K  A  P  *  K  W  K  V  H  I  L  V  I  T
 +2:  E  D  D  K  S  T  I  E  V  E  S  P  Y  S  G  N  N
 +1: R  G  *  *  K  H  H  R  S  G  K  S  I  F  W  *  *

DNA: CCCTTTGAAGTATATATGATCAAATCATATAAACAAAATGTATACATATAT
 +3:   L  *  S  I  Y  D  Q  I  I  *  T  K  C  I  H  I  Y
 +2:  P  L  K  Y  I  *  S  N  H  I  N  K  M  Y  T  Y  I
 +1: P  F  E  V  Y  M  I  K  S  Y  K  Q  N  V  Y  I  Y

DNA: ATATATATATATATATATTTTTTTTTTGTGTTGAATTTTTCCTTTATTTAAT
 +3:   I  Y  I  Y  I  F  F  F  V  L  N  F  S  F  I  *  F
 +2:  Y  I  Y  I  Y  I  F  F  C  V  E  F  F  L  Y  L  I
 +1: I  Y  I  Y  I  Y  F  F  L  C  *  I  F  P  L  F  N

DNA: TTGTACGTATAATAAAAAAATTATTAGTCAAGGAAGGACAGTTCGTAGATT
 +3:   V  G  I  I  K  K  L  L  V  K  E  G  Q  F  V  D  L
 +2:  C  R  Y  N  K  K  I  I  S  Q  G  R  T  V  R  R  F
 +1: L  *  V  *  *  K  N  Y  *  S  R  K  D  S  S  *  I

DNA: TAGACAAAGAAGTTGCCATCATCTCAATAACAGAGGTGAGATATACAAATA
 +3:   D  K  E  V  A  I  I  S  I  T  E  V  R  Y  T  N  I
 +2:  R  Q  R  S  C  H  H  L  N  N  R  G  E  I  Y  K  Y
 +1: *  T  K  K  L  P  S  S  Q  *  Q  R  *  D  I  Q  I

DNA: TAAATACTTTCTATAAATTCTATATTCTTTATTATATATGATGATTTTTTT
 +3:   N  T  F  Y  K  F  Y  I  L  Y  Y  I  *  *  F  F  Y
 +2:  K  Y  F  L  *  I  L  Y  S  L  L  Y  M  M  I  F  L
 +1: *  I  L  S  I  N  S  I  F  F  I  I  Y  D  D  F  F

DNA: ATTTTTGTTTACAGGAAAAAGATAATGAAAAAGAAAAAATAGAAGAACCTT
 +3:   F  C  L  Q  E  K  D  N  E  K  E  K  I  E  E  P  F
 +2:  F  L  F  T  G  K  R  *  *  K  R  K  N  R  R  T  F
 +1: I  F  V  Y  R  K  K  I  M  K  K  K  K  *  K  N  L

DNA: TTAAAAATAAGTAAAGAAATATATTAAGTTTGATAATATATATATATATAT
 +3:   K  N  K  *  R  N  I  L  S  L  I  I  Y  I  Y  I  Y
 +2:  *  K  *  V  K  K  Y  I  K  F  D  N  I  Y  I  Y  I
 +1: L  K  I  S  K  E  I  Y  *  V  *  *  Y  I  Y  I  Y

DNA: ATATATATATTTTTTTTTTTCTTTATATTATTAATTGTTATTATTTTATGG
 +3:   I  Y  I  F  F  F  L  Y  I  I  N  C  Y  Y  F  M  D
 +2:  Y  I  Y  F  F  F  S  L  Y  Y  *  L  L  L  F  Y  G
 +1: I  Y  I  F  F  F  F  F  I  L  L  I  V  I  I  L  W
```

Exon 3

Intron 3

Exon 4

Intron 4

Exon 5

Intron 5

```
DNA: ATTATCAGAGAAGATGAAGAAATAAATCGAGATAATATATTAATACATTAT
 +3:   Y  Q  R  R  *  R  N  K  S  R  *  Y  I  N  T  L  Y
 +2:  L  S  E  K  M  K  K  *  I  E  I  I  Y  *  Y  I  I
 +1: I  I  R  E  D  E  E  I  N  R  D  N  I  L  I  H  Y

DNA: ATAAATAAAATCAAGAAAAGTGAAGAGGGAAGAAAGTTTTTAAAAAATTTA
 +3:   K  *  N  Q  E  K  *  R  G  K  K  V  F  K  K  F  K          Exon 6
 +2:  *  I  K  S  R  K  V  K  R  E  E  S  F  *  K  I  *
 +1: I  N  K  I  K  K  S  E  E  G  R  K  F  L  K  N  L

DNA: AGGTACTCATAAAGTTTTACAATTGTTTATAAATGGGATGTTATATATATA
 +3:   V  L  I  K  F  Y  N  C  L  *  M  G  C  Y  I  Y  I
 +2:  G  T  H  K  V  L  Q  L  F  I  N  G  M  L  Y  I  Y
 +1: R  Y  S  *  S  F  T  I  V  Y  K  W  D  V  I  Y  I

DNA: TATATATATATATTTATTTATTTATTTATTTATTTATTTTTTATTATTATTT  Intron 6
 +3:   Y  I  Y  I  Y  L  F  I  Y  L  F  I  F  I  I  I  F
 +2:  I  Y  I  Y  L  F  I  Y  L  F  I  Y  F  Y  Y  Y  F
 +1: Y  I  Y  I  F  I  Y  L  F  I  Y  L  F  L  L  L  F

DNA: TTTTTTTTTATATATATATAAGTGAACAAGAAGAAAAAACACTGGAAGAGA
 +3:   F  F  Y  I  Y  I  S  E  Q  E  E  K  T  L  E  E  R
 +2:  F  F  L  Y  I  Y  K  *  T  R  R  K  N  T  G  R  E
 +1: F  F  F  I  Y  I  *  V  N  K  K  K  K  H  W  K  R

DNA: GGCTCAAATTAAATTATGAGAAATACAATAAAATTTCCAATGATCTATTCA  Exon 7
 +3:   L  K  L  N  Y  E  K  Y  N  K  I  S  N  D  L  F  R
 +2:  A  Q  I  K  L  *  E  I  Q  *  N  F  Q  *  S  I  Q
 +1: G  S  N  *  I  M  R  N  T  I  K  F  P  M  I  Y  S

DNA: GGTATATGACATATGCATAATATATACTCAATATTTAATATTCTTATAAAA
 +3:   Y  M  T  Y  A  *  Y  I  L  N  I  *  Y  S  Y  K  N
 +2:  V  Y  D  I  C  I  I  Y  T  Q  Y  L  I  F  L  *  K
 +1: G  I  *  H  M  H  N  I  Y  S  I  F  N  I  L  I  K  Intron 7

DNA: ATGTTGAAATTTATTTATTTATTTATTTATTTATTTTTTTTTATTATTATT
 +3:   V  E  I  Y  L  F  I  Y  L  F  I  F  F  Y  Y  Y  S
 +2:  C  *  N  L  F  I  Y  L  F  I  Y  F  F  L  L  L  F
 +1: M  L  K  F  I  Y  L  F  I  Y  L  F  F  F  I  I  I

DNA: CAGGTCTAGTGAAAGTACAAAAGATTATGTATTAAAAGAAAAGGAAGTAAA  Exon 8
 +3:   G  L  V  K  V  Q  K  I  M  Y  *  K  K  R  K  *  I
 +2:  R  S  S  E  S  T  K  D  Y  V  L  K  E  K  E  V  N
 +1: Q  V  *  *  K  Y  K  R  L  C  I  K  R  K  G  S  K

DNA: TACACAGAAAGATGAATATATAAATAAATTTGCAAAAATATTTTCATATAT  Intron 8
 +3:   H  R  K  M  N  I  *  I  N  L  Q  K  Y  F  H  I  Y
 +2:  T  Q  K  D  E  Y  I  N  K  F  A  K  I  F  S  Y  I
 +1: Y  T  E  R  *  I  Y  K  *  I  C  K  N  I  F  I  Y

DNA: ATATATATATATATCATATATTTTTTTCATTTGTTTGTATCTATAGAATGA
 +3:   I  Y  I  Y  H  I  F  F  S  F  V  C  I  Y  R  M  R
 +2:  Y  I  Y  I  S  Y  I  F  F  I  C  L  Y  L  *  N  E
 +1: I  Y  I  Y  I  I  Y  F  F  H  L  F  V  S  I  E  *

DNA: GAGCCAATATGAAATGGTGTTACCATCTGCATCAGAGTTGATGAGACAAAA  Exon 9
 +3:   A  N  M  K  W  C  Y  H  L  H  Q  S  *  *  D  K  T
 +2:  S  Q  Y  E  M  V  L  P  S  A  S  E  L  M  R  Q  N
 +1: E  P  I  *  N  G  V  T  I  C  I  R  V  D  E  T  K

DNA: CAAATTAAATCCAAAGGATATAACGTAATATATATATATATATTTATATTT
 +3:   N  *  I  Q  R  I  *  R  N  I  Y  I  Y  I  Y  I  Y  Intron 9
 +2:  K  L  N  P  K  D  I  T  *  Y  I  Y  I  Y  L  Y  L
 +1: Q  I  K  S  K  G  Y  N  V  I  Y  I  Y  I  F  I  F

DNA: ATATATATTTATATATTTATATTTATACCTTTTATAGAAACAGGAAAACAC
 +3:   I  Y  L  Y  I  Y  I  Y  T  F  Y  R  N  R  K  T  P
 +2:  Y  I  F  I  Y  L  Y  L  Y  L  L  *  K  Q  E  N  T
 +1: I  Y  I  Y  I  F  I  F  I  P  F  I  E  T  G  K  H

DNA: CCAATCGTATAACCTATGAAGACGTTGATGCATTTTTAAATGGACATAAAA
 +3:   N  R  I  T  Y  E  D  V  D  A  F  L  N  G  H  K  N
 +2:  Q  S  Y  N  L  *  R  R  *  C  I  F  K  W  T  *  K  Exon 10
 +1: P  I  V  *  P  M  K  T  L  M  H  F  *  M  D  I  K

DNA: ATAATTCTACTAATGTTACTTACTGCGAAAAACCAAAAGTTGAAACAATTG
 +3:   N  S  T  N  V  T  Y  C  E  K  P  K  V  E  T  I  E
 +2:  *  F  Y  *  C  Y  L  L  R  K  T  K  S  *  N  N  *
 +1: I  I  L  L  M  L  L  T  A  K  N  Q  K  L  K  Q  L
```

```
DNA: AATATGGAGATCCAAAAACTGTCGATATGACAAATATACAGAAATCAATAA
 +3:   Y  G  D  P  K  T  V  D  M  T  N  I  Q  K  S  I  K
 +2:  I  W  R  S  K  N  C  R  Y  D  K  Y  T  E  I  N  K
 +1: N  M  E  I  Q  K  L  S  I  *  Q  I  Y  R  N  Q  *

DNA: AAAATAATATGATGCTTACCTTAACCGTTCCAGTGTTCCGTGTTACTCATT
 +3:   N  N  M  M  L  T  L  T  V  P  V  F  R  V  T  H  L
 +2:  K  *  Y  D  A  Y  L  N  R  S  S  V  P  C  Y  S  F
 +1: K  I  I  *  C  L  P  *  P  F  Q  C  S  V  L  L  I

DNA: TAATAAAAACAAACGAATTACTGAAATTATATGAAAAAGTAAAACAAAAAA
 +3:   I  K  T  N  E  L  L  K  L  Y  E  K  V  K  Q  K  I
 +2:  N  K  N  K  R  I  T  E  I  I  *  K  S  K  T  K  N
 +1: *  *  K  Q  T  N  Y  *  N  Y  M  K  K  *  N  K  K

DNA: TTAGTATGAGTGTTATAATAAATAAATGTGTATCATCTGTACTATTAAATC
 +3:   S  M  S  V  I  I  N  K  C  V  S  S  V  L  L  N  H
 +2:  *  Y  E  C  Y  N  K  *  M  C  I  I  C  T  I  K  S        Exon 10
 +1: L  V  *  V  L  *  *  I  N  V  Y  H  L  Y  Y  *  I

DNA: ATCCCTTGATATACTCTACTTATATTGATAAAGATAATGGAAAAATATTAT
 +3:   P  L  I  Y  S  T  Y  I  D  K  D  N  G  K  I  L  Y
 +2:  S  L  D  I  L  Y  L  Y  *  *  R  *  W  K  N  I  I
 +1: I  P  *  Y  T  L  L  I  L  I  K  I  M  E  K  Y  Y

DNA: ATAATAAGGATGTTAATATAGGAAATGCATTAGGATTACCAGATTCTCTAT
 +3:   N  K  D  V  N  I  G  N  A  L  G  L  P  D  S  L  L
 +2:  *  *  G  C  *  Y  R  K  C  I  R  I  T  R  F  S  I
 +1: I  I  R  M  L  I  *  E  M  H  *  D  Y  Q  I  L  Y

DNA: TAACTCCTGTGTTAAAAAAAGTTGATAAAAAGGATATATATACATTGGCCA
 +3:   T  P  V  L  K  K  V  D  K  K  D  I  Y  T  L  A  N
 +2:  N  S  C  V  K  K  S  *  *  K  G  Y  I  Y  I  G  Q
 +1: *  L  L  C  *  K  K  L  I  K  R  I  Y  I  H  W  P

DNA: ATGAATGGAAGGTTTGTAAAATGCGTATATAAATTGAATTATATATATCCA
 +3:   E  W  K  V  C  K  M  R  I  *  I  E  L  Y  I  S  I
 +2:  *  M  E  G  L  *  N  A  Y  I  N  *  I  I  Y  I  H
 +1: M  N  G  R  F  V  K  C  V  Y  K  L  N  Y  I  Y  P

DNA: TTTTTTTTAATATATAAAAAGTTTATAAGAAATATTGTCTTCATGTATAAT
 +3:   F  F  N  I  *  K  V  Y  K  K  Y  C  L  H  V  *  L
 +2:  F  F  *  Y  I  K  S  L  *  E  I  L  S  S  C  I  I
 +1: F  F  L  I  Y  K  K  F  I  R  N  I  V  F  M  Y  N

DNA: TATGAACAATTTTCATCTTCCACGTAAACGGTGCACCACAATATTATATAT        Intron 10
 +3:   *  T  I  F  I  F  H  V  N  G  A  P  Q  Y  Y  I  Y
 +2:  M  N  N  F  H  L  P  R  K  R  C  T  T  I  L  Y  I
 +1: Y  E  Q  F  S  S  S  T  *  T  V  H  H  N  I  I  Y

DNA: ATATATATATATGTATATGTATATTTACATTTATATTTATTTTTTAATATG
 +3:   I  Y  I  C  I  C  I  F  T  F  I  F  I  F  *  Y  V
 +2:  Y  I  Y  M  Y  M  Y  I  Y  I  Y  I  Y  F  L  I  C
 +1: I  Y  I  Y  V  Y  V  Y  L  H  L  Y  L  F  F  N  M

DNA: TTTATTATTATTATTATTATTATTATTTTTTTTTTTTTTTTTTTTTTTT
 +3:   Y  Y  Y  Y  Y  Y  Y  Y  Y  F  F  F  F  F  F  F  F
 +2:  L  L  L  L  L  L  L  L  L  F  F  F  F  F  F  F  F
 +1: F  I  I  I  I  I  I  I  I  I  F  F  F  F  F  F  F

DNA: TTTTTTTTTTTTCCAATTTTGTAGATACTAGTTGAGAAAGGTAAAAATGGT
 +3:   F  F  F  S  N  F  V  D  T  S  *  E  R  *  K  W  S
 +2:  F  F  F  F  Q  F  C  R  Y  *  L  R  K  V  K  M  V
 +1: F  F  F  F  P  I  L  *  I  L  V  E  K  G  K  N  G

DNA: CTCTTAAGTTCTAATGATATGACAGGTAGTAATTTTTACATTTCCAATTTG
 +3:   L  K  F  *  *  Y  D  R  *  *  F  L  H  F  Q  F  G
 +2:  S  *  V  L  M  I  *  Q  V  V  I  F  T  F  P  I  W
 +1: L  L  S  S  N  D  M  T  G  S  N  F  Y  I  S  N  L

DNA: GGAATGTTCAATACTTACCAATTTGATGCAATATTGCCAAAGAATTCATCA        Exon 11
 +3:   N  V  Q  Y  L  P  I  *  C  N  I  A  K  E  F  I  M
 +2:  E  C  S  I  L  T  N  L  M  Q  Y  C  Q  R  I  H  H
 +1: G  M  F  N  T  Y  Q  F  D  A  I  L  P  K  N  S  S

DNA: TGTATCTTATCAATCGGCACAAATATTGGAAGTATTGATAACCTGGAAGAC
 +3:   Y  L  I  N  R  H  K  Y  W  K  Y  *  *  P  G  R  L
 +2:  V  S  Y  Q  S  A  Q  I  L  E  V  L  I  T  W  K  T
 +1: C  I  L  S  I  G  T  N  I  G  S  I  D  N  L  E  D
```

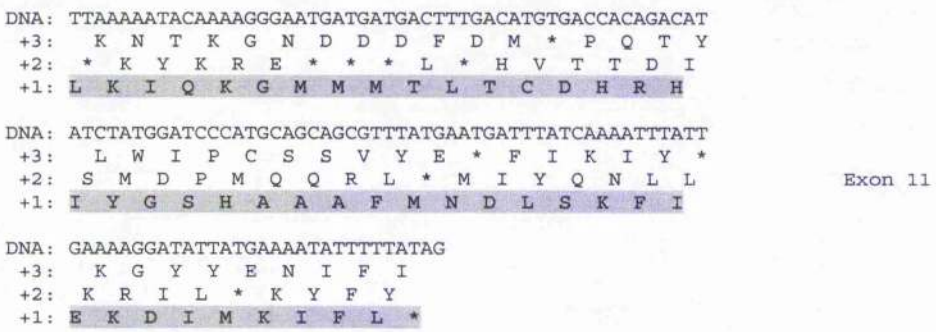


Figure 5.11 – The intron/exon boundaries of *P. falciparum* PDH-E2

The genomic DNA sequence of the *PDH-E2* gene is displayed with the resulting amino acid sequence shown in three frames below it. The predicted intron/exon boundaries from the PlasmoDB annotation (PF10_0407) are displayed in red, with resulting exons highlighted in grey.

Number	Length (bp)	A/T content (%)
Exon 1	243	77.0
Intron 1	209	92.3
Exon 2	67	61.2
Intron 2	104	82.7
Exon 3	390	72.1
Intron 3	115	83.5
Exon 4	80	68.8
Intron 4	81	84.0
Exon 5	47	83.0
Intron 5	100	90.0
Exon 6	96	80.2
Intron 6	122	91.0
Exon 7	81	71.6
Intron 7	110	86.4
Exon 8	37	78.4
Intron 8	102	85.3
Exon 9	80	66.2
Intron 9	64	93.8
Exon 10	484	73.3
Intron 10	268	85.8
Exon 11	312	69.9

Table 5.1 – Intron/exon boundaries of PDH E2

The *PDH E2* gene is predicted to contain ten introns that produce a gene product coded for by eleven exons. The size of the introns and exons are shown above along with the A/T content, and this displays that the intron sequences have a higher A/T content that the exon sequences.

```
DNA: ATGTTATACAACTTAATTATATTAATTTTTTATTTAAGATTTTCTAAATGTATTTCTAAG   60
 +1: M  L  Y  N  L  I  I  L  I  F  Y  L  R  F  S  K  *  I  S  K    20

DNA: AACAATAATTACGGTTACATCAACTTCGGTACCTTTTCAAATGTTGTAAACAACAGTAAT  120
 +1: N  N  N  Y  G  Y  I  N  F  G  T  F  S  N  V  V  N  N  S  N    40

DNA: AATTTAAGAAATAGAAAAAATGTTGTTTTTTTCAAAAATAGAAATAAAAATGCCAGCTCTA  180
 +1: N  L  R  N  R  K  N  V  V  F  S  K  I  E  I  K  M  P  A  L    60

DNA: TCTAGTACCATGACGACAGGCAAAATTGTTAAATGGAATAAAAATATAGGAGATTATGTA  240
 +1: S  S  T  M  T  T  G  K  I  V  K  W  N  K  N  I  G  D  Y  V    80

DNA: AATCTAGGTGATATTATAATGACCGTTGAAAGTGATAAAGCAGACATGGACGTGGAAGCA  300
 +1: N  L  G  D  I  I  M  T  V  E  S  D  *  A  D  M  D  V  E  A   100

DNA: TTTGATGAAGGCTTTTTAAGGGTTAAGCGTTTAGAAGATGGATGTGAAGCAAATGTTGGA  360
 +1: F  D  E  G  F  L  R  V  K  R  L  E  D  G  C  E  A  N  V  G   120

DNA: GATGTCCTAGGAGTTTTAACTACAGAGGAAAACGAAAATATGGATGAAAAGAAATATAAT  420
 +1: D  V  L  G  V  L  T  T  E  E  N  E  N  M  D  E  K  K  Y  N   140

DNA: GATGGGGACATTAACAAGACAGAGAACGAAATAAAAGTATTGAATCCTGATAAAGACAAA  480
 +1: D  G  *  I  N  K  T  E  N  E  I  K  V  L  N  P  D  K  D  K   160

DNA: AGTGAACAAATTATAAAAGAAGATATTCATTTTGTAAAGAAGCATATAAATGATGATGTA  540
 +1: S  E  Q  I  I  K  E  D  I  H  F  V  K  K  H  I  N  D  D  V   180

DNA: AATGAAGAGAAAATATTTATTCCTTTTATAAAGTGTAAAAAAAAGAAAGCAAAAATAAAT  600
 +1: N  E  E  K  I  F  I  P  F  I  K  C  K  K  K  K  A  K  I  N   200

DNA: AAATGGTTAAAAAATGAGAATGATTTTGTAAAGAAGAATGATTTGTTACTTTATGTAGAG  660
 +1: K  W  L  K  N  E  N  D  F  V  K  K  N  D  L  L  L  Y  V  E   220

DNA: GATGATAAAAGCACCATAGAAGTGGAAAGTCCATATTCTGGTATAATAAAAAAATTATTA  720
 +1: D  D  *  S  T  I  E  V  E  S  P  Y  S  G  I  I  K  K  L  L   240

DNA: GTCAAGGAAGGACAGTTCGTAGATTTAGACAAAGAAGTTGCCATCATCTCAATAACAGAG  780
 +1: V  K  E  G  Q  F  V  D  L  D  K  E  V  A  I  I  S  I  T  E   260

DNA: GAAAAAGATAATGAAAAAGAAAAAATAGAAGAACCTTTTAAAAATAAAGAAGATGAAGAA  840
 +1: E  K  D  N  E  K  E  K  I  E  E  P  F  K  N  K  E  D  E  E   280

DNA: ATAAATCGAGATAATATATTAATACATTATATAAATAAAATCAAGAAAAGTGAAGAGGGA  900
 +1: I  N  R  D  N  I  L  I  H  Y  I  N  K  I  K  K  S  E  E  G   300

DNA: AGAAAGTTTTTAAAAAATTTAAGTGAACAAGAAGAAAAAACACTGGAAGAGAGGCTCAAA  960
 +1: R  K  F  L  K  N  L  S  E  Q  E  E  K  T  L  E  E  R  L  K   320

DNA: TTAAATTATGAGAAATACAATAAAATTTCCAATGATCTATTCAGGTCTAGTGAAAGTACA 1020
 +1: L  N  Y  E  K  Y  N  K  I  S  N  D  L  F  R  S  S  E  S  T   340

DNA: AAAGATTATGTATTAAAAGAAAAGGAAAATGAGAGCCAATATGAAATGGTGTTACCATCT 1080
 +1: K  D  Y  V  L  K  E  K  E  N  E  S  Q  Y  E  M  V  L  P  S   360

DNA: GCATCAGAGTTGATGAGACAAAACAAATTAAATCCAAAGGATATAACAAACAGGAAAACA 1140
 +1: A  S  E  L  M  R  Q  N  K  L  N  P  K  D  I  T  N  R  K  T   380

DNA: CCCAATCGTATAACCTATGAAGACGTTGATGCATTTTTAAATGGACATAAAAATAATTCT 1200
 +1: P  N  R  I  T  Y  E  D  V  D  A  F  L  N  G  H  K  N  N  S   400

DNA: ACTAATGTTACTTACTGCGAAAAACCAAAAGTTGAAACAATTGAATATGGAGATCCAAAA 1260
 +1: T  N  V  T  Y  C  E  K  P  K  V  E  T  I  E  Y  G  D  P  K   420

DNA: ACTGTCGATATGACAAATATACAGAAATCAATAAAAAATAATATGATGCTTACCTTAACC 1320
 +1: T  V  D  M  T  N  I  Q  K  S  I  K  N  N  M  M  L  T  L  T   440

DNA: GTTCCAGTGTTCCGTGTTACTCATTTAATAAAAACAAACGAATTACTGAAATTATATGAA 1380
 +1: V  P  V  F  R  V  T  H  L  I  K  T  N  E  L  L  K  L  Y  E   460

DNA: AAAGTAAAACAAAAAATTAGTATGAGTGTTATAATAAATAAATGTGTATCATCTGTACTA 1440
 +1: K  V  K  Q  K  I  S  M  S  V  I  I  N  K  C  V  S  S  V  L   480

DNA: TTAAATCATCCCTTGATATACTCTACTTATATTGATAAAGATAATGGAAAAATATTATAT 1500
 +1: L  N  H  P  L  I  Y  S  T  Y  I  D  K  D  N  G  K  I  L  Y   500

DNA: AATAAGGATGTTAATATAGGAAATGCATTAGGATTACCAGATTCTCTATTAACTCCTGTG 1560
 +1: N  K  D  V  N  I  G  N  A  L  G  L  P  D  S  L  L  T  P  V   520

DNA: TTAAAAAAAGTTGATAAAAAGGATATATATACATTGGCCAATGAATGGAAGATACTAGTT 1620
 +1: L  K  K  V  D  K  K  D  I  Y  T  L  A  N  E  W  K  I  L  V   540

DNA: GAGAAAGGTAAAAATGGTCTCTTAAGTTCTAATGATATGACAGGTAGTAATTTTTACATT 1680
 +1: E  K  G  K  N  G  L  L  S  S  N  D  M  T  G  S  N  F  Y  I   560
```

```
DNA: TCCAATTTGGGAATGTTCAATACTTACCAATTTGATGCAATATTGCCAAAGAATTCATCA  1740
 +1: S  N  L  G  M  F  N  T  Y  Q  F  D  A  I  L  P  K  N  S  S    580

DNA: TGTATCTTATCAATCGGCACAAATATTGGAAGTATTGATAACCTGGAAGACTTAAAAATA  1800
 +1: C  I  L  S  I  G  T  N  I  G  S  I  D  N  L  E  D  L  K  I    600

DNA: CAAAAGGGAATGATGATGACTTTGACATGTGACCACAGACATATCTATGGATCCCATGCA  1860
 +1: Q  K  G  M  M  M  T  L  T  C  D  H  R  H  I  Y  G  S  H  A    620

DNA: GCAGCGTTTATGAATGATTTATCAAAATTTATTGAAAAGGATATTATGAAAATATTTTTA  1920
 +1: A  A  F  M  N  D  L  S  K  F  I  E  K  D  I  M  K  I  F  L    640

DNA: TAG                                                           1923
 +1: *                                                              641
```

Figure 5.12 – *Plasmodium falciparum* PDH-E2

The complementary DNA sequence of the *PDH-E2* gene is displayed with the resulting amino acid sequence shown below it. The predicted signal peptide cleavage site (Cys17) is highlighted in blue, with the possible mature protein was predicted to start around Ser51 (highlighted in yellow). The last amino acid (Asp166) in the N-terminal section (amino acids 1-166) used to construct the GFP-fusion protein is highlighted in green. The two sites of lipoylation (Lys93 and Lys223) are highlighted in red. Two residues (Tyr559 and His612) involved in the acylation reaction are highlighted in grey.

Gene	Chromosome	Prediction	Score[a]	Cleavage site
PDH E1α	11	Apicoplast	++[b]	17 or 23[d]
PDH E1β	14	Apicoplast	++[b]	n.d.
PDH E2	10	Apicoplast	++[b]	17[d]
BCKDH E1α	13	Mitochondrial	99.4%[c]	21[c]
BCKDH E1β	5	Mitochondrial	99.5%[c]	n.d.
BCKDH E2	3	Mitochondrial	92.2%[c]	14[c]
KGDH E1	8	Mitochondrial	73.8%[c]	n.d.
KGDH E2A	13	Mitochondrial	3.7 %[c]	n.d.
KGDH E2B	13	Mitochondrial	85.2%[c]	n.d.

Table 5.2 – Subcellular localisation predictions of KADH components

The sequence analyses of the KADH proteins identified in the *P. falciparum* genome predicted the subcellular localisation of the KGDH and BCKDH to the mitochondrion and PDH to the apicoplast. **a.** Predictions were performed using PLASMIT, TARGETP, SIGNALP and PLASMOAP. **b.** Predctions performed by PLASMOAP with the results based on a score of between + (poor prediction) and +++ (high confidence of apicoplast targeting). **c.** Predictions performed in PLASMIT and TARGETP were given a confidence limit, and mitochondrial targeting peptide cleavage site predictions where possible. **d.** Signal peptide cleavage was predicted using SIGNALP.

5.2.2 Subcellular localisation of KADH

The results of the sequence analyses predicted that the components of the KGDH and BCKDH would be localised in the mitochondrion and the single PDH identified would be localised in the apicoplast. These subcellular localisation predictions were experimentally investigated by producing over-expression constructs using the N-terminal targeting peptides of the *KADH* genes and fusing them in frame with the *green fluorescent protein (GFP)* gene. The transfection of these constructs into erythrocytic stages of *P. falciparum* resulted in the expression of GFP-fusion proteins, which should be targeted to the respective location of the respective protein whose targeting sequence was fused to GFP. This technique has been used frequently to analyse subcellular locations of numerous proteins in *P. falciparum* and other organisms and is thought to be a useful tool to obtain insights into the distribution and location of proteins within a cell (Foth *et al.*, 2003; Sato *et al.*, 2004; Tonkin *et al.*, 2004). *P. falciparum* parasites were electroporated to introduce the recombinant DNA into the parasites as described in section 2.2.1.5 and the transfectants were selected for using the pyrimethamine derivative WR99210 (Fidock *et al.*, 1997). The transfected plasmids were expected to be maintained by the parasites episomally as long as selective pressure with the drug was maintained and as a result express the GFP-fusion protein (driven by the HSP86 promoter). After transfection it usually took between 6 to 12 weeks before the first transfectants were observed in Giemsa stained blood smears. Once parasites reached a high enough parasitemia, they were incubated with a mitochondrial marker (MitoTracker CMXRos) as described in section 2.2.1.6 and analysed by fluorescent light microscopy.

5.2.2.1 α–Ketoglutarate dehydrogenase complex

The subcellular localisation of the KGDH was investigated using GFP-fusion constructs of the KGDH-E2 (figure 5.13). The localisation of the KGDH-E2-GFP fusion protein was investigated using three different constructs to analyse which parts of the *N*-terminus would be sufficient to target the fusion protein to its cellular location.

1. KGDH E2A-GFP was constructed using the in-frame methionine contained in intron 2. This resulted in a 165 amino acid long sequence being fused to GFP, with the last amino acid before the GFP protein being equivalent to Valine 195 in the KGDH-E2B sequence (highlighted pink in figure 5.5).

2. KGDH-E2B (C)-GFP was constructed using the KGDH-E2B initiator methionine and was amplified from cDNA. This resulted in a 114 amino acid long sequence being fused to GFP, with the last amino acid before the GFP protein being alanine 114 (highlighted pink in figure 5.5).

3. KGDH E2B (G)-GFP was constructed using the KGDH E2B initiator methionine and was amplified from gDNA. This construct resulted in the fusion of 608 bp of genomic sequence with the KGDH E2 ORF in-frame with GFP. This construct contains both possible initiation sites and would allow the parasite to splice the construct as necessary. The last amino acid before the GFP-fusion protein was equivalent to alanine 114 in the KGDH-E2B sequence (highlighted pink in figure 5.5).

The KGDH-E2A-GFP, which used the in-frame methionine located in the predicted intron 2 of the gene showed targeting to an organelle in the parasites that contained the MitoTracker staining identifying it as the mitochondrion (figure 5.13A). This was surprising because the analyses of the N-terminal amino acid sequence of KGDH-E2A did not predict the presence of a mitochondrial targeting peptide with very high confidence (3.7%). It has to be noted though that most of the parasites that were transfected with this particular construct showed no fluorescence and it appeared that only a very small number of parasites expressed the GFP-fusion protein. This low level of expression was also observed by Western blotting (with Anti-GFP antibodies) where protein samples that were extracted from parasites transfected with the KGDH-E2A-GFP construct displayed no detectable expression of GFP. This is in contrast to protein samples extracted from LipDH1-GFP transfected parasites that clearly expressed GFP (figure 5.13B).

However, the fact that those parasites that displayed GFP-fluorescence actually contained the fusion protein within an organelle which co-stained with the mitochondrial marker suggests that the targeting of a protein in *P. falciparum* to the mitochondrion is not only governed by its *N*-terminal targeting sequence but that also other parts of the amino acid sequence may be responsible for the targeting of the protein to this organelle. Similar results were described in Chapter 4 on the localisation of a LipDH2-GFP fusion protein that apparently did not contain a strong mitochondrial targeting sequence (section 4.2.4).

Transfection of *P. falciparum* erythrocytic stages and expression of KGDH-E2B-GFP containing a construct amplified from parasite cDNA resulted in targeting of the fusion protein to a parasite organelle that was co-stained with MitoTracker (figure 5.13A). Compared with the expression of the fusion protein of KGDH-E2A, this construct appeared to confer expression of the GFP-fusion protein in a greater number of parasites than KGDH-E2A-GFP and the fluorescence observed seemed more consistent and stronger than in the parasites expressing KGDH-E2A-GFP. Transfection of the parasites with KGDH-E2B-GFP generated from genomic DNA also showed targeting of the fusion protein to the mitochondrion as shown by co-staining of the organelle with MitoTracker. The image of the localisation of KGDH-E2B-GFP (G) (figure 5.13A) shows an erythrocyte infected with two trophozoites. Both parasites clearly display the staining with the mitochondrial marker MitoTracker, whereas GFP-fluorescence is only clearly visible in one of the two parasites infecting this particular erythrocyte. This image was chosen because it demonstrates a number of points that need to be made about the expression of GFP-fusion proteins in *P. falciparum* erythrocytic stages. Firstly, the intensity of GFP fluorescence which is likely to represent the level of expression of the fusion protein varies between different parasites and secondly, it seems to happen that parasites surviving the selective pressure apparently do not express the GFP-fusion protein at all. The latter observation suggests that the parasites might have some means to sustain the expression of the selectable marker at sufficiently high levels but at the same time limit the expression of the GFP-fusion protein. Further analyses of the episomes would have been necessary to further address the question what the mechanisms of this differential expression pattern are.

These results provide experimental evidence that the *P. falciparum* KGDH is localised in the mitochondrion and thus confirm the predictions made using various prediction programs (as outlined in section 2.2.2.3). Concomitantly, the mitochondrial localisation of KGDH-E2 also suggests that LipDH2, which was also found to be targeted to the parasite's mitochondrion, is the LipDH-component of this protein complex.

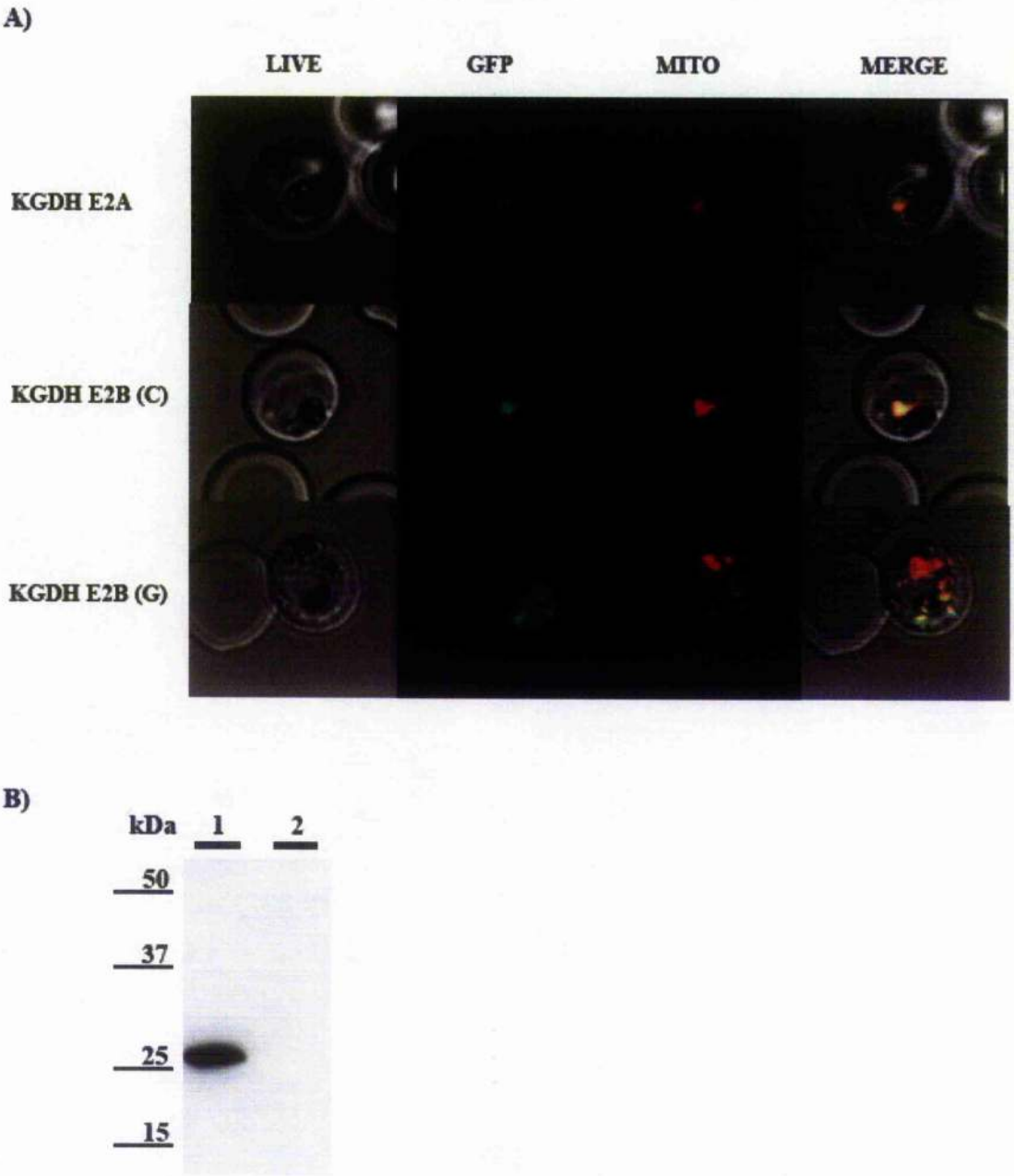


Figure 5.13 – Subcellular localisation of KGDH-E2

Panel A. The subcellular localisation of KGDH-E2 was investigated by three different GFP-fusion proteins. KGDH-E2A used the in-frame methionine within the second intron, and contained the first 165 amino acids from KGDH-E2A. KGDH-E2B (C) used the methionine in exon 1, had the introns removed and contained the first 114 amino acids of KGDH-E2B. KGDH-E2B (G) used the methionine at the start of exon 1, contained both introns and therefore the in-frame methionine. This would allow the KGDH-E2 (G) construct to be spliced as required by the parasite. Parasites mitochondria were selectively stained with Mitotracker CMXRos and images of live cells were obtained by phase contrast (LIVE) or in the GFP/FITC (GFP) or rhodamine (MITO) fluorescent channels. Overlays of the images obtained from each channel are displayed for determination of co-localisation (MERGE). All constructs produced co-localisation of GFP with mitotracker, although the image displayed for KGDH-E2B (G) contains two parasites in a single erythrocyte with only one of the parasites expressing the GFP-fusion protein to a sufficiently high level to be detectable. **Panel B** displays a western blot, using anti-GFP antibodies, of proteins extracted from parasites transfected with either LipDH1-GFP (lane 1) or KGDH-E2A-GFP (lane 2). 10 μg of extracted proteins were loaded on a 4 –12 % gradient gel, with precision all blue protein standards (Biorad) as a molecular mass marker and displays that there was no detectable KGDH-E2A-GFP expression; LipDH1-GFP extracts were loaded as a positive control.

5.2.2.2 Branched chain α–keto acid dehydrogenase complex

The subcellular localisation of the *P. falciparum* BCKDH was analysed by expressing a GFP-fusion constructs of the first 130 and 106 amino acids of the BCKDH E1β and BCKDH E2 proteins, respectively (figure 5.14). Upon fluorescent light microscopical analyses it was found that the expressed fusion proteins co-localised with the mitochondrion of the parasites that were stained with MitoTracker as described previously (section 2.2.1.6; figure 5.14). However, GFP-fluorescence was not exclusively observed to co-localise with the parasite's mitochondrion but was also found in an area that most likely represents the parasite's food vacuole (dark area in the live image of figure 5.14B; although this needs to be experimentally confirmed). This possible food vacuole targeting was observed not only in the parasites expressing BCKDH-E1β GFP-fusion protein, but also in other GFP-fusion proteins investigated in our laboratory (data not shown).

These results give experimental evidence that supports the prediction that *P. falciparum* BCKDH is localised in the mitochondrion of the parasites and suggested that LipDH2 acts as a component of the complex in the organelle. These results therefore also suggested that the single mitochondrial LipDH forms part of both KADH present in the mitochondrion of the parasites.

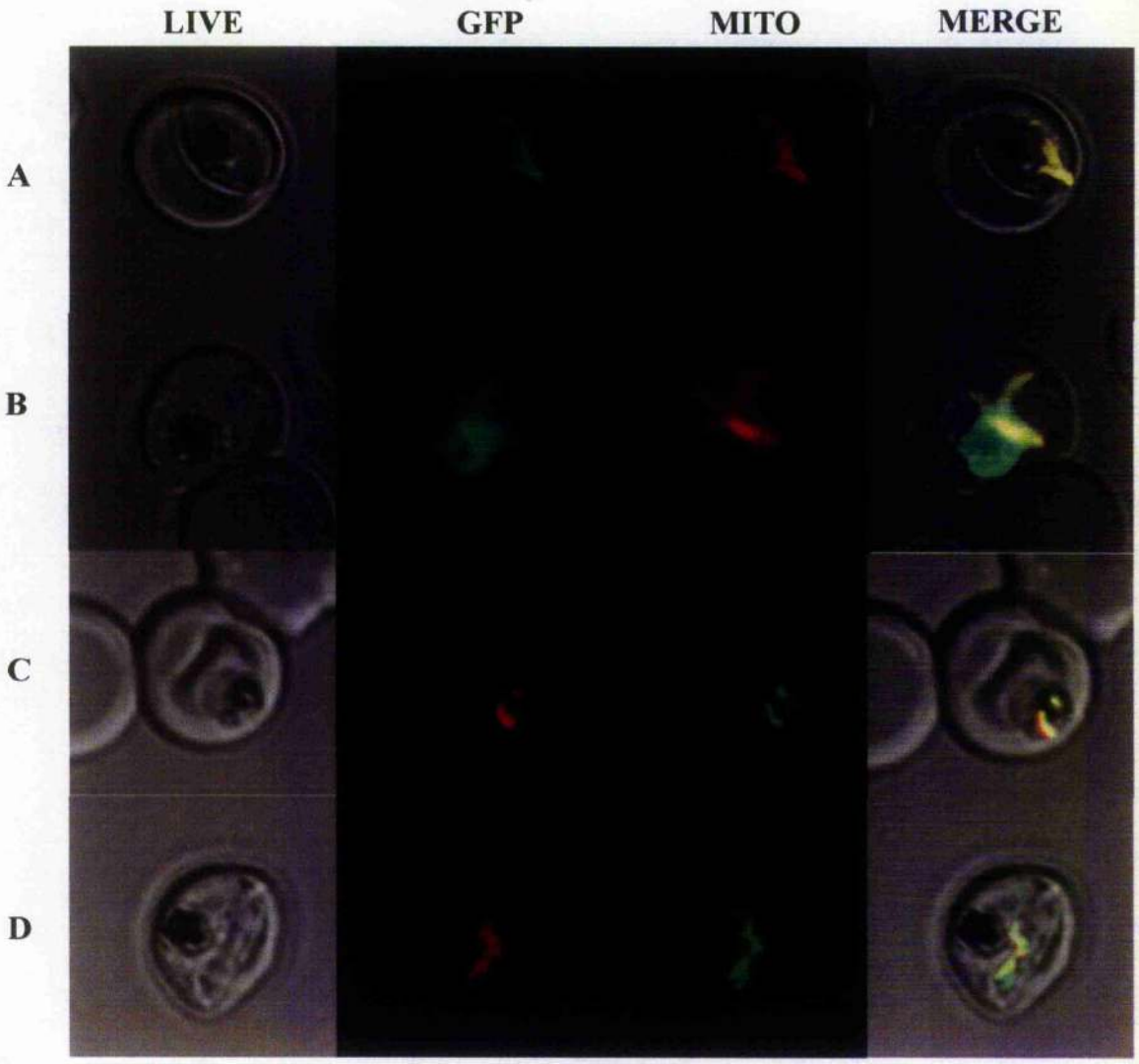


Figure 5.14 – Subcellular localisation of BCKDH

The subcellular localisation BCKDH was investigated by expressing GFP fusion proteins of the E1β and E2 components of the complex in the erythrocytic stage *P. falciparum*. The BCKDH-E1β (**panels A and B**) contained the first 130 amino acids in frame with GFP, while the BCKDH E2 (**panels C and D**) fusion protein contained the first 106 amino acids in frame with GFP. Parasites mitochondria were selectively stained with Mitotracker CMXRos and images of live cells were obtained by phase contrast (LIVE) or in the GFP/FITC (GFP) or rhodamine (MITO) fluorescent channels. Overlays of the images obtained from each channel are displayed for determination of co-localisation (MERGE). All fusion proteins display co-localisation of the GFP with mitotracker, although panel B also displays a GFP staining associated with the food vacuole.

5.2.2.3 Pyruvate dehydrogenase complex

The subcellular localisation of the components of the PDH was analysed by expressing GFP-fusions with N-terminal sequences corresponding to the first 117 and 166 amino acids of PDH-E1α and PDH-E2, respectively (figure 5.15 and figure 5.16). Expression of both fusion proteins led to the green fluorescent staining of a single organelle within the transfected parasites, which was clearly distinct from the mitochondrion but appeared to be closely associated with it. This particular localisation pattern was previously observed when GFP-fusion proteins with apicoplast targeting sequences were analysed by fluorescent light microscopy (Foth *et al.*, 2003; Tonkin *et al.*, 2004; van Dooren *et al.*, 2005). These results together with the *in silico* analyses described in section 5.2.2.3 led me to believe that the components of the single parasite PDH were located in the apicoplast rather than the mitochondrion of *Plasmodium*. This result also suggests that LipDH1, which was also found to be in an organelle closely associated with the parasite's mitochondrion form part of this apicoplast localised PDH complex.

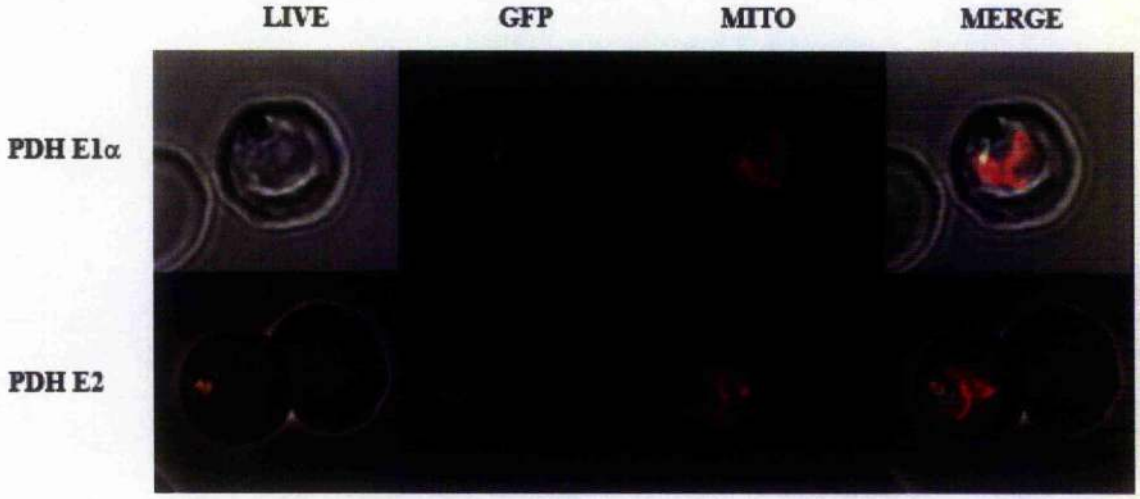


Figure 5.15 – Subcellular localisation of PDH

The subcellular localisation of PDH was investigated by GFP fusion of PDH-E1α and PDH-E2. The GFP-fusion proteins of PDH-E1α (first 117 amino acids) and PDH-E2 (first 166 amino acids) were expressed in the erythrocytic stage of the parasites. Parasites mitochondria were selectively stained with Mitotracker CMXRos and images of live cells were obtained by phase contrast (LIVE) or in the GFP/FITC (GFP) or rhodamine (MITO) fluorescent channels. Overlays of the images obtained from each channel are displayed for determination of co-localisation (MERGE). Both PDH-E1α and PDH-E2 displayed a GFP-staining that was closely associated with the mitochondria, which was consistent with localisation of the GFP to the apicoplast.

5.2.3 Organelle development during the erythrocytic stages

The potential apicoplast localisation of the single PDH of *Plasmodium* was further corroborated by following the organelle's development through the erythrocytic developmental cycle of the parasites. This was done by analysing the changes of GFP-fluorescence in parasites that were expressing the PDH-E1α-GFP (figure 5.16A – D) and LipDH1-GFP (figure 5.16E – J) fusion proteins, respectively. In parallel the development of the mitochondrion through the parasite's erythrocytic cycle was followed using the mitochondrial marker MitoTracker. Thus both the mitochondrion and the apicoplast display selective fluorescence when analysed by fluorescence light microscopy.

In the early stages of parasite development both organelles are small and rounded and appeared to be closely associated (~ between 0 and 15 hours after invasion; figure 5.16A - B). Both organelles elongated during the development to the trophozoite stage (~ between 15 and 28 hours after invasion). In the early trophozoite parasites that were analysed, it was observed that the mitochondrion elongated prior to the apicoplast as demonstrated in figure 5.16 (panels C and D). There also appeared to be less points of contact between the elongating organelles, with commonly only a single point of association (figure 5.16E - F). In the later trophozoites both organelles are elongated and began to branch in preparation for schizogony, with these branched structures exhibiting many points of association (figure 5.16G - H). In schizogony, which occurs between 28 and 48 hours after invasion, the parasite prepares to divide and the branched organelles appear to be "packaged" into the developing daughter merozoites. It is thought that the multiple associations between the two branched structures ensures that each daughter merozoite is equipped with a single mitochondrion and an apicoplast after division (figure 5.16I – J).

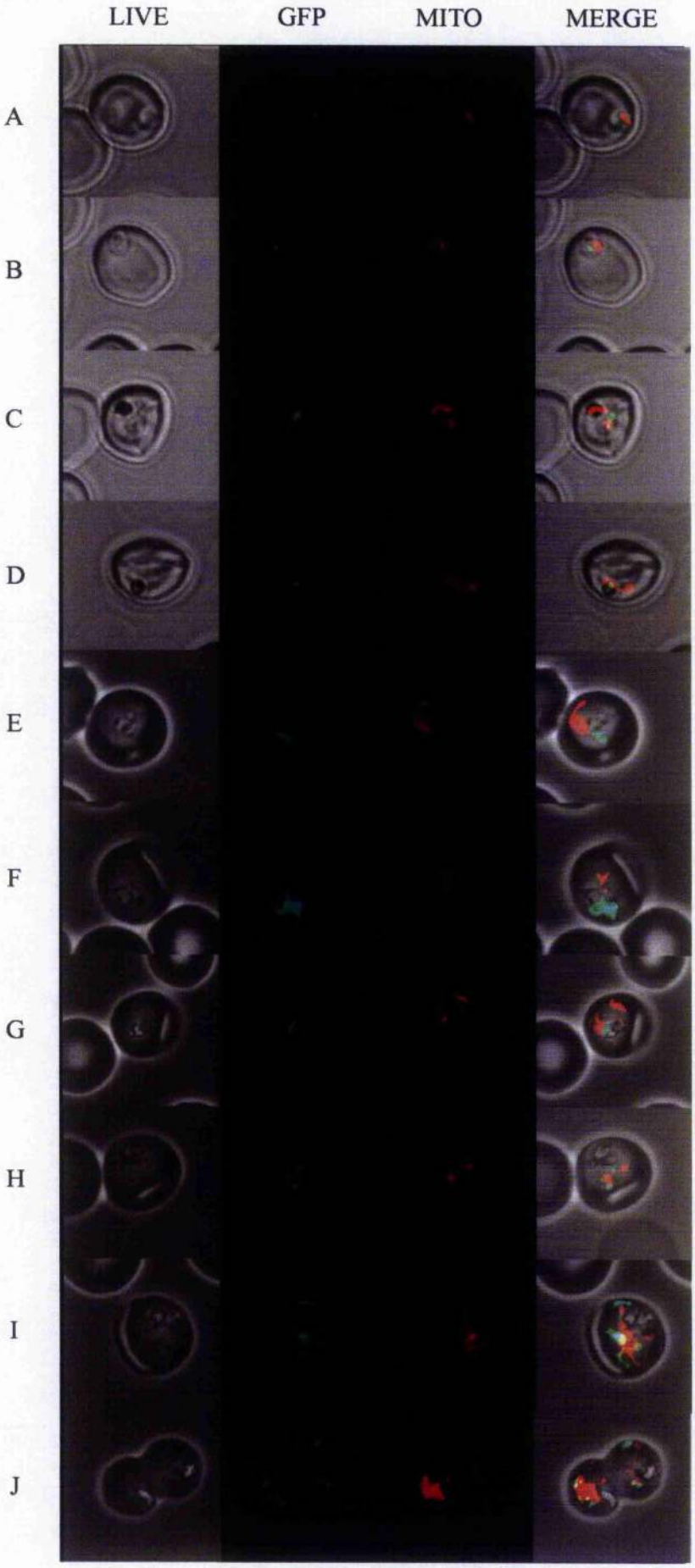

LIVE
GFP
MITO
MERGE
A
B
C
D
E
F
G
H
I
J

Figure 5.16 – Organelle development during the erythrocytic stages

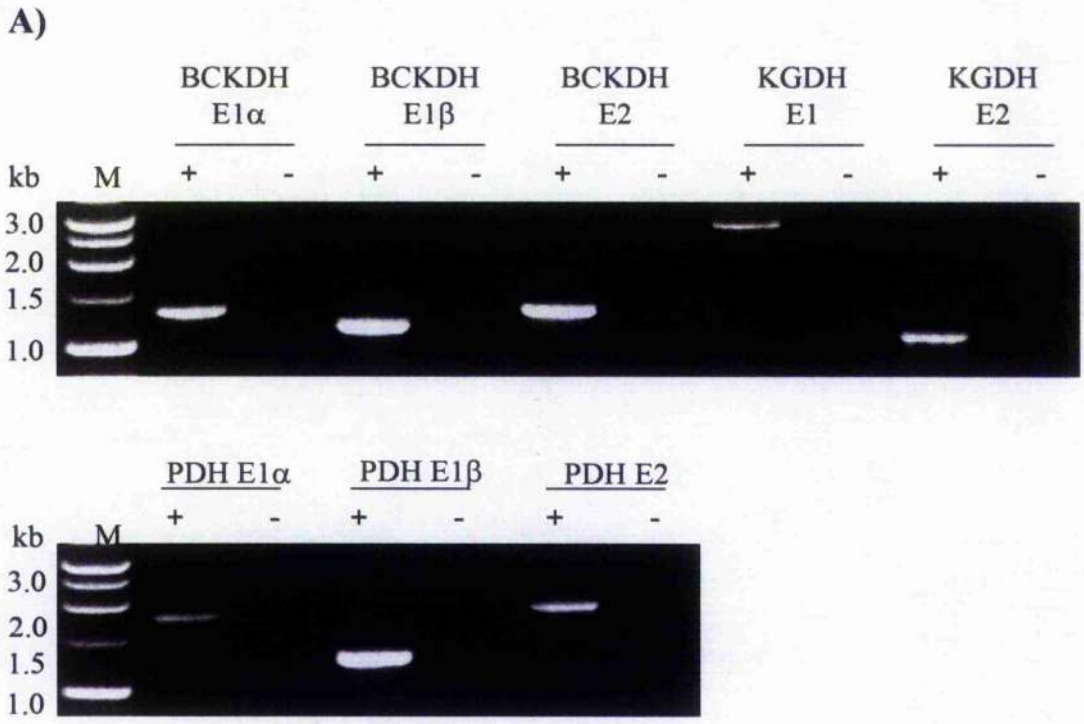
The development of the apicoplast and mitochondria was analysed through the different stages of the erythrocytic life cycle. Both organelles were small and rounded during the ring stage parasites (between 0 and 15 hours after infection) displayed in **panels A and B**; during the trophozoite stage (between 16 and 28 hours) the organelles began to elongate and eventually branching in late trophozoite stages (**panels C – H**); During schizogony (between 28 and 48 hours) the organelles are packaged into the individual daughter cells (**panels I – J**). Apicoplast development was followed by GFP-staining from the PDH-E1α (**panels A – D**) or LipDH1 fusion proteins (**panels E – J**). Development of the mitochondria was followed by staining the organelle with Mitotracker CMXRos. Images of live cells were obtained by phase contrast (LIVE), the GFP/FITC (GFP) or rhodamine (MITO) fluorescent channels. Overlays of the images from each channel are shown to display interactions between the two organelles (MERGE).

5.2.4 Components of the KADH are expressed in erythrocytic stages

In order to test whether the components of the three KADH identified in the *P. falciparum* genome are actually expressed during the erythrocytic life cycle of the parasites, I analysed the transcription and translation of the KADH components by reverse transcriptase PCR (RT-PCR) and western blot analyses.

The transcription of the genes encoding the components of the parasite KGDH, BCKDH and PDH was confirmed by RT-PCR performed on total RNA isolated from cultured parasites (figure 5.17A; lanes marked with +). All PCR products amplified were cloned into TOPO TA and their authenticity was confirmed by DNA sequencing. This experiment also allowed me to confirm the respective start-sites of LipDH2, KGDH-E2 and PDH-E2 and thus the correct annotation of the introns in these genes. Negative controls showed that there was no contamination of the RNA preparation with genomic DNA (figure 5.17A; lanes marked with -).

The protein expression of two components of the parasite KADH during their erythrocytic development was analysed by western blots of proteins isolated from cultured parasites. The antibodies utilised in these analyses were polyclonal rabbit anti-KGDH-E2 and anti-BCKDH-E2 antibodies raised against recombinantly expressed proteins (section 5.2.5 and S. Müller, unpublished data). The KGDH-E2 antibodies identified a single band of an estimated molecular size of 42 kDa, which was in good agreement with the predicted mature protein size of 43 kDa (figure 5.17B; table 5.2). The BCKDH-E2 antibodies identified two bands with estimated molecular sizes of 50 kDa and 30 kDa, respectively. The size of the larger major band was in good agreement with the predicted mature protein size of 49 kDa and the 30 kDa band might either represent a cross reacting protein or a degradation product of BCKDH-E2 that is recognised by the antiserum used (figure 5.17B; table 5.2).



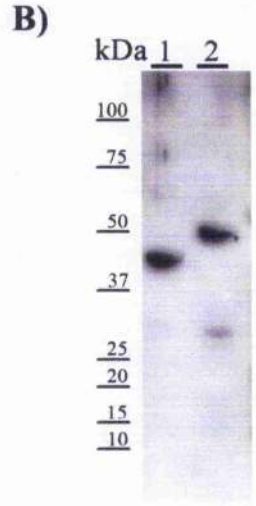


Figure 5.17 – Expression of KADH in *P. falciparum* erythrocytic stages

The transcription and translation of both LipDH genes in the erythrocytic stages was investigated by reverse transcriptase (RT-PCR) (Panel A) and western blot (Panel B). **Panel A** displays the results of RT-PCR on total RNA isolated from cultured parasites, with PCR performed in the presence (+) or absence (-) of reverse transcriptase. Samples amplified in the absence of reverse transcriptase (-) were used as negative controls, in order to test for genomic DNA contamination in RNA preparations. This displays that all the components of the KADH identified are transcribed. The products were analysed by DNA sequencing and confirmed the presence of the predicted introns. **Panel B** displays the results of western blot analyses on 10μg of parasite protein extract using polyclonal antibodies raised against recombinantly expressed KGDH-E2 (Lane 1) and BCKDH-E2 (Lane 2). The KGDH-E2 blot identified a single major band of approximately 42 kDa in size, and the BCKDH E2 blot identified two bands of 50 kDa and 30 kDa. The molecular weight markers used were precision plus all blue markers (Biorad).

5.2.5 Recombinant expression and purification of KGDH-E2

In order to start characterising the different components of the KADH and to generate antibodies against them it was attempted to recombinantly express them in *E.coli* expression systems; I started with the expression of recombinant KGDH-E2.

The recombinant expression construct of KGDH-E2 was designed so that the hydrophobic N-terminal mitochondrial targeting peptide was excluded. The rationale for this was that the mitochondrial targeting peptide potentially could cause the production of insoluble protein in the *E.coli* expression system used. However, the cleavage site for the potential mitochondrial targeting peptide of KGDH-E2B had not been predicted by any of the prediction programs used earlier (table 5.2). This led me to design an expression construct starting at exon 3, which was based on the boundary of homology with KGDH-E2 proteins from other species (data not shown). This construct was amplified by PCR from genomic DNA, cloned into TOPO TA for sequence verification and subsequently sub-cloned into the *E. coli* expression vector pET28 (figure 5.18). Cloning of this 1128 bp expression construct into pET28, which attaches a N-terminal His_6-tag, would result in an expressed protein consisting of 399 amino acids with a predicted size of 44.6 kDa (including the tag).

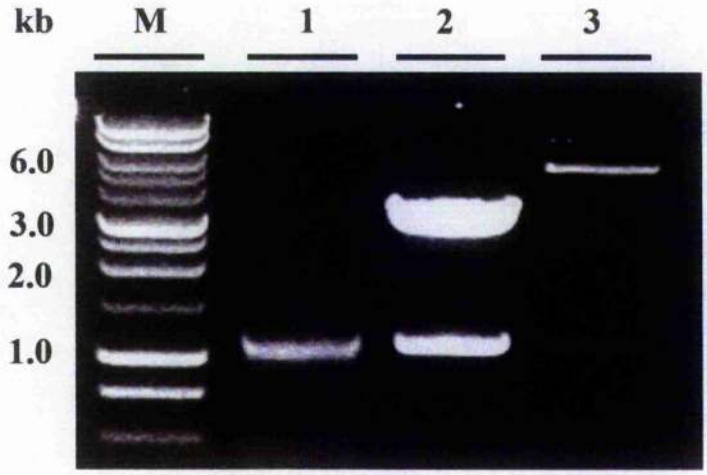


Figure 5.18 – Overview of KGDH-E2 cloning for recombinant expression

The KGDH-E2 expression construct was amplified from gDNA by Taq supermix (Invitrogen) using specific oligonucleotides (table 2.1). Lane M contains 1 kb ladder (Promega); Lane 1 contains the 1.1 kb PCR product; Lane 2 contains the TOPO TA clone used for sequence verification; Lane 3 contains the pET28 cloned expression construct. Both the TOPO TA and pET28 clones were digested with *Nhe* I / *Hind* III to remove the insert.

Test expressions of pET28 cloned KGDH-E2 were carried out to determine the parameters required for optimal recombinant expression. The expression plasmid was transformed into *E. coli* BLR(DE3) cells (Novagen) and test expressions were carried out at two different temperatures of incubation (37°C or 25°C) for varying lengths of expression (between 0 hours and overnight; after induction of expression with 1 mM IPTG). After the respective times of incubation, bacteria were harvested and pellets were lysed using "Bugbuster". Soluble protein and insoluble fractions were separated by centrifugation and recombinant expression of KGDH-E2 was analysed by western blotting using an anti-His-tag antibody. Figure 5.19 shows the western blots of the different samples and it can be seen that the major protein band that is recognised by the anti-His antibody corresponds to an approximately 45 kDa protein. Regardless of the temperature used after induction of recombinant protein expression, it was found that the highest levels of KGDH-E2 expression occurred four hours after induction. Both western blots also displayed proteins reacting with the anti-His-antibodies of higher molecular masses suggesting that KGDH-E2 forms multimers that were not dissolved by the treatment with SDS-sample buffer. This is not unusual given that the E2 subunits of KADH are known to form trimers/multimers, which are the building blocks of the KADH core (Mooney *et al.*, 2002).

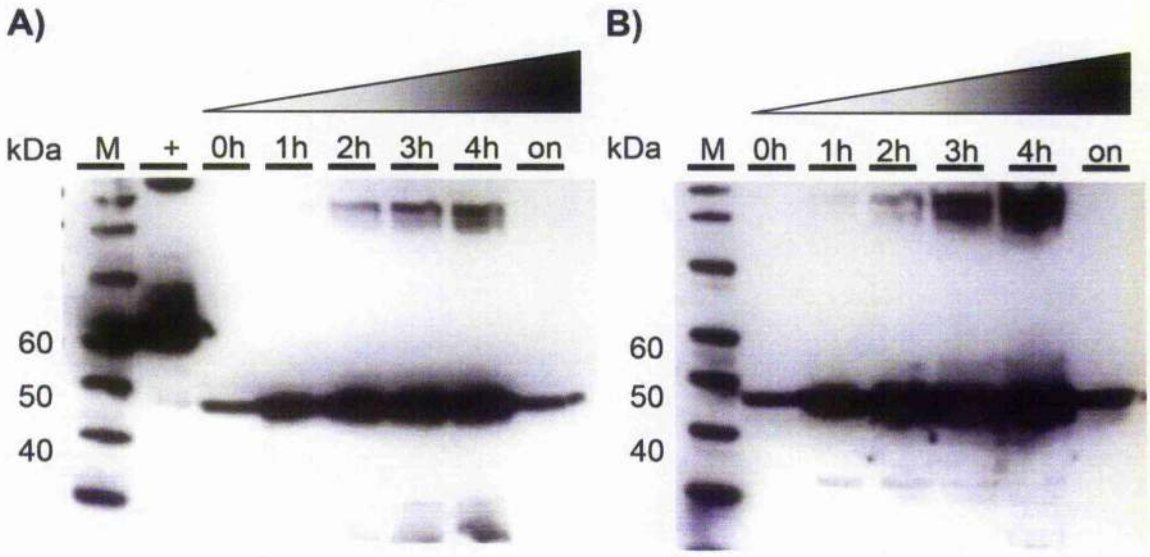


Figure 5.19 – Test expressions of KGDH-E2

Test expressions of pET28 KGDH-E2 were analysed by western blot analyses using anti His-tag antibodies. Test expressions were set up with incubation after induction at either 37°C (**Panel A**) or 25°C (**Panel B**). Samples were taken every hour after induction (0h) for four hours (1h – 4h) and again following overnight expression (on). Lane M contains Magicmark protein standards (Invitrogen); Lane + contains recombinant PfTrxR as a His-tag positive control.

In order to generate sufficient recombinant protein for purification and further characterisation, large-scale expressions were performed at 37°C for four hours following induction of expression. Bacteria were harvested and the recombinantly expressed protein was purified by Ni-NTA batch purification, with the protein eluting between 100 and 250 mM imidazole in elution buffer (figure 5.20). The fractions of the Ni-NTA purification were analysed by SDS-PAGE and western blotting. The SDS-PAGE of the elution fractions (lane 6 to 10) showed the presence of several protein bands of the following sizes: ~ 100 kDa, 53 kDa, 48 kDa and 45 kDa. The 45 kDa band probably corresponds to the recombinantly expressed KGDH-E2 protein. Western blot analysis of the elution fraction obtained with 250 mM imidazole in elution buffer shows that only a single band was reactive with the anti-His-tag antibody (lane 3) suggesting that KGDH-E2 co-purified with non-His-tagged proteins present in the *E.coli* extract (figure 5.20A & B). This result led me to hypothesise that recombinantly expressed KGDH-E2 may be able to form complexes with *E. coli* proteins which cannot be resolved by this purification method. It is known that recombinantly expressed proteins, if not properly folded in the *E. coli* system, may form complexes with bacterial chaperones to enhance their solubility. However, the sizes of the co-purifying proteins may correspond to KGDH components of *E. coli* – the 100 kDa band possibly represents KGDH-E1, the 53 kDa band may potentially be LipDH and the 48 kDa band could be KGDH-E2 from *E. coli*. This suggested that the *P. falciparum* KGDH-E2 recombinant protein might form a KGDH complex with *E. coli* KGDH-E1, E2 and LipDH proteins.

This theory was tested by applying the Ni-NTA elution fraction (250 mM imidazole) to a gel filtration column (Sephadex S-200 16/60) to attempt to separate the recombinantly expressed protein from the co-eluting proteins. It was expected that the recombinant KGDH-E2 should form multimers (most likely a 24mer) and therefore elute in the column's void volume (peak 1 in figure 5.21A; Black line). The void volume peak sample was analysed by SDS-PAGE (figure 5.21B, Lane 1) and still displayed the same protein profile as the originally loaded sample. (figure 5.21, Lane 10). However, a number of other peaks were observed in the elution profile of the protein corresponding to the following sizes: 417 kDa (peak 2), 200 kDa (peak 3) and 47 kDa (peak 5).

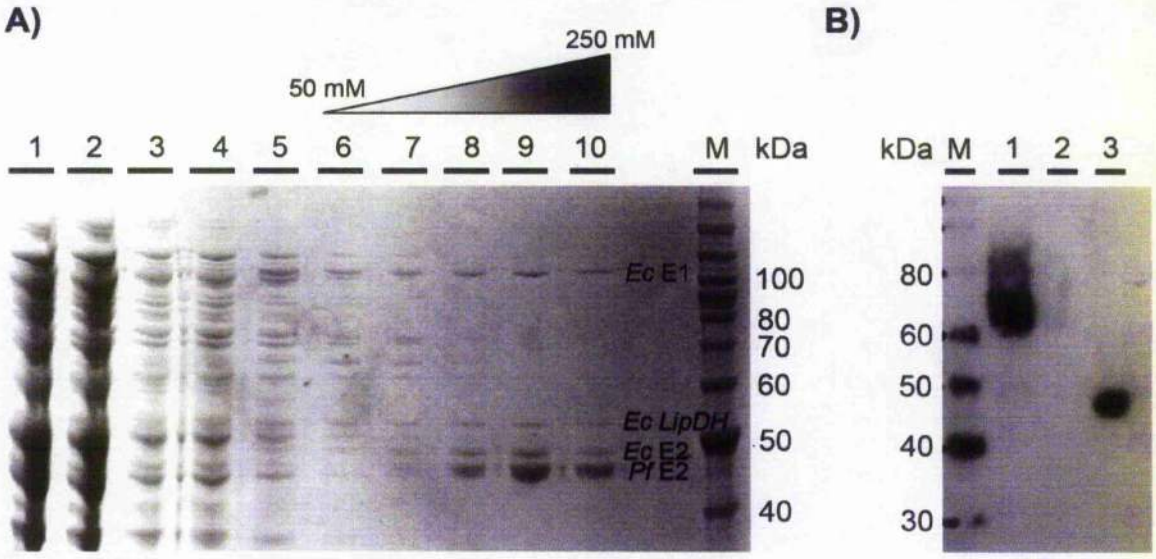


Figure 5.20 – Purification of KGDH-E2 by Ni-NTA batch binding

The purification profile for recombinantly expressed KGDH-E2 was loaded onto SDS-PAGE. **Panel A** displays a coomasie stained gel of the following samples; *E. coli* soluble fraction (Lane 1), unbound/flow through (lane 2), washes with 10 mM imidazole (Lanes 3 & 4), wash with 20 mM imidazole (Lane 5), 50 mM imidazole elution (Lane 6), 100 mM imidazole elution (Lane 7), 150 mM imidazole elution (Lane 8), 200 mM imidazole elution (Lane 9) and 250 mM imidazole elution (Lane 10). Lane M contains benchmark protein standards (Invitrogen). The 250 mM elution sample contains four main bands at 100 kDa, 53 kDa, 47 kDa and 45 kDa that correspond to *E. coli* E1 (*Ec* E1), *E. coli* LipDH (*Ec* LipDH), *E. coli* E2 (*Ec* E2) and recombinantly expressed *P. falciparum* E2 (*Pf* E2), respectively. **Panel B** displays a western blot on purification samples, using His-tag antibodies (BD biosciences). The western contains a his-tagged protein as a positive control (Lane 1; recombinant Pf TrxR), flow through fraction (Lane 2) and 250 mM imidazole elution (Lane 3). This displays that recombinantly expressed *P. falciparum* KGDH-E2 is contained in the elution fraction and most likely corresponds to the 45 kDa band in panel A.

An attempt was made to separate the complex by methods used previously to resolve individual proteins from purified KADH complexes (Pettit *et al.*, 1973). Ni-NTA elutions were incubated with 0.1M Glycine / 1M NaCl and subsequently run on a Sephadex S-200 16/60 gel filtration column (red line in figure 5.21A). The resulting void volume peak (peak 1 of red line on figure 5.21A) was analysed by SDS-PAGE, which displayed that two of the co-eluting proteins (the predicted *E. coli* α – ketoglutarate dehydrogenase and dihydrolipoamide dehydrogenase) were removed by this treatment (figure 5.21B, Lane 2). It was also observed that the resulting elution trace contained an extra peak (peak 4 of the red line in figure 5.21A) with a molecular mass of 107 kDa, which may represent the *E. coli* LipDH dimer released from the complex. I could not experimentally prove that all the components of the *E. coli* complex had been removed by this treatment and therefore another method to separate the recombinantly expressed protein from the complex formed was necessary.

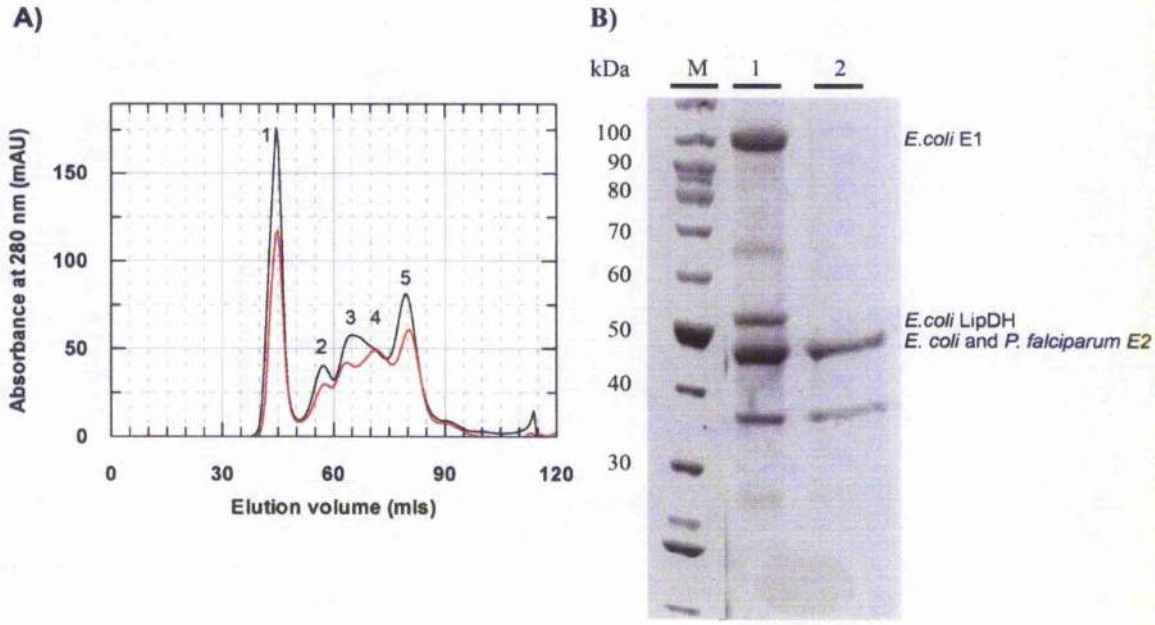


Figure 5.21 – Separation of KGDH complexes by Glycine/salt treatment

The possible formation of a complex consisting of the recombinantly expressed KGDH E2 and components of the *E. coli* KGDH was analysed by gel filtration chromatography. **Panel A** displays the elution profile of Ni-NTA purified KGDH E2 run on the sephadex S200 16/60 gel filtration column. The black line displays the run carried out in 50 mM potassium phosphate, 150 mM sodium chloride pH 7.6. The red line displays the run carried out in 0.1 mM Glycine, 1.0 M sodium chloride pH 9.0. Protein elution was followed by absorbance at 280 nm and resulted in peaks at > 1 MDa (peak 1, void volume), 417 kDa (peak 2), 200 kDa (peak 3), 107 kDa (peak 4) and 47 kDa (peak 5) **Panel B** displays the elution fraction from the gel filtration void volume fraction (peak 1 in panel A) run on SDS-PAGE. Lane 1 corresponds to the void volume from the potassium phosphate run (black line in panel A) and Lane 2 corresponds to the void volume from the glycine/salt run (Red line in panel A); Lane M contains benchmark protein standards (Invitrogen). This figure shows that the possible *E.coli* E1 and E3 bands were removed from the protein complex following treatment with 0.1M Glycine, 1M NaCl.

To further substantiate my suggestion that the recombinant *P. falciparum* protein would form complexes with the endogenous *E. coli* KGDH components, two of the co-eluted proteins (the 100 kDa and the 47 kDa proteins) were analysed by tryptic digestion followed by mass spectrometry. This identified the 100 kDa and 47 kDa co-eluting bands as the α–ketoglutarate dehydrogenase and succinyl transferase proteins from *E. coli*. These results therefore strongly suggest that the parasite KGDH-E2 indeed forms complexes with the bacterial KGDH components during the recombinant expression process. Since not all of the co-eluting bands were analysed by mass spectrometry finger printing it cannot be excluded that the recombinant protein not only forms complexes with the bacterial KGDH but possibly also with other KADH such as PDH and BCKDH. In the void volume elution fractions from the gel filtration column, bands of 66 kDa and 44 kDa could well represent components of the bacterial PDH and BCKDH proteins (PHD-E2 is a 66 kDa protein and BCKDH-E2 is a 44 kDa protein; faint bands in figure 5.21). The smaller bands observed at approximately 37 kDa and 25 kDa cannot be attributed to any of the bacterial KADH complexes.

The difficulty to obtain pure recombinant KGDH-E2 protein led me to use denaturing Ni-NTA affinity chromatography to purify the recombinant protein. Initially the denaturing purification of *P. falciparum* KGDH-E2 was carried out using a batch procedure. The denatured protein was eluted by a step-wise pH-gradient as shown in the SDS-PAGE (figure 5.22A). The elution fraction containing the recombinant protein was dialysed against non-denaturing Ni-NTA agarose loading buffer and applied to a 5 ml Ni-NTA agarose column attached to an ÄKTA FPLC system. The column was washed and the protein was eluted using a linear gradient between 10 mM and 500 mM imidazole (figure 5.22B).

Unfortunately it was not possible to continue with a biochemical characterisation of KGDH-E2 due to time constraints. However, it is anticipated that the refolding of denatured recombinant proteins from the KADH will be possible (Hanemaaijer *et al.*, 1989; Faure *et al.*, 2000) and thus the denaturing purification might present a way to obtain pure recombinant KGDH-E2 protein for future analyses.

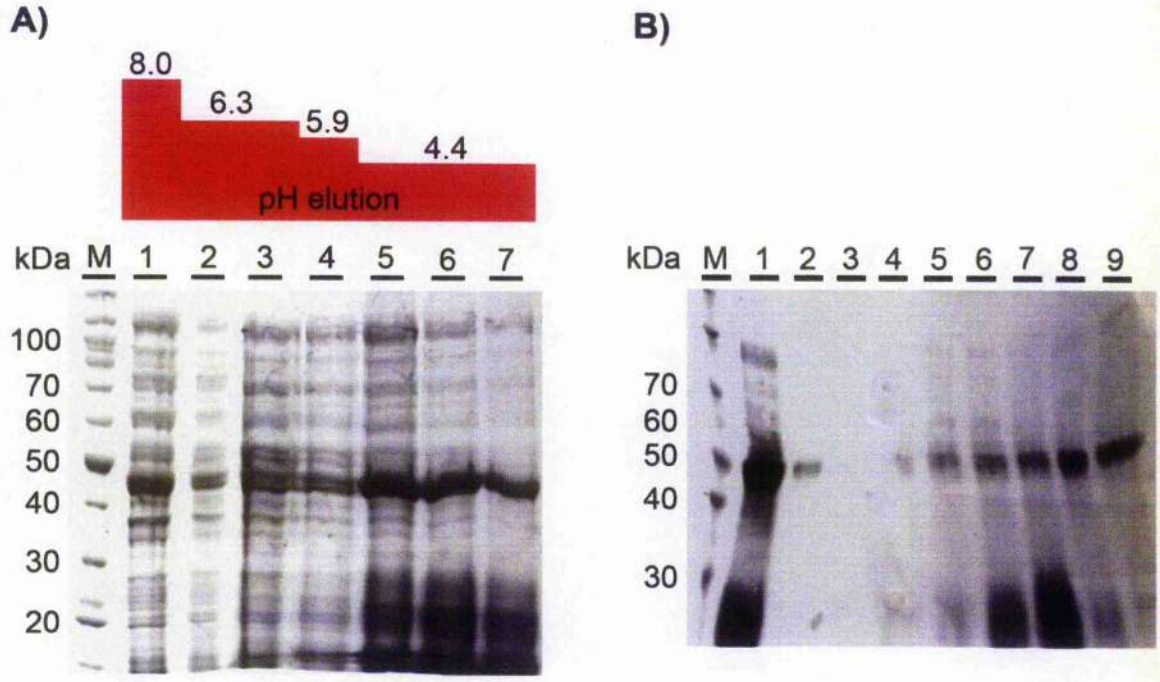


Figure 5.22 – Purification of KGDH E2 by denaturing Ni-NTA purification

The purification of recombinant KGDH E2 was carried out under denaturing conditions in order to eliminate the components of the E. coli KGDH from the purification. **Panel A** displays a SDS-PAGE of the step elutions of pET28 KGDH E2 by batch Ni-NTA purification. The majority of the recombinant protein eluted in the pH 4.4 elutions. **Panel B**; Following dialysis the pH 4.4 elutions were further purified by native Ni-NTA purification over a linear gradient of imidazole (10mM – 500 mM imidazole). Lane 1 contains the dialysed sample from denaturing Ni-NTA elution at pH 4.4; lane 2 contains the flow through fraction; lane 3 contains a 10 mM imidazole wash; lanes 4 – 7 were between 10 mM and 200 mM imidazole; lane 8 was a peak at 200 mM imidazole; lane 9 was a peak at 250 mM imidazole. Lane M in both panel A and B contains benchmark protein standards (Invitrogen).

5.3 Discussion

5.3.1 Three KADH identified in the *P. falciparum* genome

I have identified genes encoding all the components of three α–keto acid dehydrogenase complexes in the *P. falciparum* genome; a α–ketoglutarate dehydrogenase complex, a branched chain α–keto acid dehydrogenase complex and pyruvate dehydrogenase complex. Three of the genes identified (KGDH-E1, KGDH-E2 & PDH-E2) contained introns, with the *PDH E2* gene containing ten in total. These introns utilised the canonical intron/exon boundaries (GT-AG) and also exhibited a higher A/T content in the intron sequences when compared to the exon sequences as reported previously (Gardner *et al.*, 2002; Szafranski *et al.*, 2005). It was observed that all the intron containing genes that were investigated in this study (including *LipDH1* in chapter 4) contained introns primarily at the 5' end of the gene (except PDH E2), which resulted in the separation of the organelle targeting sequences from those of the mature gene product. This may allow for alternative splicing to occur, which could result in the targeting of proteins to different subcellular localisations. Although there have been no reports of differential splicing causing changes in the subcellular localisation of proteins in *P. falciparum*, genes have been identified that share a targeting sequence (van Dooren *et al.*, 2002). It was also observed during my studies that the intron/exon boundaries were mostly conserved between the *Plasmodium* species (van Lin *et al.*, 2001) and this was utilised to aid the correct prediction of intron/exon boundaries.

The presence of the genes encoding for components of all three KADH poses the question as to when these protein complexes are actively expressed in the parasites. To address this question I have performed RT-PCR analyses and found that all components of the three KADH are transcribed during the erythrocytic development of the parasites. This was somewhat surprising given that microarray analyses of *P. falciparum* throughout the erythrocytic cycle suggested that these genes would not be expressed or if so would only be expressed at extremely low levels through these developmental stages of the parasites (Bozdech *et al.*, 2003; Le Roch *et al.*, 2003). Proteome analyses of these parasite stages more or less corroborated the microarray data suggesting a minor role of the protein complexes during the erythrocytic life of the parasites (Florens *et al.*, 2002). However, analyses of protein levels of various proteins comprising KADH in the parasites (by Foth *et al.*, 2005 and this study) suggest that the genes of KADH components are not only transcribed but also present at reasonable protein levels in the

parasites (as they were detectable by Western blotting) implying a role in parasite metabolism in the erythrocytic stages of *P. falciparum* (Foth *et al.*, 2005).

Analyses of the deduced amino acid sequences displayed N-terminal extensions of various lengths, which upon *in silico* analyses suggested that the KGDH and BCKDH were mitochondrial proteins whereas the single PDH was targeted to the apicoplast of the parasites. This was a surprising finding given that in other eukaryotes all three complexes usually are mitochondrial, apart from plants which contain two distinct PDH that are mitochondrial and plastid localised (Mooney *et al.*, 2002).

5.3.2 KGDH and BCKDH are mitochondrially localised

The presence of the BCKDH and KGDH in the mitochondria, as experimentally determined in this study by expressing GFP fusion proteins in the parasites, was consistent with the observed localisation of these complexes in other organisms. The observed mis-targeting of some GFP-fusion proteins to the food vacuole has however, not been reported previously by other groups. We postulate that this may be due to the constitutive over-expression of the fusion protein in the parasites flooding the mitochondrial import machinery, which may result in mis-targeting and degradation of the GFP-fusion proteins. We have observed this mis-targeting only in mitochondrially localised proteins and it may be that the mitochondrion is more susceptible to mis-targeting than the apicoplast. The mis-targeting may also be caused by the incorrect folding of the GFP-fusion protein and it may be possible to rectify this by designing a fusion protein with a shorter leader sequence or by taking into account the domain structure of the KADH proteins.

The *P. falciparum* KGDH could provide substrates (NADH and succinyl CoA) for a number of biochemical processes including haem biosynthesis, residual TCA cycle activity, and oxidative phosphorylation. Studies with recombinantly expressed *P. falciparum* KGDH-E2 protein demonstrated that the protein formed complexes with the components of the *E. coli* KGDH. This phenomenon has not been previously reported and may be due in part to the association of the recombinant *P. falciparum* KGDH-E2 with the *E. coli* KGDH-E2. This may have been due to the reasonably high levels of amino acid sequence identity (52 % identity) in the catalytic domains (where self-association occurs) of the two proteins, which would produce a complex core that contained both E2 proteins. The association of E1 and LipDH proteins may have been due to the presence of *E. coli* KGDH-E2 in the core, as the lipoyl and subunit binding

domains of recombinant *P. falciparum* KGDH-E2 displayed much lower levels of sequence identity to the same domains of the *E. coli* KGDH-E2 (24 % identity). Although further investigations will be necessary to conclusively prove the reason for the binding of the KGDH to the recombinantly expressed *P. falciparum* KGDH-E2.

The *P. falciparum* BCKDH would be involved in the degradation of branched chain amino acids isoleucine, valine and leucine; producing α–methylbutyryl-CoA, isobutyryl-CoA and isovalyryl-CoA from the substrates α–keto–β-methylvalerate, α-ketoisovalerate and α–ketoisocaproic acid, respectively. These products may eventually lead to the production of acetyl CoA and/or succinyl CoA that could be used within the TCA cycle (Ward *et al.*, 1999; Holecek, 2002). However, the activity of the BCKDH in the parasites may be limited, as no homologue of branched chain amino acid transaminase (EC 2.6.1.42), the protein responsible for providing the BCKDH with its α – keto acid substrates, has yet to be identified in *P. falciparum* genome. Also a *P. falciparum* acyl CoA dehydrogenase homologue (EC 1.3.99.10), the protein that acts on the acyl CoA products of the BCKDH reaction, has also yet to be identified in the genome.

The BCKDH in other organisms comes under the regulation of a branched chain α – keto acid dehydrogenase kinase (EC 2.7.1.115) and phosphatase (EC 3.1.3.52). The *P. falciparum* BCKDH E2 contains a single phosphorylation site (Ser 319) although searches of the *P. falciparum* genome have yet to identify homologues for either the kinase or the phosphatase.

5.3.3 The single PDH of *P. falciparum* is apicoplast localised

I have identified a single pyruvate dehydrogenase complex in the *P. falciparum* genome and have confirmed its apicoplast localisation by expressing GFP-fusion proteins in erythrocytic stage parasites. The *P. falciparum* PDH lacked a protein X component that is required in the mitochondrial PDH for binding of the LipDH protein (Sanderson *et al.*, 1996). This was consistent with the apicoplast localisation of the PDH, as plastid localised PDH do not contain a protein X component (Mooney *et al.*, 2002). Another difference between the mitochondrial and plastid localised PDH, is that plastid PDH are not regulated by phosphorylation/dephosphorylation (Randall *et al.*, 1989). Mitochondrial PDH activity is regulated by phosphorylation of the E1α subunit by pyruvate dehydrogenase kinase (EC 2.7.1.99). This phosphorylation leads to inactivation of PDH and is implicated in control mechanisms such as glucose

homeostasis (Lydell *et al.*, 2002; Sugden *et al.*, 2003). The dephosphorylation of E1α by pyruvate dehydrogenase phosphatase (EC 3.1.3.43) activates the PDH E1α again. Mitochondrial PDH contain three possible sites of phosphorylation (Yeaman *et al.*, 1978) and although one of the phosphorylation sites is conserved in the *P. falciparum* PDH (Ser 467 in *P. falciparum* PDH E1α; figure 5.9) and plastid PDH (Johnston *et al.*, 1997), homologues of the kinase or phosphatase have not been identified (in either *P. falciparum* or plastids). It is therefore likely that the activity of plastid PDH is controlled either at the transcription/translation level, by changes in the biochemical environment (Reid *et al.*, 1975) or by product inhibition (Camp *et al.*, 1985; Camp *et al.*, 1988).

A recent study (Foth *et al.*, 2005) has further investigated the *P. falciparum* PDH with respect to phylogeny, subcellular localisation, expression in erythrocytic stages and enzyme activity. Foth *et al.* observed that the components of the *P. falciparum* PDH were consistently grouped with plastidic and cyanobacterial PDH proteins during phylogenetic analyses, confirming that they were originally encoded for in the genome of the endosymbiont that eventually became the apicoplast. Foth *et al.* also independently confirmed the apicoplast localisation of *P. falciparum* PDH by GFP fusion and immuno-fluorescence studies that demonstrated co-localisation with an apicoplast marker protein (Foth *et al.*, 2003; van Dooren *et al.*, 2005). The activity of the catalytic domain of the PDH-E2 was also investigated, demonstrating that this was active and had comparable kinetic parameters to other PDH-E2 catalytic domains tested (Yang *et al.*, 1997).

5.3.4 Consequences of KADH localisation on *P. falciparum* biochemistry

The localisation of KADH in both the mitochondrion and the apicoplast requires that *P. falciparum* has organelle specific LipDH proteins, which has been investigated and discussed in chapter 4. Another consequence of this localisation is the requirement for organelle-specific lipoylation pathways; as in order for the complexes to be active the E2 proteins are required to have a lipoyl moiety covalently attached. Studies in our (and others) laboratory have identified two separate pathways for the lipoylation of the E2 proteins of KADH; a mitochondrially localised salvage pathway and an apicoplast localised biosynthesis pathway (Thomsen-Zieger *et al.*, 2003; Wrenger *et al.*, 2004).

The major role of the apicoplast localised PDH would be to produce acetyl CoA and NADH for the type II fatty acid biosynthesis (FAS) that is performed in this organelle

(Waller *et al.*, 1998). Fatty acid biosynthesis has been shown to be essential in erythrocytic stages (Waller *et al.*, 1998; Surolia *et al.*, 2001; Waller *et al.*, 2003; Surolia *et al.*, 2004), and therefore any disruption of PDH activity should be lethal to the parasites. The presence of a mitochondrial PDH in humans requires any antimalarial inhibitors to be active specifically against the plastidic form, although at present no such inhibitors have been identified.

The absence of a mitochondrially localised PDH has wide ranging consequences on parasite biochemistry, as the mitochondrial PDH is usually the link between glycolysis and the tricarboxylic acid cycle (TCA cycle). This raises the question of what role does the TCA cycle play within the erythrocytic cycle? It has been generally accepted that erythrocytic stage *P. falciparum* do not rely on the TCA cycle for energy production, but metabolise the majority of their glucose into lactic acid by glycolysis (Bryant *et al.*, 1964). The TCA cycle must however be at least partially active in order to drive oxidative phosphorylation, which has been shown to be essential for parasite survival in these stages (Uyemura *et al.*, 2004). Although pyruvate cannot be used as a carbon source for the TCA cycle it may be possible that the TCA cycle is partially active with substrates supplied by anapleurotic reactions.

Amino acid degradation may play a key role in providing the TCA cycle with substrates, as only 20 % of the amino acids degraded during haemoglobin digestion are then used for protein synthesis (Krugliak *et al.*, 2002), suggesting that the other 80 % may be degraded for other cellular processes . The activity of the BCKDH is one possible source, as the degradation of valine, leucine and isoleucine (amino acids that are highly represented in haemoglobin) can ultimately provide the mitochondrion with sources of acetyl CoA and succinyl CoA. Another source of substrates for the TCA cycle would be the conversion of glutamate and oxaloacetate to α – ketoglutarate and aspartate via the activity of an aspartate aminotransferase that has been identified in the *P. falciparum* genome (Berger *et al.*, 2001; Gardner *et al.*, 2002). Glutamate can also be utilised to produce α – ketoglutarate via the activity of glutamate dehydrogenase, although the precise location of the three predicted glutamate dehydrogenase proteins has yet to be verified (Wagner *et al.*, 1998; Gardner *et al.*, 2002; Werner *et al.*, 2005). Another source of substrates that may be able to drive oxidative phosphorylation is the transfer of malate that has been produced from phosphoenolpyruvate (PEP), via oxaloacetate (OAA) by the activity of PEP-carboxylase (Gardner *et al.*, 2002) and a cytosolic malate dehydrogenase (Lang-Unnasch, 1992; Tripathi *et al.*, 2004). Once transported into the mitochondrion the malate can be acted on by a malate:quinone

oxidoreductase which is likely to feed electrons into oxidative phosphorylation (Gardner *et al.*, 2002; Uyemura *et al.*, 2004).

TCA cycle substrates may also be imported into the mitochondrion from the cytosol and a number of transporters have been identified that may transport α – ketoglutarate or acetyl CoA. A dicarboxylate transporter may act as an α – ketoglutarate transporter and it may be possible that acetyl CoA produced in the apicoplast could be transferred to the mitochondrion via an acetyl CoA transporter (NP_700833) that has been identified in the *P. falciparum* genome. The precise location of this transporter has not been determined, and the transfer of acetyl CoA would need to span the multiple membranes of the two organelles. If this transfer were to occur, it would further confirm the metabolic co-dependence theory that is thought to drive the close association observed between the mitochondria and the apicoplast.

5.3.5 Why are the mitochondrion and apicoplast closely associated?

Van Dooren *et al* have undertaken a study into the structure of the apicoplast and mitochondrion during the erythrocytic stages of the parasite, with similar results to those observed in my study (van Dooren *et al.*, 2005). Studies into organelle ultra-structure using electron microscopy were the first to demonstrate a close association between the *Plasmodium* mitochondrion and apicoplast (Aikawa, 1966; Prensier *et al.*, 1986; Hopkins *et al.*, 1999). This close association of the organelles has been proposed to arise from a metabolic co-dependence (Hopkins *et al.*, 1999; Ralph *et al.*, 2004) and the requirement for co-segregation of the organelles into merozoites (van Dooren *et al.*, 2005).

The possible metabolic co-dependence can be demonstrated by localisation of proteins from the haem biosynthesis pathway in the mitochondrion and the apicoplast (Sato *et al.*, 2004). This confirmed the prediction that different steps of the haem biosynthesis pathway would occur in both organelles, which suggests that metabolites must be transferred between the two organelles for the pathway to be active (Gardner *et al.*, 2002). The co-segregation of both organelles into the merozoites is essential, as both the mitochondrion and the apicoplast have been shown to be necessary for parasite survival. The mitochondrion has been proven necessary as several antimalarial drugs target mitochondrial proteins. Atovaquone has been demonstrated to inhibit the parasite's oxidative phosphorylation, by inhibiting complex III (Kessl *et al.*, 2003), destroying the organelles membrane potential (Srivastava *et al.*, 1997). Artesunate, a derivative of

artemisinin has been suggested to inhibit Complex IV activity (Zhao *et al.*, 1986). Other compounds such as Primaquine and Artemether have been suggested to act specifically against the mitochondrion by causing morphological changes (Kawai *et al.*, 1993; Vaidya *et al.*, 1993). Studies have demonstrated that parasites that lack an active apicoplast display a "delayed death", whereby they can readily develop within and escape from host cells, but are unable to re-infect to continue their life cycle (He *et al.*, 2001). Although the mechanism of this delayed death is not fully understood, this confirms the apicoplast as being essential for parasite survival. Inhibitors of apicoplast located pathways such as non-mevalonate isoprenoid biosynthesis (Fosmidomycin) and type II fatty acid biosynthesis (Triclosan) have also been demonstrated to be potent antimalarials (Jomaa *et al.*, 1999; Lell *et al.*, 2003; Waller *et al.*, 2003). Studies by Striepen *et al.* in the related apicomplexan *Toxoplasma gondii* have demonstrated that the apicoplast divides by association with centromeres (Striepen *et al.*, 2000). If division occurs by this manner in *Plasmodium* it will be further complicated by the number of daughter cells produced during schizogony, and this needs to be further investigated; along with the mode of division of the *Plasmodium* mitochondrion.

5.4 Conclusions

I have identified the genes that encode the α–keto acid dehydrogenase complexes in the *Plasmodium falciparum* genome. A number of the genes were predicted to contain introns and these were experimentally verified (by DNA sequencing) following amplification from total RNA. All of the genes identified were amplified from total RNA, which confirmed that they were transcribed in the erythrocytic stages of infection. Protein expression during these stages was also confirmed by western blotting analyses for LipDH2, BCKDH E2 and KGDH E2.

In silico analyses had predicted that the KGDH and BCKDH would be located in the mitochondrion and that the single PDH identified would be located to the apicoplast. These predicted localisations were confirmed by expressing GFP-fusion proteins in erythrocytic stage parasites, and work is ongoing to further confirm these results by immuno-fluorescence studies with organelle-specific markers. The biochemical consequences of the parasites lacking a mitochondrial PDH have been discussed and warrant further investigation. By utilising the PDH-GFP fusion proteins it was possible to visualise the development of the mitochondrion and apicoplast during the erythrocytic stages. This confirmed the pattern of close association, elongation and co-separation observed in other studies.

Analysis of the recombinantly expressed KGDH-E2 protein identified that it formed complexes with components of the *E. coli* KGDH. The recombinant protein was purified to homogeneity by denaturing purification and will allow me to further analyse the recombinant protein at a later date.

References

Adams, K. L. and Palmer, J. D. (2003). Evolution of mitochondrial gene content: gene loss and transfer to the nucleus. *Mol Phylogenet Evol* **29**: 380-395.

Aevarsson, A., Seger, K., Turley, S., Sokatch, J. R. and Hol, W. G. (1999). Crystal structure of 2-oxoisovalerate and dehydrogenase and the architecture of 2-oxo acid dehydrogenase multienzyme complexes. *Nat Struct Biol* **6**: 785-792.

Aikawa, M. (1966). The fine structure of the erythrocytic stages of three avian malarial parasites, *Plasmodium fallax*, *P. lophurae*, and *P. cathemerium*. *Am J Trop Med Hyg* **15**: 449-471.

Akerman, S. E. and Muller, S. (2003). 2-Cys peroxiredoxin PfTrx-Px1 is involved in the antioxidant defence of *Plasmodium falciparum*. *Mol Biochem Parasitol* **130**: 75-81.

Akoachere, M., Iozef, R., Rahlfs, S., Deponte, M., Mannervik, B., Creighton, D. J., Schirmer, H. and Becker, K. (2005). Characterization of the glyoxalases of the malarial parasite *Plasmodium falciparum* and comparison with their human counterparts. *Biol Chem* **386**: 41-52.

Andricopulo, A. D., Akoachere, M. B., Krogh, R., Nickel, C., McLeish, M. J., Kenyon, G. L., Arscott, L. D., Williams, C. H., Jr., Davioud-Charvet, E. and Becker, K. (2006). Specific inhibitors of *Plasmodium falciparum* thioredoxin reductase as potential antimalarial agents. *Bioorg Med Chem Lett*.

Apt, K. E., Clendennen, S. K., Powers, D. A. and Grossman, A. R. (1995). The gene family encoding the fucoxanthin chlorophyll proteins from the brown alga *Macrocystis pyrifera*. *Mol Gen Genet* **246**: 455-464.

Argyrou, A. and Blanchard, J. S. (2001). *Mycobacterium tuberculosis* lipoamide dehydrogenase is encoded by Rv0462 and not by the lpdA or lpdB genes. *Biochemistry* **40**: 11353-11363.

Argyrou, A. and Blanchard, J. S. (2004). Flavoprotein disulfide reductases: advances in chemistry and function. *Prog Nucleic Acid Res Mol Biol* **78**: 89-142.

Argyrou, A., Blanchard, J. S. and Palfey, B. A. (2002). The lipoamide dehydrogenase from *Mycobacterium tuberculosis* permits the direct observation of flavin intermediates in catalysis. *Biochemistry* **41**: 14580-14590.

Arner, E. S., Nordberg, J. and Holmgren, A. (1996). Efficient reduction of lipoamide and lipoic acid by mammalian thioredoxin reductase. *Biochem Biophys Res Commun* **225**: 268-274.

Atamna, H. and Ginsburg, H. (1993). Origin of reactive oxygen species in erythrocytes infected with *Plasmodium falciparum*. *Mol Biochem Parasitol* **61**: 231-241.

Atamna, H. and Ginsburg, H. (1995). Heme degradation in the presence of glutathione. A proposed mechanism to account for the high levels of non-heme iron found in the membranes of hemoglobinopathic red blood cells. *J Biol Chem* **270**: 24876-24883.

Baier, M. and Dietz, K. J. (2005). Chloroplasts as source and target of cellular redox regulation: a discussion on chloroplast redox signals in the context of plant physiology. *J Exp Bot* **56**: 1449-1462.

Baldwin, J., Michnoff, C. H., Malmquist, N. A., White, J., Roth, M. G., Rathod, P. K. and Phillips, M. A. (2005). High-throughput screening for potent and selective inhibitors of *Plasmodium falciparum* dihydroorotate dehydrogenase. *J Biol Chem* **280**: 21847-21853.

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

Bauer, H., Gromer, S., Urbani, A., Schnolzer, M., Schirmer, R. H. and Muller, H. M. (2003a). Thioredoxin reductase from the malaria mosquito *Anopheles gambiae*. *Eur J Biochem* **270**: 4272-4281.

Bauer, H., Massey, V., Arscott, L. D., Schirmer, R. H., Ballou, D. P. and Williams, C. H., Jr. (2003b). The mechanism of high Mr thioredoxin reductase from *Drosophila melanogaster*. *J Biol Chem* **278**: 33020-33028.

Becker, K. and Kirk, K. (2004). Of malaria, metabolism and membrane transport. *Trends Parasitol* **20**: 590-596.

Becker, K., Rahlfs, S., Nickel, C. and Schirmer, R. H. (2003). Glutathione--functions and metabolism in the malarial parasite *Plasmodium falciparum*. *Biol Chem* **384**: 551-566.

Becker, K., Tilley, L., Vennerstrom, J. L., Roberts, D., Rogerson, S. and Ginsburg, H. (2004). Oxidative stress in malaria parasite-infected erythrocytes: host-parasite interactions. *Int J Parasitol* **34**: 163-189.

Becker, K., Tilley, L., Vennerstrom, J. L., Roberts, D., Rogerson, S. and Ginsburg, H. (2004b). Oxidative stress in malaria parasite-infected erythrocytes: host-parasite interactions. *Int J Parasitol* **34**: 163-189.

Bender, A., van Dooren, G. G., Ralph, S. A., McFadden, G. I. and Schneider, G. (2003). Properties and prediction of mitochondrial transit peptides from *Plasmodium falciparum*. *Mol Biochem Parasitol* **132**: 59-66.

Benen, J., van Berkel, W., Dieteren, N., Arscott, D., Williams, C., Jr., Veeger, C. and de Kok, A. (1992). Lipoamide dehydrogenase from *Azotobacter vinelandii*: site-directed mutagenesis of the His450-Glu455 diad. Kinetics of wild-type and mutated enzymes. *Eur J Biochem* **207**: 487-497.

Benen, J., van Berkel, W., Zak, Z., Visser, T., Veeger, C. and de Kok, A. (1991). Lipoamide dehydrogenase from *Azotobacter vinelandii*: site-directed mutagenesis of the His450-Glu455 diad. Spectral properties of wild type and mutated enzymes. *Eur J Biochem* **202**: 863-872.

Berger, L. C., Wilson, J., Wood, P. and Berger, B. J. (2001). Methionine regeneration and aspartate aminotransferase in parasitic protozoa. *J Bacteriol* **183**: 4421-4434.

Berry, A., Scrutton, N. S. and Perham, R. N. (1989). Switching kinetic mechanism and putative proton donor by directed mutagenesis of glutathione reductase. *Biochemistry* **28**: 1264-1269.

Biterova, E. I., Turanov, A. A., Gladyshev, V. N. and Barycki, J. J. (2005). Crystal structures of oxidized and reduced mitochondrial thioredoxin reductase provide molecular details of the reaction mechanism. *Proc Natl Acad Sci U S A* **102**: 15018-15023.

Blanchard, J. L. and Hicks, J. S. (1999). The non-photosynthetic plastid in malarial parasites and other apicomplexans is derived from outside the green plastid lineage. *J Eukaryot Microbiol* **46**: 367-375.

Blumenstiel, K., Schoneck, R., Yardley, V., Croft, S. L. and Krauth-Siegel, R. L. (1999). Nitrofuran drugs as common subversive substrates of *Trypanosoma cruzi* lipoamide dehydrogenase and trypanothione reductase. *Biochem Pharmacol* **58**: 1791-1799.

Bocanegra, J. A., Scrutton, N. S. and Perham, R. N. (1993). Creation of an NADP-dependent pyruvate dehydrogenase multienzyme complex by protein engineering. *Biochemistry* **32**: 2737-2740.

Borges, A., Hawkins, C. F., Packman, L. C. and Perham, R. N. (1990). Cloning and sequence analysis of the genes encoding the dihydrolipoamide acetyltransferase and dihydrolipoamide dehydrogenase components of the pyruvate dehydrogenase multienzyme complex of *Bacillus stearothermophilus*. *Eur J Biochem* **194**: 95-102.

Bourguignon, J., Merand, V., Rawsthorne, S., Forest, E. and Douce, R. (1996). Glycine decarboxylase and pyruvate dehydrogenase complexes share the same dihydrolipoamide dehydrogenase in pea leaf mitochondria: evidence from mass spectrometry and primary-structure analysis. *Biochem J* **313 (Pt 1)**: 229-234.

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.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248-254.

Brandt, W. and Wessjohann, L. A. (2005). The functional role of selenocysteine (Sec) in the catalysis mechanism of large thioredoxin reductases: proposition of a swapping catalytic triad including a Sec-His-Glu state. *Chembiochem* **6**: 386-394.

Brautigam, C. A., Chuang, J. L., Tomchick, D. R., Machius, M. and Chuang, D. T. (2005). Crystal structure of human dihydrolipoamide dehydrogenase: NAD+/NADH binding and the structural basis of disease-causing mutations. *J Mol Biol* **350**: 543-552.

Brown, A. M., Kristal, B. S., Effron, M. S., Shestopalov, A. I., Ullucci, P. A., Sheu, K. F., Blass, J. P. and Cooper, A. J. (2000). Zn2+ inhibits alpha-ketoglutarate-stimulated mitochondrial respiration and the isolated alpha-ketoglutarate dehydrogenase complex. *J Biol Chem* **275**: 13441-13447.

Bryant, C., Voller, A. and Smith, M. J. (1964). The Incorporation Of Radioactivity From (14c)Glucose Into The Soluble Metabolic Intermediates Of Malaria Parasites. *Am J Trop Med Hyg* **13**: 515-519.

Bryk, R., Lima, C. D., Erdjument-Bromage, H., Tempst, P. and Nathan, C. (2002). Metabolic enzymes of *Mycobacteria* linked to antioxidant defense by a thioredoxin-like protein. *Science* **295**: 1073-1077.

Bunik, V., Raddatz, G., Lemaire, S., Meyer, Y., Jacquot, J. P. and Bisswanger, H. (1999). Interaction of thioredoxins with target proteins: role of particular structural elements and electrostatic properties of thioredoxins in their interplay with 2-oxoacid dehydrogenase complexes. *Protein Sci* **8**: 65-74.

Camp, P. J., Miernyk, J. A. and Randall, D. D. (1988). Some kinetic and regulatory properties of the pea chloroplast pyruvate dehydrogenase complex. *Biochim Biophys Acta* **933**: 269-275.

Camp, P. J. and Randall, D. D. (1985). Purification and characterisation of the pea chloroplast pyruvate dehydrogenase complex: a source of acetyl-CoA and NADH for fatty acid biosynthesis. *Plant Physiol* **77**: 571-577.

Caron, L., Douady, D., Quinet-Szely, M., de Goer, S. and Berkaloff, C. (1996). Gene structure of a chlorophyll a/c-binding protein from a brown alga: presence of an intron and phylogenetic implications. *J Mol Evol* **43**: 270-280.

Cavalier-Smith, T. (2000). Membrane heredity and early chloroplast evolution. *Trends Plant Sci* **5**: 174-182.

Cenas, N., Bironaite, D., Dickancaite, E., Anusevicius, Z., Sarlauskas, J. and Blanchard, J. S. (1994). Chinifur, a selective inhibitor and "subversive substrate" for *Trypanosoma congolense* trypanothione reductase. *Biochem Biophys Res Commun* **204**: 224-229.

Cenas, N. K., Arscott, D., Williams, C. H., Jr. and Blanchard, J. S. (1994). Mechanism of reduction of quinones by *Trypanosoma congolense* trypanothione reductase. *Biochemistry* **33**: 2509-2515.

Cerna, L., Wenchich, L., Hansikova, H., Kmoch, S., Peskova, K., Chrastina, P., Brynda, J. and Zeman, J. (2001). Novel mutations in a boy with dihydrolipoamide dehydrogenase deficiency. *Med Sci Monit* **7**: 1319-1325.

Chuang, D. T., Chuang, J. L. and Wynn, R. M. (2006). Lessons from genetic disorders of branched-chain amino acid metabolism. *J Nutr* **136**: 243S-249S.

Ciszak, E. M., Korotchkina, L. G., Dominiak, P. M., Sidhu, S. and Patel, M. S. (2003). Structural basis for flip-flop action of thiamin pyrophosphate-dependent enzymes revealed by human pyruvate dehydrogenase. *J Biol Chem* **278**: 21240-21246.

Claros, M. G. and Vincens, P. (1996). Computational method to predict mitochondrially imported proteins and their targeting sequences. *Eur J Biochem* **241**: 779-786.

Clos, J. and Brandau, S. (1994). pJC20 and pJC40--two high-copy-number vectors for T7 RNA polymerase-dependent expression of recombinant genes in *Escherichia coli*. *Protein Expr Purif* **5**: 133-137.

Collaborative, T. P. G. D. (2001). PlasmoDB: An integrative database of the *Plasmodium falciparum genome*. Tools for accessing and analyzing finished and unfinished sequence data. The *Plasmodium* Genome Database Collaborative. *Nucleic Acids Res* **29**: 66-69.

Conner, M., Krell, T. and Lindsay, J. G. (1996). Identification and purification of a distinct dihydrolipoamide dehydrogenase from pea chloroplasts. *Planta* **200**: 195-202.

Coombs, G. H., Westrop, G. D., Suchan, P., Puzova, G., Hirt, R. P., Embley, T. M., Mottram, J. C. and Muller, S. (2004). The amitochondriate eukaryote *Trichomonas vaginalis* contains a divergent thioredoxin-linked peroxiredoxin antioxidant system. *J Biol Chem* **279**: 5249-5256.

Crabb, B. S., Rug, M., Gilberger, T. W., Thompson, J. K., Triglia, T., Maier, A. G. and Cowman, A. F. (2004). Transfection of the human malaria parasite *Plasmodium falciparum*. *Methods Mol Biol* **270**: 263-276.

Cunningham, M. L., Zvelebil, M. J. and Fairlamb, A. H. (1994). Mechanism of inhibition of trypanothione reductase and glutathione reductase by trivalent organic arsenicals. *Eur J Biochem* **221**: 285-295.

Davidson, M. W., Griggs, B. G., Jr., Boykin, D. W. and Wilson, W. D. (1975). Mefloquine, a clinically useful quinolinemethanol antimalarial which does not significantly bind to DNA. *Nature* **254**: 632-634.

Davioud-Charvet, E., McLeish, M. J., Veine, D. M., Giegel, D., Arscott, L. D., Andricopulo, A. D., Becker, K., Muller, S., Schirmer, R. H., Williams, C. H., Jr. and Kenyon, G. L. (2003). Mechanism-based inactivation of thioredoxin reductase from *Plasmodium falciparum* by Mannich bases. Implication for cytotoxicity. *Biochemistry* **42**: 13319-13330.

de Kok, A., Hengeveld, A. F., Martin, A. and Westphal, A. H. (1998). The pyruvate dehydrogenase multi-enzyme complex from Gram-negative bacteria. *Biochim Biophys Acta* **1385**: 353-366.

Deitsch, K., Driskill, C. and Wellems, T. (2001). Transformation of malaria parasites by the spontaneous uptake and expression of DNA from human erythrocytes. *Nucleic Acids Res* **29**: 850-853.

Denoya, C. D., Fedechko, R. W., Hafner, E. W., McArthur, H. A., Morgenstern, M. R., Skinner, D. D., Stutzman-Engwall, K., Wax, R. G. and Wernau, W. C. (1995). A second branched-chain alpha-keto acid dehydrogenase gene cluster (bkdFGH) from *Streptomyces avermitilis*: its relationship to avermectin biosynthesis and the construction of a bkdF mutant suitable for the production of novel antiparasitic avermectins. *J Bacteriol* **177**: 3504-3511.

Deonarain, M. P., Berry, A., Scrutton, N. S. and Perham, R. N. (1989). Alternative proton donors/acceptors in the catalytic mechanism of the glutathione reductase of *Escherichia coli*: the role of histidine-439 and tyrosine-99. *Biochemistry* **28**: 9602-9607.

Deonarain, M. P., Scrutton, N. S. and Perham, R. N. (1992). Engineering surface charge. 2. A method for purifying heterodimers of *Escherichia coli* glutathione reductase. *Biochemistry* **31**: 1498-1504.

DeRocher, A., Hagen, C. B., Froehlich, J. E., Feagin, J. E. and Parsons, M. (2000). Analysis of targeting sequences demonstrates that trafficking to the *Toxoplasma gondii* plastid branches off the secretory system. *J Cell Sci* **113 (Pt 22)**: 3969-3977.

Dhanasekaran, S., Chandra, N. R., Chandrasekhar Sagar, B. K., Rangarajan, P. N. and Padmanaban, G. (2004). Delta-aminolevulinic acid dehydratase from *Plasmodium falciparum*: indigenous versus imported. *J Biol Chem* **279**: 6934-6942.

Diefenbach, R. J., Candy, J. M., Mattick, J. S. and Duggleby, R. G. (1992). Effects of substitution of aspartate-440 and tryptophan-487 in the thiamin diphosphate binding region of pyruvate decarboxylase from *Zymomonas mobilis*. *FEBS Lett* **296**: 95-98.

Dong, Y. and Vennerstrom, J. L. (2003). Mechanisms of in situ activation for peroxidic antimalarials. *Redox Rep* **8**: 284-288.

Douce, R., Bourguignon, J., Neuburger, M. and Rebeille, F. (2001). The glycine decarboxylase system: a fascinating complex. *Trends Plant Sci* **6**: 167-176.

Dry, I. B. and Wiskich, J. T. (1987). 2-Oxoglutarate dehydrogenase and pyruvate dehydrogenase activities in plant mitochondria: interaction via a common coenzyme a pool. *Arch Biochem Biophys* **257**: 92-99.

Eckstein-Ludwig, U., Webb, R. J., Van Goethem, I. D., East, J. M., Lee, A. G., Kimura, M., O'Neill, P. M., Bray, P. G., Ward, S. A. and Krishna, S. (2003). Artemisinins target the SERCA of *Plasmodium falciparum*. *Nature* **424**: 957-961.

Egan, T. J., Combrinck, J. M., Egan, J., Hearne, G. R., Marques, H. M., Ntenteni, S., Sewell, B. T., Smith, P. J., Taylor, D., van Schalkwyk, D. A. and Walden, J. C. (2002). Fate of haem iron in the malaria parasite *Plasmodium falciparum*. *Biochem J* **365**: 343-347.

Egan, T. J., Mavuso, W. W., Ross, D. C. and Marques, H. M. (1997). Thermodynamic factors controlling the interaction of quinoline antimalarial drugs with ferriprotoporphyrin IX. *J Inorg Biochem* **68**: 137-145.

Else, A. J., Clarke, J. F., Willis, A., Jackman, S. A., Hough, D. W. and Danson, M. J. (1994). Dihydrolipoamide dehydrogenase in the *trypanosoma* subgenus, trypanozoon. *Mol Biochem Parasitol* **64**: 233-239.

Emanuelsson O, N. H., Brunak S, von Heijne G. (2000). Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J Mol Biol.* **300**: 1005-1016.

Fairlamb, A. H., Carter, N. S., Cunningham, M. and Smith, K. (1992). Characterisation of melarsen-resistant *Trypanosoma brucei brucei* with respect to cross-resistance to other drugs and trypanothione metabolism. *Mol Biochem Parasitol* **53**: 213-222.

Farber, P. M., Arscott, L. D., Williams, C. H., Jr., Becker, K. and Schirmer, R. H. (1998). Recombinant *Plasmodium falciparum* glutathione reductase is inhibited by the antimalarial dye methylene blue. *FEBS Lett* **422**: 311-314.

Faure, M., Bourguignon, J., Neuburger, M., MacHerel, D., Sieker, L., Ober, R., Kahn, R., Cohen-Addad, C. and Douce, R. (2000). Interaction between the lipoamide-containing H-protein and the lipoamide dehydrogenase (L-protein) of the glycine decarboxylase multienzyme system 2. Crystal structures of H- and L-proteins. *Eur J Biochem* **267**: 2890-2898.

Fichera, M. E. and Roos, D. S. (1997). A plastid organelle as a drug target in apicomplexan parasites. *Nature* **390**: 407-409.

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**: 10931-10936.

Fisher, C. R., Chuang, J. L., Cox, R. P., Fisher, C. W., Star, R. A. and Chuang, D. T. (1991). Maple syrup urine disease in Mennonites. Evidence that the Y393N mutation in E1 alpha impedes assembly of the E1 component of branched-chain alpha-keto acid dehydrogenase complex. *J Clin Invest* **88**: 1034-1037.

Flint, D. H., Emptage, M. H. and Guest, J. R. (1992). Fumarase a from *Escherichia coli*: purification and characterization as an iron-sulfur cluster containing enzyme. *Biochemistry* **31**: 10331-10337.

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

Foth, B. J., Ralph, S. A., Tonkin, C. J., Struck, N. S., Fraunholz, M., Roos, D. S., Cowman, A. F. and McFadden, G. I. (2003). Dissecting apicoplast targeting in the malaria parasite *Plasmodium falciparum*. *Science* **299**: 705-708.

Foth, B. J., Stimmler, L. M., Handman, E., Crabb, B. S., Hodder, A. N. and McFadden, G. I. (2005). The malaria parasite *Plasmodium falciparum* has only one pyruvate dehydrogenase complex, which is located in the apicoplast. *Mol Microbiol* **55**: 39-53.

Fry, M. and Beesley, J. E. (1991). Mitochondria of mammalian *Plasmodium* spp. *Parasitology* **102 Pt 1**: 17-26.

Fujiwara, N., Fujii, T., Fujii, J. and Taniguchi, N. (2001). Roles of N-terminal active cysteines and C-terminal cysteine-selenocysteine in the catalytic mechanism of mammalian thioredoxin reductase. *J Biochem (Tokyo)* **129**: 803-812.

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.

Gardner, M. J., Tettelin, H., Carucci, D. J., Cummings, L. M., Smith, H. O., Fraser, C. M., Venter, J. C. and Hoffman, S. L. (1999). The malaria genome sequencing project: complete sequence of *Plasmodium falciparum* chromosome 2. *Parassitologia* **41**: 69-75.

Gazaryan, I. G., Krasnikov, B. F., Ashby, G. A., Thorneley, R. N., Kristal, B. S. and Brown, A. M. (2002). Zinc is a potent inhibitor of thiol oxidoreductase activity and stimulates reactive oxygen species production by lipoamide dehydrogenase. *J Biol Chem* **277**: 10064-10072.

Gelhaye, E., Rouhier, N., Navrot, N. and Jacquot, J. P. (2005). The plant thioredoxin system. *Cell Mol Life Sci* **62**: 24-35.

Ghisla, S., Massey, V., Lhoste, J. M. and Mayhew, S. G. (1974). Fluorescence and optical characteristics of reduced flavines and flavoproteins. *Biochemistry* **13**: 589-597.

Gibson, G. E., Park, L. C., Sheu, K. F., Blass, J. P. and Calingasan, N. Y. (2000). The alpha-ketoglutarate dehydrogenase complex in neurodegeneration. *Neurochem Int* **36**: 97-112.

Gilberger, T. W., Bergmann, B., Walter, R. D. and Muller, S. (1998). The role of the C-terminus for catalysis of the large thioredoxin reductase from *Plasmodium falciparum*. *FEBS Lett* **425**: 407-410.

Gilberger, T. W., Walter, R. D. and Muller, S. (1997). Identification and characterization of the functional amino acids at the active site of the large thioredoxin reductase from *Plasmodium falciparum*. *J Biol Chem* **272**: 29584-29589.

Ginsburg, H. and Golenser, J. (2003). Glutathione is involved in the antimalarial action of chloroquine and its modulation affects drug sensitivity of human and murine species of *Plasmodium*. *Redox Rep* **8**: 276-279.

Gladyshev, V. N., Jeang, K. T. and Stadtman, T. C. (1996). Selenocysteine, identified as the penultimate C-terminal residue in human T-cell thioredoxin reductase, corresponds to TGA in the human placental gene. *Proc Natl Acad Sci U S A* **93**: 6146-6151.

Gladyshev, V. N., Krause, M., Xu, X. M., Korotkov, K. V., Kryukov, G. V., Sun, Q. A., Lee, B. J., Wootton, J. C. and Hatfield, D. L. (1999). Selenocysteine-containing thioredoxin reductase in *C. elegans*. *Biochem Biophys Res Commun* **259**: 244-249.

Grafakou, O., Oexle, K., van den Heuvel, L., Smeets, R., Trijbels, F., Goebel, H. H., Bosshard, N., Superti-Furga, A., Steinmann, B. and Smeitink, J. (2003). Leigh syndrome due to compound heterozygosity of dihydrolipoamide dehydrogenase gene mutations. Description of the first E3 splice site mutation. *Eur J Pediatr* **162**: 714-718.

Gray, M. W. (1992). The endosymbiont hypothesis revisited. *Int Rev Cytol* **141**: 233-357.

Greenwood, B. and Mutabingwa, T. (2002). Malaria in 2002. *Nature* **415**: 670-672.

Griffin, T. A., Wynn, R. M. and Chuang, D. T. (1990). Expression and assembly of mature apotransacylase (E2b) of bovine branched-chain alpha-keto acid dehydrogenase complex in *Escherichia coli*. Demonstration of transacylase activity and modification by lipoylation. *J Biol Chem* **265**: 12104-12110.

Grinblat, L., Sreider, C. M. and Stoppani, A. O. (1991). Superoxide anion production by lipoamide dehydrogenase redox-cycling: effect of enzyme modifiers. *Biochem Int* **23**: 83-92.

Gromer, S., Johansson, L., Bauer, H., Arscott, L. D., Rauch, S., Ballou, D. P., Williams, C. H., Jr., Schirmer, R. H. and Arner, E. S. (2003). Active sites of thioredoxin reductases: why selenoproteins? *Proc Natl Acad Sci U S A* **100**: 12618-12623.

Gromer, S., Wissing, J., Behne, D., Ashman, K., Schirmer, R. H., Flohe, L. and Becker, K. (1998). A hypothesis on the catalytic mechanism of the selenoenzyme thioredoxin reductase. *Biochem J* **332 (Pt 2)**: 591-592.

Grossman, A., Manodori, A. and Snyder, D. (1990). Light-harvesting proteins of diatoms: their relationship to the chlorophyll a/b binding proteins of higher plants and their mode of transport into plastids. *Mol Gen Genet* **224**: 91-100.

Gudi, R., Bowker-Kinley, M. M., Kedishvili, N. Y., Zhao, Y. and Popov, K. M. (1995). Diversity of the pyruvate dehydrogenase kinase gene family in humans. *J Biol Chem* **270**: 28989-28994.

Gutierrez-Correa, J., Krauth-Siegel, R. L. and Stoppani, A. O. (2003). Phenothiazine radicals inactivate *Trypanosoma cruzi* dihydrolipoamide dehydrogenase: enzyme protection by radical scavengers. *Free Radic Res* **37**: 281-291.

Hamada, M., Koike, K., Nakaula, Y., Hiraoka, T. and Koike, M. (1975). A kinetic study of the alpha-keto acid dehydrogenase complexes from pig heart mitochondria. *J Biochem (Tokyo)* **77**: 1047-1056.

Hanemaaijer, R., Westphal, A. H., Berg, A., Van Dongen, W., de Kok, A. and Veeger, C. (1989). The gene encoding dihydrolipoyl transacetylase from *Azotobacter vinelandii*. Expression in *Escherichia coli* and activation and isolation of the protein. *Eur J Biochem* **181**: 47-53.

Harris, R. A., Bowker-Kinley, M. M., Wu, P., Jeng, J. and Popov, K. M. (1997). Dihydrolipoamide dehydrogenase-binding protein of the human pyruvate dehydrogenase complex. DNA-derived amino acid sequence, expression, and reconstitution of the pyruvate dehydrogenase complex. *J Biol Chem* **272**: 19746-19751.

Harwaldt, P., Rahlfs, S. and Becker, K. (2002). Glutathione S-transferase of the malarial parasite *Plasmodium falciparum*: characterization of a potential drug target. *Biol Chem* **383**: 821-830.

Harwood, J. L. (1996). Recent advances in the biosynthesis of plant fatty acids. *Biochim Biophys Acta* **1301**: 7-56.

Hawkins, C. F., Borges, A. and Perham, R. N. (1989). A common structural motif in thiamin pyrophosphate-binding enzymes. *FEBS Lett* **255**: 77-82.

He, C. Y., Shaw, M. K., Pletcher, C. H., Striepen, B., Tilney, L. G. and Roos, D. S. (2001). A plastid segregation defect in the protozoan parasite *Toxoplasma gondii*. *Embo J* **20**: 330-339.

Henderson, C. E., Perham, R. N. and Finch, J. T. (1979). Structure and symmetry of *B. stearothermophilus* pyruvate dehydrogenase multienzyme complex and implications for eucaryote evolution. *Cell* **17**: 85-93.

Hirt, R. P., Muller, S., Embley, T. M. and Coombs, G. H. (2002). The diversity and evolution of thioredoxin reductase: new perspectives. *Trends Parasitol* **18**: 302-308.

Ho, L., Wexler, I. D., Kerr, D. S. and Patel, M. S. (1989). Genetic defects in human pyruvate dehydrogenase. *Ann N Y Acad Sci* **573**: 347-359.

Hodges, M., Yikilmaz, E., Patterson, G., Kasvosve, I., Rouault, T. A., Gordeuk, V. R. and Loyevsky, M. (2005). An iron regulatory-like protein expressed in *Plasmodium falciparum* displays aconitase activity. *Mol Biochem Parasitol*.

Holecek, M. (2002). Relation between glutamine, branched-chain amino acids, and protein metabolism. *Nutrition* **18**: 130-133.

Holmgren, A. (1977). Bovine thioredoxin system. Purification of thioredoxin reductase from calf liver and thymus and studies of its function in disulfide reduction. *J Biol Chem* **252**: 4600-4606.

Holmgren, A. (1979). Thioredoxin catalyzes the reduction of insulin disulfides by dithiothreitol and dihydrolipoamide. *J Biol Chem* **254**: 9627-9632.

Hong, Y. S., Kerr, D. S., Craigen, W. J., Tan, J., Pan, Y., Lusk, M. and Patel, M. S. (1996). Identification of two mutations in a compound heterozygous child with dihydrolipoamide dehydrogenase deficiency. *Hum Mol Genet* **5**: 1925-1930.

Hong, Y. S., Kerr, D. S., Liu, T. C., Lusk, M., Powell, B. R. and Patel, M. S. (1997). Deficiency of dihydrolipoamide dehydrogenase due to two mutant alleles (E340K and G101del). Analysis of a family and prenatal testing. *Biochim Biophys Acta* **1362**: 160-168.

Hopkins, J., Fowler, R., Krishna, S., Wilson, I., Mitchell, G. and Bannister, L. (1999). The plastid in *Plasmodium falciparum* asexual blood stages: a three-dimensional ultrastructural analysis. *Protist* **150**: 283-295.

Hoppe, H. C., Verschoor, J. A. and Louw, A. I. (1991). *Plasmodium falciparum*: a comparison of synchronisation methods for in vitro cultures. *Exp Parasitol* **72**: 464-467.

Huang, B., Gudi, R., Wu, P., Harris, R. A., Hamilton, J. and Popov, K. M. (1998). Isoenzymes of pyruvate dehydrogenase phosphatase. DNA-derived amino acid sequences, expression, and regulation. *J Biol Chem* **273**: 17680-17688.

Hyde, J. E. (2005). Exploring the folate pathway in *Plasmodium falciparum*. *Acta Trop* **94**: 191-206.

Igamberdiev, A. U., Bykova, N. V., Ens, W. and Hill, R. D. (2004). Dihydrolipoamide dehydrogenase from porcine heart catalyzes NADH-dependent scavenging of nitric oxide. *FEBS Lett* **568**: 146-150.

Jackson-Constan, D. and Keegstra, K. (2001). Arabidopsis genes encoding components of the chloroplastic protein import apparatus. *Plant Physiol* **125**: 1567-1576.

Jacob, J., Schirmer, R. H. and Gromer, S. (2005). The conserved histidine 106 of large thioredoxin reductases is likely to have a structural role but not a base catalyst function. *FEBS Lett* **579**: 745-748.

Jocelyn, P. C. (1967). The standard redox potential of cysteine-cystine from the thiol-disulphide exchange reaction with glutathione and lipoic acid. *Eur J Biochem* **2**: 327-331.

Johnson, D. C., Dean, D. R., Smith, A. D. and Johnson, M. K. (2005). Structure, function, and formation of biological iron-sulfur clusters. *Annu Rev Biochem* **74**: 247-281.

Johnson, M. T., Yang, H. S., Magnuson, T. and Patel, M. S. (1997). Targeted disruption of the murine dihydrolipoamide dehydrogenase gene (Dld) results in perigastrulation lethality. *Proc Natl Acad Sci U S A* **94**: 14512-14517.

Johnston, M. L., Luethy, M. H., Miernyk, J. A. and Randall, D. D. (1997). Cloning and molecular analyses of the *Arabidopsis thaliana* plastid pyruvate dehydrogenase subunits. *Biochim Biophys Acta* **1321**: 200-206.

Johnston, M. L., Miernyk, J. A. and Randall, D. D. (2000). Import, processing, and assembly of the alpha- and beta-subunits of chloroplast pyruvate dehydrogenase. *Planta* **211**: 72-76.

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**: 1573-1576.

Kabe, Y., Ando, K., Hirao, S., Yoshida, M. and Handa, H. (2005). Redox regulation of NF-kappaB activation: distinct redox regulation between the cytoplasm and the nucleus. *Antioxid Redox Signal* **7**: 395-403.

Kanzok, S. M., Schirmer, R. H., Turbachova, I., Iozef, R. and Becker, K. (2000). The thioredoxin system of the malaria parasite *Plasmodium falciparum*. Glutathione reduction revisited. *J Biol Chem* **275**: 40180-40186.

Kawai, S., Kano, S. and Suzuki, M. (1993). Morphologic effects of artemether on *Plasmodium falciparum* in *Aotus trivirgatus*. *Am J Trop Med Hyg* **49**: 812-818.

Ke, J., Behal, R. H., Back, S. L., Nikolau, B. J., Wurtele, E. S. and Oliver, D. J. (2000). The role of pyruvate dehydrogenase and acetyl-coenzyme A synthetase in fatty acid synthesis in developing *Arabidopsis* seeds. *Plant Physiol* **123**: 497-508.

Kessl, J. J., Lange, B. B., Merbitz-Zahradnik, T., Zwicker, K., Hill, P., Meunier, B., Palsdottir, H., Hunte, C., Meshnick, S. and Trumpower, B. L. (2003). Molecular basis for atovaquone binding to the cytochrome bc1 complex. *J Biol Chem* **278**: 31312-31318.

Kim, H., Liu, T. C. and Patel, M. S. (1991). Expression of cDNA sequences encoding mature and precursor forms of human dihydrolipoamide dehydrogenase in *Escherichia coli*. Differences in kinetic mechanisms. *J Biol Chem* **266**: 9367-9373.

Kim, H. and Patel, M. S. (1992). Characterization of two site-specifically mutated human dihydrolipoamide dehydrogenases (His-452----Gln and Glu-457----Gln). *J Biol Chem* **267**: 5128-5132.

Kohler, S., Delwiche, C. F., Denny, P. W., Tilney, L. G., Webster, P., Wilson, R. J., Palmer, J. D. and Roos, D. S. (1997). A plastid of probable green algal origin in Apicomplexan parasites. *Science* **275**: 1485-1489.

Koike, K. (1998). Cloning, structure, chromosomal localization and promoter analysis of human 2-oxoglutarate dehydrogenase gene. *Biochim Biophys Acta* **1385**: 373-384.

Komaki-Yasuda, K., Kawazu, S. and Kano, S. (2003). Disruption of the *Plasmodium falciparum* 2-Cys peroxiredoxin gene renders parasites hypersensitive to reactive oxygen and nitrogen species. *FEBS Lett* **547**: 140-144.

Koshkin, A., Knudsen, G. M. and Ortiz De Montellano, P. R. (2004). Intermolecular interactions in the AhpC/AhpD antioxidant defense system of *Mycobacterium tuberculosis*. *Arch Biochem Biophys* **427**: 41-47.

Krauth-Siegel, R. L., Bauer, H. and Schirmer, R. H. (2005). Dithiol proteins as guardians of the intracellular redox milieu in parasites: old and new drug targets in trypanosomes and malaria-causing plasmodia. *Angew Chem Int Ed Engl* **44**: 690-715.

Krauth-Siegel, R. L. and Schoneck, R. (1995). Flavoprotein structure and mechanism. 5. Trypanothione reductase and lipoamide dehydrogenase as targets for a structure-based drug design. *Faseb J* **9**: 1138-1146.

Kriek, N., Tilley, L., Horrocks, P., Pinches, R., Elford, B. C., Ferguson, D. J., Lingelbach, K. and Newbold, C. I. (2003). Characterization of the pathway for transport of the cytoadherence-mediating protein, PfEMP1, to the host cell surface in malaria parasite-infected erythrocytes. *Mol Microbiol* **50**: 1215-1227.

Krnajski, Z., Gilberger, T. W., Walter, R. D., Cowman, A. F. and Muller, S. (2002). Thioredoxin reductase is essential for the survival of *Plasmodium falciparum* erythrocytic stages. *J Biol Chem* **277**: 25970-25975.

Krnajski, Z., Gilberger, T. W., Walter, R. D. and Muller, S. (2001). The malaria parasite *Plasmodium falciparum* possesses a functional thioredoxin system. *Mol Biochem Parasitol* **112**: 219-228.

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

Krungkrai, J. (1995). Purification, characterization and localization of mitochondrial dihydroorotate dehydrogenase in *Plasmodium falciparum*, human malaria parasite. *Biochim Biophys Acta* **1243**: 351-360.

Krungkrai, J. (2004). The multiple roles of the mitochondrion of the malarial parasite. *Parasitology* **129**: 511-524.

Krungkrai, J., Kanchanarithisak, R., Krungkrai, S. R. and Rochanakij, S. (2002). Mitochondrial NADH dehydrogenase from *Plasmodium falciparum* and *Plasmodium berghei*. *Exp Parasitol* **100**: 54-61.

Krungkrai, J., Krungkrai, S. R. and Bhumiratana, A. (1993). *Plasmodium berghei*: partial purification and characterization of the mitochondrial cytochrome c oxidase. *Exp Parasitol* **77**: 136-146.

Krungkrai, J., Webster, H. K. and Yuthavong, Y. (1990). Folate and cobalamin metabolism in *Plasmodium falciparum*. *Parasitol Today* **6**: 388-391.

Kuhara, T., Shinka, T., Inoue, Y., Matsumoto, M., Yoshino, M., Sakaguchi, Y. and Matsumoto, I. (1983). Studies of urinary organic acid profiles of a patient with dihydrolipoyl dehydrogenase deficiency. *Clin Chim Acta* **133**: 133-140.

Kuriyan, J., Krishna, T. S., Wong, L., Guenther, B., Pahler, A., Williams, C. H., Jr. and Model, P. (1991). Convergent evolution of similar function in two structurally divergent enzymes. *Nature* **352**: 172-174.

Kyes, S., Horrocks, P. and Newbold, C. (2001). Antigenic variation at the infected red cell surface in malaria. *Annu Rev Microbiol* **55**: 673-707.

Lang-Unnasch, N. (1992). Purification and properties of *Plasmodium falciparum* malate dehydrogenase. *Mol Biochem Parasitol* **50**: 17-25.

Lang-Unnasch, N. (1995). *Plasmodium falciparum*: antiserum to malate dehydrogenase. *Exp Parasitol* **80**: 357-359.

Lantwin, C. B., Schlichting, I., Kabsch, W., Pai, E. F. and Krauth-Siegel, R. L. (1994). The structure of *Trypanosoma cruzi* trypanothione reductase in the oxidized and NADPH reduced state. *Proteins* **18**: 161-173.

Lawlis, V. B. and Roche, T. E. (1981). Inhibition of bovine kidney alpha-ketoglutarate dehydrogenase complex by reduced nicotinamide adenine dinucleotide in the presence or absence of calcium ion and effect of adenosine 5'-diphosphate on reduced nicotinamide adenine dinucleotide inhibition. *Biochemistry* **20**: 2519-2524.

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

Learngaramkul, P., Petmitr, S., Krungkrai, S. R., Prapunwattana, P. and Krungkrai, J. (1999). Molecular characterization of mitochondria in asexual and sexual blood stages of *Plasmodium falciparum*. *Mol Cell Biol Res Commun* **2**: 15-20.

Lee, S. R., Bar-Noy, S., Kwon, J., Levine, R. L., Stadtman, T. C. and Rhee, S. G. (2000). Mammalian thioredoxin reductase: oxidation of the C-terminal cysteine/selenocysteine active site forms a thioselenide, and replacement of selenium with sulfur markedly reduces catalytic activity. *Proc Natl Acad Sci U S A* **97**: 2521-2526.

Leed, A., DuBay, K., Ursos, L. M., Sears, D., De Dios, A. C. and Roepe, P. D. (2002). Solution structures of antimalarial drug-heme complexes. *Biochemistry* **41**: 10245-10255.

Leichus, B. N., Bradley, M., Nadeau, K., Walsh, C. T. and Blanchard, J. S. (1992). Kinetic isotope effect analysis of the reaction catalyzed by *Trypanosoma congolense* trypanothione reductase. *Biochemistry* **31**: 6414-6420.

Lell, B., Ruangweerayut, R., Wiesner, J., Missinou, M. A., Schindler, A., Baranek, T., Hintz, M., Hutchinson, D., Jomaa, H. and Kremsner, P. G. (2003). Fosmidomycin, a novel chemotherapeutic agent for malaria. *Antimicrob Agents Chemother* **47**: 735-738.

Lennon, B. W., Williams, C. H., Jr. and Ludwig, M. L. (2000). Twists in catalysis: alternating conformations of *Escherichia coli* thioredoxin reductase. *Science* **289**: 1190-1194.

Lill, R. and Muhlenhoff, U. (2005). Iron-sulfur-protein biogenesis in eukaryotes. *Trends Biochem Sci* **30**: 133-141.

Lindsay, H., Beaumont, E., Richards, S. D., Kelly, S. M., Sanderson, S. J., Price, N. C. and Lindsay, J. G. (2000). FAD insertion is essential for attaining the assembly competence of the dihydrolipoamide dehydrogenase (E3) monomer from *Escherichia coli*. *J Biol Chem* **275**: 36665-36670.

Liu, T. C., Kim, H., Arizmendi, C., Kitano, A. and Patel, M. S. (1993). Identification of two missense mutations in a dihydrolipoamide dehydrogenase-deficient patient. *Proc Natl Acad Sci U S A* **90**: 5186-5190.

Lohrer, H. and Krauth-Siegel, R. L. (1990). Purification and characterization of lipoamide dehydrogenase from *Trypanosoma cruzi*. *Eur J Biochem* **194**: 863-869.

Loria, P., Miller, S., Foley, M. and Tilley, L. (1999). Inhibition of the peroxidative degradation of haem as the basis of action of chloroquine and other quinoline antimalarials. *Biochem J* **339 (Pt 2)**: 363-370.

Luthman, M. and Holmgren, A. (1982). Rat liver thioredoxin and thioredoxin reductase: purification and characterization. *Biochemistry* **21**: 6628-6633.

Lutziger, I. and Oliver, D. J. (2000). Molecular evidence of a unique lipoamide dehydrogenase in plastids: analysis of plastidic lipoamide dehydrogenase from *Arabidopsis thaliana*. *FEBS Lett* **484**: 12-16.

Lutziger, I. and Oliver, D. J. (2001). Characterization of two cDNAs encoding mitochondrial lipoamide dehydrogenase from *Arabidopsis*. *Plant Physiol* **127**: 615-623.

Lydell, C. P., Chan, A., Wambolt, R. B., Sambandam, N., Parsons, H., Bondy, G. P., Rodrigues, B., Popov, K. M., Harris, R. A., Brownsey, R. W. and Allard, M. F. (2002). Pyruvate dehydrogenase and the regulation of glucose oxidation in hypertrophied rat hearts. *Cardiovasc Res* **53**: 841-851.

Macasev, D., Whelan, J., Newbigin, E., Silva-Filho, M. C., Mulhern, T. D. and Lithgow, T. (2004). Tom22', an 8-kDa trans-site receptor in plants and protozoans, is a conserved feature of the TOM complex that appeared early in the evolution of eukaryotes. *Mol Biol Evol* **21**: 1557-1564.

Maeng, C. Y., Yazdi, M. A., Niu, X. D., Lee, H. Y. and Reed, L. J. (1994). Expression, purification, and characterization of the dihydrolipoamide dehydrogenase-binding protein of the pyruvate dehydrogenase complex from *Saccharomyces cerevisiae*. *Biochemistry* **33**: 13801-13807.

Mannervik, B. (1973). A branching reaction mechanism of glutathione reductase. *Biochem Biophys Res Commun* **53**: 1151-1158.

Marsac, C., Stansbie, D., Bonne, G., Cousin, J., Jehenson, P., Benelli, C., Leroux, J. P. and Lindsay, G. (1993). Defect in the lipoyl-bearing protein X subunit of the pyruvate dehydrogenase complex in two patients with encephalomyelopathy. *J Pediatr* **123**: 915-920.

Martin, W., Stoebe, B., Goremykin, V., Hapsmann, S., Hasegawa, M. and Kowallik, K. V. (1998). Gene transfer to the nucleus and the evolution of chloroplasts. *Nature* **393**: 162-165.

Matsuzaki, M., Kikuchi, T., Kita, K., Kojima, S. and Kuroiwa, T. (2001). Large amounts of apicoplast nucleoid DNA and its segregation in *Toxoplasma gondii*. *Protoplasma* **218**: 180-191.

Mattevi, A., Obmolova, G., Sokatch, J. R., Betzel, C. and Hol, W. G. (1992). The refined crystal structure of *Pseudomonas putida* lipoamide dehydrogenase complexed with NAD+ at 2.45 A resolution. *Proteins* **13**: 336-351.

Mattevi, A., Schierbeek, A. J. and Hol, W. G. (1991). Refined crystal structure of lipoamide dehydrogenase from *Azotobacter vinelandii* at 2.2 A resolution. A comparison with the structure of glutathione reductase. *J Mol Biol* **220**: 975-994.

Matuda, S., Nakano, K., Ohta, S., Saheki, T., Kawanishi, Y. and Miyata, T. (1991). The alpha-ketoacid dehydrogenase complexes. Sequence similarity of rat pyruvate dehydrogenase with Escherichia coli and *Azotobacter vinelandii* alpha-ketoglutarate dehydrogenase. *Biochim Biophys Acta* **1089**: 1-7.

McCartney, R. G., Sanderson, S. J. and Lindsay, J. G. (1997). Refolding and reconstitution studies on the transacetylase-protein X (E2/X) subcomplex of the mammalian pyruvate dehydrogenase complex: evidence for specific binding of the dihydrolipoamide dehydrogenase component to sites on reassembled E2. *Biochemistry* **36**: 6819-6826.

McFadden, G. I. (2001). Chloroplast origin and integration. *Plant Physiol* **125**: 50-53.

McFadden, G. I. and Roos, D. S. (1999). Apicomplexan plastids as drug targets. *Trends Microbiol* **7**: 328-333.

McFadden, G. I. and Waller, R. F. (1997). Plastids in parasites of humans. *Bioessays* **19**: 1033-1040.

McLeod, R., Muench, S. P., Rafferty, J. B., Kyle, D. E., Mui, E. J., Kirisits, M. J., Mack, D. G., Roberts, C. W., Samuel, B. U., Lyons, R. E., Dorris, M., Milhous, W. K. and Rice, D. W. (2001). Triclosan inhibits the growth of *Plasmodium falciparum* and *Toxoplasma gondii* by inhibition of apicomplexan Fab I. *Int J Parasitol* **31**: 109-113.

Meng, M. and Chuang, D. T. (1994). Site-directed mutagenesis and functional analysis of the active-site residues of the E2 component of bovine branched-chain alpha-keto acid dehydrogenase complex. *Biochemistry* **33**: 12879-12885.

Meshnick, S. R. (1998). Artemisinin antimalarials: mechanisms of action and resistance. *Med Trop (Mars)* **58**: 13-17.

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

Miranda-Vizuete, A., Damdimopoulos, A. E. and Spyrou, G. (2000). The mitochondrial thioredoxin system. *Antioxid Redox Signal* **2**: 801-810.

Moini, H., Packer, L. and Saris, N. E. (2002a). Antioxidant and prooxidant activities of alpha-lipoic acid and dihydrolipoic acid. *Toxicol Appl Pharmacol* **182**: 84-90.

Moini, H., Tirosh, O., Park, Y. C., Cho, K. J. and Packer, L. (2002b). R-alpha-lipoic acid action on cell redox status, the insulin receptor, and glucose uptake in 3T3-L1 adipocytes. *Arch Biochem Biophys* **397**: 384-391.

Mooney, B. P., Miernyk, J. A. and Randall, D. D. (2002). The complex fate of alpha-ketoacids. *Annu Rev Plant Biol* **53**: 357-375.

Mu, J., Ferdig, M. T., Feng, X., Joy, D. A., Duan, J., Furuya, T., Subramanian, G., Aravind, L., Cooper, R. A., Wootton, J. C., Xiong, M. and Su, X. Z. (2003). Multiple transporters associated with malaria parasite responses to chloroquine and quinine. *Mol Microbiol* **49**: 977-989.

Muller, S. (2004). Redox and antioxidant systems of the malaria parasite *Plasmodium falciparum*. *Mol Microbiol* **53**: 1291-1305.

Mulrooney, S. B. and Williams, C. H., Jr. (1994). Potential active-site base of thioredoxin reductase from *Escherichia coli*: examination of histidine245 and aspartate139 by site-directed mutagenesis. *Biochemistry* **33**: 3148-3154.

Murphy, A. D., Doeller, J. E., Hearn, B. and Lang-Unnasch, N. (1997). *Plasmodium falciparum*: cyanide-resistant oxygen consumption. *Exp Parasitol* **87**: 112-120.

Nakamoto, H. and Bardwell, J. C. (2004). Catalysis of disulfide bond formation and isomerization in the *Escherichia coli* periplasm. *Biochim Biophys Acta* **1694**: 111-119.

Neagle, J., De Marcucci, O., Dunbar, B. and Lindsay, J. G. (1989). Component X of mammalian pyruvate dehydrogenase complex: structural and functional relationship to the lipoate acetyltransferase (E2) component. *FEBS Lett* **253**: 11-15.

Nielsen, H., Engelbrecht, J., Brunak, S. and von Heijne, G. (1997). Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng* **10**: 1-6.

Nzila, A., Ward, S. A., Marsh, K., Sims, P. F. and Hyde, J. E. (2005a). Comparative folate metabolism in humans and malaria parasites (part I): pointers for malaria treatment from cancer chemotherapy. *Trends Parasitol* **21**: 292-298.

Nzila, A., Ward, S. A., Marsh, K., Sims, P. F. and Hyde, J. E. (2005b). Comparative folate metabolism in humans and malaria parasites (part II): activities as yet untargeted or specific to *Plasmodium*. *Trends Parasitol* **21**: 334-339.

Odessey, R. (1982). Purification of rat kidney branched-chain oxo acid dehydrogenase complex with endogenous kinase activity. *Biochem J* **204**: 353-356.

Ono, K., Hakozaki, M., Suzuki, T., Mori, T., Hata, H. and Kochi, H. (2001). cDNA cloning of the chicken branched-chain alpha-keto acid dehydrogenase complex. Chicken-specific residues of the acyltransferase affect the overall activity and the interaction with the dehydrogenase. *Eur J Biochem* **268**: 727-736.

Packer, L., Witt, E. H. and Tritschler, H. J. (1995). alpha-Lipoic acid as a biological antioxidant. *Free Radic Biol Med* **19**: 227-250.

Packman, L. C., Perham, R. N. and Roberts, G. C. (1984). Domain structure and 1H-n.m.r. spectroscopy of the pyruvate dehydrogenase complex of *Bacillus stearothermophilus*. *Biochem J* **217**: 219-227.

Pai, E. F. and Schulz, G. E. (1983). The catalytic mechanism of glutathione reductase as derived from x-ray diffraction analyses of reaction intermediates. *J Biol Chem* **258**: 1752-1757.

Pancic, P. G. and Strotmann, H. (1993). Structure of the nuclear encoded gamma subunit of CF0CF1 of the diatom *Odontella sinensis* including its presequence. *FEBS Lett* **320**: 61-66.

Papworth, C., Bauer, J.C., Braman, J., and Wright, D.A. (1996). Site-directed mutagenesis in one day with greater than 80% efficiency. *Strategies* **9**: 3-4.

Patel, M., Kerr, D. S. and Wexler, I. D. (1992). Biochemical and molecular aspects of pyruvate dehydrogenase comlex deficiency. *Int Pediatr* **7**: 16 - 22.

Patel, M. P. and Blanchard, J. S. (1999). Expression, purification, and characterization of *Mycobacterium tuberculosis* mycothione reductase. *Biochemistry* **38**: 11827-11833.

Perham, R. N. (1991). Domains, motifs, and linkers in 2-oxo acid dehydrogenase multienzyme complexes: a paradigm in the design of a multifunctional protein. *Biochemistry* **30**: 8501-8512.

Perham, R. N. (2000). Swinging arms and swinging domains in multifunctional enzymes: catalytic machines for multistep reactions. *Annu Rev Biochem* **69**: 961-1004.

Perham, R. N., Leistler, B., Solomon, R. G. and Guptasarma, P. (1996). Protein engineering of domains in flavoprotein disulphide oxidoreductases: contributions to folding and assembly. *Biochem Soc Trans* **24**: 61-66.

Peterson, D. S., Walliker, D. and Wellems, T. E. (1988). Evidence that a point mutation in dihydrofolate reductase-thymidylate synthase confers resistance to pyrimethamine in *falciparum* malaria. *Proc Natl Acad Sci U S A* **85**: 9114-9118.

Pettit, F. H., Hamilton, L., Munk, P., Namihira, G., Eley, M. H., Willms, C. R. and Reed, L. J. (1973). Alpha-keto acid dehydrogenase complexes. XIX. Subunit structure of the *Escherichia coli* alpha-ketoglutarate dehydrogenase complex. *J Biol Chem* **248**: 5282-5290.

Pfanner, N. and Geissler, A. (2001). Versatility of the mitochondrial protein import machinery. *Nat Rev Mol Cell Biol* **2**: 339-349.

Pfanner, N., Hoeben, P., Tropschug, M. and Neupert, W. (1987). The carboxyl-terminal two-thirds of the ADP/ATP carrier polypeptide contains sufficient information to direct translocation into mitochondria. *J Biol Chem* **262**: 14851-14854.

Pick, U., Haramaki, N., Constantinescu, A., Handelman, G. J., Tritschler, H. J. and Packer, L. (1995). Glutathione reductase and lipoamide dehydrogenase have opposite stereospecificities for alpha-lipoic acid enantiomers. *Biochem Biophys Res Commun* **206**: 724-730.

Popov, K. M., Kedishvili, N. Y., Zhao, Y., Shimomura, Y., Crabb, D. W. and Harris, R. A. (1993). Primary structure of pyruvate dehydrogenase kinase establishes a new family of eukaryotic protein kinases. *J Biol Chem* **268**: 26602-26606.

Popov, K. M., Zhao, Y., Shimomura, Y., Kuntz, M. J. and Harris, R. A. (1992). Branched-chain alpha-ketoacid dehydrogenase kinase. Molecular cloning, expression, and sequence similarity with histidine protein kinases. *J Biol Chem* **267**: 13127-13130.

Powis, G. and Montfort, W. R. (2001). Properties and biological activities of thioredoxins. *Annu Rev Biophys Biomol Struct* **30**: 421-455.

Prensier, G. and Slomianny, C. (1986). The karyotype of *Plasmodium falciparum* determined by ultrastructural serial sectioning and 3D reconstruction. *J Parasitol* **72**: 731-736.

Press, W. H., Teukolsky, S.A., Vetterling, W.T., and Flannery, B.P. (1992). Numerical recipes in C, the art of scientific writing. 2nd Edition., 683-688, Cambridge University Press, New York.

Przyborski, J. M. and Lanzer, M. (2005). Protein transport and trafficking in *Plasmodium falciparum*-infected erythrocytes. *Parasitology* **130**: 373-388.

Rakauskiene, G. A., Cenas, N. K. and Kulys, J. J. (1989). A 'branched' mechanism of the reverse reaction of yeast glutathione reductase. An estimation of the enzyme standard potential values from the steady-state kinetics data. *FEBS Lett* **243**: 33-36.

Ralph, S. A., D'Ombrain, M. C. and McFadden, G. I. (2001). The apicoplast as an antimalarial drug target. *Drug Resist Updat* **4**: 145-151.

Ralph, S. A., van Dooren, G. G., Waller, R. F., Crawford, M. J., Fraunholz, M. J., Foth, B. J., Tonkin, C. J., Roos, D. S. and McFadden, G. I. (2004). Tropical infectious diseases: metabolic maps and functions of the *Plasmodium falciparum* apicoplast. *Nat Rev Microbiol* **2**: 203-216.

Randall, D. D., Miernyk, J. A., Fang, T. K., Budde, R. J. and Schuller, K. A. (1989). Regulation of the pyruvate dehydrogenase complexes in plants. *Ann N Y Acad Sci* **573**: 192-205.

Randall, D. D., Rubin, P. M. and Fenko, M. (1977). Plant pyruvate dehydrogenase complex purification, characterization and regulation by metabolites and phosphorylation. *Biochim Biophys Acta* **485**: 336-349.

Reed, L. J., Koike, M., Levitch, M. E. and Leach, F. R. (1958). Studies on the nature and reactions of protein-bound lipoic acid. *J Biol Chem* **232**: 143-158.

Reeder, J. C., Rieckmann, K. H., Genton, B., Lorry, K., Wines, B. and Cowman, A. F. (1996). Point mutations in the dihydrofolate reductase and dihydropteroate synthetase genes and in vitro susceptibility to pyrimethamine and cycloguanil of *Plasmodium falciparum* isolates from Papua New Guinea. *Am J Trop Med Hyg* **55**: 209-213.

Reid, E. E., Lyttle, C. R., Canvin, D. T. and Dennis, D. T. (1975). Pyruvate dehydrogenase complex activity in proplastids and mitochondria of developing castor bean endosperm. *Biochem Biophys Res Commun* **62**: 42-47.

Richter, S. and Lamppa, G. K. (1998). A chloroplast processing enzyme functions as the general stromal processing peptidase. *Proc Natl Acad Sci U S A* **95**: 7463-7468.

Rietveld, P., Arscott, L. D., Berry, A., Scrutton, N. S., Deonarain, M. P., Perham, R. N. and Williams, C. H., Jr. (1994). Reductive and oxidative half-reactions of glutathione reductase from *Escherichia coli*. *Biochemistry* **33**: 13888-13895.

Robinson, B. H. (1995). Lactic acidemia: disorders of pyruvate carboxylase, pyruvate dehydrogenase. New York, McGraw-Hill.

Roche, T. E., Baker, J. C., Yan, X., Hiromasa, Y., Gong, X., Peng, T., Dong, J., Turkan, A. and Kasten, S. A. (2001). Distinct regulatory properties of pyruvate dehydrogenase kinase and phosphatase isoforms. *Prog Nucleic Acid Res Mol Biol* **70**: 33-75.

Rudzinska, M. A. (1969). The fine structure of malaria parasites. *Int Rev Cytol* **25**: 161-199.

Rujan, T. and Martin, W. (2001). How many genes in *Arabidopsis* come from cyanobacteria? An estimate from 386 protein phylogenies. *Trends Genet* **17**: 113-120.

Russell, G. C., Machado, R. S. and Guest, J. R. (1992). Overproduction of the pyruvate dehydrogenase multienzyme complex of *Escherichia coli* and site-directed substitutions in the E1p and E2p subunits. *Biochem J* **287 (Pt 2)**: 611-619.

Sahlman, L. and Williams, C. H., Jr. (1989). Lipoamide dehydrogenase from *Escherichia coli*. Steady-state kinetics of the physiological reaction. *J Biol Chem* **264**: 8039-8045.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). Molecular Cloning: a Laboratory Manual. New york, Cold Spring Harbor.

Sandalova, T., Zhong, L., Lindqvist, Y., Holmgren, A. and Schneider, G. (2001). Three-dimensional structure of a mammalian thioredoxin reductase: implications for mechanism and evolution of a selenocysteine-dependent enzyme. *Proc Natl Acad Sci U S A* **98**: 9533-9538.

Sanderson, S. J., Miller, C. and Lindsay, J. G. (1996). Stoichiometry, organisation and catalytic function of protein X of the pyruvate dehydrogenase complex from bovine heart. *Eur J Biochem* **236**: 68-77.

Sato, S., Clough, B., Coates, L. and Wilson, R. J. (2004). Enzymes for heme biosynthesis are found in both the mitochondrion and plastid of the malaria parasite *Plasmodium falciparum*. *Protist* **155**: 117-125.

Sato, S., Rangachari, K. and Wilson, R. J. (2003). Targeting GFP to the malarial mitochondrion. *Mol Biochem Parasitol* **130**: 155-158.

Sato, S. and Wilson, R. J. (2005). Organelle-specific cochaperonins in apicomplexan parasites. *Mol Biochem Parasitol* **141**: 133-143.

Schatz, G. and Dobberstein, B. (1996). Common principles of protein translocation across membranes. *Science* **271**: 1519-1526.

Scheibel, L. W. and Pflaum, W. K. (1970). Carbohydrate metabolism in *Plasmodium knowlesi*. *Comp. Biochem. Physiol.* **37**: 543 - 553.

Schierbeek, A. J., Swarte, M. B., Dijkstra, B. W., Vriend, G., Read, R. J., Hol, W. G., Drenth, J. and Betzel, C. (1989). X-ray structure of lipoamide dehydrogenase from *Azotobacter vinelandii* determined by a combination of molecular and isomorphous replacement techniques. *J Mol Biol* **206**: 365-379.

Schoneck, R., Billaut-Mulot, O., Numrich, P., Ouaissi, M. A. and Krauth-Siegel, R. L. (1997). Cloning, sequencing and functional expression of dihydrolipoamide dehydrogenase from the human pathogen *Trypanosoma cruzi*. *Eur J Biochem* **243**: 739-747.

Schulz, G. E. and Schirmer, H. (1990). Principles of protein structure. New York, Springer.

Schulze, E., Westphal, A. H., Veenhuis, M. and de Kok, A. (1992). Purification and cellular localization of wild type and mutated dihydrolipoyltransacetylases from *Azotobacter vinelandii* and *Escherichia coli* expressed in *E. coli*. *Biochim Biophys Acta* **1120**: 87-96.

Scrutton, N. S., Berry, A., Deonarain, M. P. and Perham, R. N. (1990). Active site complementation in engineered heterodimers of *Escherichia coli* glutathione reductase created in vivo. *Proc Biol Sci* **242**: 217-224.

Shaag, A., Saada, A., Berger, I., Mandel, H., Joseph, A., Feigenbaum, A. and Elpeleg, O. N. (1999). Molecular basis of lipoamide dehydrogenase deficiency in Ashkenazi Jews. *Am J Med Genet* **82**: 177-182.

Shany, E., Saada, A., Landau, D., Shaag, A., Hershkovitz, E. and Elpeleg, O. N. (1999). Lipoamide dehydrogenase deficiency due to a novel mutation in the interface domain. *Biochem Biophys Res Commun* **262**: 163-166.

Sherman, I. W. (1998). A brief history of malaria and discovery of the parasites life cycle. Washington, D.C., ASM Press.

Sienkiewicz, N., Daher, W., Dive, D., Wrenger, C., Viscogliosi, E., Wintjens, R., Jouin, H., Capron, M., Muller, S. and Khalife, J. (2004). Identification of a mitochondrial superoxide dismutase with an unusual targeting sequence in *Plasmodium falciparum*. *Mol Biochem Parasitol* **137**: 121-132.

Sinden, R. E. (1998). Gametocytes and sexual development. Washington, D.C., ASM Press.

Snow, R. W., Guerra, C. A., Noor, A. M., Myint, H. Y. and Hay, S. I. (2005). The global distribution of clinical episodes of *Plasmodium falciparum* malaria. *Nature* **434**: 214-217.

Sreider, C. M., Grinblat, L. and Stoppani, A. O. (1992). Reduction of nitrofuran compounds by heart lipoamide dehydrogenase: role of flavin and the reactive disulfide groups. *Biochem Int* **28**: 323-334.

Srivastava, I. K., Rottenberg, H. and Vaidya, A. B. (1997). Atovaquone, a broad spectrum antiparasitic drug, collapses mitochondrial membrane potential in a malarial parasite. *J Biol Chem* **272**: 3961-3966.

Stan, T., Brix, J., Schneider-Mergener, J., Pfanner, N., Neupert, W. and Rapaport, D. (2003). Mitochondrial protein import: recognition of internal import signals of BCS1 by the TOM complex. *Mol Cell Biol* **23**: 2239-2250.

Stephens, P. E., Darlison, M. G., Lewis, H. M. and Guest, J. R. (1983). The pyruvate dehydrogenase complex of *Escherichia coli* K12. Nucleotide sequence encoding the pyruvate dehydrogenase component. *Eur J Biochem* **133**: 155-162.

Striepen, B., Crawford, M. J., Shaw, M. K., Tilney, L. G., Seeber, F. and Roos, D. S. (2000). The plastid of *Toxoplasma gondii* is divided by association with the centrosomes. *J Cell Biol* **151**: 1423-1434.

Sugden, M. C. and Holness, M. J. (2003). Recent advances in mechanisms regulating glucose oxidation at the level of the pyruvate dehydrogenase complex by PDKs. *Am J Physiol Endocrinol Metab* **284**: E855-862.

Suraveratum, N., Krungkrai, S. R., Leangaramgul, P., Prapunwattana, P. and Krungkrai, J. (2000). Purification and characterization of *Plasmodium falciparum* succinate dehydrogenase. *Mol Biochem Parasitol* **105**: 215-222.

Surolia, A., Ramya, T. N., Ramya, V. and Surolia, N. (2004). 'FAS't inhibition of malaria. *Biochem J* **383**: 401-412.

Surolia, N. and Surolia, A. (2001). Triclosan offers protection against blood stages of malaria by inhibiting enoyl-ACP reductase of *Plasmodium falciparum*. *Nat Med* **7**: 167-173.

Szafranski, K., Lehmann, R., Parra, G., Guigo, R. and Glockner, G. (2005). Gene organization features in A/T-rich organisms. *J Mol Evol* **60**: 90-98.

Takashima, E., Takamiya, S., Takeo, S., Mi-ichi, F., Amino, H. and Kita, K. (2001). Isolation of mitochondria from *Plasmodium falciparum* showing dihydroorotate dependent respiration. *Parasitol Int* **50**: 273-278.

Takeo, S., Kokaze, A., Ng, C. S., Mizuchi, D., Watanabe, J. I., Tanabe, K., Kojima, S. and Kita, K. (2000). Succinate dehydrogenase in *Plasmodium falciparum* mitochondria: molecular characterization of the SDHA and SDHB genes for the catalytic subunits, the flavoprotein (Fp) and iron-sulfur (Ip) subunits. *Mol Biochem Parasitol* **107**: 191-205.

Thekkumkara, T. J., Ho, L., Wexler, I. D., Pons, G., Liu, T. C. and Patel, M. S. (1988). Nucleotide sequence of a cDNA for the dihydrolipoamide acetyltransferase component of human pyruvate dehydrogenase complex. *FEBS Lett* **240**: 45-48.

Thieme, R., Pai, E. F., Schirmer, R. H. and Schulz, G. E. (1981). Three-dimensional structure of glutathione reductase at 2 A resolution. *J Mol Biol* **152**: 763-782.

Thomsen-Zieger, N., Schachtner, J. and Seeber, F. (2003). Apicomplexan parasites contain a single lipoic acid synthase located in the plastid. *FEBS Lett* **547**: 80-86.

Thorpe, C. and Williams, C. H., Jr. (1976). Differential reactivity of the two active site cysteine residues generated on reduction of pig heart lipoamide dehydrogenase. *J Biol Chem* **251**: 3553-3557.

Tian, J., Bryk, R., Shi, S., Erdjument-Bromage, H., Tempst, P. and Nathan, C. (2005). *Mycobacterium tuberculosis* appears to lack alpha-ketoglutarate dehydrogenase and encodes pyruvate dehydrogenase in widely separated genes. *Mol Microbiol* **57**: 859-868.

Tonkin, C. J., van Dooren, G. G., Spurck, T. P., Struck, N. S., Good, R. T., Handman, E., Cowman, A. F. and McFadden, G. I. (2004). Localization of organellar proteins in *Plasmodium falciparum* using a novel set of transfection vectors and a new immunofluorescence fixation method. *Mol Biochem Parasitol* **137**: 13-21.

Tovar-Mendez, A., Miernyk, J. A. and Randall, D. D. (2003). Regulation of pyruvate dehydrogenase complex activity in plant cells. *Eur J Biochem* **270**: 1043-1049.

Toyoda, T., Suzuki, K., Sekiguchi, T., Reed, L. J. and Takenaka, A. (1998). Crystal structure of eukaryotic E3, lipoamide dehydrogenase from yeast. *J Biochem (Tokyo)* **123**: 668-674.

Trager, W. and Jensen, J. B. (1976). Human malaria parasites in continuous culture. *Science* **193**: 673-675.

Trenholme, C. M., Williams, R. L., Desjardins, R. E., Frischer, H., Carson, P. E., Rieckmann, K. H. and Canfield, C. J. (1975). Mefloquine (WR 142,490) in the treatment of human malaria. *Science* **190**: 792-794.

Tripathi, A. K., Desai, P. V., Pradhan, A., Khan, S. I., Avery, M. A., Walker, L. A. and Tekwani, B. L. (2004). An alpha-proteobacterial type malate dehydrogenase may complement LDH function in *Plasmodium falciparum*. Cloning and biochemical characterization of the enzyme. *Eur J Biochem* **271**: 3488-3502.

Truscott, K. N., Brandner, K. and Pfanner, N. (2003). Mechanisms of protein import into mitochondria. *Curr Biol* **13**: R326-337.

Truscott, K. N., Pfanner, N. and Voos, W. (2001). Transport of proteins into mitochondria. *Rev Physiol Biochem Pharmacol* **143**: 81-136.

Tsai, C. S. (1980). Kinetic studies of multifunctional reactions catalysed by lipoamide dehydrogenase. *Int J Biochem* **11**: 407-413.

Tsai, C. S. and Wand, A. J. (1992). pH dependent kinetic studies of lipoamide dehydrogenase catalysis. *Int J Biochem* **24**: 1801-1806.

Umlas, J. and Fallon, J. N. (1971). New thick-film technique for malaria diagnosis. Use of saponin stromatolytic solution for lysis. *Am J Trop Med Hyg* **20**: 527-529.

Uyemura, S. A., Luo, S., Moreno, S. N. and Docampo, R. (2000). Oxidative phosphorylation, Ca(2+) transport, and fatty acid-induced uncoupling in malaria parasites mitochondria. *J Biol Chem* **275**: 9709-9715.

Uyemura, S. A., Luo, S., Vieira, M., Moreno, S. N. and Docampo, R. (2004). Oxidative phosphorylation and rotenone-insensitive malate- and NADH-quinone oxidoreductases in *Plasmodium yoelii yoelii* mitochondria in situ. *J Biol Chem* **279**: 385-393.

Vaidya, A. B., Lashgari, M. S., Pologe, L. G. and Morrisey, J. (1993). Structural features of *Plasmodium* cytochrome b that may underlie susceptibility to 8-aminoquinolines and hydroxynaphthoquinones. *Mol Biochem Parasitol* **58**: 33-42.

van Dooren, G. G., Marti, M., Tonkin, C. J., Stimmler, L. M., Cowman, A. F. and McFadden, G. I. (2005). Development of the endoplasmic reticulum, mitochondrion and apicoplast during the asexual life cycle of *Plasmodium falciparum*. *Mol Microbiol* **57**: 405-419.

van Dooren, G. G., Schwartzbach, S. D., Osafune, T. and McFadden, G. I. (2001). Translocation of proteins across the multiple membranes of complex plastids. *Biochim Biophys Acta* **1541**: 34-53.

van Dooren, G. G., Su, V., D'Ombrain, M. C. and McFadden, G. I. (2002). Processing of an apicoplast leader sequence in *Plasmodium falciparum* and the identification of a putative leader cleavage enzyme. *J Biol Chem* **277**: 23612-23619.

van Lin, L. H., Pace, T., Janse, C. J., Birago, C., Ramesar, J., Picci, L., Ponzi, M. and Waters, A. P. (2001). Interspecies conservation of gene order and intron-exon structure in a genomic locus of high gene density and complexity in *Plasmodium*. *Nucleic Acids Res* **29**: 2059-2068.

Varadharajan, S., Sagar, B. K., Rangarajan, P. N. and Padmanaban, G. (2004). Localization of ferrochelatase in *Plasmodium falciparum*. *Biochem J* **384**: 429-436.

Volkman, S. and Wirth, D. (1998). Functional analysis of pfmdr1 gene of *Plasmodium falciparum*. *Methods Enzymol* **292**: 174-181.

von Heijne, G., Steppuhn, J. and Herrmann, R. G. (1989). Domain structure of mitochondrial and chloroplast targeting peptides. *Eur J Biochem* **180**: 535-545.

Vothknecht, U. C. and Soll, J. (2000). Protein import: the hitchhikers guide into chloroplasts. *Biol Chem* **381**: 887-897.

Wagner, J. T., Ludemann, H., Farber, P. M., Lottspeich, F. and Krauth-Siegel, R. L. (1998). Glutamate dehydrogenase, the marker protein of *Plasmodium falciparum*--cloning, expression and characterization of the malarial enzyme. *Eur J Biochem* **258**: 813-819.

Wahlgren, M., Barragan, A., Chen, Q., Fernandez, V., Hagblom, P., Heddini, A., Schlichtherle, M., Scholander, C., Sundstrom, A., Treutiger, C. J., von Euller, A. and Carlson, J. (1998). Rosetting of malaria-infected erythrocytes: Ligands, host receptors, and the involvement of serum proteins. Washington, D.C., ASM Press.

Waller, R. F., Keeling, P. J., Donald, R. G., Striepen, B., Handman, E., Lang-Unnasch, N., Cowman, A. F., Besra, G. S., Roos, D. S. and McFadden, G. I. (1998). Nuclear-encoded proteins target to the plastid in Toxoplasma gondii and *Plasmodium falciparum*. *Proc Natl Acad Sci U S A* **95**: 12352-12357.

Waller, R. F., Ralph, S. A., Reed, M. B., Su, V., Douglas, J. D., Minnikin, D. E., Cowman, A. F., Besra, G. S. and McFadden, G. I. (2003). A type II pathway for fatty acid biosynthesis presents drug targets in *Plasmodium falciparum*. *Antimicrob Agents Chemother* **47**: 297-301.

Waller, R. F., Reed, M. B., Cowman, A. F. and McFadden, G. I. (2000). Protein trafficking to the plastid of *Plasmodium falciparum* is via the secretory pathway. *Embo J* **19**: 1794-1802.

Wang, P. F., Arscott, L. D., Gilberger, T. W., Muller, S. and Williams, C. H., Jr. (1999). Thioredoxin reductase from *Plasmodium falciparum*: evidence for interaction between the C-terminal cysteine residues and the active site disulfide-dithiol. *Biochemistry* **38**: 3187-3196.

Wang, X., Connor, M., Smith, R., Maciejewski, M. W., Howden, M. E., Nicholson, G. M., Christie, M. J. and King, G. F. (2000). Discovery and characterization of a family of insecticidal neurotoxins with a rare vicinal disulfide bridge. *Nat Struct Biol* **7**: 505-513.

Ward, D. E., Ross, R. P., van der Weijden, C. C., Snoep, J. L. and Claiborne, A. (1999). Catabolism of branched-chain alpha-keto acids in *Enterococcus faecalis*: the bkd gene cluster, enzymes, and metabolic route. *J Bacteriol* **181**: 5433-5442.

Werner, C., Stubbs, M. T., Krauth-Siegel, R. L. and Klebe, G. (2005). The crystal structure of *Plasmodium falciparum* glutamate dehydrogenase, a putative target for novel antimalarial drugs. *J Mol Biol* **349**: 597-607.

Wexler, I. D., Hemalatha, S. G. and Patel, M. S. (1991). Sequence conservation in the alpha and beta subunits of pyruvate dehydrogenase and its similarity to branched-chain alpha-keto acid dehydrogenase. *FEBS Lett* **282**: 209-213.

White, N. J. (1998). Malaria pathophysiology. Washington, D.C., ASM Press.

White, S. (1996). Malaria. London, W.B. Saunders company.

Williams, C., Jr (1976). Flavin containing dehydrogenases. New York, Academic press.

Williams, C. H., Arscott, L. D., Muller, S., Lennon, B. W., Ludwig, M. L., Wang, P. F., Veine, D. M., Becker, K. and Schirmer, R. H. (2000). Thioredoxin reductase two modes of catalysis have evolved. *Eur J Biochem* **267**: 6110-6117.

Williams, C. H., Jr. (1992). Lipoamide dehydrogenase, glutathione reductase, thioredoxin reductase, and mercuric ion reductase - A family of flavoenzyme transhydrogenases. Boca Raton, CRC Press.

Williams, C. H., Jr., Allison, N., Russell, G. C., Prongay, A. J., Arscott, L. D., Datta, S., Sahlman, L. and Guest, J. R. (1989). Properties of lipoamide dehydrogenase and thioredoxin reductase from *Escherichia coli* altered by site-directed mutagenesis. *Ann N Y Acad Sci* **573**: 55-65.

Williams, C. H., Jr., Arscott, L. D. and Schulz, G. E. (1982). Amino acid sequence homology between pig heart lipoamide dehydrogenase and human erythrocyte glutathione reductase. *Proc Natl Acad Sci U S A* **79**: 2199-2201.

Williamson, D. H., Denny, P. W., Moore, P. W., Sato, S., McCready, S. and Wilson, R. J. (2001). The in vivo conformation of the plastid DNA of *Toxoplasma gondii*: implications for replication. *J Mol Biol* **306**: 159-168.

Williamson, D. H., Gardner, M. J., Preiser, P., Moore, D. J., Rangachari, K. and Wilson, R. J. (1994). The evolutionary origin of the 35 kb circular DNA of *Plasmodium falciparum*: new evidence supports a possible rhodophyte ancestry. *Mol Gen Genet* **243**: 249-252.

Wilson, R. J., Denny, P. W., Preiser, P. R., Rangachari, K., Roberts, K., Roy, A., Whyte, A., Strath, M., Moore, D. J., Moore, P. W. and Williamson, D. H. (1996). Complete gene map of the plastid-like DNA of the malaria parasite *Plasmodium falciparum*. *J Mol Biol* **261**: 155-172.

Wilson, R. J., Williamson, D. H. and Preiser, P. (1994). Malaria and other Apicomplexans: the "plant" connection. *Infect Agents Dis* **3**: 29-37.

Wrenger, C. and Muller, S. (2003). Isocitrate dehydrogenase of *Plasmodium falciparum*. *Eur J Biochem* **270**: 1775-1783.

Wrenger, C. and Muller, S. (2004). The human malaria parasite *Plasmodium falciparum* has distinct organelle-specific lipoylation pathways. *Mol Microbiol* **53**: 103-113.

Wu, Y., Sifri, C. D., Lei, H. H., Su, X. Z. and Wellems, T. E. (1995). Transfection of *Plasmodium falciparum* within human red blood cells. *Proc Natl Acad Sci U S A* **92**: 973-977.

Xia, L., Bjornstedt, M., Nordman, T., Eriksson, L. C. and Olsson, J. M. (2001). Reduction of ubiquinone by lipoamide dehydrogenase. An antioxidant regenerating pathway. *Eur J Biochem* **268**: 1486-1490.

Yang, D., Song, J., Wagenknecht, T. and Roche, T. E. (1997). Assembly and full functionality of recombinantly expressed dihydrolipoyl acetyltransferase component of the human pyruvate dehydrogenase complex. *J Biol Chem* **272**: 6361-6369.

Yano, K., Komaki-Yasuda, K., Kobayashi, T., Takemae, H., Kita, K., Kano, S. and Kawazu, S. (2005). Expression of mRNAs and proteins for peroxiredoxins in *Plasmodium falciparum* erythrocytic stage. *Parasitol Int* **54**: 35-41.

Yap, M. W., Kara, U. A., ten Heggeler-Bordier, B., Ting, R. C. and Tan, T. M. (1997). Partial nucleotide sequence and organisation of extrachromosomal plastid-like DNA in *Plasmodium berghei*. *Gene* **200**: 91-98.

Yeaman, S. J., Hutcheson, E. T., Roche, T. E., Pettit, F. H., Brown, J. R., Reed, L. J., Watson, D. C. and Dixon, G. H. (1978). Sites of phosphorylation on pyruvate dehydrogenase from bovine kidney and heart. *Biochemistry* **17**: 2364-2370.

Zhang, J., Krugliak, M. and Ginsburg, H. (1999). The fate of ferriprotorphyrin IX in malaria infected erythrocytes in conjunction with the mode of action of antimalarial drugs. *Mol Biochem Parasitol* **99**: 129-141.

Zhao, Y., Hanton, W. K. and Lee, K. H. (1986). Antimalarial agents, 2. Artesunate, an inhibitor of cytochrome oxidase activity in *Plasmodium berghei*. *J Nat Prod* **49**: 139-142.

Zhong, L. and Holmgren, A. (2000). Essential role of selenium in the catalytic activities of mammalian thioredoxin reductase revealed by characterization of recombinant enzymes with selenocysteine mutations. *J Biol Chem* **275**: 18121-18128.

Zhu, G. and Keithly, J. S. (2002). Alpha-proteobacterial relationship of apicomplexan lactate and malate dehydrogenases. *J Eukaryot Microbiol* **49**: 255-261.

"Here with a Loaf of Bread beneath the Bough,
A Flask of Wine, a Book of Verse - and Thou
Beside me singing in the Wilderness -
And Wilderness is Paradise now."

Poem by Edward Fitzgerald